Counting rational points over function fields

by

Jakob Glas

September, 2024

A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Committee in charge:

Carrie Bernecky, Chair
Timothy Browning
Pankaj Vishe
Tamas Hausel

Institute of Science and Technology Austria

The thesis of Jakob Glas, titled *Counting rational points over function fields*, is approved by:

Supervisor: Timothy Browning, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Committee Member: Pankaj Vishe, Durham University, Durham, United Kingdom

Signature: ______________________________

Committee Member: Tamas Hausel, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Defense Chair: Carrie Bernecky, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Signed page is on file

ISTA Thesis, ISSN: 2663-337X

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I accept full responsibility for the content and factual accuracy of this work, including the data and their analysis and presentation, and the text and citation of other work.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: ______________________________

Jakob Glas
September, 2024

Signed page is on file

Abstract

In this thesis, we are dealing with both arithmetic and geometric problems coming from the study of rational points with a particular focus on function fields over finite fields:

(1) Using the circle method we produce upper bounds for the number of rational points of bounded height on diagonal cubic surfaces and fourfolds over $\mathbb{F}_q(t)$. This is based on joint work with Leonhard Hochfilzer.

(2) We study rational points on smooth complete intersections X defined by cubic and quadratic hypersurfaces over $\mathbb{F}_q(t)$. We refine the Farey dissection of the "unit square" developed by Vishe [202] and use the circle method with a Kloosterman refinement to establish an asymptotic formula for the number of rational points of bounded height on X when $\dim(X) \geq 23$. Under the same hypotheses, we also verify weak approximation.

(3) In joint work with Hochfilzer, we obtain upper bounds for the number of rational points of bounded height on del Pezzo surfaces of low degree over any global field. Our approach is to take hyperplane sections, which reduces the problem to uniform estimates for the number of rational points on curves.

(4) We develop a version of the circle method capable of counting $\mathbb{F}_q$-points on jet schemes of moduli spaces of rational curves on hypersurfaces. Combining this with a spreading out argument and a result of Mustaţă [150], this allows us to show that these moduli spaces only have canonical singularities under suitable assumptions on the degree and the dimension.

In addition, we give an overview of guiding questions and conjectures in the field of rational points and explain the basic mechanism underlying the circle method.

Acknowledgements

The past four years have been an incredibly joyful journey and this was only possible because of the people that were part of it.

First and foremost, I want to express my deepest gratitude to my supervisor Tim Browning: for being so generous with his time and ideas; for sharing his excellent taste in mathematics; and for being a teacher, a mentor and a friend over the past four years. Without his help and support this thesis would not have been possible.

I was fortunate to always be part of an amazing working group. I want to thank Margaret, Dante, Tal, Zhizhong, Julian, Alec and Florian as well as Stephanie, Yijie, Loïs, Nick, Matteo, Victor, Lena and Shuntaro for all the fun times and the mathematics I learnt from them.

During my PhD, I had the pleasure to collaborate with many brilliant people. Special thanks to Leo, Tim, Victor, Kevin, Sam, Rosa and Matthew. Their insights and cooperation have significantly enriched my work.

In the last months of 2022, I spent several weeks in Göttingen and Bath. I am grateful to Damaris and Dan for hosting me. In addition, I want to thank all the people in your groups that made my stays so enjoyable. In particular, thanks to Leo and Julian for all the lunches, dinners and beers.

I also want to express my appreciation to Shuntaro Yamagishi for carefully reading parts of my thesis. In addition, I want to thank Tamas Hausel and Pankaj Vishe for being on my committee and for helpful comments on my thesis.

For Chapter 5, I want to thank Damaris Schindler for helpful comments. I would also like to thank Efthymios Sofos for bringing Davenport's question to Leo's and my attention and Keith Matthews for providing us with scanned copies of the original correspondence. For Chapter 6, I would like to thank Pankaj Vishe for useful comments and conversations. Moreover, I am grateful to Dante Bonolis and Julian Lyczak for sharing their expertise in exponential sums and geometry. For Chapter 7 would like to thank Dan Loughran for helpful comments. For Chapter 8, I would like to thank Matthew Hase-Liu for helpful comments and discussions.

Outside of my mathematical life I want to thank all my friends for all the time we spent together and the experiences we share. In addition, I am indebted to my partner Martina for being such a wonderful part of my life. Lastly, I want to thank my family, Peter, Petra and Stefan, for always supporting me and my ambitions. Without their unwavering believe in me I could not have come so far.

Funding. While working on this thesis I was supported by a FWF grant (DOI 10.55776/P36278).

About the Author

Jakob Glas obtained a BSc in Mathematics at the University of Salzburg in 2018. Before joining ISTA in 2020, he completed a MSc in Pure Mathematics at ETH Zürich. Since he joined the Browning group in 2021, he has been mostly working on problems in the context of rational points with a particular focus on function fields and intersections with geometry.

List of Collaborators and Publications

J. Glas and L. Hochfilzer, On a question of Davenport and diagonal cubic forms over $\mathbb{F}_q(t)$. *Preprint*, 2022. (arXiv:2208.05422)

J. Glas, Rational points on complete intersections of cubic and quadric hypersurfaces over $\mathbb{F}_q(t)$. *Preprint, to appear in Journal of the LMS*, 2023. (arXiv:2306.02718)

J. Glas and L. Hochfilzer, Rational points on del Pezzo surfaces of low degree. *Preprint*, 2024. (arXiv:2401.04759)

J. Glas, Canonical singularities on moduli spaces of rational curves via the circle method. *Preprint*, 2024. (arXiv:2405.16648)

Table of Contents

List of Tables

CHAPTER 1

Introduction

Given polynomials $F_1, \ldots, F_R \in \mathbb{Z}[x_1, \ldots, x_n]$, seeking integer or rational solutions to the system of equations $F_1(\boldsymbol{x}) = \cdots = F_R(\boldsymbol{x}) = 0$ is one of the most foundational problems in number theory. Such equations are called *Diophantine equations*, named after the Greek mathematician Diophantus of Alexandria, who studied them more than 2000 years ago. However, the investigation of these equations already started before the times of Diophantus. For example, it has already known to the Babylonians how to compute Pythagorean triples in 1800 BC. In most cases it is already highly non-trivial to decide whether a given system of Diophantine equations has a solution. For example, for an integer $n \geq 1$, let

$$F_n(x, y, z) = x^n + y^n - z^n.$$

When $n = 1$, the equation $F_n = 0$ is linear and all the infinitely many solutions can be described quite easily. For $n = 2$ there are infinitely many solutions, the so-called "Pythagorean triples" that are known since antiquity. However, as soon as $n \geq 3$, the picture suddenly changes. Fermat claimed in the 1670s that the trivial solutions satisfying $xyz = 0$ are the only ones and remarked "*I have discovered a truly marvelous proof of this, which this margin is too narrow to contain.*" This innocuous assertion became the motivation for one of the biggest achievements in number theory and all of mathematics in the 20th century. Andrew Wiles together with Richard Taylor [208, 193] proved the modularity theorem in 1995, more than 300 years after Fermat, and thus completed the proof of Fermat's claim.

However, Diophantine equations are not only interesting for their own sake, or because of their impact on number theory. One the one hand, mathematicians have borrowed tools from diverse areas of mathematics such as algebraic geometry and logic to tackle Diophantine equations. On the other hand, Diophantine equations have had a massive impact on other areas such as algebraic geometry. For example, the igniting spark for the revolution of algebraic geometry and the development of étale cohomology by Grothendieck and his collaborators was an attempt to solve the Weil conjectures. The connection between Diophantine equations and geometry is given by the fact that the system of equations $F_1(\boldsymbol{x}) = \cdots = F_R(\boldsymbol{x}) = 0$ defines an algebraic variety inside the affine space $\mathbb{A}^n$, and if the F_i are all homogeneous, inside the projective space $\mathbb{P}^{n-1}$. Therefore, integer solutions to a system of homogeneous equations correspond to rational points on the variety defined by that system.

One point of view that crystallized over the last century is that the geometry of the underlying variety governs the behaviour of its set of rational points. We shall expand upon this perspective in Chapter 2. In particular, we introduce some of the guiding questions in the field of rational points and explain how certain geometric properties can be used to (at least conjecturally) answer them.

This entails qualitative aspects, such as criteria for the existence of solutions, but also finer measures for the quantitative distribution of rational points. Given a parameter $B \geq 0$, we can for example consider the counting function

$$N_{F_1,\dots,F_R}(B) = \#\{\boldsymbol{x} \in \mathbb{Z}^n \colon F_1(\boldsymbol{x}) = \cdots = F_R(\boldsymbol{x}) = 0, |\boldsymbol{x}| < B\}$$

and try understand to how it behaves as $B \to \infty$. Manin and his collaborators [76] put forward a conjecture for an analogue of this counting function for rational points of bounded height on Fano varieties that we explain in Section 2.3. For the particular case of complete intersections in projective space, one of the most versatile tools to study this counting function is the *circle method* that goes back to the pioneering work of Hardy and Ramanujan on the partition function [93]. The starting point is the simple identity

$$\int_0^1 e^{2\pi i\alpha n}\mathrm{d}\alpha = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \in \mathbb{Z} \setminus \{0\}. \end{cases}$$

When $F \in \mathbb{Z}[x_1, \dots, x_n]$ and $\boldsymbol{x} \in \mathbb{Z}^n$, we can apply this identity to $n = F(\boldsymbol{x})$. This allows us express the counting function $N_F(B)$ as the integral

$$\int_0^1 S(\alpha)\mathrm{d}\alpha,$$

where $S(\alpha)$ is the exponential sum

$$S(\alpha) = \sum_{\substack{\boldsymbol{x} \in \mathbb{Z}^n \\ |\boldsymbol{x}| < B}} e^{2\pi i\alpha F(\boldsymbol{x})}.$$

Via this approach one can study Diophantine equations through the lens of harmonic analysis. We explain the basic mechanism underlying the circle method in the setting of function fields in Chapter 4.

In the specific context that we are working in, a function field will typically refer to $\mathbb{F}_q(t)$, where $\mathbb{F}_q$ is a finite field with q elements. The function field $\mathbb{F}_q(t)$ and the rational numbers share many of similarities and both fall under the more general context of global fields. The key advantage of function fields is that their geometric nature allows for new tools to be brought into play, with the effect that many big open problems over the rationals are in fact theorems over function fields. Roughly speaking, this thesis is mostly dealing with problems related to counting rational points over function fields. A key role is played by the circle method and to set it up, we develop the basic theory for harmonic analysis over function fields in Chapter 3.

Let us now discuss the genuinely new contributions of this thesis. In Chapter 5 we use the circle method to study the analogue of the counting function $N_F(B)$ over function fields when F is a diagonal cubic form in 4 or 6 variables, which is based on joint work with Hochfilzer [88]. This investigation parallels work of Hooley [111] and Heath-Brown [103] over the integers, but the crucial difference is that our work is *unconditional*, whereas they rely on unproved hypotheses about Hasse–Weil L-functions coming from cubic hypersurfaces. In addition, we give applications to the asymptotic Waring problem for cubes and weak approximation for cubic hypersurfaces defined by diagonal forms in at least 7 variables.

Chapter 6 is concerned with producing an asymptotic formula for the counting function associated to smooth complete intersections of cubic and quadratic hypersurfaces over $\mathbb{F}_q(t)$ and is based on the author's work [85]. The quantitative result about the distribution of rational points is also used to establish weak approximation for these complete intersections. Our main tool is again the circle method. The key technical advancement is a refinement of a Farey dissection developed by Vishe [202] that opens up the possibility for a Kloosterman refinement.

Chapter 7 departs from the circle method and uses more geometric methods. Here we present work that stems from a collaboration with Hochfilzer [89]. We produce upper bounds for the number of rational points of bounded height on del Pezzo surfaces of degree at most 5 over arbitrary global fields. Our approach generalises the work of Heath-Brown [102] on cubic surfaces over $\mathbb{Q}$. Simply speaking, the idea is to reduce the problem to producing uniform estimates for the number of rational points of bounded height on elliptic curves. In characteristic 0, the bounds we require follow from a deep conjecture relating the rank of an elliptic curve to its conductor, whereas in positive characteristic our bounds are unconditional thanks to work of Brumer [46].

Over function fields, rational points of bounded height can be understood more geometrically. Specifically, if $X \subset \mathbb{P}^{n-1}$ is the hypersurface defined as the vanishing locus of $F \in \mathbb{F}_q[x_1, \dots, x_n]$, then $\mathbb{F}_q(t)$-rational points of bounded height can be related to the $\mathbb{F}_q$-points of the variety $\mathrm{Mor}_e(\mathbb{P}^1, X)$ of degree e morphisms from $\mathbb{P}^1$ to X. We will explain this in more detail in Section 2.4. Via a combination of spreading out arguments and the Lang–Weil estimate, this can be used to deduce geometric information about $\mathrm{Mor}_e(\mathbb{P}^1, X)$ from sufficiently uniform point count estimates over finite fields. This was heavily exploited in work of Browning and Vishe [44], where they showed irreducibility and determined the dimension of the moduli space $\mathcal{M}_{0,0}(X, e)$ of rational curves of degree e lying on a smooth hypersurface $X \subset \mathbb{P}^{n-1}$ of degree d, providing n is sufficiently large with respect to d. In Chapter 8 we develop a suitable version of the circle method that allows us to count $\mathbb{F}_q[t, s]/(s^{m+1})$ points on hypersurfaces defined over $\mathbb{F}_q$ of bounded degree in t. In conjunction with the work of Mustaţă [150] relating singularities to jet schemes, this enables us to deduce that $\mathcal{M}_{0,0}(X, e)$ only has canonical singularities if n is sufficiently large with respect to d and e.

Conventions

We will make use of the big-O and Vinogradov's notation $O, o, \ll\gg, \asymp, \dots$. Moreover, we will indicate a dependence of the implied constant on certain parameters by subscripts, unless specified otherwise. The letter ε will denote an arbitrarily small real number whose exact value may change from one occurrence to the next. Its appearance together with one of the symbols $O, o, \ll, \gg, \dots$ indicates that the implied constant depends on ε.

CHAPTER 2

Rational points on varieties

In 1900 Hilbert posed a famous list of 23 problems, of which the tenth asks for the existence of an algorithm that can decide whether a given Diophantine equation has integer solutions. It took mathematicians more than 70 years to find an answer to this problem: Matiyasevich [140] proved that such an algorithm *cannot* possibly exist! So despite their innocuous appearance, finding solutions to Diophantine equations or proving their non-existence, is an incredibly difficult problem in general.

It is natural to study polynomial equations not just over the integers or the rationals, but also over other rings or fields of arithmetic interest. The main bulk of this thesis is concerned with precisely this problem and we shall now delve into the basic questions underlying this field.

A guiding principle in the study of polynomial equations is that *geometry determines arithmetic*, as we will explain in the subsequent sections based on various examples and conjectures, with a particular focus on hypersurfaces and complete intersections. There is no need to restrict ourselves to the rational numbers or the integers, and so we shall take K to be a global field. Given $F_1, \dots, F_R \in K[x_1, \dots, x_n]$ we assume for simplicity that all the F_i's are homogeneous and consider the projective variety $X = V(F_1, \dots, F_R) \subset \mathbb{P}^{n-1}$ defined as the common vanishing locus of the F_i's. Studying the set of non-trivial solutions of $F_1 = \cdots = F_R = 0$ therefore amounts to understanding the set $X(K)$ of K-rational points of X. To simplify the following discussion, we shall henceforth assume that X is smooth and geometrically integral.

2.1 Kodaira dimension and Zariski density

We begin with a very coarse question about the distribution of rational points.

Question 2.1.1. *Suppose that $X(K) \neq \emptyset$. Is $X(K)$ Zariski dense in X?*

In terms of geometry, the simplest case is when X defines a curve. In this case Zariski density is equivalent to $X(K)$ being infinite. We can attach a geometric invariant $g \in \mathbb{Z}_{\geq 0}$ to X, called the *genus*. For example, if $g = 0$, then either $X(K) = \emptyset$ or X is K-isomorphic to $\mathbb{P}^1$ and hence there are infinitely many K-rational points on X. Table 2.1 explains the trichotomy for the set of K-rational on curves in terms of the genus.

When $g = 1$, then either $X(K) = \emptyset$ or X is an *elliptic curve*. The Mordell–Weil theorem asserts that the set of K-rational points is a finitely generated abelian group. In particular,

genus	rational points
$g = 0$	$X(K) = \emptyset$ or $\#X(K) = \infty$
$g = 1$	$\#X(K) < \infty$ or $\#X(K) = \infty$
$g \geq 2$	$\#X(K) < \infty$

Table 2.1: Rational points on curves.

$X(K)$ is infinite if and only if the underlying group has positive rank. When $g \geq 2$ and K is a number field, then the finiteness of $X(K)$ was conjectured by Mordell [148] in 1922 and proved by Faltings [74] 61 years later. An analogous result in positive characteristic was established by Samuel [170]. In particular, we see that for curves the answer to Question 2.1.1 is completely determined by their genus.

Turning to the higher dimensional situation, a rough classification of the geometry of X is provided by its *Kodaira dimension*, which measures the positivity of the canonical divisor K_X of X. Given any integer $m \geq 1$, we can consider the associated rational map $\phi_m \colon X \dashrightarrow \mathbb{P}^N$ induced by the linear system $|mK_X|$ and the Kodaira dimension of X is defined to be

$$\kappa(X) = \sup_{m \geq 1} \dim \phi_m(X).$$

If $|mK_X| = \emptyset$ for all $m \geq 1$, then we set $\kappa(X) = -\infty$. With this convention it is clear that $\kappa(X) \in \{-\infty\} \cup \{0, \dots, \dim X\}$. To get an idea of what the Kodaira dimension measures, we will now give some examples.

Example 2.1.2. If X is a curve, then

$$\kappa(X) = \begin{cases} -\infty & \text{if } g = 0, \\ 0 & \text{if } g = 1, \\ 1 & \text{if } g \geq 2. \end{cases}$$

In particular, we could have rephrased Table 7.1 in terms of the Kodaira dimension.

Example 2.1.3. If $X = V(F_1, \dots, F_R) \subset \mathbb{P}^{n-1}$ is a complete intersection defined by forms F_i of degree d_i for $i = 1, \dots, R$, then

$$\kappa(X) = \begin{cases} -\infty & \text{if } n > d_1 + \cdots + d_R, \\ 0 & \text{if } n = d_1 + \cdots + d_R, \\ n - 1 - R & \text{else.} \end{cases}$$

In some sense varieties with $\kappa(X) = -\infty$ are special, while the majority of varieties will have maximal Kodaira dimension.

Definition 2.1.4. We say that X is of

1. *general type* if $\kappa(X) = \dim X$,
2. *intermediate type* if $0 \leq \kappa(X) < \dim X$.

An important class of varieties of negative Kodaira dimension is formed by those whose anti-canonical divisor is ample.

Definition 2.1.5. A variety X is called a *Fano* variety if its anti-canonical divisor $-K_X$ is ample.

Example 2.1.6. A curve is Fano if and only if it has genus 0. A complete intersection inside projective space is Fano if and only if it has negative Kodaira dimension.

Recall that a variety X of dimension n is called *uniruled* over K if there exists a variety Y of dimension $n-1$ and a dominant K-rational map $\phi\colon Y \times \mathbb{P}^1 \dashrightarrow X$. Moreover, we say that X is *separably uniruled* if there exists such a ϕ which is separable. These two notions agree in characteristic 0. If X is uniruled over K and $X(K) \neq \emptyset$, then $X(K)$ is in fact Zariski dense. We know that if X is separably uniruled, then $\kappa(X) = -\infty$ [124, Chapter IV, Corollary 1.10] and in fact, it is conjectured $\kappa(X) = -\infty$ implies that X is uniruled [124, Chapter IV, Conjecture 1.12]. This is known if $\dim X \leq 3$ and $\mathrm{char}(K) = 0$ due to work of Miyaoka [147]. However, if $\kappa(X) = -\infty$, it is not necessarily true that X is *separably* uniruled, as was shown by Sato [171]. It is therefore natural to conjecture that rational points of X are at least *potentially* Zariski dense if $\kappa(X) = -\infty$, which means that $X(L)$ is Zariski dense for some finite field extension L/K.

At the other and of the spectrum, we have the following conjecture generalising Mordell's conjecture to higher dimensions.

Conjecture 2.1.7 (Bombieri–Lang)**.** *Suppose that X is a variety of general type over a number field. Then $X(K)$ is not Zariski dense.*

The conjecture seems very difficult to prove at present and only very few cases are known. In addition, it also has strong consequences for rational points on curves: Caporaso, Harris and Mazur [49] have shown that the Bombieri–Lang conjecture implies for any number field K and any integer $g \geq 2$ the existence of a constant $C(K, g) > 0$ such that any smooth genus g curve over K has at most $C(K, g)$ rational points.

In positive characteristic the analogue of the Bombieri–Lang conjecture is false. Indeed, if $\mathrm{char}(K) = p$, then the Fermat hypersurface $\sum_{i=1}^{n} x_i^d$ is uniruled and smooth whenever $d = p^r + 1$ for any $r \geq 1$. In particular, rational points are dense, although the hypersurface is of general type if r is large with respect to n. This example seems to be first considered systematically by Shioda [184]. The analogue of the result by Caporaso, Harris and Mazur does not hold either. Assuming $p > 3$, Conceição, Ulmer and Voloch [55] give an example of an infinite family $\mathcal{F}$ of non-isotrivial genus g curves C over $\mathbb{F}_p(t)$ whose number of rational points is unbounded as C varies over $\mathcal{F}$.

We have thus seen that – apart from pathological examples in positive characteristic – varieties of general type are expected to have few rational points, while varieties of negative Kodaira dimension are expected to have an abundance. We have not said anything about varieties of intermediate type and the reason is that anything can happen in this case. This can already be seen by considering curves, where intermediate type varieties with a rational point correspond to elliptic curves. So the set of rational points can be dense or not.

2.2 The Hasse principle and weak approximation

The previous section was concerned with a coarse classification of the abundance of rational points in terms of Zariski density, with the guiding invariant the Kodaira dimension. We will now turn to the more difficult question of finding criteria guaranteeing the existence of rational points and introduce a finer measure for the distribution of rational points.

Let Ω_K be the set of places of K for $\nu \in \Omega_K$ denote by K_ν the completion of K at ν. Since $K \subset K_\nu$, an obvious necessary condition for $X(K)$ to be non-empty is that $X(K_\nu)$ is non-empty for every place ν. Writing $\mathbb{A}_K$ for the ring of adeles of K, this is equivalent to $X(\mathbb{A}_K) \neq \emptyset$. The Hasse principle asks about the converse.

Definition 2.2.1. We say that a family of varieties $\mathfrak{X}$ satisfies the Hasse principle if $X(\mathbb{A}_K) \neq \emptyset$ implies $X(K) \neq \emptyset$ for any $X \in \mathfrak{X}$.

The name of the Hasse principle goes back to pioneering work of Hasse [96] and Minkowski [146], who proved the Hasse principle for quadrics. However, in general the Hasse principle is too much to hope for.

Example 2.2.2. Let C be the curve given by $3x_1^3+4x_2^3+5x_3^3 = 0$ inside $\mathbb{P}^2$. Then Selmer [181] has shown that C has solutions over $\mathbb{Q}_p$ for every prime p and over $\mathbb{R}$, but not over $\mathbb{Q}$.

Example 2.2.3. The surface defined as the intersection of two quadrics $x_1x_2 = x_3^2 - 5x_4^2$ and $(x_1 + x_2)(x_1 + 2x_2) = x_3^2 - 5x_5^2$ fails the Hasse principle over $\mathbb{Q}$. This example is due to Birch and Swinnerton-Dyer [15].

To explain failures of the Hasse principle, Manin in his 1970 ICM address proposed the following construction. Let $\mathrm{Br}(X) = H^2_{\text{ét}}(X, \mathbb{G}_m)$ be the Brauer group of X. Class field theory provides us with a natural paring

$$\langle\ ,\ \rangle\colon X(\mathbb{A}_K) \times \mathrm{Br}(X) \to \mathbb{Q}/\mathbb{Z}$$

and we define the *Brauer–Manin set*

$$X(\mathbb{A}_K)^{\mathsf{Br}} = \{x \in X(\mathbb{A}_K)\colon \langle x, \alpha\rangle = 0 \text{ for all } \alpha \in \mathrm{Br}(X)\}.$$

The Albert–Brauer–Hasse–Noether theorem implies that $X(K) \subset X(\mathbb{A}_K)^{\mathrm{Br}}$. This prompts the following definition.

Definition 2.2.4. We say that there is a *Brauer–Manin obstruction* to the Hasse principle if $X(\mathbb{A}_K) \neq \emptyset$, but $X(\mathbb{A}_K)^{\mathrm{Br}} = \emptyset$.

Both Examples 2.2.2 and 2.2.3 have a Brauer–Manin obstruction to the Hasse principle. However, Skorobogatov [187] has given examples of bi-elliptic surfaces whose failure of the Hasse principle cannot be explained by a Brauer–Manin obstruction.

Recall that $X(K)$ injects into $X(\mathbb{A}_K)$. As $\mathbb{A}_K$ is a locally compact ring with respect to the restricted product topology, we can endow the adelic points $X(\mathbb{A}_K)$ with a natural topology.

Definition 2.2.5. We say that X satisfies *weak approximation* if $X(K)$ is dense in $X(\mathbb{A}_K)$.

It is clear that if X satisfies weak approximation, then the Hasse principle holds as well. Moreover, weak approximation also implies Zariski density. It turns out that $X(\mathbb{A}_K)^{\mathrm{Br}}$ is closed in $X(\mathbb{A}_K)$ and so

$$\overline{X(K)} \subset X(\mathbb{A}_K)^{\mathrm{Br}},$$

where $\overline{X(K)}$ denotes the topological closure of $X(K)$ inside $X(\mathbb{A}_K)$.

Definition 2.2.6. We say that there is a *Brauer–Manin obstruction* to weak approximation if $X(\mathbb{A}_K)^{\mathrm{Br}} \subsetneq X(\mathbb{A}_K)$.

Definition 2.2.7. The Brauer–Manin obstruction is the only one to the Hasse principle (respectively weak approximation) if $X(\mathbb{A}_K)^{\mathrm{Br}} \neq \emptyset$ (respectively $X(\mathbb{A}_K)^{\mathrm{Br}} = X(\mathbb{A}_K)$) implies $X(K) \neq \emptyset$ (respectively $\overline{X(K)} = X(\mathbb{A}_K)$).

Both the Hasse principle and weak approximation are invariant under birational equivalence for smooth projective K-varieties. Moreover, as in the case of Zariski density, there is a conjectural geometric description of a large class of varieties for which the Brauer–Manin obstruction is expected to be the only one to the Hasse principle and weak approximation.

Definition 2.2.8. A variety X is (separably) rationally connected if any two general points of $\overline{X}$ can be joined by a (separable) rational curve.

Examples of rationally connected varieties are Fano varieties. This was independently proved by Campana [48] and Kollár–Miyaoka–Mori [125]. In characteristic 0 the notions of being separably rationally connected and rationally connected agree. It is still an open problem whether Fano varieties are separably rationally connected.

We then have the following conjecture about the Hasse principle and weak approximation put forward by Colliot-Thélène [54] for number fields.

Conjecture 2.2.9. *Let X be a separably rationally connected variety. Then the Brauer–Manin obstruction is the only one to the Hasse principle and weak approximation.*

Colliot-Thélène's conjecture lies very deep and is far from being proven. For instance, an affirmative resolution would solve the inverse Galois problem [183, Chapter 3.5]. Originally, Colliot-Thélène stated his conjecture only for number fields. We have chosen to state it here for *separably* rationally connected varieties in positive characteristic to avoid pathological behaviour.

In fact, this conjecture has been extended by Skorobogatov [188] to a particular class of surfaces of intermediate type.

Conjecture 2.2.10. *Let X be a K3 surface. Then the Brauer–Manin obstruction is the only one to the Hasse principle and weak approximation.*

We have already mentioned that quadrics over global fields satisfy the Hasse principle. Moreover, since any quadric with a rational point is birational to projective space, they also satisfy weak approximation. Indeed, satisfying the weak approximation property is invariant under birational maps between smooth projective varieties [183, Lemma 3.5.5].

Let now $X \subset \mathbb{P}^{n-1}$ be a Fano smooth complete intersection of dimension $n - 1 - R$. If $\dim X \geq 3$ and $\mathrm{char}(K) = 0$, then the Brauer–Manin obstruction is vacuous [160, Appendix A] and so Conjecture 2.2.10 implies that X satisfies weak approximation. A foundational result is provided by work of Birch [16], in which he proves the Hasse principle for smooth complete intersections $X \subset \mathbb{P}^{n-1}$ defined by R forms of degree d over $\mathbb{Q}$ provided $n > R + (d-1)2^d R(R+1)$. This was later extended to number fields by Skinner [186], in which he also establishes the weak approximation property. Birch's work has subsequently been generalized to complete intersections of not necessarily the same degree by Browning and Heath-Brown [40].

In the function field setting the Hasse principle turns out to be a rather easy consequence of the Lang–Tsen theory if $n > d_1^2 + \cdots + d_R^2$. Indeed, with this condition [91, Theorem 3.6] implies the existence of a rational point on X, thereby verifying the Hasse principle. However, establishing weak approximation remains a substantial challenge. Building on the work of Kubota [127], Lee [131] showed that Birch's work can be translated to the function field setting and further demonstrated that weak approximation holds under the same conditions on the number of variables when $\mathrm{char}(K) > d$. All the results mentioned thus far make use of the *circle method*, which we will explain in more detail in Chapter 4. Moreover, if the degree of the hypersurface is small, there are substantially stronger results available. We will give an overview of the results obtained using the circle method in Section 4.5.

2.3 Manin's conjecture

Instead of just studying *qualitative* questions about $X(K)$, we can try to *quantify* the distribution of rational points on $X(K)$. If $\#X(K) = \infty$, then of course we cannot just simply count all points and so we have to make this problem more precise.

First of all, we want to begin with a simple heuristic for hypersurfaces. Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be a form of degree d. Then for any $B \geq 1$, there are approximately B^n available $\boldsymbol{x} \in \mathbb{Z}^n$ with $|\boldsymbol{x}| < B$. Moreover, for such $\boldsymbol{x}$ we will have $|F(\boldsymbol{x})| \ll_F B^d$. Assuming that the values of F are evenly distributed, this leads to the expectation that the "probability" for $F(\boldsymbol{x}) = 0$ to hold should be of order B^{-d}. In particular, it is reasonable to expect that

$$\#\{\boldsymbol{x} \in \mathbb{Z}^n \colon F(\boldsymbol{x}) = 0, |\boldsymbol{x}| < B\} \asymp B^{n-d}. \tag{2.3.1}$$

Therefore, if $n > d$ — which corresponds precisely to $X = V(F) \subset \mathbb{P}^{n-1}$ being Fano — there is a reason to believe that there should be many solutions to $F(\boldsymbol{x}) = 0$ (provided that there are any). Strictly speaking we are not counting rational points on X, but going from the counting function above to only counting rational points is straightforward using Möbius inversion.

We will now explain how to extend this counting mechanism to any projective variety, beginning with the case of projective space. Given $x \in \mathbb{P}^{n-1}(\mathbb{Q})$, we can always find a representative $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n \setminus \{\boldsymbol{0}\}$ of x such that $\gcd(x_1, \dots, x_n) = 1$. In this case we define the *height* of x to be $H(x) = \max\{|x_1|, \dots, |x_n|\}$. This construction generalises to arbitrary global fields and is explained in more detail in Chapter 7. Moreover, given a variety X together with an ample line bundle L, a suitable multiple $L^{\otimes m}$ will be very ample and hence corresponds to an embedding $\varphi_{L^{\otimes m}} \colon X \to \mathbb{P}^N$ for some N. We can then define the associated height function

$$H_L \colon X(K) \to \mathbb{R}_{\geq 0}, \quad x \mapsto H(\varphi_{L^{\otimes m}}(x))^{1/m}.$$

Note that the height function depends on the choice of the embedding.

For $W \subset X(K)$, we are then interested in the counting function

$$N_{L,W}(B) \coloneqq \#\{x \in W \colon H_L(x) \leq B\}. \tag{2.3.2}$$

Since we assume that L is ample, the height function H_L satisfies the *Northcott property*, meaning that $N_{L,X(K)}(B)$ is finite for any $B \geq 1$.

Example 2.3.1. If $X = V(F) \subset \mathbb{P}^{n-1}$ is a hypersurface of degree d over $\mathbb{Q}$ with $n > d$, then its anti-canonical divisor corresponds to the very ample line bundle $\omega_X^{-1} = O_X(n-d)$ and the corresponding height function is $H_{\omega_X^{-1}}(x) = \max\{|x_1|, \dots, |x_n|\}^{n-d}$, where $(x_1, \dots, x_n) \in \mathbb{Z}^n$ is a representative for $x \in \mathbb{P}^{n-1}(\mathbb{Q})$ with $\gcd(x_1, \dots, x_n) = 1$.

If $W \subset X(K)$ is infinite, then it is natural to ask for an asymptotic formula for $N_{L,W}(B)$ as $B \to \infty$. For Fano varieties the anti-canonical bundle ω_X^{-1} gives a natural choice of ample line bundle and as explained in Section 2.2, the set of rational points is expected to be dense as soon as it is non-empty. Manin and his collaborators [76] put forward a conjectural asymptotic formula for the number of rational points on Fano varieties over number fields, which was later refined and extended to the case of positive characteristic by Peyre [156, 157]. Before stating it, let us recall the notion of a thin set.

Definition 2.3.2. Let X be a variety over K and $A \subset X(K)$. If A is not Zariski dense in X, then A is a called a *thin set of type I*. If there exists a dominant K-morphism $\varphi\colon Y \to X$ with Y geometrically integral, $\dim Y = \dim X$ and $\deg \varphi \geq 2$ such that $A \subset \varphi(Y(K))$, then A is a *thin set of type II*. In general, a subset of $X(K)$ is *thin* if it is contained in a finite union of thin sets of type I and II.

Conjecture 2.3.3 (Manin–Peyre)**.** *Let X be a Fano variety over a global field K. Then there exists a thin set $Z \subset X(K)$ such that for $W = X(K) \setminus Z$, we have*

$$N_{\omega_X^{-1},W}(B) \sim c_{\textsf{Peyre}} B (\log B)^{\operatorname{rk}\operatorname{Pic}(X)-1} \quad \text{as } B \to \infty,$$

where $c_{\textsf{Peyre}}$ is Peyre's constant and $\operatorname{rk}\operatorname{Pic}(X)$ is the rank of the Picard group of X. Moreover, if $\operatorname{char}(K) > 0$, we require B to run through the values of $H_{\omega_X^{-1}}(X(K))$.

The constant in the Manin–Peyre conjecture is defined as

$$c_{\textsf{Peyre}} = \alpha(X)\beta(X)\tau(X),$$

where

$$\begin{aligned}
\alpha(X) &= \frac{1}{(\textsf{rk}\operatorname{Pic}(X)-1)!}\int_{C_{\textsf{eff}}(X)^\vee} e^{-\langle \omega_X^{-1}, y\rangle} \mathrm{d}y,\\
\beta(X) &= \#H^1(\mathbb{Q}, \operatorname{Pic}(\overline{X})),\\
\tau(X) &= \omega(X(\mathbb{A}_K)^{\textsf{Br}}).
\end{aligned}$$

Here $C_{\textsf{eff}}(X)^\vee$ denotes the dual cone of the cone of effective divisors on X and ω is the *Tamagawa measure* on the adelic points of X. Roughly speaking, $\alpha(X)$ measures the size of the line bundle, while $\tau(X)$ measures the size of the closure of the rational points of X inside its adelic points.

Remark. The reason for removing a thin set is that sometimes special accumulating subsets can dominate the point count. Indeed, let $X \subset \mathbb{P}^3$ be a cubic surface over $\mathbb{Q}$ containing a line M, which is defined over $\mathbb{Q}$. Then it is easy to see that rational points on M will already contribute $\gg B^2$ to $N_{\omega_X^{-1},X(\mathbb{Q})}(B)$ and $M(\mathbb{Q})$ is a thin set of type I.

Remark. In its original form Manin suggested that the removal of a thin set of type I of $X(K)$ might be sufficient for the asymptotic formula in Conjecture 2.3.3 to hold. However, Batyrev and Tschinkel [7] provide the following counterexample over number fields containing third roots of unity. Let $X \subset \mathbb{P}^3 \times \mathbb{P}^3$ be the bi-degree (1,3) hypersurface given by $\sum_{i=0}^3 x_i y_i^3 = 0$ and denote by $\pi\colon X \to \mathbb{P}^3$ the projection onto the first factor. The adjunction formula implies that $\omega_X^{-1} = O(3,1)$, so that an anti-canonical height function for X is given by $H(\boldsymbol{x})^3 H(\boldsymbol{y})$, where H is the naive height on $\mathbb{P}^3$. An application of the Lefschetz hyperplane theorem shows that $\operatorname{rk}\operatorname{Pic}(X) = 2$, so that for some complement of a thin set $W \subset X(K)$ Manin's conjecture predicts $N_{\omega_X^{-1},W}(B) \sim cB\log B$ as $B \to \infty$. However, if we let $\varphi\colon \mathbb{P}^3 \to \mathbb{P}^3$ be the morphism given by $[x_0, \dots, x_3] \mapsto [x_0^3, \dots, x_3^3]$, then for any $\boldsymbol{x} \in \varphi(\mathbb{P}^3(K))$, the assumption that K contains the third roots of unity implies that $\pi^{-1}(\boldsymbol{x})$ is a split cubic surface. A split cubic surface has Picard rank 7 and Batyrev and Tschinkel managed to show that it contains $\gg B(\log B)^3$ rational points of anti-canonical height at most B. The set of $(\boldsymbol{x}, \boldsymbol{y}) \in X(K)$ for which $\boldsymbol{x} \in \varphi(\mathbb{P}^3(K))$ is Zariski dense and so Manin's conjecture cannot hold if one only allows for the removal of a thin set of type I.

Since then more counterexamples have been found over any global field [130, 138, 19] and Batyrev and Tschinkel's counterexample has been extended to number fields not containing third roots of unity [77], but of all them can be explained by the removal of a thin set.

If $X \subset \mathbb{P}^{n-1}$ is a Fano hypersurface with $n \geq 5$, then it follows from the Lefschetz hyperplane theorem that $\operatorname{rk}\operatorname{Pic}(X) = 1$. In view of Example 2.3.1 the order of growth presented in our naive heuristic (2.3.1) therefore matches the order of growth predicted by Manin's conjecture.

Remark. Instead of working with heights as introduced above, we could have worked with *Arakelov heights*, which are constructed using locally metrized line bundles. Peyre [156] states his conjecture in this more general setting and for different choices of line bundles. This can change the exponents of both B and $\log(B)$ and the Tamagawa measure ω depends on the choice of the metrized line bundle.

Manin's conjecture for projective spaces is a classical result of Schanuel [175] over number fields and also known over function fields [182, Section 2.5]. Moreover, in the previous section we have already mentioned work of Birch [16], Skinner [186] and Lee [131] establishing the Hasse principle and weak approximation for complete intersections $X \subset \mathbb{P}^{n-1}$ of low degree. In fact, what they do is to produce an asymptotic formula for the counting function $N_{\omega_X^{-1},X}(B)$ and show that the constant is positive if the necessary local conditions are satisfied. Birch's result predates the conjectures of Manin and Peyre and was one of the guiding results in formulating them.

When the degrees of the defining equations are small, one can sometimes do substantially better than Birch's result. In particular, working over $\mathbb{F}_q(t)$ in Chapters 5 and 6 we will prove asymptotic formulae for the number of rational points of bounded height on diagonal cubic hypersurfaces and complete intersections of cubic and quadratic hypersurfaces respectively.

Let us now turn to varieties of small dimension. If X is a Fano curve with $X(K) \neq \emptyset$, then X is K-isomorphic to $\mathbb{P}^1$ and so Manin's conjecture holds as well. For surfaces the situation is already more mysterious and Manin's conjecture appears to be out of reach in most cases. In Chapter 7 we shall thus tackle the problem of producing *upper bounds* for the number of rational points of bounded height on Fano surfaces and give a more detailed overview of the current state of results.

Remark. An aesthetic and conceptual weakness of the Manin–Peyre conjecture is that there is no description of the exceptional thin set to be removed. This issue has been addressed by Lehmann, Sengupta and Tanimoto [132, 133], who proposed a geometric description of the exceptional set. In another direction, Peyre [158, 159] suggested two alternatives to bypass the necessity of removing a thin set. The "freeness" approach only counts rational points whose tangent space is sufficiently well-rounded, while the "all the heights" approach takes into account a tuple of height functions associated to a basis of the Picard group. However, Sawin [172] has shown that le Rudulier's work [130] also provides a counterexample to Peyre's freeness alternative. In addition, smooth cubic threefolds with an infinite family of lines defined over the ground field are most likely not compatible with the all the heights approach.

2.4 Moduli spaces of rational curves

We will now explain the connection between rational points on varieties over global function fields and moduli spaces of rational curves. We have already mentioned the guiding principle that geometry determines arithmetic. This section is concerned about the converse: sometimes tools from number theory can be used to shed light on problems in algebraic geometry.

To streamline the exposition, we assume that $K = \mathbb{F}_q(t)$ is the function field of $\mathbb{P}^1$ over $\mathbb{F}_q$ and take $X \subset \mathbb{P}^{n-1}$ to be a hypersurface defined over $\mathbb{F}_q$. It would be possible to work in the more general setting where K is the function field of a smooth projective curve C of genus $g \geq 1$ and X is defined over K and not just over its field of constants.

As in the setting over $\mathbb{Q}$, any rational point $x \in \mathbb{P}^{n-1}(\mathbb{F}_q(t))$ has a representative $\boldsymbol{x} \in \mathbb{F}_q[t]^n$ such that $\gcd(x_1, \dots, x_n) = 1$. In this case the height with respect to the line bundle $O(1)$ is just $H(\boldsymbol{x}) = \max_{i=1,\dots,n} q^{\deg x_i}$. By definition, a K-rational point on X is a morphism $\operatorname{Spec} K \to X$. One the one hand, since $\mathbb{P}^1$ is smooth and X is proper, the valuative criterion of properness implies that any morphism $\operatorname{Spec} K \to X$ extends to a morphism $\mathbb{P}^1 \to X$ over $\mathbb{F}_q$. On the other hand, given a morphism $\mathbb{P}^1 \to X$ over $\mathbb{F}_q$, the composition $\operatorname{Spec} K \to \mathbb{P}^1 \to X$ produces a K-rational point of X. In particular, if we define

$$\operatorname{Mor}_e(\mathbb{P}^1, X) = \{f \colon \mathbb{P}^1 \to X \colon \deg f = e\}$$

to be the variety of morphisms of degree e from $\mathbb{P}^1$ to X, then we obtain a natural bijection between $\operatorname{Mor}_e(\mathbb{P}^1, X)(\mathbb{F}_q)$ and the set

$$\{\boldsymbol{x} \in \mathbb{P}^{n-1}(K) \colon F(\boldsymbol{x}) = 0 \text{ and } H(\boldsymbol{x}) = q^e\}.$$

In the context of Manin's conjecture, this means that understanding the counting function $N_{\omega_X^{-1}, X}(q^e)$ as $e \to \infty$, is equivalent to understanding the number of $\mathbb{F}_q$-points of $\operatorname{Mor}_e(\mathbb{P}^1, X)$ as $e \to \infty$.

In algebraic geometry when trying to understand the geometry of a variety, it is often useful to understand its rational curves. For example, a variety is separably unirational if and only if there exists a free rational curve on X and separably rationally connected if and only if there is a very free rational curve [64, Corollary 4.17]. For this reason it bears fruit to study the geometry of $\operatorname{Mor}_e(\mathbb{P}^1, X)$. This space does not precisely parameterise rational curves and so algebraic geometers instead often work with the stack $\mathcal{M}_{0,0}(X, e)$ of rational curves of degree e on X or its Kontsevich compactification $\overline{\mathcal{M}}_{0,0}(X, e)$.

Over finite fields a key link between the geometry of a variety and its arithmetic is established via the Lang–Weil estimate [129].

Theorem 2.4.1 (Lang–Weil)**.** *Let $X \subset \mathbb{P}^{n-1}$ be a variety over $\mathbb{F}_q$. Then*

$$\#X(\mathbb{F}_q) = c_X(\mathbb{F}_q) q^{\dim X}(1 + O(q^{-1/2}))$$

uniformly in q, where $c_X(\mathbb{F}_q)$ is the number of top dimensional geometrically irreducible components of X defined over $\mathbb{F}_q$.

In particular, if we start with a variety X over $\mathbb{F}_q$, then there exists a positive integer l_0 such that its (top dimensional) geometrically irreducible components are all defined over $\mathbb{F}_{q^{l_0}}$ and hence we can determine their number if we get an asymptotic for $\#X(\mathbb{F}_{q^{kl_0}})$ as $k \to \infty$.

A crucial difference between number fields and global function fields is that we can consider a "horizontal problem" by letting q vary. For example, if X is a scheme defined over $\mathbb{Z}$, we can consider its reduction modulo any prime p and thus also over any field extension of $\mathbb{F}_p$. It is possible to apply a similar procedure in more geometric situations, which sometimes allows one to reduce geometric problems to the analogous problems in a family of finite fields.

This was heavily exploited in work of Browning and Vishe [44], in which they used tools from analytic number theory to count $\mathbb{F}_q$-points on $\mathrm{Mor}_e(\mathbb{P}^1, X)$ for hypersurfaces X of small degree and then deduced via the Lang–Weil estimate that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is irreducible and determined its dimension. In Chapter 8 we will make use of a similar strategy and count $\mathbb{F}_q$-points on the jet schemes of $\mathrm{Mor}_e(\mathbb{P}^1, X)$ to deduce that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ only has canonical singularities if n is large compared to e and the degree of the hypersurface X. This will be complemented by a thorough discussion of existing results.

It turns out that the horizontal problem of letting the size of the finite field tend to infinity is somewhat easier than the vertical problem of letting the degree e go to infinity. This phenomenon has been observed in many occasions, such as Katz and Sarnak's work on the connection between random matrix theory and L-functions [120], or the resolution of the Bateman–Horn conjecture by Entin [73]. In particular, in the context of counting rational points, it seems that whenever we can prove Manin's conjecture for a class of varieties over $\mathbb{F}_q(t)$, then we can also prove that the corresponding moduli spaces of rational curves are irreducible and of the expected dimension.

Typically methods from algebraic geometry only give information for *generic* hypersurfaces, which is of little use for counting points. However, if they work for concrete examples, these results can be used to deduce at least upper bounds for the number of rational points, as we shall now explain. Let $X \subset \mathbb{P}^{n-1}$ be a hypersurface of degree d over $\mathbb{F}_q$. Then we can realise $\mathrm{Mor}_e(\mathbb{P}^1, X)$ as an open subset of a closed subset defined by $de+1$ equations of degree d inside $\mathbb{P}^{n(e+1)-1}$, leading to the expectation that

$$\dim \mathrm{Mor}_e(\mathbb{P}^1, X) = n(e+1) - de - 2 = e(n-d) + n - 2.$$

In particular, if $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is of the expected dimension, then it is an open subset of a complete intersection of degree d^{de+1}.

Given a variety $Y \subset \mathbb{A}^n$ over $\overline{\mathbb{F}}_q$, we define $\delta(Y) = \sum_{i=1}^N \deg(Y_i)$, where $Y = \cup_{i=1}^N Y_i$ is the decomposition of Y into its geometrically irreducible components. Note that if Y is equidimensional, then $\delta(Y) = \deg(Y)$. In addition, the Bézout inequality $\delta(Y \cap Z) \leq \delta(Y)\delta(Z)$ always holds. We then have the following simple estimate due to Cafure and Matera [47, Lemma 2.1].

Lemma 2.4.2. *Let $Y \subset \mathbb{A}^n$ be a variety over $\overline{\mathbb{F}}_q$. Then*

$$\#Y(\mathbb{F}_q) \leq \delta(Y) q^{\dim Y}.$$

Note that strictly speaking Lemma 2.1 of [47] is only stated for varieties defined over $\mathbb{F}_q$, but an inspection of the proof reveals that it remains valid for varieties defined over $\overline{\mathbb{F}}_q$.

As explained by Debarre [64, Theorem 2.6], Riemann–Roch implies that any geometrically irreducible component of $\mathrm{Mor}_e(\mathbb{P}^1, X)$ has at least the expected dimension. The number of $\mathbb{F}_q$-points of $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is clearly bounded by the number of $\mathbb{F}_q$-points on its affine cone. Consequently, we immediately obtain the following estimate from Lemma 2.4.2.

Corollary 2.4.3. *Let $X \subset \mathbb{P}^{n-1}$ be a hypersurface of degree d over $\mathbb{F}_q$ and suppose that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is of the expected dimension $e(n-d)+n-2$. Then*

$$N_X(e) := \#\{\boldsymbol{x} \in X(\mathbb{F}_q(t))\colon H(\boldsymbol{x}) = q^e\} \ll d^{de} q^{e(n-d)},$$

where the implied constant only depends on d, n and q.

Example 2.4.4. If X is a cubic threefold over $\mathbb{C}$, Coskun and Starr [57] have shown that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ has 2 irreducible components and is of the expected dimension for any $e \geq 1$. In particular, if their methods continue to hold in positive characteristic, then we would obtain

$$N_X(e) \ll 27^e q^{2e}.$$

Manin's conjecture suggests that $N_X(e) \sim cq^{2e}$. If $\varepsilon > 0$ is given, and $q > 27^{1/\varepsilon}$, then $N_X(e) \ll q^{e(2+\varepsilon)}$, so that we could get arbitrarily close to the expected growth upon taking q sufficiently large. The most powerful tool available to deal with $N_X(e)$ is the circle method and it seems that without the injection of radical new ideas, we cannot do better than $N_X(e) \ll q^{9e/4}$. Even the implementation of the ratios conjecture into the circle method by Wang [204], which was carried over to function fields by Browning, Wang and the author [39], might only be able to produce the conditional estimate $q^{e(9/4-\delta)}$ for a small value of $\delta > 0$.

Example 2.4.5. If X is a smooth cubic surface over $\mathbb{F}_q$, then it follows from work of Beheshti, Lehmann, Riedl and Tanimoto [8] that the only irreducible components of the space $\mathcal{M}_{0,0}(X, e)$ that do not have the expected dimension parameterise multiple covers of lines. These correspond to $\mathbb{F}_q(t)$-points lying on lines, and so letting $U \subset X$ be the complement of lines, in the notation of Section 2.3 we obtain the upper bound

$$N_{U,\omega_X^{-1}}(q^e) \ll 27^e q^e.$$

In particular, if q is sufficiently large we again come arbitrarily close to the linear growth predicted by Conjecture 2.3.3. This estimate should be compared with $N_{U,\omega_X^{-1}}(q^e) \ll q^{e(4/3+\varepsilon)}$ that we establish in Chapter 7. In addition, the work [8] applies to del Pezzo surfaces of *any* degree, and so a modified version of Lemma 2.4.2 would imply an analogous bound. However, note that our bound from Chapter 7 also applies to cubic surfaces over $\mathbb{F}_q(t)$ (and in fact any function field over $\mathbb{F}_q$), and not just to those defined over the constant field $\mathbb{F}_q$.

Remark. It is natural to try to use the Lang–Weil estimate to also produce *lower bounds* for $N_X(e)$ or $N_{U,\omega_X^{-1}}(q^e)$. However, applied to $\mathrm{Mor}_e(\mathbb{P}^1, X)$ the best explicit bounds we have either require q to grow in the degree [47] or the implied constant in the error term has order of growth e^e [84]. A possible route to bypass this issue would be a uniform homological stability result, which in conjunction with the Grothendieck–Lefschetz trace formula could yield sharper estimates.

CHAPTER 3

Harmonic analysis over function fields

While number theory is a subject of mathematics with a long history, its extension to function fields over finite fields is quite recent. It has already been known in the 19th century that there is a strong analogy between number fields and the theory of Riemann surfaces, which inspired Dedekind and Weber [65] to re-build the theory of Riemann surfaces from a completely algebraic point of view via valuations. A big step towards the incorporation of function fields over finite fields into number theory was taken by Artin in his thesis [4], where he worked on quadratic extensions of function fields in close parallel to the case of number fields. This was continued with the work of Hasse [97] on the Riemann hypothesis over finite fields, with important contributions from F. K. Schmidt [179], that eventually culminated in a proof of the Riemann hypothesis for curves by Weil [206]. Subsequently, Weil proposed a vast generalisation of his achievements, commonly referred to as the "Weil conjectures". These were the driving force in Grothendieck's attempt to put algebraic geometry on a modern footing and the development of étale cohomology, allowing him and his collaborators to prove all of Weil's conjectures but one. What remained was a general form of the Riemann hypothesis that was eventually resolved by Deligne [66]. It should be mentioned though that Dwork [70] was the first to prove the rationality of the zeta function, which is part of the Weil conjectures.

The geometric nature of function fields gives access to tools from geometry that are not available over number fields. For this reason, function fields have been a prominent testing ground for conjectures whose counterpart over number fields is beyond the reach of current methods. The Riemann hypothesis already mentioned is one such example. To name just a few recent instances, let us mention the work of Ellenberg, Venkatesh and Westerland [72] on the Cohen–Lenstra heuristics using homological stability properties of Hurwitz spaces; of Sawin and Shusterman resolving the twin prime conjecture [174] and the quadratic Bateman–Horn conjecture [173] in a strong quantitative sense; and the work of Bergström, Diaconu, Petersen and Westerland [9] that reduces the moment problem of quadratic Dirichlet L-functions to a (uniform) homological stability property of representations of the Braid group with twisted coefficients, that was later proved by Miller, Patzt, Petersen and Randall-Williams [145].

In most chapters of this thesis we tackle number theoretic problems with tools from analysis. We will therefore focus on building the infrastructure needed to conduct harmonic analysis. This will be of especial importance for Chapters 5 and 6. Since we are only working over the function field $K = \mathbb{F}_q(t)$ in these chapters, we limit ourselves to that case, although it would not present much difficulty to extend the theory to other global function fields. All of the material presented here can be found in [35, Chapter 5] or [161]. A systematic treatment of harmonic analysis over global fields was first given by Tate in his thesis [192] to unify the treatment of Hecke L-functions.

Norms and places

Let $K = \mathbb{F}_q(t)$ and $\mathcal{O} = \mathbb{F}_q[t]$. The set of places of K is in bijection with monic irreducible polynomials $\varpi \in \mathbb{F}_q[t]$ and the infinite place t^{-1}. Unlike in characteristic 0, every place of $\mathbb{F}_q(t)$ is non-archimedean. For any place ν of K, let $v_\nu \colon K \to \mathbb{Z}$ be the associated valuation that we normalise in such a way that it is surjective. We then obtain a norm

$$|\cdot|_\nu \colon K \to q^{\mathbb{Z}}, \quad x \mapsto q^{\deg(\nu) v_\nu(x)}.$$

If $\nu = t^{-1}$, then we will use the shorthand notation $|\cdot| = |\cdot|_\infty$. For $\nu = t^{-1}$ this means explicitly the following. Let $x \in K$ be given. Then we can write $x = a/r$ with $a, r \in \mathbb{F}_q[t]$ coprime and obtain

$$|x| = q^{\deg a - \deg r},$$

with the convention $|0| = 0$. The field $K_\infty = \mathbb{F}_q((t^{-1}))$ of Laurent series in t^{-1} arises naturally as the completion of K with respect to $|\cdot|$. Any non-zero $\alpha \in K_\infty$ can be written uniquely as

$$\alpha = \sum_{i \le M} a_i t^i, \tag{3.0.1}$$

for some $M \in \mathbb{Z}$, where $a_i \in \mathbb{F}_q$ and $a_M \neq 0$. With this representation the extension of $|\cdot|$ to K_∞ is then given by $|\alpha| = q^M$. Although the field K_∞ can be thought of as a positive characteristic analogue of $\mathbb{R}$, the norm $|\cdot|$ behaves very differently to the usual absolute value on $\mathbb{R}$. As we have already mentioned, the norm is non-archimedean, meaning that it satisfies the ultrametric property

$$|\alpha + \beta| \le \max\{|\alpha|, |\beta|\} \text{ for any } \alpha, \beta \in K_\infty.$$

In many occasions this heavily simplifies the analysis compared to the situation in characteristic 0.

For any $n \ge 1$, we can extend the norm to K_∞^n by setting $|\boldsymbol{\alpha}| = \max_{1 \le i \le n} |\alpha_i|$ whenever $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in K_\infty^n$. For $\alpha \in K_\infty$, in the notation of (3.0.1) we define its *integer part* and *fractional part* to be

$$[\alpha] = \sum_{i \ge 0} a_i t^i \quad \text{and} \quad \{\alpha\} = \sum_{i \le -1} a_i t^i,$$

respectively. In particular, $\alpha = \{\alpha\} + [\alpha]$ and $[\alpha] \in \mathbb{F}_q[t]$. The *distance to the nearest integer* is then given by $\|\alpha\| = |\{\alpha\}|$.

Characters

Let $e_q \colon \mathbb{F}_q \to \mathbb{C}^\times$ be the standard additive character given by

$$e_q(x) = \exp\left(\frac{2\pi i \operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}{p}\right),$$

where $\operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p} \colon \mathbb{F}_q \to \mathbb{F}_p \simeq \mathbb{Z}/p\mathbb{Z}$ is the standard field trace and $p = \operatorname{char}(\mathbb{F}_q)$. The function $\psi \colon K_\infty \to \mathbb{C}^\times$ given by

$$\psi(\alpha) = e_q(a_{-1})$$

in the notation of (3.0.1) then defines an additive character of K_∞ that is trivial on the subgroup $\{\alpha \in K_\infty \colon |\alpha| < q^{-1}\}$, but not on the "unit interval"

$$\mathbb{T} = \{\alpha \in K_\infty \colon |\alpha| < 1\}.$$

Integration

The field K_∞ is a locally compact Hausdorff group with respect to addition. In particular, it comes with a Haar measure $\mathrm{d}\alpha$ that is unique if we require

$$\int_{\mathbb{T}} 1\mathrm{d}\alpha = 1.$$

The measure extends naturally to K_∞^n by $\mathrm{d}\boldsymbol{\alpha} = \mathrm{d}\alpha_1 \cdots \mathrm{d}\alpha_n$ for $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in K_\infty^n$ for any $n \geq 1$. Given a measureable subset $A \subset K_\infty^n$, we will use the shorthand notation

$$\mathrm{vol}(A) = \int_A 1\mathrm{d}\alpha.$$

A key property of the Haar measure is the following change of variables formula established by Igusa for any local field [118, Lemma 7.4.2].

Lemma 3.0.1. *Let $\Gamma \subset K_\infty^n$ be a measurable set and $g\colon \Gamma \to \mathbb{C}$ be a measurable function. Then for any $M \in \mathrm{GL}_n(K_\infty)$ we have*

$$\int_\Gamma g(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha} = |\det M| \int_{M\boldsymbol{\beta}\in\Gamma} g(M\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta}.$$

Having introduced the most basic concepts, it is helpful to compare them with their counterparts over $\mathbb{Q}$.

	$\mathbb{Q}$	$\mathbb{F}_q(t)$
ring of integers	$\mathbb{Z}$	$\mathbb{F}_q[t]$
finite places	primes p	monic irreducible polynomials ϖ
infinite place	absolute value on $\mathbb{R}$	$\lvert\cdot\rvert$ induced by t^{-1}
Pontryagin dual of ring of integers	$[0,1) = \mathbb{R}/\mathbb{Z}$	$\mathbb{T} = K_\infty/\mathbb{F}_q[t]$
Haar measure	Lebesgue measure	$\mathrm{d}\alpha$ such that $\int_{\mathbb{T}} 1\mathrm{d}\alpha = 1$
additive character	$\mathbb{R} \ni x \mapsto e^{2\pi i x}$	$\psi\colon K_\infty \to \mathbb{C}^\times$

Remark. We have already mentioned that in contrast to characteristic 0, there is nothing special about the place at infinity t^{-1}. In fact, our choice of t was already arbitrary and we could have instead replaced t with a different degree 1 element of K. More generally, if one works with a function field K in positive characteristic, there is no natural candidate for "the" ring of integers and there is always a choice of places involved.

Orthogonality relations

In Chapter 4 we will explain how the circle method works. The starting point for the circle method is the orthogonality relation

$$\int_{\mathbb{T}} \psi(\alpha x)\mathrm{d}\alpha = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \in \mathbb{F}_q[t] \setminus \{0\}, \end{cases} \tag{3.0.2}$$

which expresses the characteristic function of $\{0\}$ on $\mathbb{F}_q[t]$ as an oscillatory integral over its Pontryagin dual $\mathbb{T} = K_\infty/\mathbb{F}_q[t]$. It will be convenient to have the following more general orthogonality at our disposal, where for any $R \in \mathbb{R}$ we adopt the notation $\widehat{R} = q^R$.

Lemma 3.0.2. *Let $N \in \mathbb{Z}$ and $x \in K_\infty$. Then*

$$\int_{|\alpha|<\widehat{N}^{-1}} \psi(\alpha x)\mathrm{d}\alpha = \begin{cases} \widehat{N}^{-1} & \text{if } |x| < \widehat{N}, \\ 0 & \text{else.} \end{cases}$$

Similarly, we have the following orthogonality relation for exponential sums.

Lemma 3.0.3. *For any $\alpha \in K_\infty$ and integer $N \geq 1$, we have*

$$\sum_{\substack{x\in\mathcal{O} \\ |x|<\widehat{N}}} \psi(\alpha x) = \begin{cases} \widehat{N} & \text{if } \|\alpha\| < \widehat{N}^{-1}, \\ 0 & \text{else.} \end{cases}$$

Poisson summation

In harmonic analysis over $\mathbb{R}$, one works with $S(\mathbb{R}^n)$, the space of Schwartz functions consisting of smooth functions with the property that all their derivatives decay rapidly. Thanks to the non-archimedean nature, over K_∞ one again works with a simpler notion.

Definition 3.0.4. Let $n \geq 1$ and $f \colon K_\infty^n \to \mathbb{C}$ be a function. We say that f is a *Schwartz–Bruhat function* if it is locally constant and has compact support. We denote by $S(K_\infty^n)$ the space of all Schwartz–Bruhat functions on K_∞^n.

Being a locally compact field, K_∞ is *self-dual*, meaning that we can identify it with its Pontryagin dual. For $f \in S(K_\infty^n)$, we can therefore define its *Fourier transform* to be a function $\mathcal{F}(f) \colon K_\infty^n \to \mathbb{C}$ given by

$$\mathcal{F}(f)(\boldsymbol{v}) = \int_{\mathbb{T}^n} f(\boldsymbol{x})\psi(\boldsymbol{x}\cdot\boldsymbol{v})\mathrm{d}\boldsymbol{x},$$

where $\boldsymbol{x}\cdot\boldsymbol{v}$ denotes the usual inner product. A key property of the Fourier transform is that it is a linear automorphism of $S(K_\infty^n)$. In addition, with our normalisation of the Haar measure and the choice of the character ψ, the Fourier inversion formula $\mathcal{F}(\mathcal{F}(f)) = f$ holds. A crucial input in Chapters 5 and 6 to analyse the exponential sums involved is the *Poisson summation formula*.

Proposition 3.0.5. *Let $f \in S(K_\infty^n)$. Then*

$$\sum_{\boldsymbol{x}\in\mathcal{O}^n} f(\boldsymbol{x}) = \sum_{\boldsymbol{v}\in\mathcal{O}^n} \mathcal{F}(f)(\boldsymbol{v}).$$

Remark. More generally, one can prove an adelic version of the Poisson summation formula over any global field. Over function fields this seemingly purely analytic statement implies the Riemann–Roch theorem, a fundamental result in the theory of algebraic curves.

CHAPTER 4

The circle method

The goal of this chapter is to introduce the circle method, which serves as the main technical apparatus for Chapters 5, 6 and 8. The circle method goes back to pioneering work of Hardy and Ramanujan [93] on the partition function. Its modern incarnation via exponential sums is based on the work of Vinogradov on Waring's problem in 1926 and found its role as a central player in the study of rational points through the agency of Davenport [62, 59, 60]. Within its century of existence, the circle method has seen spectacular applications in a diverse area of mathematics and was the subject of intense research. This ranges from a sheaf theoretic version by Browning and Sawin [41] to a motivic version suitable for stabilisation results in the Grothendieck ring of varieties by Bilu and Browning [13], both of which have applications to the study of moduli spaces of rational curves on hypersurfaces. In Chapter 8 we shall develop an extension of the circle method to show that under favourable circumstances these moduli spaces only have canonical singularities. A more thorough discussion and historical account of the circle method can be found in the classical books of Davenport [61] and Vaughan [200] or the more recent treatise of Browning [35].

In this chapter we explain the basic mechanism underlying the circle method by means of applying it to count rational points on hypersurfaces over $K = \mathbb{F}_q(t)$. In particular, we will present a proof of Weyl's inequality and indicate some of the refinements that have been developed in the context of the circle method. We will then end this chapter by giving a summary of results in the realm of rational points that have been obtained using the circle method.

Kubota [127] was the first to develop a version of the circle method over $\mathbb{F}_q(t)$ and applied it to Waring's problem. The setting we are dealing with here loosely follows work of Lee [131]. We will write $\mathcal{O} = \mathbb{F}_q[t]$ and define $\mathcal{O}^+$ to be the collection of monic polynomials in $\mathcal{O}$. We also make use of the notation of Chapter 3, so that K_∞ is the completion at the infinite place, $\mathbb{T} = K_\infty/\mathcal{O}$ is the Pontryagin dual of $\mathcal{O}$ and $\psi\colon K_\infty \to \mathbb{C}^\times$ is a non-trivial character that is non-trivial on $\mathbb{T}$, but not on any smaller compact subgroup. In addition, $|\cdot|\colon K_\infty \to \mathbb{R}_{\geq 0}$ denotes the norm induced by the infinite place. The igniting spark in the circle method is the orthogonality relation

$$\int_{\mathbb{T}} \psi(\alpha x)\mathrm{d}\alpha = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \in \mathcal{O} \setminus \{0\}, \end{cases} \tag{4.0.1}$$

that was already stated in (3.0.2). Given $F \in \mathcal{O}[x_1, \dots, x_n]$ homogeneous of degree d and

$P \in \mathcal{O}$, we can consider the counting function

$$N_F(P) = \#\{\boldsymbol{x} \in \mathcal{O}^n \colon |\boldsymbol{x}| < |P|, F(\boldsymbol{x}) = 0\}$$

as $|P| \to \infty$ that we already encountered in Section 2.3 in the context of Manin's conjecture. Strictly speaking $N_F(P)$ does not count rational points on the hypersurface $X = V(F) \subset \mathbb{P}^{n-1}$, but it is straightforward to pass from an asymptotic formula for $N_F(P)$ to one for

$$N_X(P) = \#\{\boldsymbol{x} \in \mathbb{P}^{n-1}(K) \colon F(\boldsymbol{x}) = 0, H(\boldsymbol{x}) < |P|\},$$

where $H \colon \mathbb{P}^{n-1}(K) \to \mathbb{R}_{\geq 0}$ denotes the naive height on projective space. Recall that if $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathcal{O}^n$ is a representative for $x \in \mathbb{P}^{n-1}(K)$ with $\gcd(x_1, \dots, x_n) = 1$, then $H(x) = |\boldsymbol{x}|$. In addition, x has precisely $|\mathbb{F}_q^\times| = q - 1$ such representatives. Suppose that

$$N_F(P) = c_F|P|^{n-d} + O(|P|^{n-d-\kappa})$$

for some $\kappa > 0$. Then we may use the function field analogue of the Möbius function $\mu \colon \mathbb{F}_q[t] \to \{\pm 1\}$ to detect the coprimality condition $\gcd(x_1, \dots, x_n) = 1$. This gives

$$\begin{aligned}
N_X(P) &= \frac{1}{q-1} \sum_{\substack{\boldsymbol{x} \in \mathcal{O}^n \setminus \{\boldsymbol{0}\} \colon |\boldsymbol{x}| < |P| \\ F(\boldsymbol{x}) = 0}} \sum_{\substack{k \in \mathcal{O}^+ \\ k | \gcd(x_1, \dots, x_n)}} \mu(k) \\
&= \frac{1}{q-1} \sum_{\substack{k \in \mathcal{O}^+ \\ |k| < |P|}} \mu(k) N_F(P/k) + O(|P|) \\
&= \frac{c_F}{q-1} |P|^{n-d} \sum_{\substack{k \in \mathcal{O}^+ \\ |k| < |P|}} \frac{\mu(k)}{|k|^{n-d}} + O(|P|^{n-d-\kappa}) \\
&= \frac{c_F}{(q-1)\zeta_K(n-d)} |P|^{n-d} + O(|P|^{n-d-\kappa}),
\end{aligned}$$

assuming $n - d > 1$ and where $\zeta_K(s) = (1 - q^{1-s})^{-1}$ is the zeta function of $K = \mathbb{F}_q(t)$. If $n = d + 1$, the situation is a bit more delicate, see for example the discussion before Corollary 2 [100] for quadrics. However, apart from the case $d = 2$ we are very far from producing an asymptotic in this case.

Applying (4.0.1) to $N_F(P)$, we immediately obtain the identity

$$N_F(P) = \int_{\mathbb{T}} S(\alpha) \mathrm{d}\alpha, \tag{4.0.2}$$

where

$$S(\alpha) := \sum_{\substack{\boldsymbol{x} \in \mathcal{O}^n \\ |\boldsymbol{x}| < |P|}} \psi(\alpha F(\boldsymbol{x})).$$

The basic strategy now is to divide $\mathbb{T}$ into a set of major arcs, on which we wish to asymptotically evaluate $S(\alpha)$, and a set of minor arcs, on which we want to show that $|S(\alpha)|$ is sufficiently small. The workhorse in the circle method is Weyl's inequality, for which we present a proof in the next section.

4.1 Weyl differencing

A versatile tool for estimating $S(\alpha)$ is through the process of *Weyl differencing*. The key idea behind Weyl differencing is to successively square $|S(\alpha)|$ and reduce the degree of the polynomial involved to end up with a linear exponential sum that we understand well. We will explain this when $F \in \mathcal{O}[x_1, \ldots, x_n]$ is a polynomial of degree d over $K = \mathbb{F}_q(t)$ with $\mathrm{char}(\mathbb{F}_q) > d$. We will see that the quality of our estimates only depends on the degree d homogeneous part of F, which we shall denote by F_d. We can thus write

$$F_d(\boldsymbol{x}) = \sum_{1 \le i_1, \ldots, i_d \le n} f_{i_1, \ldots, i_d} x_{i_1} \cdots x_{i_d},$$

where the coefficients $f_{i_1,\ldots,i_d} \in \mathcal{O}$ are symmetric in the indices. It will be convenient to introduce the multilinear forms

$$L_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) = d! \sum_{1 \le i_1, \ldots, i_{d-1} \le n} f_{i_1, \ldots, i_{d-1}, i} x_{i_1}^{(1)} \cdots x_{i_{d-1}}^{(d-1)} \tag{4.1.1}$$

for $i = 1, \ldots, n$, where $\boldsymbol{x}^{(j)} = (x_1^{(j)}, \ldots, x_n^{(j)})$ for $j = 1, \ldots, d-1$. Observe that we have the identity

$$\frac{\partial F_d}{\partial x_i}(\boldsymbol{x}) = \frac{1}{(d-1)!} L_i(\boldsymbol{x}, \ldots, \boldsymbol{x}). \tag{4.1.2}$$

Given $\boldsymbol{y} \in \mathcal{O}^n$, we define the differencing operator $\Delta_{\boldsymbol{y}} \colon K_\infty[x_1, \ldots, x_n] \to K_\infty[x_1, \ldots, x_n]$ via its action on $G \in K_\infty[x_1, \ldots, x_n]$ by

$$\Delta_{\boldsymbol{y}}(G)(\boldsymbol{x}) = G(\boldsymbol{x} + \boldsymbol{y}) - G(\boldsymbol{x}).$$

For $k \ge 2$ we define recursively $\Delta_{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}(G) = \Delta_{\boldsymbol{y}_k}(\Delta_{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}(G))$. The key property of the differencing operator is that if $\deg G = e$ and $\mathrm{char}(K) \nmid e$, then $\deg \Delta_{\boldsymbol{y}}(G) = e - 1$ for $\boldsymbol{y} \neq \boldsymbol{0}$. Squaring $|S(\alpha)|$ gives

$$|S(\alpha)|^2 = \sum_{|\boldsymbol{x}_1| < |P|} \sum_{|\boldsymbol{x}_2| < |P|} \psi(\alpha(F(\boldsymbol{x}_1) - F(\boldsymbol{x}_2))).$$

We can now make the change of variables $\boldsymbol{x} = \boldsymbol{x}_2$ and $\boldsymbol{y} = \boldsymbol{x}_1 - \boldsymbol{x}_2$. Then by the ultrametric property we have $|\boldsymbol{x}_2| < |P|$ if and only $|\boldsymbol{y}| < |P|$, so that

$$|S(\alpha)|^2 = \sum_{|\boldsymbol{y}| < |P|} \sum_{|\boldsymbol{x}| < |P|} \psi(\alpha(\Delta_{\boldsymbol{y}}(F)(\boldsymbol{x}))). \tag{4.1.3}$$

Inductively we obtain via an application of Cauchy–Schwarz

$$\begin{aligned}
|S(\alpha)|^{2^{d-1}} &\le \left(|P|^{(2^{d-2}-(d-1))n} \sum_{|\boldsymbol{x}^{(1)}| < |P|} \cdots \sum_{|\boldsymbol{x}^{(d-2)}| < |P|} \left| \sum_{|\boldsymbol{x}| < |P|} \psi(\alpha(\Delta_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(d-2)}}(F)(\boldsymbol{x}))) \right| \right)^2 \\
&\le |P|^{(2^{d-1}-d)n} \sum_{|\boldsymbol{x}^{(1)}| < |P|} \cdots \sum_{|\boldsymbol{x}^{(d-2)}| < |P|} \left| \sum_{|\boldsymbol{x}| < |P|} \psi(\alpha(\Delta_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(d-2)}}(F)(\boldsymbol{x}))) \right|^2 \\
&\le |P|^{(2^{d-1}-d)n} \sum_{|\boldsymbol{x}^{(1)}| < |P|} \cdots \sum_{|\boldsymbol{x}^{(d-1)}| < |P|} \sum_{|\boldsymbol{x}| < |P|} \psi(\alpha(\Delta_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(d-1)}}(F)(\boldsymbol{x}))).
\end{aligned}$$

We now observe that

$$\Delta_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(d-1)}}(F)(\boldsymbol{x}) = \sum_{i=1}^{n} L_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) x_i + \psi,$$

where ψ is a term that is independent of $\boldsymbol{x}$. In particular, after bringing absolute values inside, we have obtained the following result.

Lemma 4.1.1. *Let $F \in \mathcal{O}[x_1, \dots, x_n]$ be a polynomial of degree d. Then*

$$|S(\alpha)|^{2^{d-1}} \leq |P|^{(2^{d-1}-d)n} \sum_{|\boldsymbol{x}^{(1)}|<|P|} \cdots \sum_{|\boldsymbol{x}^{(d-1)}|<|P|} \left| \prod_{i=1}^{n} \sum_{|x|<|P|} \psi(\alpha L_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)})x) \right|.$$

Remark. The factor $d!$ in the definition of L_i is the reason for our assumption $\mathrm{char}(K) > d$. Indeed, if $\mathrm{char}(K) \leq d$, then $L_i = 0$ and so the inequality of Lemma 4.1.1 reduces to the trivial estimate $|S(\alpha)|^{2^{d-1}} \leq |P|^{2^{d-1}n}$.

The crucial point at this stage is that each expression in the sum above is linear in x and linear exponential sums are easy to understand. Let $0 \leq s \leq \deg P$ be an integer and define

$$K_s(P) = \# \left\{ (\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) \in \mathcal{O}^{(d-1)n} \colon \begin{array}{l} |\boldsymbol{x}^{(j)}| < |P|q^{-s} \text{ for } 1 \leq j \leq d-1, \\ \left\| \alpha L_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) \right\| < |P|^{-1}q^{-s(d-1)} \end{array} \right\}.$$

Then upon combining Lemma 4.1.1 with the orthogonality relation of Lemma 3.0.3, we immediately arrive at the following result.

Corollary 4.1.2. *We have*

$$|S(\alpha)|^{2^{d-1}} \leq |P|^{(2^{d-1}-(d-1))n} K_0(P).$$

To estimate $K_0(P)$, Davenport came up with an ingenious idea that relates $K_0(P)$ to $K_s(P)$, where s is eventually chosen in such a way that the Diophantine inequality

$$\left\| \alpha L_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) \right\| < |P|^{-1}q^{-s(d-1)}$$

forces $L_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)})$ to vanish. The key ingredient is his "shrinking lemma", whose function field version is for example proven in Lemma 6.4 of [41].

Lemma 4.1.3. *Let $\gamma \in \mathrm{Mat}_{n\times n}(K_\infty)$ be a symmetric matrix. Given integers a, c, s such that $c > 0$ and $s \geq 0$, let $N_{\gamma,a,c}$ denote the number of $\boldsymbol{x} \in \mathcal{O}^n$ such that $\|\gamma \boldsymbol{x}\| < q^{-c}$ and $|\boldsymbol{x}| < q^a$. Then*

$$\frac{N_{\gamma,a,c}}{N_{\gamma,a-s,c+s}} \leq q^{ns+n\max\{\lfloor \frac{a-c}{2} \rfloor, 0\}}.$$

Corollary 4.1.4. *Suppose that $0 \leq s \leq \deg P$. Then*

$$K_0(P) \leq q^{s(d-1)n} K_s(P).$$

Proof. We can fix all but one of $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}$ in the definition of $K_0(P)$ and apply Lemma 4.1.3 to each $\boldsymbol{x}^{(i)}$ individually for $1 \leq i \leq d-1$. At the ith step, we have $c = \deg P + s(i-1)$ and $a = \deg P$, so that for any $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(i-1)}, \boldsymbol{x}^{(i+1)}, \dots, \boldsymbol{x}^{(d-1)} \in \mathcal{O}^n$, we have

$$\begin{aligned} &\#\{\boldsymbol{x}^{(i)} \in \mathcal{O}^n \colon |\boldsymbol{x}^{(i)}| < |P|, \left\|\alpha L_j(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)})\right\| < |P|^{-1}q^{-s(i-1)} \text{ for } 1 \leq j \leq n\} \\ &\leq q^{sn} \#\{\boldsymbol{x}^{(i)} \in \mathcal{O}^n \colon |\boldsymbol{x}^{(i)}| < |P|q^{-s}, \left\|\alpha L_j(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)})\right\| < |P|^{-1}q^{-si} \text{ for } 1 \leq j \leq n\} \end{aligned}$$

by Lemma 4.1.3. Indeed, $a - c = -s(i-1) \leq 0$ and the matrix underlying the linear map $\boldsymbol{x}^{(i)} \mapsto \alpha(L_1(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}), \dots, L_n(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}))$ is symmetric since the coefficients in the definition (4.1.1) of the L_i's are symmetric in the indices. □

Our goal is now to apply Corollary 4.1.4 and choose s in such a way that the Diophantine inequality in the definition of $n_s(P)$ forces $L_i(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)})=0$ for $1\leq i\leq n$. Before doing so, we need an elementary estimate for the number of integral points on affine varieties.

Lemma 4.1.5. *Let $W\subset\mathbb{A}^N$ be an equi-dimensional variety of degree D. Then*

$$\#\{\boldsymbol{x}\in W(\mathcal{O})\colon |\boldsymbol{x}|<\widehat{Z}\}\ll \widehat{Z}^{\dim W},$$

where the implied constant only depends on N and D.

Proof. We proceed by induction on $\dim W$. If $\dim W=0$, then $W(\mathcal{O})$ is a union of at most D points, so that the estimate holds trivially. Suppose now that $\dim W\geq 1$. This implies that there exists an index $1\leq i_0\leq N$ such that the hyperplane $\{x_{i_0}=a\}$ intersects W properly for any $a\in\overline{\mathbb{F}_q(t)}$. Thus we have

$$\#\{\boldsymbol{x}\in W(\mathcal{O})\colon |\boldsymbol{x}|<\widehat{Z}\}\leq \sum_{\substack{a\in\mathcal{O}\\ |a|<\widehat{Z}}}\#\{\boldsymbol{x}\in W(\mathcal{O})\colon x_{i_0}=a, |\boldsymbol{x}|<\widehat{Z}\}.$$

The intersection $W\cap\{x_{i_0}=a\}$ has dimension at most $\dim W-1\geq 0$ and degree at most D, because W is equi-dimensional. Thus we can apply the induction hypothesis to deduce that

$$\#\{\boldsymbol{x}\in W(\mathcal{O})\colon |\boldsymbol{x}|<\widehat{Z}\}\ll \sum_{\substack{a\in\mathcal{O}\\ |a|<\widehat{Z}}}\widehat{Z}^{\dim W-1}\leq \widehat{Z}^{\dim W}$$

as desired. □

Let σ denote the dimension of the singular locus of the affine hypersurface $V(F_d)\subset\mathbb{A}^n$. In particular, $\sigma=0$ if and only if the projective hypersurface defined by $F_d=0$ inside $\mathbb{P}^{n-1}$ is smooth.

Lemma 4.1.6. *Let $Z\subset\mathbb{A}^{(d-1)n}$ be the variety defined by $L_i(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)})=0$ for $i=1,\dots,n$ and suppose that $\mathrm{char}(K)>d$. Then any irreducible component of Z has dimension at most $(d-2)n+\sigma$.*

Proof. Let $\Delta\subset\mathbb{A}^{n(d-1)}$ be the diagonal defined by

$$\Delta=\{(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)})\in\mathbb{A}^{n(d-1)}\colon \boldsymbol{x}^{(1)}=\dots=\boldsymbol{x}^{(d-1)}\},$$

so that Δ is irreducible of dimension n. If $\underline{\boldsymbol{x}}=(\boldsymbol{x},\dots,\boldsymbol{x})$ belongs to $\Delta\cap Z$, then as we assume $\mathrm{char}(K)>d$ by (4.1.2) we must have $\nabla F(\boldsymbol{x})=0$. This forces $\boldsymbol{x}$ to lie on the singular locus of the affine hypersurface defined by $F=0$, which by assumption has dimension σ. Thus we have

$$\sigma\geq\dim\Delta\cap Z\geq\dim\Delta+\dim Z-(d-1)n$$

and hence $\dim Z\leq(d-2)n+\sigma$. □

We now have everything at hand to establish the main estimate for $S(\alpha)$. For a polynomial $G\in K_\infty[x_1,\dots,x_n]$, let H_G be the maximum of the absolute values of its coefficients.

Proposition 4.1.7. *Let $F \in \mathcal{O}[x_1, \ldots, x_n]$ be of degree d and let F_d be its homogeneous degree d part. Suppose that $\alpha = a/r + \theta$, where $\alpha, \theta \in \mathbb{T}$ and $a, r \in \mathcal{O}$ are coprime with $|a| < |r|$. If σ is the dimension of the singular locus of the hypersurface $V(F_d) \subset \mathbb{A}^n$ and $\operatorname{char}(K) > d$, then*

$$|S(\alpha)| \ll |P|^n \left(\frac{|P| + H_{F_d}|P^d\theta r| + q^{d-1}|r|}{|P|^d} + \frac{q^{d-1}}{q^{d-1}H_{F_d}|r| + |r\theta P^d|} \right)^{\frac{n-\sigma}{(d-1)2^{d-1}}},$$

where the implied constant only depends on n and d, but not on q.

Proof. Let $0 \le s \le \deg P$ be an integer satisfying

(1) $q^{-(d-1)}H_{F_d}|r\theta P^{d-1}| < q^{s(d-1)}$,

(2) $|rP^{-1}| < q^{s(d-1)}$

(3) (i) $q^{-(d-1)}H_F|r|^{-1}|P|^{d-1} < q^{s(d-1)}$ or
    (ii) $|r\theta P|^{-1} < q^{s(d-1)}$.

Then by Corollaries 4.1.2 and 4.1.4 we have

$$|S(\alpha)|^{2^{d-1}} \le |P|^{(2^{d-1}-(d-1))n} q^{s(d-1)n} K_s(P).$$

Suppose that $(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) \in \mathcal{O}^{n(d-1)}$ is counted by $K_s(P)$. Let $1 \le i \le n$ and set $m = L_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)})$. Our goal is to show that $m = 0$. First, observe that $H_{L_i} \le H_{F_d}$ for $1 \le i \le n$. Second, by the definition of $K_s(P)$ we have $|\boldsymbol{x}^{(j)}| \le |P|q^{-s-1}$ for $1 \le j \le d-1$, so that by the ultrametric property $|m| \le H_{F_d}(|P|q^{-s-1})^{d-1}$. In particular, (1) implies that

$$|\theta m| \le H_{F_d}|\theta|(|P|q^{-s-1})^{d-1} < |r|^{-1} \le 1$$

and hence $\|\theta m\| = |\theta m|$. Therefore, by the definition of $K_s(P)$, we obtain

$$\begin{aligned} \left\| \frac{am}{r} \right\| &\le \max\{\|\theta m\|, \|\alpha m\|\} \\ &< \max\{|r|^{-1}, |P|^{-1}q^{-s(d-1)}\} \\ &\le |r|^{-1} \end{aligned}$$

by (2). This implies that $r \mid m$. If the first alternative of (3) holds, then $|m| < |r|$ and thus $m = 0$. If the second alternative is true, then $r \mid m$ implies

$$|\theta m| = \|\theta m\| = \|\alpha m\| < |P|^{-1}q^{-s(d-1)}$$

and so $|m| < |\theta P|^{-1}q^{-s(d-1)} \le |r|$, which again implies $m = 0$. The choice

$$1 + \left\lfloor \frac{\log_q \max\{q^{-(d-1)}, q^{-(d-1)}H_{F_d}|r\theta P^{d-1}|, |rP^{-1}|, |r|^{-1}\min\{q^{-(d-1)}H_{F_d}|P|^{d-1}, |\theta P|^{-1}\}\}}{d-1} \right\rfloor$$

for s satisfies (1)–(3) and $s \ge 0$. In particular, every $(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)})$ counted by $K_s(P)$ must satisfy $L_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) = 0$ for $i = 1, \ldots, n$. Assuming $s \le \deg P$ we obtain

$$|S(\alpha)|^{2^{d-1}} \le |P|^{(2^{d-1}-(d-1))n} q^{s(d-1)n} \#\{\underline{\boldsymbol{x}} \in Z(\mathcal{O}) \colon |\underline{\boldsymbol{x}}| < |P|q^{-s}\} \ll |P|^{2^{d-1}n} \left(\frac{q^s}{|P|} \right)^{n-\sigma} \tag{4.1.4}$$

by Lemmas 4.1.5 and 4.1.6. If $s > \deg P$, then since $\sigma \leq n$, this estimate is worse than the trivial bound and so the estimate holds in any case. The term q^s is at most

$$q\left(\max\{q^{-(d-1)}, q^{-(d-1)}H_{F_d}|r\theta P^{d-1}|, |rP^{-1}|, |r|^{-1}\min\{q^{-(d-1)}H_{F_d}|P|^{d-1}, |\theta P|^{-1}\}\right)^{\frac{1}{d-1}}$$
$$\leq |P|\left(\frac{|P| + H_{F_d}|r\theta P^d| + q^{d-1}|r|}{|P|^d} + \frac{q^{d-1}}{q^{d-1}H_{F_d}|r| + |r\theta P^d|}\right)^{\frac{1}{d-1}},$$

which upon inserting into (4.1.4) completes the proof. □

For some applications, such as establishing weak approximation, it is convenient to have an estimate available for a slight generalisation of the sum $S(\alpha)$. Let $\boldsymbol{x}_0 \in K_\infty^n$, $M \in \mathcal{O} \setminus \{0\}$ and $\boldsymbol{b} \in \mathcal{O}^n$ be given. For a parameter $L \geq 0$ and $\alpha \in \mathbb{T}$, we then define

$$\widetilde{S}(\alpha) = \sum_{\substack{\boldsymbol{x}\in\mathcal{O}^n \\ |\boldsymbol{x}/P - \boldsymbol{x}_0| < \widehat{L}^{-1} \\ \boldsymbol{x}\equiv \boldsymbol{b} \bmod M}} \psi(\alpha F(\boldsymbol{x})).$$

Corollary 4.1.8. *For $|P| > |M|\widehat{L}$, under the same assumptions as in Proposition 4.1.7, we have*

$$|\widetilde{S}(\alpha)| \ll \left(\frac{|P|}{\widehat{L}|M|}\right)^n \times \left(\frac{(\widehat{L}|M|)^{d-1}|P| + H_{F_d}|M|^d|r\theta P^d| + q^{d-1}|r|(\widehat{L}|M|)^d}{|P|^d}\right.$$
$$\left. + \frac{q^{d-1}(\widehat{L}|M|)^d}{q^{d-1}H_{F_d}(\widehat{L}|M|^2)^d|r| + |r\theta|}\right)^{\frac{n-\sigma}{(d-1)2^{d-1}}},$$

where the implied constant only depends only on n and d, but not on q.

Proof. Without loss of generality, we may assume that $|\boldsymbol{b}| < |M|$. We can write $P\boldsymbol{x}_0 = \boldsymbol{y}_0 + \boldsymbol{z}_0$ for some $\boldsymbol{y}_0 \in \mathbb{T}^n$ and $\boldsymbol{z}_0 \in \mathcal{O}^n$. As $|\boldsymbol{b}| < |M|$, we then have $|\boldsymbol{x}/P - \boldsymbol{x}_0| < \widehat{L}^{-1}$ and $\boldsymbol{x} \equiv \boldsymbol{b} \bmod M$ if and only if $\boldsymbol{x} = M\boldsymbol{y} + \boldsymbol{z}_0 + \boldsymbol{b}$ for some $\boldsymbol{y} \in \mathcal{O}^n$ with $|\boldsymbol{y}| < \frac{|P|}{|M|\widehat{L}}$. Upon defining $G(\boldsymbol{y}) = F(M\boldsymbol{y} + \boldsymbol{z}_0 + \boldsymbol{b})$, we thus have

$$\widetilde{S}(\alpha) = \sum_{|\boldsymbol{y}| < \frac{|P|}{|M|\widehat{L}}} \psi(\alpha G(\boldsymbol{y})). \tag{4.1.5}$$

The polynomial $G(\boldsymbol{y})$ has coefficients in $\mathcal{O}$ and its homogeneous degree d part is $M^d F_d(\boldsymbol{y})$, so that $H_{G_d} = H_{F_d}|M|^d$. In particular, the corollary follows from Proposition 4.1.7 applied to the sum on the right hand side of (4.1.5). □

A variant of the squaring and differencing process explained above is *van der Corput differencing.* Let $\mathcal{H} \subset \mathcal{O}^n$ be a finite subset. For a function $f\colon K_\infty^n \to \mathbb{C}$ that is supported on the set $\{\boldsymbol{x} \in K_\infty^n \colon |\boldsymbol{x}| < |P|\}$, the starting point is the identity

$$\#\mathcal{H}\sum_{\boldsymbol{x}\in\mathcal{O}^n} f(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathcal{H}}\sum_{\boldsymbol{x}\in\mathcal{O}^n} f(\boldsymbol{x}+\boldsymbol{h}) = \sum_{\boldsymbol{x}\in\mathcal{O}^n}\sum_{\boldsymbol{h}\in\mathcal{H}} f(\boldsymbol{x}+\boldsymbol{h}).$$

Applying Cauchy–Schwarz yields

$$\#\mathcal{H}^2\left|\sum_{\boldsymbol{x}\in\mathcal{O}^n} f(\boldsymbol{x})\right|^2 \leq |P|^n \sum_{\boldsymbol{h}\in\mathcal{H}} N(\boldsymbol{h}) \sum_{\boldsymbol{x}\in\mathcal{O}^n} f(\boldsymbol{x}+\boldsymbol{h})\overline{f(\boldsymbol{x})},$$

where $N(\boldsymbol{h}) = \#\{(\boldsymbol{h}_1, \boldsymbol{h}_2) \in \mathcal{H}^2 \colon \boldsymbol{h}_1 - \boldsymbol{h}_2 = \boldsymbol{h}\}$. Taking

$$f(\boldsymbol{x}) = \begin{cases} \psi(\alpha F(\boldsymbol{x})) & \text{if } |\boldsymbol{x}| < |P|, \\ 0 & \text{else,} \end{cases}$$

thus gives

$$|S(\alpha)|^2 \leq |P|^n \# \mathcal{H}^{-2} \sum_{\boldsymbol{h} \in \mathcal{H}} N(\boldsymbol{h}) \sum_{|\boldsymbol{x}|, |\boldsymbol{x}+\boldsymbol{h}| < |P|} \psi(\alpha(F(\boldsymbol{x}+\boldsymbol{h}) - F(\boldsymbol{x}))).$$

Note that by choosing $\mathcal{H} = \{\boldsymbol{x} \in \mathcal{O}^n \colon |\boldsymbol{x}| < |P|\}$ we can recover the inequality (4.1.3) obtained by the standard squaring and differencing process, so nothing has been lost. However, the advantage is that one can freely choose the set $\mathcal{H}$, thereby allowing for greater flexibility. Heath-Brown [105] has introduced *averaged* van der Corput differencing that allowed him to establish that any cubic form over $\mathbb{Q}$ in at least 14 variables possesses a non-trivial zero. This improves upon an old result of Davenport [60], who showed that having at least 16 variables is sufficient for the existence of a non-trivial zero.

4.2 Major and minor arcs

Having established a sufficiently strong estimate for $|S(\alpha)|$ in Proposition 4.1.7, we will now proceed to explain the main steps in the circle method. It is clear that the quality of the estimate in Proposition 4.1.7 is controlled by "how well" α can be approximated by a fraction $a/r \in \mathbb{F}_q(t)$. The following result is the function field analogue of Dirichlet's approximation theorem and quantifies the quality of a rational approximation. A proof can be found in [35, Lemma 5.7].

Proposition 4.2.1. *Let $Q \geq 1$ be an integer and $\alpha \in \mathbb{T}$. Then there exist $a \in \mathcal{O}$, $r \in \mathcal{O}^+$ such that $|a| < |r| \leq \widehat{Q}$, $\gcd(a,r) = 1$ and*

$$\left| \alpha - \frac{a}{r} \right| < \frac{1}{|r| \widehat{Q}}.$$

Remark. An important difference to characteristic 0 is that Dirichlet's approximation already provides us with an exact decomposition of the unit interval. What we mean by this is that the intervals in the union

$$\mathbb{T} = \bigsqcup_{\substack{r \in \mathcal{O}^+ \\ |r| \leq \widehat{Q}}} \bigsqcup_{\substack{|a| < |r| \\ \gcd(a,r)=1}} \{\alpha \in \mathbb{T} \colon |\alpha - a/r| < |r|^{-1} \widehat{Q}^{-1}\}$$

are non-overlapping thanks to the ultrametric property of the norm, for any positive integer Q. This analogue of a Farey dissection drastically simplifies the possibility of exhibiting extra cancellation when averaging $S(a/r + \theta)$ over the numerators a or even the denominators r. We will explain this in more detail in Section 4.4

Let $\alpha \in \mathbb{T}$ and suppose that $\alpha = a/r + \theta$, where $a, r \in \mathcal{O}$ are coprime, $|a| < |r|$ and r is monic. Dirichlet's approximation theorem guarantees the existence of such a, r and θ with $|r| \leq |P|^{d/2}$ and $|\theta| < |r|^{-1} |P|^{-d/2}$. In this case, Proposition 4.1.7 simplifies to

$$|S(\alpha)| \ll |P|^n |r|^{-\frac{n-\sigma}{(d-1)2^{d-1}}} \max\{1, |\theta P^d|\}^{-\frac{n-\sigma}{(d-1)2^{d-1}}}. \tag{4.2.1}$$

In particular, we see that we get a power saving over the trivial bound $|S(\alpha)| \leq |P|^n$ if $|r| \geq |P|^\Delta$ or $|\theta| \geq |P|^{\Delta-d}$ for some fixed $\Delta > 0$. Motivated by this, for a small parameter $0 < \Delta < 1/2$ we define the *major arcs* to be

$$\mathfrak{M}(\Delta) = \bigcup_{\substack{r \in \mathcal{O}^+ \\ |r| \leq |P|^\Delta}} \bigcup_{\substack{a \in \mathcal{O} \\ |a| < |r| \\ \gcd(a,r)=1}} \left\{\alpha \in \mathbb{T} \colon |\alpha - a/r| < |P|^{\Delta-d}\right\} \tag{4.2.2}$$

and our goal is to evaluate $S(\alpha)$ asymptotically for $\alpha \in \mathfrak{M}(\Delta)$. The exact value of Δ is of no importance and only affects the quality of the error term.

The *minor arcs* are defined to be the complement $\mathfrak{m}(\Delta) = \mathbb{T} \setminus \mathfrak{M}(\Delta)$ and we wish to show that the contribution from them makes up an acceptable error term.

Proposition 4.2.2. *Assuming* $n > \sigma + (d-1)2^d$ *and* $\mathrm{char}(K) > d$, *we have*

$$\int_{\mathfrak{m}(\Delta)} |S(\alpha)| \mathrm{d}\alpha \ll |P|^{n-d-\kappa}$$

for some $\kappa > 0$.

Proof. Let $\alpha \in \mathfrak{m}(\Delta)$. Then by Proposition 4.2.1 there exist coprime $a, r \in \mathcal{O}$ such that $|a| < |r| \leq |P|^{d/2}$, r is monic and $|\alpha - a/r| < |r|^{-1}|P|^{-d/2}$. As $\alpha \in \mathfrak{m}(\Delta)$, we must either have $|r| > |P|^\Delta$ or $|r| \leq |P|^\Delta$ and $|\alpha - a/r| \geq |P|^{\Delta-d}$. It follows that

$$\int_{\mathfrak{m}(\Delta)} |S(\alpha)| \leq \Sigma_1 + \Sigma_2,$$

where

$$\Sigma_1 = \sum_{\substack{r \in \mathcal{O}^+ \\ |P|^\Delta < |r| \leq |P|^{d/2}}} \sum_{\substack{|a| < |r| \\ \gcd(a,r)=1}} \int_{|\theta| < |r|^{-1}|P|^{-d/2}} |S(a/r + \theta)| \mathrm{d}\theta$$

and

$$\Sigma_2 = \sum_{\substack{r \in \mathcal{O}^+ \\ |r| \leq |P|^\Delta}} \sum_{\substack{|a| < |r| \\ \gcd(a,r)=1}} \int_{|P|^{\Delta-d} \leq |\theta| < |r|^{-1}|P|^{-d/2}} |S(a/r + \theta)| \mathrm{d}\theta.$$

In particular, under our assumption $n > \sigma + (d-1)2^d$, by (4.2.1) we have

$$\begin{aligned}
\Sigma_1 &\ll |P|^n \sum_{|P|^\Delta < |r| \leq |P|^{d/2}} |r|^{1 - \frac{n-\sigma}{(d-1)2^{d-1}}} \int_{|\theta| < |r|^{-1}|P|^{-d/2}} \max\{1, |\theta P^d|\}^{-\frac{n-\sigma}{(d-1)2^{d-1}}} \mathrm{d}\theta \\
&\ll |P|^{n-d} \sum_{|P|^\Delta < |r|} |r|^{1 - \frac{n-\sigma}{(d-1)2^{d-1}}} \\
&\ll |P|^{n-d-\delta},
\end{aligned}$$

where $\delta = -\Delta(2 - \frac{n-\sigma}{(d-1)2^{d-1}}) > 0$. Similarly, for Σ_2 we have

$$\begin{aligned}
\Sigma_2 &\ll |P|^n \sum_{|r| \leq |P|^\Delta} |r|^{1 - \frac{n-\sigma}{(d-1)2^{d-1}}} \int_{|P|^{\Delta-d} \leq |\theta| < |r|^{-1}|P|^{-d/2}} |\theta P^d|^{-\frac{n-\sigma}{(d-1)2^{d-1}}} \mathrm{d}\theta \\
&\ll |P|^{n-d-\delta'},
\end{aligned}$$

where $\delta' = -\Delta(1 - \frac{n-\sigma}{(d-1)2^{d-1}}) > 0$. □

We now commence with our analysis of the major arcs. Let $\alpha = a/r + \theta \in \mathfrak{M}(\Delta)$, so that $a, r \in \mathcal{O}$ are coprime satisfying $|a| < |r| \leq |P|^{\Delta}$ with r monic and $|\theta| < |P|^{\Delta - d}$. Any $\boldsymbol{x} \in \mathcal{O}^n$ can be written uniquely as $\boldsymbol{x} = r\boldsymbol{y} + \boldsymbol{z}$ with $\boldsymbol{y}, \boldsymbol{z} \in \mathcal{O}^n$ satisfying $|\boldsymbol{y}| < |P||r|^{-1}$ and $|\boldsymbol{z}| < |r|$. Thus

$$S(\alpha) = \sum_{|\boldsymbol{z}|<|r|} \psi\left(\frac{aF(\boldsymbol{z})}{r}\right) \sum_{|\boldsymbol{y}|<\frac{|P|}{|r|}} \psi(\theta F(r\boldsymbol{y} + \boldsymbol{z})). \tag{4.2.3}$$

Using Taylor expansion, we see that

$$\begin{aligned} |\theta(F(r\boldsymbol{y}+\boldsymbol{z}) - F(r\boldsymbol{y}))| &< H_F |P|^{\Delta-d} \max_{1\leq i\leq d} \max\{|\boldsymbol{z}|^i (|r||\boldsymbol{y}|)^{d-i}\} \\ &< H_F |P|^{\Delta-d} \max_{1\leq i\leq d} \{|r|^i |P|^{d-i}\} \\ &\leq H_F |P|^{\Delta-d} \max_{1\leq i\leq d} \{|P|^{d-i(1-\Delta)}\} \\ &\leq H_F |P|^{2\Delta-1}, \end{aligned} \tag{4.2.4}$$

where we used that the maximum in the penultimate line occurs at $i = 1$, because $\Delta < 1/2$. In particular, as $\Delta < 1/2$, we have $|\theta(F(r\boldsymbol{y}+\boldsymbol{z}) - F(r\boldsymbol{y}))| < q^{-1}$ for $|P|$ sufficiently large and hence $\psi(\theta F(r\boldsymbol{y}+\boldsymbol{z})) = \psi(\theta F(r\boldsymbol{y}))$ for $\alpha \in \mathfrak{M}(\Delta)$. Taking $a = 0$ and $r = 1$ in (4.2.3) shows that

$$S(\theta) = |r|^n \sum_{|\boldsymbol{y}|<\frac{|P|}{|r|}} \psi(\theta F(r\boldsymbol{y})).$$

If we define

$$S_{a,r} = \sum_{|\boldsymbol{x}|<|r|} \psi\left(\frac{aF(\boldsymbol{x})}{r}\right)$$

for $a \in \mathcal{O}$, we can thus conclude

$$S(\alpha) = |r|^{-n} S_{a,r} S(\theta) \tag{4.2.5}$$

for $\alpha \in \mathfrak{M}(\Delta)$.

The next step is to replace $S(\theta)$ by an integral. Using the fact that $|\theta| < |P|^{\Delta-d}$, an argument parallel to (4.2.4) shows that $\psi(\theta F(\boldsymbol{y})) = \psi(\theta F(\boldsymbol{y}+\boldsymbol{z}))$ for any $\boldsymbol{z} \in \mathbb{T}^n$. Integrating $\boldsymbol{z}$ over $\mathbb{T}^n$ implies

$$S(\theta) = \int_{\mathbb{T}^n} \sum_{|\boldsymbol{y}|<|P|} \psi(\theta F(\boldsymbol{y}+\boldsymbol{z}))\mathrm{d}\boldsymbol{z} = \int_{|\boldsymbol{x}|<|P|} \psi(\theta F(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \tag{4.2.6}$$

upon making the change of variables $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$. In addition, after two obvious changes of variables we have

$$\begin{aligned} \int_{|\theta|<|P|^{\Delta-d}} \int_{|\boldsymbol{x}|<|P|} \psi(\theta F(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}\theta &= |P|^n \int_{|\theta|<|P|^{\Delta-d}} \int_{\mathbb{T}^n} \psi(\theta P^d F(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}\theta \\ &= |P|^{n-d} \int_{|\theta|<|P|^{\Delta}} \int_{\mathbb{T}^n} \psi(\theta F(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}\theta, \end{aligned} \tag{4.2.7}$$

where we used that $F(P\boldsymbol{x}) = P^d F(\boldsymbol{x})$, which holds because F is homogeneous of degree d.

For $r \in \mathcal{O}^+$, let us define the complete exponential sum

$$S_r = \sum_{\substack{|a|<|r| \\ \gcd(a,r)=1}} S_{a,r}$$

and for $\theta \in K_\infty$ the integral

$$I(\theta) = \int_{\mathbb{T}^n} \psi(\theta F(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}.$$

As explained in Remark 4.2 the union in the definition (4.2.2) is disjoint, so that combining (4.2.5), (4.2.6) and (4.2.7) hands us the identity

$$\int_{\mathfrak{M}(\Delta)} S(\alpha)\mathrm{d}\alpha = |P|^{n-d}\mathfrak{S}(|P|^{\Delta})\sigma_\infty(|P|^{\Delta}), \tag{4.2.8}$$

where for $R > 0$,

$$\mathfrak{S}(R) = \sum_{\substack{r \in \mathcal{O}^+ \\ |r| \leq R}} |r|^{-n} S_r$$

is the truncated *singular series* and

$$\sigma_\infty(R) = \int_{|\theta|<R} I(\theta)\mathrm{d}\theta$$

is the truncated *singular integral*.

Observe that taking $|P| = |r|$ and $\theta = 0$ in Proposition 4.1.7 hands us the estimate

$$|S_r| \ll |r|^{1+n-\frac{n-\sigma}{(d-1)2^{d-1}}}. \tag{4.2.9}$$

In addition, the identity (4.2.6) may be used to derive the upper bound

$$I(\theta) \ll (1+|\theta|)^{-\frac{n-\sigma}{(d-1)2^{d-1}}}. \tag{4.2.10}$$

We can use (4.2.9) to deduce that the completed singular series

$$\mathfrak{S} = \sum_{r \in \mathcal{O}^+} |r|^{-n} S_r$$

converges absolutely for $n > \sigma + (d-1)2^d$ and that

$$|\mathfrak{S} - \mathfrak{S}(|P|^{\Delta})| \ll |P|^{-\delta} \tag{4.2.11}$$

for some $\delta > 0$. Similarly, (4.2.10) shows that the completed singular integral

$$\sigma_\infty = \int_{K_\infty} I(\theta)\mathrm{d}\theta$$

converges absolutely if $n > \sigma + (d-1)2^d$ and that

$$|\sigma_\infty - \sigma_\infty(|P|^{\Delta})| \ll |P|^{-\delta'} \tag{4.2.12}$$

for some $\delta' > 0$. Combining (4.2.8) with (4.2.11) and (4.2.12) therefore yields the following result.

Proposition 4.2.3. *Let $n > \sigma + (d-1)2^d$ and suppose that $\mathrm{char}(K) > d$. Then for $0 < \Delta < 1/2$ we have*

$$\int_{\mathfrak{M}(\Delta)} S(\alpha)\mathrm{d}\alpha = \sigma_\infty \mathfrak{S}|P|^{n-d} + O(|P|^{n-d-\kappa})$$

for some $\kappa > 0$.

In particular, Propositions 4.2.2 and 4.2.3 in conjunction with the identity (4.0.2) deliver us the sought-after asymptotic formula

$$N_F(P) = \sigma_\infty \mathfrak{S}|P|^{n-d} + O(|P|^{n-d-\kappa}).$$

To understand when the main term in the asymptotic formula is positive, we have to take a closer look at the constants σ_∞ and $\mathfrak{S}$. We begin with the latter. A straightforward application of the Chinese remainder theorem shows that

$$S_{r_1 r_2} = S_{r_1} S_{r_2}$$

whenever $r_1, r_2 \in \mathcal{O}^+$ are coprime. In particular, the singular series factors into the Euler product

$$\mathfrak{S} = \prod_{\varpi} \sigma_\varpi,$$

where the product runs over all monic irreducible polynomials $\varpi \in \mathcal{O}^+$ and σ_ϖ denotes the ϖ-adic density

$$\sigma_\varpi = \sum_{k\geq 0} |\varpi|^{-kn} S_{\varpi^k}.$$

Note that for $k \geq 1$, we have

$$\begin{aligned} S_{\varpi^k} &= \sum_{|\boldsymbol{x}|<|\varpi|^k} \left(\sum_{|a|<|\varpi|^k} \psi\left(\frac{aF(\boldsymbol{x})}{\varpi^k}\right) - \sum_{|a|<|\varpi|^{k-1}} \psi\left(\frac{aF(\boldsymbol{x})}{\varpi^{k-1}}\right)\right) \\ &= |\varpi|^k v(\varpi^k) - |\varpi|^{k-1+n} v(\varpi^{k-1}), \end{aligned}$$

where for $r \in \mathcal{O}^+$ we have defined

$$v(r) = \#\{\boldsymbol{x} \in (\mathcal{O}/r\mathcal{O})^n \colon F(\boldsymbol{x}) \equiv 0 \bmod r\}.$$

Therefore, for any $m \geq 1$ we have a telescoping sum

$$\begin{aligned} \sum_{k=0}^{m} |\varpi|^{-kn} S_{\varpi^k} &= 1 + \sum_{k=1}^{m} |\varpi|^{k(1-n)} v(\varpi^k) - \sum_{k=1}^{m} |\varpi|^{(k-1)(1-n)} v(\varpi^{k-1}) \\ &= |\varpi|^{m(1-n)} v(\varpi^m) \end{aligned}$$

which shows that

$$\sigma_\varpi = \lim_{k\to\infty} |\varpi|^{k(1-n)} v(\varpi^k).$$

The exponent $1-n$ should be understood as a normalising factor which comes from the fact that the dimension of $V(F) \subset \mathbb{A}^n$ is $n-1$. We can interpret σ_ϖ thus as measuring the density of ϖ-adic solutions to $F(\boldsymbol{x}) = 0$. In particular, with the help of Hensel's lemma one can show that $\sigma_\varpi > 0$ providing there exists a non-singular solution $\boldsymbol{x} \in \mathcal{O}_\varpi^n$ to $F(\boldsymbol{x}) = 0$, where $\mathcal{O}_\varpi$ is the ring of integers of K_ϖ. If this holds for every monic irreducible polynomial ϖ, then in view of the absolute convergence of $\mathfrak{S}$ this is enough to guarantee that $\mathfrak{S} > 0$. If n is small, then in some cases the truncated singular series $\mathfrak{S}(R)$ grows like a power of log. This is related to the logarithmic factor in Manin's conjecture and may be interpreted roughly as saying that a large Picard rank leads to "many" solutions modulo primes.

Turning to the singular integral, the orthogonality relation from Lemma 3.0.2 and Fubini's theorem imply that

$$\begin{aligned} \sigma_\infty &= \lim_{|P|\to\infty} \int_{\mathbb{T}^n} \int_{|\theta|<|P|} \psi(\theta F(\boldsymbol{x})) \mathrm{d}\theta \mathrm{d}\boldsymbol{x} \\ &= \lim_{|P|\to\infty} |P| \operatorname{vol}(\{\boldsymbol{x} \in \mathbb{T}^n \colon |F(\boldsymbol{x})| < |P|^{-1}\}). \end{aligned}$$

In complete analogy to the finite places one can show that $\sigma_\infty > 0$ if there exists a non-singular solution $\boldsymbol{x} \in \mathbb{T}^n$ to $F(\boldsymbol{x}) = 0$. In other words, the circle method allows us to not only establish an asymptotic formula, but also the Hasse principle for the smooth locus of the projective hypersurface $V(F) \subset \mathbb{P}^{n-1}$. Interpreted as a product of local densities, the term $\sigma_\infty \mathfrak{S}$ matches (up to a constant factor coming from the fact that we count all solutions and not just primitive ones) the Tamagawa measure of the associated hypersurface that is part of Peyre's constant in Conjecture 2.3.3.

Theorem 4.2.4. *Let $n > \sigma + (d-1)2^d$ and suppose that $\mathrm{char}(K) > d$. Then*

$$N_F(P) = \sigma_\infty \mathfrak{S} |P|^{n-d} + O(|P|^{n-d-\kappa})$$

for some $\kappa > 0$. In addition, $\sigma_\infty > 0$ if there exists a non-singular solution $\boldsymbol{x} \in \mathbb{T}^n$ to $F(\boldsymbol{x}) = 0$ and $\mathfrak{S} > 0$ if there exists a non-singular solution $\boldsymbol{x} \in \mathcal{O}_\varpi^n$ to $F(\boldsymbol{x}) = 0$ for every monic irreducible $\varpi \in \mathcal{O}^+$.

Theorem 4.2.4 was first proved by Lee in his PhD thesis [131]. It is the function field analogue of the seminal work by Birch [16] over $\mathbb{Q}$. In fact, both Birch and Lee work in the more general context of complete intersections $X \subset \mathbb{A}^n$ that are defined by R homogeneous forms of degree d. In this case the asymptotic formula has leading term $\sigma_\infty \mathfrak{S} |P|^{n-dR}$, where σ_∞ and $\mathfrak{S}$ are the corresponding singular integral and singular series for the complete intersection $F_1 = \cdots = F_R = 0$. For complete intersections one has to replace σ by

$$\sigma_X = \dim\{\boldsymbol{x} \in \mathbb{A}^n \colon \mathrm{rk}(J(\boldsymbol{x})) < R\},$$

where $J(\boldsymbol{x}) = (\frac{\partial F_i}{\partial x_j})_{1 \le i \le R, 1 \le j \le n}$ is the Jacobian matrix of X. Note that unless X is a hypersurface, σ_X does not necessarily agree with the dimension of the singular locus of X. This has led to a lot of confusion in the literature and is for example incorrectly stated in Lee's thesis [131]. Birch and Lee are able to produce an asymptotic formula for the counting function associated to $F_1, \ldots, F_R$ providing that

$$n \geq \sigma_X + 1 + R(R+1)(d-1)2^{d-1}.$$

Skinner [186] generalised Birch's result to general number fields and in fact Lee [131] also proved the analogous result over arbitrary global function fields, assuming $\mathrm{char}(K) > d$.

One of the main features of the circle method is that when it works, it allows one to establish the smooth Hasse principle for the underlying variety. In addition, by imposing extra congruence conditions one can also verify the weak approximation property. This was done by both Lee [131] and Skinner [186]. Over function fields, the Hasse principle is less interesting in the range of variables where the circle method applies. Indeed, the theory of Lang–Tsen fields implies that any variety $Y \subset \mathbb{P}^{n-1}$ defined by forms $F_1, \ldots, F_R$ of degrees $d_1, \ldots, d_R$ has a rational point over a global function field providing only that $n > d_1^2 + \cdots + d_R^2$. However, establishing weak approximation or an asymptotic formula for the number of rational points of bounded height remains a challenging problem.

In most applications of the circle method the bottleneck lies within the estimation of the minor arcs and requires n to be very large with respect to d. Over $\mathbb{Q}$, the asymptotic evaluation of the major arcs in Proposition 4.2.3 is already known for hypersurfaces to hold as soon as $n > \sigma + 4(d-1)$ thanks to work of Nguyen [152]. For most $\alpha \in \mathbb{T}$, we expect the values $\psi(\alpha F(\boldsymbol{x}))$ to be uniformly distributed among the unit circle for $\boldsymbol{x} \in \mathcal{O}^n$, so that an application of the central limit theorem suggests that typically $|S(\alpha)| \approx |P|^{n/2}$. As the expected main term is of order $|P|^{n-d}$, with the circle method alone it appears to be very difficult to establish an asymptotic formula for $N_F(P)$ unless $n > 2d$, even though Manin's conjecture suggests an asymptotic already for $n > d$. In practice, we usually stay very far from the "square root barrier". However, there are examples where one can get very close to the theoretical limit and one such instance will be the subject of Chapter 5.

4.3 Diagonal forms

A form of the shape $F(\boldsymbol{x}) = \sum_{i=1}^n a_i x_i^d \in \mathcal{O}[x_1, \dots, x_n]$ with $a_1 \cdots a_n \neq 0$ is called a *diagonal form*. In applications of the circle method, diagonal forms have a special place, because their particular shape allows for new arguments to be brought into play. An important ingredient consists of employing *mean value estimates*, as we shall now explain. Let

$$T(\alpha) = \sum_{\substack{x \in \mathcal{O} \\ |x| < |P|}} \psi(\alpha x^d),$$

so that

$$N_F(P) = \int_{\mathbb{T}} \prod_{i=1}^n T(a_i \alpha) \mathrm{d}\alpha.$$

Remark. At this moment it is good to see why the circle method is particularly powerful when the number of variables is large compared to the degree. Indeed, if we achieve an upper bound of the form $|T(\alpha)| \ll |P|^{1-\delta}$ for $\alpha \in \mathfrak{m}(\Delta)$, then since $|a_i| \ll 1$ we immediately get $\prod_{i=1}^n |T(a_i\alpha)|^n \ll |P|^{n-n\delta}$, so that increasing the number of variables amplifies the saving we get in the estimates for the exponential sum involved.

For any $0 \leq s \leq \lfloor \frac{n}{2} \rfloor$, an application of Hölder's inequality yields

$$\begin{aligned} \int_{\mathfrak{m}} \prod_{i=1}^n |T(a_i\alpha)| \mathrm{d}\alpha &\leq \prod_{i=1}^n \left(\int_{\mathfrak{m}} |T(a_i\alpha)|^n \mathrm{d}\alpha \right)^{1/n} \\ &\leq \sup_{\substack{\alpha \in \mathfrak{m} \\ i=1,\dots,n}} |T(a_i\alpha)|^{n-2s} \prod_{i=1}^n \left(\int_{\mathbb{T}} |T(a_i\alpha)|^{2s} \mathrm{d}\alpha \right)^{1/n}. \end{aligned} \tag{4.3.1}$$

Upon defining

$$K_d(P) = \#\{\boldsymbol{x} \in \mathcal{O}^{2s} \colon |\boldsymbol{x}| < |P|, x_1^d + \cdots + x_s^d = x_{s+1}^d + \cdots + x_{2s}^d\}$$

and using the orthogonality relation (4.0.1), the integral above may be evaluated as

$$\int_{\mathbb{T}} |T(a_i\alpha)|^{2s} \mathrm{d}\alpha = K_d(P).$$

If $\alpha \in \mathfrak{m}(\Delta)$, then Weyl's inequality in Proposition 4.1.7 gives

$$|T(a_i\alpha)| \ll |P|^{1-\delta} \tag{4.3.2}$$

for some $\delta > 0$. Thus, if we have an upper bound of the expected order of magnitude

$$K_d(P) \ll |P|^{2s-d+\varepsilon}, \tag{4.3.3}$$

we can show that the contribution from the minor arcs to $N_F(P)$ is negligible. The main ingredient to achieve (4.3.3) is *Vinogradov's mean value theorem.* For any $d, s \geq 1$, let

$$J_{s,d}(P) = \#\{\boldsymbol{x} \in \mathcal{O}^{2s} \colon |\boldsymbol{x}| < |P|, x_1^k + \cdots + x_s^k = x_{s+1}^k + \cdots + x_{2s}^k \text{ for } 1 \leq k \leq d\}.$$

By taking $x_i = x_{s+i}$ for $i = 1, \ldots, s$, we easily obtain the lower bound

$$J_{s,d}(P) \gg |P|^s. \tag{4.3.4}$$

Alternatively, for $\boldsymbol{h} \in \mathcal{O}^d$, we may consider $J_{s,d}(P, \boldsymbol{h})$, which counts solutions $\boldsymbol{x} \in \mathcal{O}^{2s}$ with $|\boldsymbol{x}| < |P|$ to the system of equations

$$\sum_{i=1}^{s} (x_i^k - x_{s+i}^k) = h_k \quad \text{for } k = 1, \ldots, d.$$

Upon defining

$$U(\boldsymbol{\alpha}) = \sum_{\substack{x \in \mathcal{O} \\ |x| < |P|}} \psi(\alpha_1 x + \cdots + \alpha_d x^d)$$

for $\boldsymbol{\alpha} \in \mathbb{T}^d$, by (4.0.1) we have

$$J_{s,d}(P, \boldsymbol{h}) = \int_{\mathbb{T}^d} |U(\boldsymbol{\alpha})|^{2s} \psi(-\boldsymbol{h} \cdot \boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha} \leq J_{s,d}(P).$$

A necessary condition for the system of equations in the definition of $J_{s,d}(P, \boldsymbol{h})$ to be soluble is clearly $|h_k| \ll |P|^k$ for $k = 1, \ldots, d$ and the number of such $\boldsymbol{h} \in \mathcal{O}^d$ is $O(|P|^{d(d+1)/2})$, so that

$$\sum_{\substack{\boldsymbol{h} \in \mathcal{O}^d \\ |h_k| \ll |P|^k}} J_{s,d}(P, \boldsymbol{h}) \ll |P|^{d(d+1)/2} J_{s,d}(P).$$

The left hand side is easily seen to be $\gg |P|^{2s}$, so that

$$J_{s,d}(P) \gg |P|^{2s - d(d+1)/2}. \tag{4.3.5}$$

The main conjecture of Vinogradov asserts that there is a matching upper bound to (4.3.4) and (4.3.5):

$$J_{s,d}(P) \ll_\varepsilon |P|^{s+\varepsilon} + |P|^{2s - d(d+1)/2 + \varepsilon} \tag{4.3.6}$$

for any $s, d \geq 1$. After a great body of work, the conjectured estimate (4.3.6) has been resolved in full by Bourgain, Demeter and Guth [20] and Wooley [211, 212]. The former work is only concerned about $\mathbb{Q}$, while Wooley's method of nested efficient congruencing is flexible enough to establish (4.3.6) over any global field K as long as $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) > d$.

The connection between $J_{s,d}(P)$ and $K_d(P)$ is revealed through the following simple observation. If $|\boldsymbol{x}| < |P|$, then certainly $|\sum_{i=1}^s (x_i^k - x_{s+i}^k)| \ll |P|^k$, so that

$$\begin{aligned} K_d(P) &= \sum_{\substack{\boldsymbol{h} \in \mathcal{O}^{d-1} \\ |h_k| \ll |P|^k}} \int_{\mathbb{T}^d} |U(\boldsymbol{\alpha})|^{2s} \psi(-\alpha_1 h_1 + \cdots - \alpha_{d-1} h_{d-1}) \mathrm{d}\boldsymbol{\alpha} \\ &\ll |P|^{d(d-1)/2} J_{s,d}(P). \end{aligned}$$

In particular, taking $s = d(d+1)/2$, it follows from (4.3.6) that

$$K_d(P) \ll |P|^{d(d+1) - d + \varepsilon} = |P|^{2s - d + \varepsilon}.$$

Together with (4.3.1) and (4.3.2) this shows that

$$\int_{\mathfrak{m}(\Delta)} \prod_{i=1}^{n} |T(a_i\alpha)| \mathrm{d}\alpha \ll |P|^{n-d-\delta}$$

for some $\delta > 0$ if $\mathrm{char}(K) > d$ and $n > d(d+1)$. In addition, for a diagonal form it is known that

$$\int_{\mathfrak{M}(\Delta)} \prod_{i=1}^{n} T(a_i\alpha) \mathrm{d}\alpha = \sigma_\infty \mathfrak{S}|P|^{n-d} + O(|P|^{n-d-\delta})$$

assuming only that $n > 2d$ and $\mathrm{char}(K) \nmid d$ by [127, Theorem 30]. In particular, for a diagonal form we can obtain an asymptotic formula, and likewise establish weak approximation by adding extra congruence conditions whenever $n > d(d+1)$ and $\mathrm{char}(K) > d$. This should be compared with Theorem 4.2.4, where the number of variables grows exponentially with d.

An alternative to Vinogradov's mean value theorem is *Hua's inequality*. It states that

$$\int_{\mathbb{T}} |T(\alpha)|^{2^d} \mathrm{d}\alpha \ll_\varepsilon |P|^{2^d-d+\varepsilon}$$

whenever $\mathrm{char}(K) > d$. It is a classical result over $\mathbb{Q}$ and was extended to $\mathbb{F}_q(t)$ by Kubota in his PhD thesis [127]. Unlike Vinogradov's mean value theorem, Hua's inequality can be proved fairly easily, only relying on the divisor estimate. It allows one to establish an asymptotic formula for $N_F(P)$ when F is diagonal whenever $n > 2^d$. Observe that this is superior to what can be achieved using Vinogradov's mean value theorem only when $d \in \{2, 3, 4\}$.

4.4 Delta method

A major innovation in the circle method is found in the work of Kloosterman [123], in which he essentially uses the Farey dissection of the unit interval together with the Poisson summation formula to obtain cancellations in the sum $S(a/r+\theta)$ when averaged over a modulo r such that $\gcd(a, r) = 1$. Exploiting this extra averaging is called a *Kloosterman refinement*. A classical implementation of the circle method for non-singular quadratic forms $F \in \mathcal{O}[x_1, \dots, x_n]$ requires $n \geq 5$ to get an asymptotic formula for $N_F(P)$, while Kloosterman's approach is capable of handling $n = 4$.

Building on work of Duke, Friedlander and Iwaniec [69], Heath-Brown [100] used a smooth decomposition of the delta symbol to develop an alternative to the classical circle method that is particularly powerful when dealing with forms of small degree. Following work of Browning and Vishe [37], we will now explain the delta method in more detail in the case $K = \mathbb{F}_q(t)$, where it is much simpler to set up. In fact, it turns out to be a direct consequence of combining Dirichlet's approximation theorem with the Poisson summation formula, so that it is disputable whether it should really be called "delta method" over $\mathbb{F}_q(t)$.

Let $Q \geq 1$ be an integer. Then thanks to the ultrametric property of the norm, Proposition 4.2.1 already gives a partition of the unit interval. Indeed, we have

$$\mathbb{T} = \bigsqcup_{\substack{r \in \mathcal{O}^+ \\ |r| \leq \widehat{Q}}} \bigsqcup_{\substack{a \in \mathcal{O} \\ |a| < |r| \\ \gcd(a,r)=1}} \left\{ \alpha \in \mathbb{T} \colon \left| \alpha - \frac{a}{r} \right| < \frac{1}{\widehat{Q}|r|} \right\},$$

which via the orthogonality relation (4.0.1) implies

$$\sum_{\substack{r\in\mathcal{O}^+\\|r|\leq\widehat{Q}}}\sum_{\substack{a\in\mathcal{O}\\|a|<|r|\\\gcd(a,r)=1}}\psi\left(\frac{ax}{r}\right)\int_{|\theta|<\widehat{Q}^{-1}|r|^{-1}}\psi(\theta x)\mathrm{d}\theta=\begin{cases}1 & \text{if } x=0,\\ 0 & \text{if } x\in\mathcal{O}\setminus\{0\}.\end{cases}$$

The analogue of this identity is the starting point in Heath-Brown's delta method and paves the way for a *double* Kloosterman refinement, meaning that one also exploits cancellation when averaging over r. Over the integers the integral over θ is replaced by a more complicated function. The reason why the set-up in positive characteristic is so much simpler can be attributed to the fact that the indicator functions of intervals are already smooth in an appropriate sense, while over the integers one has to "smoothen" them first.

The next step is to evaluate the exponential sum $S(a/r+\theta)$ via the Poisson summation formula as stated in Proposition 3.0.5, which after maneuvering some of the terms and a change of variables leads to an identity of the form

$$N_F(P)=|P|^n\sum_{\substack{r\in\mathcal{O}^+\\|r|\leq\widehat{Q}}}|r|^{-n}\int_{|\theta|<\widehat{Q}^{-1}|r|^{-1}}\sum_{\boldsymbol{v}\in\mathcal{O}^n}S_r(\boldsymbol{v})I_r(\boldsymbol{v},\theta)\mathrm{d}\theta, \tag{4.4.1}$$

where

$$S_r(\boldsymbol{v})=\sum_{\substack{a \bmod (r)\\(a,r)=1}}\sum_{\boldsymbol{x} \bmod (r)}\psi\left(\frac{aF(\boldsymbol{x})+\boldsymbol{v}\cdot\boldsymbol{x}}{r}\right)$$

and

$$I_r(\boldsymbol{v},\theta)=\int_{\mathbb{T}^n}\psi\left(\theta F(P\boldsymbol{x})+\frac{P\boldsymbol{v}\cdot\boldsymbol{x}}{r}\right)\mathrm{d}\boldsymbol{x}$$

are a complete exponential sum modulo r and an oscillatory integral respectively.

A crucial difference to the classical circle method is that there is no division into major and minor arcs. Instead, via Poisson summation we are aiming at evaluating the exponential sum $S(\alpha)$ asymptotically on the whole unit interval, with the expectation that the main term is coming from $\boldsymbol{v}=\boldsymbol{0}$ and the remaining contribution makes up a negligible error term. This is not always true, however. In Chapter 5 we will see that for cubic surfaces, zeros of the dual form make up a main term in the asymptotic formula that corresponds precisely to the contribution coming from rational points on lines.

Let us now explain how to deal with the expression (4.4.1). Firstly, via the method of non-stationary phase, one can show that $I_r(\boldsymbol{v},\theta)=0$ unless $|\boldsymbol{v}|\ll|r|/|P|$. Moreover, if one goes through the analysis carefully, it transpires that the optimal choice of $\widehat{Q}$ is of size $|P|^{d/2}$. In particular, in the generic case $|r|=\widehat{Q}$, we can truncate the sum to $|\boldsymbol{v}|\ll|P|^{d/2-1}$. This already makes the sum shorter if $d\leq 3$, but offers no advantage if $d\geq 4$. Secondly, the sums $S_r(\boldsymbol{v})$ are multiplicative with respect to r. To estimate them, we can thus reduce to the case of prime powers $r=\varpi^k$. The most difficult case is when $k=1$, in which case $S_\varpi(\boldsymbol{v})$ is a complete exponential sum over a finite field that can be estimated efficiently using Deligne's resolution of the Weil conjectures [66, 67], providing the singular locus of the hypersurface under consideration is not too large. Thirdly, to estimate $S_{\varpi^k}(\boldsymbol{c})$ when $k\geq 2$, one typically uses more elementary techniques that resemble the stationary phase approach that one can use to estimate $I_r(\boldsymbol{c},\theta)$. However, this does not always lead to upper bounds of the desired strength. Instead one can estimate the average $\sum_{|\boldsymbol{v}|\ll|r|/|P|}|S_r(\boldsymbol{v})|$ when r is square-full, which turns out to be rather involved.

When estimating the exponential sum $S_r(\boldsymbol{c})$ one can already exhibit a Kloosterman refinement by taking into account the sum over the numerator. Moreover, in the case of quadratic forms, for generic square-free r, they can be evaluated explicitly through Gauss sums. Via Mellin inversion this allows one to relate the average $\sum_r S_r(\boldsymbol{c})$ to certain Dirichlet L-functions and gain extra cancellation when summing over r, which is precisely the double Kloosterman refinement mentioned earlier. For cubic forms, the same strategy leads to a family of Hasse–Weil L-functions associated to hyperplane sections of the corresponding cubic hypersurface. Little is known about these L-functions in the integer setting, but assuming standard conjectures such as the generalised Riemann hypothesis, analytic continuation and the existence of a functional equation, Hooley [116] was able to deal with cubic forms in 8 variables. Over $\mathbb{F}_q(t)$ these standard conjectures about Hasse–Weil L-functions are known thanks to work of Grothendieck and Deligne's resolution of the Riemann hypothesis. Browning and Vishe [37] exploited this fact and proved the analogue of Hooley's result unconditionally over $\mathbb{F}_q(t)$. In Chapter 5 we will follow a similar strategy to obtain almost optimal upper bounds for diagonal cubic forms in 6 variables over $\mathbb{F}_q(t)$.

So far we have only talked about applications of the circle method to counting rational points on hypersurfaces. In fact, with the classical circle method one can already deal with complete intersections, as demonstrated by Birch [16]. If one would like to use an analogue of the delta method, the situation becomes more complicated. Munshi [149] used a nested version of the delta method to produce an asymptotic formula for the number of rational points of bounded height on smooth complete intersections of two quadrics in 11 variables over the integers. A major step forward was made by Vishe [202]. He developed a Farey dissection of $\mathbb{T}^2$ over $\mathbb{F}_q(t)$, thereby allowing for a double Kloosterman refinement. This enabled him to deal with smooth complete intersections of two quadrics in 9 variables. Chapter 6 is concerned with generalising his approach to handle intersections of a cubic and a quadratic hypersurface. This requires extending his Farey dissection to lopsided boxes to take into account the different degrees of the defining forms.

4.5 Results obtained using the circle method

Having explained the basic mechanism underlying the circle method and some of its extensions, we conclude this chapter by a survey of results in the context of rational points that have been obtained using the circle method. Let K be a global field and $F_1, \dots, F_R \in K[x_1, \dots, x_n]$ be homogeneous forms of degree $d_1, \dots, d_R$ respectively. We shall write $\boldsymbol{d} = (d_1, \dots, d_R)$.

We denote by $X \subset \mathbb{P}^{n-1}$ the projective variety cut out by $F_1 = \cdots = F_R = 0$ and assume that it is a complete intersection. We write $U \subset \mathbb{A}^n$ for the affine cone of X. Recall the Birch singular locus

$$\mathsf{BSing}(X) = \{\boldsymbol{x} \in \mathbb{A}^n \colon \operatorname{rk}(J(\boldsymbol{x})) < R\},$$

where $J(\boldsymbol{x}) = (\frac{\partial F_i}{\partial x_j}(\boldsymbol{x}))_{1 \le i \le R, 1 \le j \le n}$ is the Jacobian matrix of U. Let σ_X be the dimension of $\mathsf{BSing}(X)$ and define

$$B_{d,n} = 1 + R(R+1)(d-1)2^{d-1}.$$

Birch [16] has obtained an asymptotic formula for

$$N_{F_1,\dots,F_R}(P) = \#\{\boldsymbol{x} \in \mathbb{Z}^n \colon F_1(\boldsymbol{x}) = \cdots = F_R(\boldsymbol{x}) = 0, |\boldsymbol{x}| < P\}$$

as $P \to \infty$ precisely when $F_1, \dots, F_R \in \mathbb{Z}[x_1, \dots, x_n]$ are forms of degree d and $n \ge \sigma_X + B_{d,n}$. We will use this result as a benchmark, which we will compare other results to. Let us begin by reviewing some general results in Table 4.1 for complete intersections when $d = d_1 = \cdots = d_R$ that improve upon Birch's result.

number of variables	*field*	*conditions*	*author(s)*
$n \geq \sigma_X + (d - \frac{1}{2}\sqrt{d})2^d$	$\mathbb{Q}$	$R = 1$	Browning–Prendiville [31]
$n \geq 1 + \sigma_X + \frac{3}{4}d2^d - 2d$	$\mathbb{Q}$	$3 \leq d \leq 9$ and $R = 1$	Browning–Prendiville [31]
$n \geq d^2 - d + 2\lfloor\sqrt{2d+2}\rfloor - 1$	$\mathrm{char}(K) = 0$ or $\mathrm{char}(K) > d$	diagonal form	Wooley [212]
$n \geq 1 + R(d2^d + 1)$	$\mathbb{Q}$	$F_1, \ldots, F_R$ generic when $d > 3$	Rydin Myerson [165, 166, 167]
$n \geq 1 + R(d(d-1)2^d + 1)$	$\mathbb{F}_q(t)$, $q > (d-1)^n$ and $\mathrm{char}(\mathbb{F}_q) > d$	$F_1, \ldots, F_R \in \mathbb{F}_q[x_1, \ldots, x_n]$ and X smooth	Browning–Vishe–Yamagishi [45]

Table 4.1: Improvements on Birch's result for arbitrary degrees.

The result of Browning and Prendiville [31] essentially uses a combination of all the tools we have at our disposal to estimate $|S(\alpha)|$. This includes $d-3$ applications of van der Corput differencing, which is followed by the Poisson summation formula to estimate the resulting cubic exponential sum, Weyl differencing or a combination of both. Wooley's work uses the advances on Vinogradov's mean value theorem, that we have already mentioned in Section 4.3. Rydin Myerson result is particularly powerful compared to Birch's result when R is large, since the number of variables is only required to grow *linearly* in R compared to the quadratic growth stipulated by Birch. To achieve this, he establishes a sort of "respulsion principle" for the exponential sums involved, from which he deduces that the measure of the set on which the exponential sum is large must be small. This involves a certain auxiliary inequality, which he can only show to hold for generic systems of forms. When $d \in \{2, 3\}$, he is able to remove the genericity assumption [166, 167]. The work of Browning, Vishe and Yamagishi is the function field analogue of Rydin Myerson's approach. They are able to remove the genericity assumption for *any* degree. The price they pay is that they require q to be somewhat larger than in usual applications of the circle method, that the forms are already defined over the constant field $\mathbb{F}_q$ and slightly stronger assumptions on the number of variables involved.

One of the disadvantages of Birch's theorem is that it only applies to systems of forms of equal degree. A priori it is not clear whether a generalisation to systems of forms of different degrees is even possible, since the process of Weyl differencing eliminates the appearance of any form of small degree. Schmidt [180] was the first one to find a way that overcomes this difficulty. However, the bound he obtains on the required number of variables grows quite rapidly and is difficult to compute. This was substantially improved when Browning and Heath-Brown [40] revisited the problem in 2014. Although their most general result is too complicated to state here, we illustrate the strength of their result with two examples. Let $X \subset \mathbb{P}^m$ be a smooth variety. Then they are able to obtain an asymptotic formula for the number of rational points of bounded height on X and establish the Hasse principle whenever

$$\dim(X) \geq (\deg(X) - 1)2^{\deg(X)} - 1.$$

What is surprising about this result is that apart from the smoothness assumption, there is no extra assumption about the geometry of X. In particular, it is not even required to be a complete intersection. If $U \subset \mathbb{A}^n$ is a complete intersection defined by two forms of degrees d_1, d_2 with $d_1 > d_2$ such that $X \subset \mathbb{P}^{n-1}$ is smooth, then they obtain an asymptotic formula for the counting function involved as soon as

$$n > (2 + d_2)(d_1 - 1)2^{d_1 - 1} + d_2 2^{d_2 - 1}.$$

Frei and Madritsch [78] worked out the number field analogue of Browning and Heath-Brown's work.

An alternative point of view is through the h-invariant of a form. Recall that the h-invariant of a form $F \in K[x_1, \ldots, x_n]$ is the least positive integer h such that F can be written as

$$F = G_1 H_1 + \cdots + G_h H_h,$$

where $G_1, H_1, \ldots, G_h, H_h \in K[x_1, \ldots, x_n]$ are forms of positive degree. Schmidt's result [180] already mentioned adopts this point of view and produces an asymptotic formula for the relevant counting function if the h-invariant is large enough compared to the degree. In general it seems difficult to compare the h-invariant to other geometric invariants of F. In his recent work, Bernert [10] established an asymptotic formula for $N_F(P)$ over $\mathbb{Q}$ when $F \in \mathbb{Z}[x_1, \ldots, x_n]$ is a cubic form whose h-invariant is at least 14.

There is a great deal of work in the literature on forms of small degree. Table 4.2 summarises some of the results, in which "non-singular" refers to smoothness of the projective variety X. A few remarks are in order to explain its content. Of course, we have already mentioned

d	$B_{d,n}$	*number of variables*	*field*	*conditions*	*author(s)*
2	5	$n \geq 4$	$\mathbb{Q}$	non-singular	Kloosterman [123]
		$n \geq 3$	$\mathbb{Q}$	non-singular	Heath-Brown [100]
		$n \geq 4$ even	$\mathrm{char}(K) = 0$	non-singular	Getz [83]
3	17	$n \geq 10$	$\mathbb{Q}$	non-singular	Heath-Brown [98]
		$n \geq 10$	$\mathrm{char}(K) = 0$	non-singular	Browning–Vishe [33]
		$n \geq 9$	$\mathbb{Q}$	at worst isolated singularities	Hooley [112, 113, 114, 115]
		$n = 8$	$\mathbb{Q}$	non-singular, conditional	Hooley [116]
		$n = 8$	$\mathbb{F}_q(t)$, $\mathrm{char}(\mathbb{F}_q) > 3$	non-singular	Browning–Vishe [34]
		$n \geq 8$	$\mathbb{Q}$	F diagonal	Vaughan [201]
		$n \geq 7$	$\mathbb{Q}$	F diagonal, conditional	Hooley [111]
		$n \geq 7$	$\mathbb{F}_q(t)$, $\mathrm{char}(\mathbb{F}_q) \neq 3$	F diagonal	Glas–Hochfilzer [88]
		$n = 6$	$\mathbb{Q}$	$F = \sum_{i=1}^6 x_i^3$, conditional	Wang [204]
		$n = 6$	$\mathbb{F}_q(t)$, $\mathrm{char}(\mathbb{F}_q) > 3$	$F = \sum_{i=1}^6 x_i^3$, conditional	Browning–Glas–Wang [38]
4	49	$n \geq 41$	$\mathbb{Q}$	non-singular	Browning–Heath-Brown [29]
		$n \geq 40$	$\mathbb{Q}$	non-singular	Hanselmann [92]
		$n \geq 30$	$\mathbb{Q}$	non-singular	Marmon–Vishe [139]
(2,2)	13	11	$\mathbb{Q}$	non-singular	Munshi [149]
		10	$\mathbb{Q}$	X contains 2 singular points conjugate over an imaginary quadratic field	Arala [2]
		9	$\mathbb{Q}$	X contains two singular points conjugate over $\mathbb{Q}(i)$	Browning–Munshi [30]
		9	$\mathbb{F}_q(t)$, $\mathrm{char}(\mathbb{F}_q) > 2$	non-singular	Vishe [202]
(3,2)	-	29	$\mathbb{Q}$	non-singular	Browning–Dietmann–Heath-Brown [37]
		26	$\mathbb{F}_q(t)$, $\mathrm{char}(\mathbb{F}_q) > 3$	non-singular	Glas [85]
(3,3)	49	39	$\mathbb{Q}$	non-singular	Northey–Vishe [153]

Table 4.2: Improvements on Birch's result for small degrees.

Kloosterman's seminal work [123] on quadratic forms introducing what is now called a "Kloosterman refinement". Based on a smooth decomposition of the delta symbol by Duke, Friedlander and Iwaniec [69], Heath-Brown [100] systematically developed the "delta method" that we discussed in Section 4.4 and applied it to quadratic forms. Getz [83] pushed his methods further by extending Heath-Brown's work to the number field setting and even exhibits a secondary main term in the asymptotic formula for quadratic forms in an even number of variables. His student Tran [197] also worked out a secondary main term for diagonal quadratic forms over $\mathbb{Q}$ in an odd number of variables.

The "conditional" in the work of Hooley [116] refers to standard conjectures about Hasse-Weil L-functions attached to cubic hypersurfaces that we already discussed in Section 4.4. The work of Wang [204] additionally assumes a square-free sieve conjecture and most importantly a form of the ratios conjecture for averages of Hasse-Weil L-functions parameterised by hyperplane sections of the Fermat cubic fivefold. Browning, Wang and the author [38] have adapted Wang's approach in the function field setting and were able to remove all conditional assumptions except for the ratios conjecture.

Aside from the results listed in Table 4.2, there is other important work established using the circle method, even though it falls short of producing an asymptotic formula. Most notably, in a body of work Davenport [62, 59, 60] showed that any cubic form in at least 16 variables over $\mathbb{Q}$ admits a non-trivial zero. His argument considers two alternatives: either the cubic form has a zero for "geometric reasons", or there is enough extra structure which makes the circle method applicable. This was subsequently improved upon in Heath-Brown's seminal work [105] on cubic forms in 14 variables, in which he adds averaged van der Corput differencing into the circle method's arsenal. Bernert and Hochfilzer [11] have translated his approach to imaginary quadratic fields and it would be interesting to see whether one can make it work over arbitrary number fields.

Browning and Heath-Brown's treatise [29] on quartic forms in 41 variables is a precursor of Browning and Prendiville's work [31] on forms in arbitrary degree, whose method we already explained above. Hanselmann [92] was able to save an extra variable by introducing Heath-Brown's *averaged* version of van der Corput differencing compared to the pointwise van der Corput differencing employed by Browning and Heath-Brown. This was substantially improved by Marmon and Vishe [139], who managed to save 10 more variables by incorporating a Kloosterman refinement.

Considering the case of complete intersections, most effort has been concentrated on the intersection of two quadrics. Munshi [149] has developed a nested version of the delta method, allowing him to obtain an asymptotic formula for as few as 11 variables. Browning and Munshi [30] are even able to reduce this to 9 variables in very special circumstances. This was generalised by Arala [2], whose most general result requires 10 variables, where the loss of one variable is solely attributed to the case when an imaginary quadratic number field does not have class number one. In positive characteristic, Vishe [202] has made substantial progress by developing a Farey dissection of the "unit square" $\mathbb{T}^2$, which leads to an identity that should be thought of as analogous to a two-dimensional version of the delta method. In forthcoming work of Li, Rydin Myerson and Vishe this approach is adapted to the rational numbers.

Browning, Dietmann and Heath-Brown [37] combined Poisson summation with a sophisticated version of Weyl differencing to treat non-singular intersections of cubic and quadratic forms over $\mathbb{Q}$ in 29 variables or more. By refining Vishe's Farey dissection of the unit square, the author [85] could incorporate a Kloosterman refinement over $\mathbb{F}_q(t)$, thereby reducing the number of variables required to 26.

For intersections of two cubic forms, Northey and Vishe [153] use an averaged version of van der Corput differencing followed by Kloosterman refinement to save 10 variables compared to Birch's result.

Finally, we want to end this chapter by highlighting applications of the circle method to the study of rational points that go beyond the setting of complete intersections inside projective space. Schindler [176] generalised Birch's result to bi-projective hypersurfaces inside $\mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$ over $\mathbb{Q}$ of bi-degree (d_1, d_2). Whereas a naive adaptation of Weyl differencing would require $d_1 + d_2 - 1$ differencing steps, the key observation here is that $d_1 - 1$ iterations of differencing for one variable and $d_2 - 1$ for the other suffice. Hochfilzer [109] obtained results for systems of forms in bi-projective space by combining Schindler's approach with Rydin Myerson's version of the circle method.

Another generalisation is due to Brandes [23]. In her work, she is able to asymptotically count $\mathbb{Q}$-linear subspaces contained in complete intersections that go beyond a naive adaptation of Birch's approach. Flores [75] has shown how to adapt the circle method to quartic forms in weighted projective space, which under favourable circumstances allow him to achieve the theoretical limit of the circle method that is suggested by the "square root barrier". Another instance where one can get close to the "square root barrier" is that of special forms that are built out of norm forms, as demonstrated by the work of Birch, Davenport and Lewis [14] that was generalised to number fields independently by Schindler and Skorobogatov [178] and Swarbrick-Jones [191].

There has also been some activity surrounding the Birch singular locus. In particular, the work of Schindler [177], Dietmann [68] and Yamagishi [213] is concerned with replacing it by another invariant that in some instances can be strictly smaller than σ_X.

Recently Getz and his collaborators [3] introduced a "non-abelian" version of the delta method that is capable of handling forms of low degree over central division algebras.

CHAPTER 5

Diagonal cubic forms over function fields

This chapter is based on joint work with Hochfilzer [88].

5.1 Introduction

Given a non-singular cubic form $F \in K[x_1, \ldots, x_n]$ with coefficients in a global field K, we are interested in the counting function

$$N(P) = \#\{\boldsymbol{x} \in \mathcal{O}^n \colon |\boldsymbol{x}| < |P|, F(\boldsymbol{x}) = 0\}, \tag{5.1.1}$$

where $\mathcal{O} \subset K$ is the ring of integers, $P \in \mathcal{O}$ and $|\cdot|$ is a suitable absolute value on K. For $n \geq 5$, one generally expects an asymptotic formula of the form

$$N(P) \sim c|P|^{n-3} \tag{5.1.2}$$

as $|P| \to \infty$ for some constant $c \geq 0$. For large values of n, this has been successfully achieved using the Hardy–Littlewood–Ramanujan circle method. For $K = \mathbb{Q}$, the current state of the art is due to Hooley [112], who showed that $n \geq 9$ suffices for (5.1.2) to hold. In fact, conditional on unproved hypotheses about certain Hasse–Weil L-functions, in [116] he pushed his approach further with the outcome that $n \geq 8$ is enough. For $K = \mathbb{F}_q(t)$, using the fact that the analogous hypotheses are in fact theorems by virtue of Deligne's work [67], Browning–Vishe [34] proved unconditionally the asymptotic formula (5.1.2) for $n \geq 8$ and $\mathrm{char}(K) > 3$. However, for small values of n, an asymptotic remains largely out of reach. Assuming F to be non-singular and diagonal, which means

$$F(\boldsymbol{x}) = \sum_{i=1}^{n} F_i x_i^3, \quad F_i \in \mathcal{O} \setminus \{0\}, \tag{5.1.3}$$

Heath-Brown [106] has provided an upper bound of the form $N(P) \ll |P|^{3+\varepsilon}$ for $n = 6$ and $K = \mathbb{Q}$, matching the predicted asymptotic up to a factor of $|P|^{\varepsilon}$. However, his work relies on deep unproven conjectures about certain Hasse–Weil L-functions.

Our first goal of this work is to prove the analogous result unconditionally for $K = \mathbb{F}_q(t)$. One of the main novelties of our work is that we also obtain results when $\mathrm{char}(K) = 2$. Usually the circle method breaks down in small characteristic due to a Weyl differencing process. We manage to bypass this issue by applying Poisson summation instead, along with a recursion argument regarding the density of solutions of the dual form F^* of F.

From now on we write $\mathcal{O} = \mathbb{F}_q[t]$ and we work with the absolute value given by $|P| = q^{\deg P}$ for $P \in \mathcal{O}$. By abuse of notation we also write $|\boldsymbol{x}| := \max_i |x_i|$ for $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathcal{O}^n$.

Theorem 5.1.1. *Let $K = \mathbb{F}_q(t)$ with $\mathrm{char}(K) \neq 3$. Suppose F is given by* (5.1.3)*. Then for $n = 6$ we have*

$$N(P) \ll |P|^{3+\varepsilon}.$$

As explained in Section 4.3, in applications of the circle method one frequently uses upper bounds for the counting function

$$M(P) = \#\left\{\boldsymbol{x} \in \mathcal{O}^6 \colon x_1^3 + x_2^3 + x_3^3 = x_4^3 + x_5^3 + x_6^3 \colon |\boldsymbol{x}| < |P|\right\}$$

to estimate the contribution from the minor arcs. Until now the strongest estimate followed from Hua's lemma, which gives $M(P) \ll |P|^{7/2+\varepsilon}$. In a 1964 letter to Keith Matthews [63] Davenport asked whether one could achieve the bound $M(P) \ll |P|^{3+\varepsilon}$. Theorem 5.1.1 provides an affirmative answer to his question. In addition, in recent work Browning, Wang and the author [39] were able to remove the epsilon for a certain weighted version of $N(P)$, conditional on a suitable form of the ratios conjecture.

For $n = 4$ the situation is more complicated and one does not expect (5.1.2) to hold in general. The cubic surface $X \subset \mathbb{P}^3$ might contain rational lines and any such will contribute $\gg |P|^2$ rational points to the counting function $N(P)$. According to Manin's conjecture [76], one expects

$$N^\circ(P) \sim c|P|(\log|P|)^{\rho-1}, \tag{5.1.4}$$

where $N^\circ(P)$ only counts rational points that do not lie on any rational line contained in X and ρ is the rank of the Picard group of X.

Over $K = \mathbb{Q}$, partial progress was made by Heath-Brown [106], who showed how to isolate the contribution to $N(P)$ coming from points on rational lines when F is diagonal in the delta method. He also managed to give an upper bound of the form $N^\circ(P) \ll |P|^{3/2+\varepsilon}$, again only conditionally on certain conjectures about Hasse–Weil L-functions. As for $n = 6$, working over $K = \mathbb{F}_q(t)$ allows us to establish the estimates unconditionally and we also succeed in isolating the contribution coming from points on rational lines under certain restrictions on the characteristic of K.

Theorem 5.1.2. *Suppose F is given by* (5.1.3)*. If $\mathrm{char}(K) > 3$, then for $n = 4$, we have*

$$N^\circ(P) \ll |P|^{3/2+\varepsilon},$$

where $N^\circ(P)$ is defined as $N(P)$ with the extra condition that $\boldsymbol{x}$ does not lie on any rational line contained in the surface $F = 0$. These lines, if they exist, are of the form

$$b_i x_i + b_j x_j = b_k x_k + b_l x_l = 0,$$

for some $b_i, b_j, b_k, b_l \in K$ such that

$$\left(\frac{b_i}{b_j}\right)^3 = \frac{F_i}{F_j}, \quad \text{and} \quad \left(\frac{b_k}{b_l}\right)^3 = \frac{F_k}{F_l},$$

where $\{i,j,k,l\} = \{1,2,3,4\}$.

While if $\mathrm{char}(K) = 2$*, then for* $n = 4$ *we have*

$$N(P) \ll |P|^{2+\varepsilon}.$$

In characteristic 2 the shape of the dual form of F prevents us from isolating the contribution coming from rational points on rational lines to $N(P)$. However, we still manage to give a non-trivial upper bound for the counting function $N(P)$, thereby providing evidence that the main contribution to $N(P)$ comes from points on rational lines.

In Chapter 7, using different methods we will establish the estimate $N^\circ(P) \ll |P|^{4/3+\varepsilon}$ for any smooth cubic surface over any global function field whose characteristic exceeds 3.

Our work also shares some similarity with the recent findings of Wang. In [204] he established an asymptotic formula for $N(P)$ for diagonal cubic forms over $\mathbb{Q}$ when $n = 6$ conditional on conjectures about mean values of ratios of L-functions and the square-free sieve. His approach required to isolate the contribution coming from rational points on rational linear subspaces, which he achieved in [205], similar to Heath-Brown's [106] treatment when $n = 4$. Recently, Browning, Wang and the author [39] translated Wang's work to $\mathbb{F}_q(t)$ and were able to remove all conditional hypotheses except for the ratios conjecture.

So far we have ignored the constant c appearing in the asymptotic formula (5.1.2), despite its arithmetic significance. It encapsulates information about the existence of rational points on X and has received a conjectural interpretation as an adelic volume by Peyre [156]. For $n \geq 5$ it is expected to be positive as soon as $X(K_\nu) \neq \emptyset$ for all completions K_ν of K, or in other words, it reflects that X is expected to satisfy the Hasse principle. A key feature of the circle method is that when it provides an asymptotic formula, it automatically confirms the Hasse principle. So in particular, thanks to Hooley [112], we know that the Hasse principle holds for non-singular cubic forms in $n \geq 9$ variables over $\mathbb{Q}$ and the work of Browning–Vishe establishes the Hasse principle for non-singular cubic forms over $\mathbb{F}_q(t)$ in at least 8 variables.

In fact, by imposing further congruence conditions on $\boldsymbol{x}$ in the definition of $N(P)$ in (5.1.1) Browning–Vishe show that X satisfies weak approximation. Recall from Chapter 2 that this means that under the diagonal embedding

$$X(K) \longrightarrow \prod_\nu X(K_\nu)$$

the image of $X(K)$ is dense with respect to the product topology. Using Theorem 5.1.1 as a mean value estimate for the minor arc contribution, we can apply a classical version of the circle method to draw the same conclusions for diagonal cubic forms in $n \geq 7$ variables.

Theorem 5.1.3. *Let* $K = \mathbb{F}_q(t)$ *with* $\mathrm{char}(K) > 3$ *and* F *be a diagonal cubic form in* $n \geq 7$ *variables. Then the hypersurface* $X \subset \mathbb{P}^{n-1}$ *cut out by* F *satisfies the Hasse principle and weak approximation.*

One reason for being able to deal with fewer variables than Browning–Vishe is that when F is diagonal we have better control over the exponential sums involved and that we get stronger estimates for the density of solutions of bounded height of the dual form F^* of F. However, this alone along with the estimates by Browning–Vishe on averages of exponential sums would not be sufficient to prove Theorem 5.1.1–5.1.3. We additionally make use of slightly better estimates through an argument that enables us to bypass the lack of a convenient form of partial summation over K.

It should be noted that the Hasse principle over $K = \mathbb{F}_q(t)$ is an easy consequence of the Lang–Tsen theory of C_i fields for $n \geq 10$, which in fact establishes that $X(K) \neq \emptyset$ in this case. For smaller values of n, only little is known about the Hasse principle or weak approximation over $\mathbb{F}_q(t)$. Colliot-Thélène [54] has established the Hasse principle for diagonal cubic forms in $n \geq 5$ variables when $q \equiv 2 \mod 3$ and for $n = 4$ for the same range of q under some additional combinatorial constraints on the coefficients of F. Furthermore, for arbitrary non-singular cubic hypersurfaces $X \subset \mathbb{P}^{n-1}$ Tian [195] has shown that the Hasse principle holds when $\mathrm{char}(K) > 5$ and $n \geq 6$. Assuming the existence of a rational point, Tian–Zhang [196] have also verified that X satisfies weak approximation at places of good reduction whose residue fields have at least 11 elements as soon as $n \geq 4$. In fact, the results by Colliot-Thélène, Tian and Tian–Zhang were all shown to hold for any global function field K of a smooth curve over a finite field.

As a further application of Theorem 5.1.1, we are able to improve the asymptotic version of Waring's problem over $\mathbb{F}_q(t)$ for cubes. Waring's problem in degree d in this context is concerned with finding the smallest value of n such that

$$P = x_1^d + \cdots + x_n^d$$

has a solution in $\boldsymbol{x} \in \mathcal{O}^n$ for every $P \in \mathcal{O}$ with sufficiently large degree. Over $\mathbb{F}_q(t)$, in contrast to the integer setting, there might be global obstructions for P to be representable as a sum of d-th powers, for example if its leading coefficient is not a sum of n d-th powers in $\mathbb{F}_q$. Therefore, one usually restricts to $P \in \mathbb{J}_q^d[t]$, which is defined as the additive closure of d-th powers in $\mathbb{F}_q[t]$. In order to avoid cancellation in the x_i variables coming from the terms of degree larger than $\deg P$, it is more natural to consider the *strict Waring problem*. There, one is concerned with finding the minimal number $G_q(d) = n$ such that every sufficiently large polynomial $P \in \mathbb{J}_q^d[t]$ can be written as

$$P = x_1^d + \cdots + x_n^d,$$

where $\deg x_i \leq \left\lceil \frac{\deg P}{d} \right\rceil$. In order to study a more refined version of Waring's problem, we introduce the quantity $\widetilde{G}_q(d)$, which is the smallest number n such that we obtain an asymptotic formula for

$$R_n(P) = \#\{\boldsymbol{x} \in \mathcal{O}^n \colon |\boldsymbol{x}| \leq q^{\left\lceil \frac{\deg(P)}{d} \right\rceil},\ x_1^d + \cdots + x_n^d = P\},$$

for $P \in \mathbb{J}_q^d[t]$ as $\deg(P) \to \infty$. In his PhD thesis [127] Kubota tackled the asymptotic strict Waring problem over $\mathbb{F}_q(t)$ and showed $\widetilde{G}_q(d) \leq 2^d + 1$ whenever $\mathrm{char}(\mathbb{F}_q) > d$. The restriction in Kubota's work on the characteristic comes from Weyl differencing, producing a factor of $d!$ and hence rendering trivial bounds when estimating exponential sums if $\mathrm{char}(\mathbb{F}_q) \leq d$. For degrees $d \geq 4$ this was improved by Liu–Wooley [136] by replacing Weyl differencing with an application of the large sieve to also obtain results for $\mathrm{char}(\mathbb{F}_q) \leq d$.

Returning to the case of cubes, in characteristic 2 the current state of the art is due to Car–Cherly [51] who showed $\widetilde{G}_{2^h}(3) \leq 11$. They managed to avoid Weyl differencing with an application of Poisson summation along with a version of Weyl's inequality in characteristic 2 developed in [50].

Further, work by Gallardo [81] and Car–Gallardo [52] shows

$$G_q(3) \leq \begin{cases} 7, & \text{if } q \notin \{7, 13, 16\} \\ 8, & \text{if } q \in \{13, 16\} \\ 9, & \text{if } q = 7. \end{cases}$$

Rather than using a circle method approach, the last set of bounds are achieved using elementary arguments. As a result these methods do not produce an asymptotic formula, hence do not yield new bounds for $\widetilde{G}_q(3)$.

We can again use Theorem 5.1.1 as a minor arc mean value estimate in order to improve the current best known bound for $\widetilde{G}_q(3)$ for any q not divisible by 3 as well as for $G_7(3)$, $G_{13}(3)$ and $G_{16}(3)$. Our work on Waring's problem for cubes constitutes a significant improvement on the current state of the art. In particular, our result improves the previously best known upper bound of $\widetilde{G}_q(3)$ by 4 variables if q is even and by 2 variables if q is odd.

Theorem 5.1.4. *If $\mathrm{char}(\mathbb{F}_q) \neq 3$, then we have $\widetilde{G}_q(3) \leq 7$ and thus also $G_q(3) \leq 7$.*

This theorem is the function field counterpart of a result by Hooley [111], who proved the asymptotic Waring problem for cubes over integers in $n \geq 7$ variables conditional on hypotheses on certain Hasse–Weil L-functions. We also obtain a power saving error term in the asymptotic formula for $R_n(P)$. The best unconditional result in the integer setting is due to Vaughan [201], who resolved the asymptotic Waring problem for cubes in 8 variables, although he obtained only log savings in the error term.

To deduce Theorem 5.1.4 from Theorem 5.1.1, we require a power saving when estimating a certain Weyl sum. For Waring's problem this has been carried out by Car [50], which allows us to establish Theorem 5.1.4 in characteristic 2. Although it would be possible to adapt the work of Car adequately to handle the Weyl sums appearing in the treatment of weak approximation and thus extend Theorem 5.1.3 to the case $\mathrm{char}(K) = 2$, we have decided against including such an adaptation here given the length of our paper .

While the techniques used to prove Theorems 5.1.1 – 5.1.4 are not applicable when $\mathrm{char}(K) = 3$, one can almost trivially deal with the problems directly. In fact, studying the solutions to the diagonal cubic equation (5.1.3) reduces to solving a system of linear equations. In particular, the Hasse principle and weak approximation hold trivially. Further it is easy to see that $\widetilde{G}_q(3) = 1$ holds when $\mathrm{char}(K) = 3$.

Outline

To prove Theorem 5.1.1 and Theorem 5.1.2 we employ a technique known as the *delta method* over $\mathbb{F}_q(t)$ developed by Browning–Vishe [34], but which is much simpler than the version of Heath-Brown [106] invoked over the integers. The starting point of the delta method is a smooth decomposition of the Kronecker delta function, a technique that goes back to Duke–Friedlander–Iwaniec [69]. Over $\mathbb{F}_q(t)$, indicator functions of intervals are smooth in an appropriate sense and so this decomposition is essentially rendered trivial.

In Section 5.2, we set up the circle method and arrive at an expression of the form

$$N(w, P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \sum_{c \in \mathcal{O}^n} S_r(\boldsymbol{c}) I_r(\boldsymbol{c}),$$

for a weighted version of the main counting function, involving certain exponential sums $S_r(\boldsymbol{c})$ and oscillatory integrals $I_r(\boldsymbol{c})$.

In Sections 5.3 and 5.4, we estimate the integrals $I_r(\boldsymbol{c})$ and the exponential sums $S_r(\boldsymbol{c})$, respectively. More precisely, we obtain cancellations when averaging $S_r(\boldsymbol{c})$ over r giving essentially optimal bounds. These estimates are possible due to work by Deligne [67] and the required analysis of the relevant L-functions has been carried out in [34, Section 3]. The quality of the estimates of the exponential sums is connected to the dual form of the cubic form. This prompts us to study its rational solutions in Section 5.5.

Classically, to combine these estimates one would use partial summation, a tool that is not available in a useful form to us in the function field setting. In [34] this causes significant difficulty, and in fact the approach by Browning–Vishe comes with a slight loss in the estimates rendering them insufficient for our purposes. We can resolve this issue with Lemma 5.3.5, where we show that $I_r(\boldsymbol{c})$ only depends on the absolute value of r and so via q-adic summation we can separate the quantities without any loss.

In Section 5.6, we combine the estimates using this new approach and finish our treatment in the case $n = 6$, thereby proving Theorem 5.1.1. In the case $\mathrm{char}(K) = 2$, it turns out that the dual form F^* of F is again a non-singular cubic form. For this reason, in Section 5.6.3, we can introduce a self-improving process in the proof of Theorem 5.1.1 and the second part of Theorem 5.1.2 that turns any saving into the desired upper bound. Finally, we use Theorem 5.1.1 as a mean value estimate in an application of the classical circle method to deal with the asymptotic Waring's problem for cubes and weak approximation for diagonal cubic hypersurfaces in $n \geq 7$ variables in Section 5.7.

If $n = 4$ and $\mathrm{char}(K) > 3$ we need to deal separately with the terms coming from *special solutions* of the dual form. This is the content of Section 5.8, where we show that these terms correspond to points coming from rational lines on X.

We follow the convention that ε denotes and arbitraryily small positive real number whose exact value might change from one line to the next. All of the implied constants throughout this chapter are allowed to depend on ε, the cardinality of the constant field q and on the form F.

5.2 Setting up the circle method

We will use the notation and material from Chapter 3.

Given a polynomial $F \in \mathcal{O}[x_1, \dots, x_n]$ and $w \in S(K_\infty^n)$, we are interested in the counting function

$$N(w, P) = \sum_{\substack{\boldsymbol{x} \in \mathcal{O}^n \\ F(\boldsymbol{x})=0}} w\left(\frac{\boldsymbol{x}}{P}\right).$$

For estimating the integrals appearing in our work, it is necessary to work with such a weighted counting function, since we require ∇F to be bounded away from 0 on $\mathrm{supp}(w)$. To estimate our original counting function defined in (5.1.1), it suffices to take w to be the characteristic function of the set $\{\boldsymbol{x} \in \mathbb{T} \colon |\boldsymbol{x}| = q^{-1}\}$. Indeed, it follows that

$$N(w, P) = \#\{\boldsymbol{x} \in \mathcal{O}^n \colon F(\boldsymbol{x}) = 0, |\boldsymbol{x}| = q^{-1}|P|\},$$

so that an upper bound of the shape $N(P, w) \ll |P|^k$ yields $N(P) \ll |P|^k$ by summing over $|P|$.

For a fixed parameter $Q \geq 1$ to be specified later, we deduce from Proposition 4.2.1 and the remark directly afterwards together with the orthogonality relation from Lemma 3.0.2 that

$$N(w,P) = \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \sideset{}{'}\sum_{|a|<|r|} \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} S(a/r+\theta)\mathrm{d}\theta,$$

where $\sum'_{|a|<|r|}$ means that we sum over $a \in \mathcal{O}$ with $(a,r)=1$ only and

$$S(\alpha) = \sum_{\boldsymbol{x}\in\mathcal{O}^n} \psi(\alpha F(\boldsymbol{x}))w(\boldsymbol{x}/P)$$

for $\alpha \in \mathbb{T}$. As explained in [34, Chapter 4], since w is a Schwartz-Bruhat function we can evaluate $S(\theta + a/r)$ using the Poisson summation formula from Proposition 3.0.5 to obtain

$$N(w,P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \sum_{\boldsymbol{c}\in\mathcal{O}^n} S_r(\boldsymbol{c})I_r(\theta,\boldsymbol{c})\mathrm{d}\theta, \tag{5.2.1}$$

where

$$S_r(\boldsymbol{c}) = \sideset{}{'}\sum_{|a|<|r|} \sum_{|\boldsymbol{x}|<|r|} \psi\left(\frac{aF(\boldsymbol{x})+\boldsymbol{c}\cdot\boldsymbol{x}}{r}\right) \tag{5.2.2}$$

and

$$I_r(\theta,\boldsymbol{c}) = \int_{K_\infty^n} w(\boldsymbol{x})\psi\left(\theta P^3 F(\boldsymbol{x}) + \frac{P\boldsymbol{c}\cdot\boldsymbol{x}}{r}\right)\mathrm{d}\boldsymbol{x}. \tag{5.2.3}$$

The expression (5.2.1) is the starting point for our work and from now on we will mostly be concerned about estimating the integrals $I_r(\theta,\boldsymbol{c})$ and the sums $S_r(\boldsymbol{c})$.

5.3 Integral estimates

For $f \in K_\infty[x_1,\dots,x_n]$, we denote by H_f its height, that is, the maximum of the absolute values of its coefficients. Given $\gamma \in K_\infty$, $\boldsymbol{w} \in K_\infty^n$ and $f \in K_\infty[x_1,\dots,x_n]$, integrals of the form

$$J_f(\gamma,\boldsymbol{w}) := \int_{K_\infty^n} w(\boldsymbol{x})\psi(\gamma f(\boldsymbol{x}) + \boldsymbol{w}\cdot\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{5.3.1}$$

appear quite frequently in our work. We shall now collect the required estimates for them. Upon noting that $w(\boldsymbol{x}) = \chi_{\mathbb{T}}(\boldsymbol{x}) - \chi_{t^{-1}\mathbb{T}}(\boldsymbol{x})$, the next lemma follows directly from [34, Lemma 2.4].

Lemma 5.3.1. *Let $\gamma \in K_\infty$ and $\boldsymbol{w} \in K_\infty^n$ be such that $|\boldsymbol{w}| > q$ and $|\boldsymbol{w}| \geq H_f|\gamma|$. Then $J_f(\gamma,\boldsymbol{w}) = 0$.*

The next result [34, Lemma 2.7] is the main ingredient for estimating the integrals $J_f(\gamma,\boldsymbol{w})$.

Lemma 5.3.2. *We have*

$$\int_{\mathbb{T}^n\setminus\Omega} \psi(\gamma f(\boldsymbol{x}) + \boldsymbol{w}\cdot\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 0,$$

where $\Omega \subset \mathbb{T}^n$ is given by

$$\Omega = \left\{\boldsymbol{x}\in\mathbb{T}^n \colon |\gamma\nabla f(\boldsymbol{x}) + \boldsymbol{w}| \leq H_f \max\left\{1, |\gamma|^{1/2}\right\}\right\}.$$

In our setting, this leads to the following estimate.

Lemma 5.3.3. *Suppose $F \in K_\infty[x_1, \ldots, x_n]$ is a non-singular cubic form. Let $\gamma \in K_\infty$ and $\boldsymbol{w} \in K_\infty^n \setminus \{\boldsymbol{0}\}$ be such that $|\boldsymbol{w}| \gg 1$. Then $J_F(\gamma, \boldsymbol{w}) = 0$, unless*

$$|\boldsymbol{w}| \ll |\gamma| \ll |\boldsymbol{w}|,$$

in which case

$$J_F(\gamma, \boldsymbol{w}) \ll \operatorname{vol}(\{\boldsymbol{x} \in \operatorname{supp}(w) \colon |\gamma \nabla F(\boldsymbol{x}) + \boldsymbol{w}| \ll |\boldsymbol{w}|^{1/2}\}).$$

Proof. First note $J_F(\gamma, \boldsymbol{w}) = 0$ if $|\boldsymbol{w}| > \max\{q, H_F|\gamma|\}$ by Lemma 5.3.1. Since by assumption $1 \ll |\boldsymbol{w}|$, we may thus assume $1 \ll |\boldsymbol{w}| \ll |\gamma|$. For $\boldsymbol{a} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$, let

$$w_{\boldsymbol{a}}(\boldsymbol{x}) = \begin{cases} 1 & \text{if } |\boldsymbol{x} - \boldsymbol{a}t^{-1}| < |t|^{-1}, \\ 0 & \text{else.} \end{cases} \tag{5.3.2}$$

We can then write $w(\boldsymbol{x}) = \sum_{\boldsymbol{a} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}} w_{\boldsymbol{a}}(\boldsymbol{x})$, so that

$$\begin{aligned} J_F(\gamma, \boldsymbol{w}) &= \sum_{\boldsymbol{a} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}} \int_{\mathbb{T}^n} w_{\boldsymbol{a}}(\boldsymbol{x}) \psi(\gamma F(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}) \mathrm{d}\boldsymbol{x} \\ &= \sum_{\boldsymbol{a} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}} q^{-n} \psi(t^{-1}\boldsymbol{w} \cdot \boldsymbol{a}) \int_{\mathbb{T}^n} \psi(\gamma G_{\boldsymbol{a}}(\boldsymbol{y}) + t^{-1}\boldsymbol{w} \cdot \boldsymbol{y}) \mathrm{d}\boldsymbol{y}, \end{aligned} \tag{5.3.3}$$

where we performed the change of variables $\boldsymbol{y} = t\boldsymbol{x} - \boldsymbol{a}$ and wrote $G_{\boldsymbol{a}}(\boldsymbol{y}) = F((\boldsymbol{y} + \boldsymbol{a})t^{-1})$. From Lemma 5.3.2 we deduce that each inner integral is bounded by

$$\operatorname{vol}(\{\boldsymbol{y} \in \mathbb{T}^n \colon |\gamma \nabla G_{\boldsymbol{a}}(\boldsymbol{y}) + t^{-1}\boldsymbol{w}| \ll H_{G_{\boldsymbol{a}}}|\gamma|^{1/2}\}),$$

which in turn may be bounded from above by

$$\operatorname{vol}(\{\boldsymbol{x} \in \operatorname{supp}(w_{\boldsymbol{a}}) \colon |\gamma \nabla F(\boldsymbol{x}) + \boldsymbol{w}| \ll H_F|\gamma|^{1/2}\}), \tag{5.3.4}$$

since $H_{G_{\boldsymbol{a}}} \leq H_F$. Denote the set in (5.3.4) by $\Omega_{\boldsymbol{a}}$. Note that since F is assumed to be non-singular, we have $\nabla F(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \Omega_{\boldsymbol{a}}$. Since $\operatorname{supp}(w_{\boldsymbol{a}})$ is compact for every $\boldsymbol{a}$, this implies $\nabla F(\boldsymbol{x}) \gg_w 1$ for all $\boldsymbol{x} \in \Omega_{\boldsymbol{a}}$. In particular, unless $|\boldsymbol{w}| \gg |\gamma \nabla F(\boldsymbol{x})| \gg |\gamma|$ the sets $\Omega_{\boldsymbol{a}}$ are all empty and the integral vanishes. Finally the Lemma follows upon noting

$$\operatorname{vol}(\Omega_{\boldsymbol{a}}) \ll \operatorname{vol}(\{\boldsymbol{x} \in \operatorname{supp}(w) \colon |\gamma \nabla F(\boldsymbol{x}) + \boldsymbol{w}| \ll |\boldsymbol{w}|^{1/2}\}),$$

for any $\boldsymbol{a} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ and substituting this into (5.3.3). □

Since we work with a diagonal cubic form $F(\boldsymbol{x}) = \sum_{i=1}^n F_i x_i^3$ with $F_i \in \mathcal{O} \setminus \{0\}$, we have $\nabla F(\boldsymbol{x}) = (3F_1 x_1^2, \ldots, 3F_n x_n^2)$. Therefore in order to find an upper bound for $J_F(\gamma, \boldsymbol{w})$ the following lemma will be useful.

Lemma 5.3.4. *Let $a, b \in K_\infty$ and consider the set*

$$P_{a,b} = \{x \in \mathbb{T} \colon |x^2 - a| < |b|\}.$$

Then we have

$$\operatorname{vol}(P_{a,b}) \ll \min\{|b|^{1/2}, |b||a|^{-1/2}\}.$$

Proof. Note first that the result is trivial if $a=0$ or $b=0$. Hence we may write

$$a=\sum_{i\le K} a_i t^i, \quad \text{and} \quad b=\sum_{j\le M} b_j t^j,$$

where $a_K, b_M \neq 0$. We will proceed in two cases.

Case 1: $|a|<|b|$. Then via the ultrametric triangle inequality we note

$$|x^2-a|<|b| \iff |x|^2<|b|,$$

for any $x\in\mathbb{T}$. Thus $\operatorname{vol}(P_{a,b}) \ll |b|^{1/2}=\min\{|b|^{1/2},|b||a|^{-1/2}\}$.

Case 2: $|a|\ge|b|$. Let $x=\sum_{i\le -1} x_i t^i\in\mathbb{T}$. Then $|x^2-a|<|b|$ can only hold if $|x|^2=|a|$. In particular K must be even, $K\le -1$ must hold and $x_{K/2+1}=\cdots=x_{-1}=0$. Write

$$x^2=\sum_{\ell\le K} X_\ell t^\ell,$$

where $X_\ell=\sum_{i+j=\ell} x_i x_j$. Then, requiring

$$|x^2-a|<|b|=q^M$$

implies $X_\ell=a_\ell$ for $\ell=M,\dots,K$. Now $X_K=x_{K/2}^2$, so the condition $X_K=a_K$ yields at most two possible solutions for $x_{K/2}$. Further, since

$$X_{K-r}=2x_{K/2}x_{K/2-r}+\sum_{\substack{i+j=K-r\\ K/2-r<i,j<K/2}} x_i x_j,$$

we find inductively that a solution to $x_{K/2}^2=a_K$ uniquely determines $x_{K/2-r}$ for $r=1,\dots,M+K$. To summarise, in this case, there are at most two possibilities for the values of the coefficients $x_{-1},\dots,x_{M-K/2}$. Therefore we obtain

$$\operatorname{vol}(P_{a,b}) \ll \operatorname{vol}\left(t^{M-K/2}\mathbb{T}\right)=q^{M-K/2}=|b||a|^{-1/2}.$$

Finally, noticing that $|b||a|^{-1/2}\le |b|^{1/2}$ if $|a|\ge|b|$ finishes the proof of this lemma. □

In light of Lemma 5.3.4 we thus find

$$\operatorname{vol}(\{\boldsymbol{x}\in\operatorname{supp}(w)\colon |\gamma\nabla F(\boldsymbol{x})+\boldsymbol{w}| \ll |\boldsymbol{w}|^{1/2}\}) \ll \prod_{i=1}^n \min\{|\boldsymbol{w}|^{-1/4},|w_i|^{-1/2}\}$$

if F is a diagonal cubic form. Noting that the expression on the right hand side is $\gg_q 1$ if $|\boldsymbol{w}|\ll 1$ we infer from Lemma 5.3.3

$$J_F(\gamma,\boldsymbol{w}) \ll \prod_{i=1}^n \min\{|\boldsymbol{w}|^{-1/4},|w_i|^{-1/2}\}, \tag{5.3.5}$$

for all $\gamma\in K_\infty$ and all $\boldsymbol{w}\in K_\infty^n\setminus\{\boldsymbol{0}\}$.

We will also have to deal with averages of $I_r(\theta,\boldsymbol{c})$ over θ, which are of the form

$$I_r(\boldsymbol{c}) := \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} I_r(\theta,\boldsymbol{c})\mathrm{d}\theta.$$

While we do not have a convenient form of partial summation available in the function field setting, the next lemma will be crucial in replacing this tool.

Lemma 5.3.5. *Let $r_1, r_2 \in \mathcal{O}$ be such that $|r_1| = |r_2|$. Then $I_{r_1}(\boldsymbol{c}) = I_{r_2}(\boldsymbol{c})$.*

Proof. Write $r = r_1$ for brevity. We shall show that $I_r(\boldsymbol{c})$ only depends on the absolute value of r. Indeed, recalling (5.2.3), for $\boldsymbol{c}$ fixed we have

$$\begin{aligned} I_r(\boldsymbol{c}) &= \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \int_{K_\infty^n} w(\boldsymbol{x})\psi\left(\theta P^3 f(\boldsymbol{x}) + \frac{P\boldsymbol{c}\cdot\boldsymbol{x}}{r}\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\theta \\ &= |r|^n \int_{K_\infty^n} w(r\boldsymbol{y})\psi(P\boldsymbol{c}\cdot\boldsymbol{y}) \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \psi(\theta P^3 r^3 f(\boldsymbol{y}))\mathrm{d}\theta\mathrm{d}\boldsymbol{y}, \end{aligned} \tag{5.3.6}$$

where we used Fubini's theorem and applied the change of variables $\boldsymbol{y} = \boldsymbol{x}r^{-1}$. It follows from Lemma 3.0.2 that

$$\int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \psi(\theta P^3 r^3 f(\boldsymbol{y}))\mathrm{d}\theta = \begin{cases} (|r|\widehat{Q})^{-1} & \text{if } |P^3 f(\boldsymbol{y})| < |r|^{-2}\widehat{Q}, \\ 0 & \text{else.} \end{cases}$$

We conclude that the value of the inner integral in (5.3.6) only depends on $|r|$ for $\boldsymbol{y}$ and $\boldsymbol{c}$ fixed. The claim now follows, since w only depends on the absolute value of its argument. □

By the previous lemma, to estimate $I_r(\boldsymbol{c})$ it thus suffices to consider the case $r = t^Y$ for integers $Y \geq 0$. In the notation above, for $r \in \mathcal{O}\setminus\{0\}$, $\boldsymbol{c} \in \mathcal{O}^n$, $\theta \in \mathbb{T}$ and $P \in \mathcal{O}$ we have

$$I_r(\theta, \boldsymbol{c}) = J_F\left(P^3\theta, \frac{P}{r}\boldsymbol{c}\right).$$

Since $I_r(\theta, \boldsymbol{c})$ vanishes unless $\frac{|P||\boldsymbol{c}|}{|r|} \ll |\theta||P|^3 \ll \frac{|P||\boldsymbol{c}|}{|r|}$, we deduce from (5.3.5) the following integral estimate.

Lemma 5.3.6. *Let $Y \geq 0$, $\boldsymbol{c} \in \mathcal{O}^n \setminus \{\boldsymbol{0}\}$, and $P \in \mathcal{O}$. Then*

$$I_{t^Y}(\boldsymbol{c}) \ll \min\left\{\frac{|\boldsymbol{c}|}{\widehat{Y}|P|^2}, \widehat{Y}^{-1}\widehat{Q}^{-1}\right\} \prod_{i=1}^n \min\left\{\left(\frac{|P||\boldsymbol{c}|}{\widehat{Y}}\right)^{-1/4}, \left(\frac{|P||c_i|}{\widehat{Y}}\right)^{-1/2}\right\}.$$

So far we have not yet achieved any non-trivial estimates for $I_{t^Y}(\boldsymbol{0})$ and in fact we will have to do slightly better than the trivial bound for our treatment.

Lemma 5.3.7. *Assume $n \geq 4$. Let $P \in \mathcal{O}\setminus\{\boldsymbol{0}\}$. Then for any $Y \geq 1$ we have*

$$I_{t^Y}(\boldsymbol{0}) \ll |P|^{-3}.$$

Proof. Note that (5.3.3) together with (5.3.4) imply that

$$I_{t^Y}(\theta, \boldsymbol{0}) \ll \sum_{\boldsymbol{a}\in\mathbb{F}_q^n\setminus\{\boldsymbol{0}\}} \mathrm{vol}(\Omega_{\boldsymbol{a}}),$$

where $w_{\boldsymbol{a}}$ is defined in (5.3.2) and

$$\Omega_{\boldsymbol{a}} = \{\boldsymbol{x} \in \mathrm{supp}(w_{\boldsymbol{a}})\colon |\theta\nabla F(\boldsymbol{x})| \ll \max\{1, |\theta P^3|\}^{1/2}\}.$$

As $\nabla F(\boldsymbol{x}) \neq \boldsymbol{0}$ for any $\boldsymbol{x} \in w_{\boldsymbol{a}}$ and $\mathrm{supp}(w_{\boldsymbol{a}})$ is compact, we have $|\nabla F(\boldsymbol{x})| \gg 1$ for any $\boldsymbol{x} \in \mathrm{supp}(w_{\boldsymbol{a}})$. As there are $q^n - 1 = O(1)$ possibilities for $\boldsymbol{a}$, this implies that the set above is empty unless $|\theta| \ll |P|^{-3}$. As we have the trivial estimate $\mathrm{vol}(\Omega_{\boldsymbol{a}}) \ll 1$, the claim of the lemma easily follows. □

5.4 Exponential sum estimates

We want to estimate the sum

$$\begin{aligned} S_r(\boldsymbol{c}) &= \sideset{}{'}\sum_{|a|<|r|} \sum_{|\boldsymbol{x}|<|r|} \psi\left(\frac{aF(\boldsymbol{x}) + \boldsymbol{c}\cdot\boldsymbol{x}}{r}\right) \\ &= \sideset{}{'}\sum_{|a|<|r|} \prod_{i=1}^{n} \sum_{|x|<|r|} \psi\left(\frac{aF_i x^3 + c_i x}{r}\right), \end{aligned} \tag{5.4.1}$$

where $F(\boldsymbol{x}) = \sum_{i=1}^n F_i x_i^3$. The corresponding sum over the integers has already been subject to thorough investigation by Heath-Brown [98] and Hooley [111]. Browning–Vishe [34] have translated many of the properties to the function field setting, some of which we shall record here.

The quality of our estimates is intimately connected to the dual form F^* of F, which is an absolutely irreducible polynomial of degree $2^{n-2} \times 3$ whose zero locus parameterises hyperplanes that have a singular intersection with the projective hypersurface cut out by F. As explained by Wang [203, Appendix D], if F is diagonal and $\mathrm{char}(K) > 3$, we can take

$$F^*(\boldsymbol{c}) = \left(\prod_{i=1}^n F_i\right)^{2^{n-2}} \prod \left((F_1^{-1}c_1^3)^{1/2} \pm \cdots \pm (F_n^{-1}c_n^3)^{1/2}\right), \tag{5.4.2}$$

where the inner product runs through all possible combinations of $\pm$. In fact, in [203] this is only shown for $K = \mathbb{Q}$, but one can check that the requirement $\mathrm{char}(K) > 3$ is sufficient for (5.4.2) to hold. In characteristic 2, we have the following result.

Lemma 5.4.1. *Let K be a field of characteristic 2 and $F(\boldsymbol{x}) = \sum_{i=1}^n F_i x_i^3 \in K[x_1, \dots, x_n]$ be a non-singular cubic form. Then the dual form of F is given by*

$$F^*(\boldsymbol{c}) = \left(\prod_{i=1}^n F_i\right) \sum_{i=1}^n F_i^{-1} c_i^3.$$

Proof. By definition the zero locus $V(F^*) \subset \mathbb{P}^{n-1}$ parameterises points $\boldsymbol{c} \in \mathbb{P}^{n-1}$ such that the hyperplane $\boldsymbol{c}\cdot\boldsymbol{x} = 0$ has a singular intersection with $V(F^*)$. This means, that there exists $\boldsymbol{x} \in \mathbb{P}^{n-1}(\overline{K})$ such that

$$\mathrm{rk}\begin{pmatrix} \nabla F(\boldsymbol{x}) \\ \boldsymbol{c} \end{pmatrix} = 1, \quad \boldsymbol{c}\cdot\boldsymbol{x} = 0 \quad \text{and} \quad F(\boldsymbol{x}) = 0. \tag{5.4.3}$$

Since we assume F to be non-singular, the rank condition implies that $\boldsymbol{c}$ is proportional to $\nabla F(\boldsymbol{x})$, that is, $x_i^2 = \lambda F_i^{-1} c_i$ for some $\lambda \in \overline{K}^\times$ and $i = 1, \dots, n$. Any pair $(\boldsymbol{x}, \boldsymbol{c})$ having this property then satisfies $F(\boldsymbol{x}) = 0$ if and only if $\boldsymbol{c}\cdot\boldsymbol{x} = 0$. Moreover, the third condition in (5.4.3) is equivalent to

$$\sum_{i=1}^n F_i^{-1/2} c_i^{3/2} = 0,$$

where we used that every element of $\overline{K}$ has a unique square-root as $\mathrm{char}(K) = 2$. However, again since we are in characteristic 2, this is is equivalent to

$$\sum_{i=1}^n F_i^{-1} c_i^3 = 0.$$

The result now follows after clearing denominators. ☐

Note that if $r_1, r_2 \in \mathcal{O}$ are coprime, then

$$S_{r_1 r_2}(\boldsymbol{c}) = S_{r_1}(\boldsymbol{c}) S_{r_2}(\boldsymbol{c}), \tag{5.4.4}$$

which follows readily from the Chinese remainder theorem. This essentially reduces the task of estimating $S_r(\boldsymbol{c})$ to prime power moduli. Indeed, suppose $S_{\varpi^k}(\boldsymbol{c}) \leq C|\varpi|^{k\alpha}$ for some $\alpha > 0$ and some absolute constant C. Let $\omega(r)$ be the number of prime divisors of r. Then by multiplicativity of $S_r(\boldsymbol{c})$ we have

$$S_r(\boldsymbol{c}) = \prod_{\varpi^k \| r} S_{\varpi^k}(\boldsymbol{c}) \leq \prod_{\varpi^k \| r} C|\varpi|^{k\alpha} = C^{\omega(r)}|r|^\alpha \ll \tau(r)|r|^\alpha \ll |r|^{\alpha+\varepsilon}$$

by the usual estimate for the divisor function $\tau(r)$, see [35, Lemma 5.9].

Further, if ϖ is irreducible such that $\varpi \nmid F^*(\boldsymbol{c})$, then Browning–Vishe [34, Section 5] show

$$S_{\varpi^k}(\boldsymbol{c}) = 0 \quad \text{for } k \geq 2. \tag{5.4.5}$$

5.4.1 Square-free moduli contribution

Deligne's resolution of the Weil conjectures [66] shows that we get square-root cancellation for the sums $S_\varpi(\boldsymbol{c})$ whenever ϖ is suitably generic:

$$S_\varpi(\boldsymbol{c}) \ll |\varpi|^{(n+1)/2} |(\varpi, \nabla F^*(\boldsymbol{c}))|^{1/2}. \tag{5.4.6}$$

However, this is not sufficient for our purposes. In the integer setting Hooley [111] was the first to achieve an extra saving when averaging the sums $S_r(\boldsymbol{c})$ over r by appealing to certain hypotheses about Hasse–Weil L-functions associated to cubic threefolds. By virtue of Deligne's proof of the Weil conjectures [67] these hypotheses are in fact theorems in the function field setting. This enabled Browning–Vishe [34, Lemma 8.5] to establish the following result unconditionally.

Lemma 5.4.2. *Suppose n is even and $F^*(\boldsymbol{c}) \neq 0$. Then for any $Z \geq 0$ and $\varepsilon > 0$, we have*

$$\sum_{\substack{|r| \leq \widehat{Z} \\ (r, \Delta_F F^*(\boldsymbol{c}))=1}} \frac{S_r(\boldsymbol{c})}{|r|^{(n+1)/2}} \ll |\boldsymbol{c}|^\varepsilon \widehat{Z}^{1/2+\varepsilon},$$

where Δ_F is the discriminant of F and by virtue of (5.4.5) *r ranges over square-free values only.*

Remark. In fact Browning–Vishe have to consider averages of $S_r(\boldsymbol{c})$ twisted by a Dirichlet character of K_∞ since they were unable to separate the integral $I_r(\theta, \boldsymbol{c})$ from summation. However, we can resolve this issue with Lemma 5.3.5 allowing us to combine Lemma 5.4.2 with the integral bounds from Lemma 5.3.6 more efficiently.

5.4.2 Pointwise estimates

For $B \in \mathcal{O}$ fixed and $a, r \in \mathcal{O} \setminus \{0\}$ with $(a, r) = 1$, let

$$S_r(a, c) = \sum_{|x| < |r|} \psi\left(\frac{aBx^3 + cx}{r}\right).$$

In view of (5.4.1) upper bounds for $S_r(a,c)$ directly transform into estimates for $S_r(\boldsymbol{c})$. Moreover, by (5.4.4) it suffices to consider the case $r = \varpi^k$, where ϖ is irreducible. Hooley [111] has provided upper bounds for the integer-analogue of the sum $S_{\varpi^k}(a,c)$ whenever $\varpi \nmid B$. As explained by Heath-Brown [106], these estimates also hold if $\varpi \mid B$ when we allow the implied constant to depend on B. Hooley's and Heath-Brown's proofs of these results go through almost verbatim in the function field setting and so we spare the reader from the tedious exercise of reproducing them here. To state the final outcome, we need some notation. First, we set $\{\varpi^k, c\} = (\varpi^k, c)$ for $k = 2$ and for $k \geq 3$, we define $\{\varpi^k, c\} = |\varpi|^{-1}$ if $\varpi \parallel c$ and $\{\varpi^k, c\} = (\varpi^k, c)$ else. For later use, we generalise this to square-full r by setting

$$\{r, c\} \coloneqq \prod_{\varpi^k \| r} \{\varpi^k, c\}.$$

We then have

$$S_{\varpi^k}(a,c) \ll |\varpi|^{k/2} |\{\varpi^k, c\}|^{1/4} \quad \text{for } k \geq 2. \tag{5.4.7}$$

We shall also use an estimate of Hua [117, Lemma 1.1], whose proof, again, readily translates to the function field setting. If $g(x) = \sum_{i=0}^d g_i x^i \in \mathcal{O}[x]$, then for any $\varpi \in \mathcal{O}$ irreducible we have

$$\sum_{|x| < |\varpi|^k} \psi\left(\frac{g(x)}{\varpi^k}\right) \ll |\varpi|^{k(1-1/d)} |(\varpi^k, g_0, \ldots, g_d)|^{1/d}, \tag{5.4.8}$$

where the constant depends only on ε and d. Originally this was stated in the case when $\varpi \nmid (g_0, \ldots, g_d)$, but the factor $|(\varpi^k, g_0, \ldots, g_d)|^{1/d}$ in the estimate accounts for the possibility of $\varpi \mid (g_0, \ldots, g_d)$. Therefore we obtain

$$S_{\varpi^k}(a,c) \ll |\varpi|^{2k/3},$$

where the implied constant depends on ε but crucially not on a since we assumed $\varpi \nmid a$. Using (5.4.1), we can immediately deduce the following lemma from (5.4.7) and (5.4.8), which is the analogue of [106, Lemma 5.1.].

Lemma 5.4.3. *It holds that*

$$S_{\varpi^2}(\boldsymbol{c}) \ll |\varpi|^{2+n}.$$

In addition, if $(\varpi^k, \boldsymbol{c}) = H_\varpi$ *and there at least* m *indices* i *such that* $(\varpi^k, c_i) = H_\varpi$*, then*

$$S_{\varpi^k}(\boldsymbol{c}) \ll |\varpi|^{k+2(n-m)k/3+mk/2} |H_\varpi|^{m/4}.$$

5.4.3 Averages over square-full moduli

Suppose we are given a set of t indices $\mathcal{T} \subset \{1, \ldots, n\}$ and positive integers C_i for $i \in \mathcal{T}$. For $\boldsymbol{C} \coloneqq (C_i)_{i \in \mathcal{T}}$ we define $\mathcal{R}(\boldsymbol{C}) \subset \mathcal{O}^n$ to be the set of tuples $\boldsymbol{c} = (c_1, \ldots, c_n)$ such that $|c_i| = \widehat{C}_i$ if $i \in \mathcal{T}$ and $c_j = 0$ whenever $j \notin \mathcal{T}$. Given $Y \in \mathbb{Z}_{>0}$, we are interested in averages of the form

$$\mathcal{A}(\mathcal{R}(\boldsymbol{C}), \widehat{Y}) \coloneqq \sum_{\substack{\boldsymbol{c} \in \mathcal{R}(\boldsymbol{C}) \\ F^*(\boldsymbol{c}) \neq 0}} \sum_{\substack{r \in \mathcal{O} \\ |r| = \widehat{Y}}} |S_r(\boldsymbol{c})|, \tag{5.4.9}$$

where r is restricted to square-full polynomials.

Lemma 5.4.4. *With the notation from above, we have*

$$\mathcal{A}(\mathcal{R}(\boldsymbol{C}), \widehat{Y}) \ll_\varepsilon \widehat{Y}^{1+n/2+(n-t)/6}(\widehat{Y}\widehat{C})^\varepsilon \#\mathcal{R}(\boldsymbol{C}),$$

where $\widehat{C} = \max_{i\in\mathcal{T}} \widehat{C}_i$.

The proof of Lemma 5.4.4 is along the same lines as that of [106, Lemma 5.2], and so we shall be brief.

Proof. First of all, we introduce some notation. Fix $\boldsymbol{c} \in \mathcal{R}(\boldsymbol{C})$. For $r \in \mathcal{O}$ monic square-full, we write

$$r = r_* \prod_{i\in\mathcal{T}} r_i, \tag{5.4.10}$$

where the various coprime factors r_*, r_i are defined as follows. We let r_* be the product of those monic prime powers ϖ^k such that $\varpi^k \parallel r$ and $k = 2$ or $\varpi \nmid c_i$ for $i \in \mathcal{T}$. Moreover, for $i \in \mathcal{T}$, we define r_i to be the product of monic prime powers $\varpi^k \parallel r$ such that $\varpi \mid c_i$, but $\varpi \nmid c_j$ for any $j \in \mathcal{T}$ with $j < i$. In particular, any r_i is cube-full. Since all the factors in (5.4.10) are coprime, it follows from (5.4.4) that

$$S_r(\boldsymbol{c}) = S_{r_*}(\boldsymbol{c}) \prod_{i\in\mathcal{T}} S_{r_i}(\boldsymbol{c}).$$

Using the fact that $S_{\varpi^k}(\boldsymbol{c}) = 0$ if $\varpi \nmid F^*(\boldsymbol{c})$ for $k \geq 2$ and the estimates (5.4.7) and (5.4.8), we deduce that

$$S_r(\boldsymbol{c}) \ll \eta(r, \boldsymbol{c})|r|^{1+n/2+(n-t)/6+\varepsilon} \prod_{i,j\in\mathcal{T}} |\{r_i, c_j\}|^{1/4},$$

where $\eta(r, \boldsymbol{c}) = 1$ if $\varpi \mid F^*(\boldsymbol{c})$ for all primes $\varpi \mid r_*$ and $\eta(r, \boldsymbol{c}) = 0$ else. Let us now fix the absolute values of r_* and of the various r_i's, say $|r_*| = \widehat{Y}_*$ and $|r_i| = \widehat{Y}_i$, and denote their contribution to $\mathcal{A}(\mathcal{R}(\boldsymbol{C}), \widehat{Y})$ by $\mathcal{A}(Y_*, \boldsymbol{Y})$, where $\boldsymbol{Y} = (Y_i)_{i\in\mathcal{T}}$. We then have

$$\mathcal{A}(Y_*, \boldsymbol{Y}) \ll \widehat{Y}^{1+n/2+(n-t)/6+\varepsilon} \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} \sum_{\substack{|r_i|=\widehat{Y}_i\\ i\in\mathcal{T}}} \prod_{i,j\in\mathcal{T}} |\{r_i, c_j\}|^{1/4} S_{\boldsymbol{c}},$$

where we have suppressed the dependence of r_* and of the r_i's on $\boldsymbol{c}$ in the notation and where

$$S_{\boldsymbol{c}} = \sum_{|r_*|=\widehat{Y}_*} \eta(r, \boldsymbol{c}).$$

Heath-Brown's argument for estimating $S_{\boldsymbol{c}}$ goes through almost verbatim in our setting and gives $S_{\boldsymbol{c}} \ll (\widehat{Y}\widehat{C})^\varepsilon$. Therefore, we have

$$\mathcal{A}(Y_*, \boldsymbol{Y}) \ll \widehat{Y}^{1+n/2+(n-t)/6+\varepsilon}(\widehat{Y}\widehat{C})^\varepsilon \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} \sum_{\substack{|r_i|=\widehat{Y}_i\\ i\in\mathcal{T}}} \prod_{i,j\in\mathcal{T}} |\{r_i, c_j\}|^{1/4}.$$

To achieve the desired upper bound, we shall now only require that each r_i is cube-full and that $\varpi \mid c_i$ whenever $\varpi \mid r_i$, so that in particular the r_i's do not depend on $\boldsymbol{c}$ anymore. Thus, after setting

$$S(j) = \sum_{|c_j|=\widehat{C}_j} \prod_{i\in\mathcal{T}} |\{r_i, c_j\}|^{1/4},$$

we obtain

$$\mathcal{A}(Y_*, \boldsymbol{Y}) \ll \widehat{Y}^{1+n/2+(n-t)/6+\varepsilon}(\widehat{Y}\widehat{C})^{\varepsilon} \sum_{\substack{|r_i|=\widehat{C}_i \\ i\in\mathcal{T}}} \prod_{j\in\mathcal{T}} S(j). \tag{5.4.11}$$

It is again straightforward to verify that Heath-Brown's argument continues to hold in our setting, yielding

$$\sum_{\substack{|r_i|=\widehat{C}_i \\ i\in\mathcal{T}}} \prod_{j\in\mathcal{T}} S(j) \ll \widehat{Y}^{(n+1)\varepsilon} \#\mathcal{R}(\boldsymbol{C}).$$

With a new choice of ε, we conclude

$$\mathcal{A}(Y_*, \boldsymbol{Y}) \ll \widehat{Y}^{1+n/2+(n-t)/6}(\widehat{Y}\widehat{C})^{\varepsilon} \#\mathcal{R}(\boldsymbol{C}),$$

so that the statement of the lemma follows from the fact that there are only $\widehat{Y}^{\varepsilon}$ possibilities for admissible tuples $(Y_*, \boldsymbol{Y})$. ☐

5.5 Rational points on the dual hypersurface

In this section we study roots of the dual form F^* of F that was defined in (5.4.2). Our first goal is to find an upper bound for the number of solutions $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| \leq \widehat{C}$ when $\mathrm{char}(K) > 3$. In order to achieve this we closely follow the strategy of Heath-Brown [98, Section 7]. The result of Lemma 5.5.2 is standard over the rational numbers, however we could not find a proof in the literature for our setting and so we included a proof here.

If $n = 4$ and $\mathrm{char}(K) > 3$ we call a solution $\boldsymbol{c}$ to $F^*(\boldsymbol{c}) = 0$ *special* if $c_1, \dots, c_4 \neq 0$ and there are indices i, j, k, l such that $\{i, j, k, l\} = \{1, 2, 3, 4\}$ and

$$(F_i^{-1}c_i^3)^{1/2} + (F_j^{-1}c_j^3)^{1/2} = (F_k^{-1}c_k^3)^{1/2} + (F_l^{-1}c_l^3)^{1/2} = 0$$

holds for a suitable choice of square roots. We call a solution $\boldsymbol{c}$ to $F^*(\boldsymbol{c}) = 0$ *ordinary* if it is not special. In particular, if $\mathrm{char}(K) = 2$ every solution is ordinary.

Lemma 5.5.1. *Assume $\mathrm{char}(K) > 3$. If $n = 6$, then the number of solutions to $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| \leq \widehat{C}$ is bounded by $O(\widehat{C}^{3+\varepsilon})$. Moreover, if $n = 4$, then the number of ordinary solutions to $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| \leq \widehat{C}$ is bounded by $O(\widehat{C}^{1+\varepsilon})$.*

Before we can begin with the proof of this lemma, we need an auxiliary result. In the following we fix $\zeta \in \mathbb{F}_q^{\times}$ to be a representative of a non-trivial element in $\mathbb{F}_q^{\times}/\mathbb{F}_q^{\times,2}$. If $\mathrm{char}(\mathbb{F}_q) > 2$ this certainly exists — we may for example pick ζ to be a primitive root of $\mathbb{F}_q^{\times}$.

Lemma 5.5.2. *Suppose $\mathrm{char}(K) > 3$. Let $m_1, \dots, m_n \in \mathcal{O}$ be a collection of distinct square-free polynomials such that each m_i is either monic or has leading coefficient ζ. Then $\{\sqrt{m_1}, \dots, \sqrt{m_n}\}$ is a linearly independent set over K.*

Proof. We will prove the result by induction on n. The cases $1 \leq n \leq 3$ can easily be verified directly, so suppose $n \geq 4$. Assume for a contradiction that $\lambda_1, \dots, \lambda_n \in K$ not all zero are such that

$$\sum_{k=1}^{n} \lambda_k \sqrt{m_k} = 0.$$

Note that we may assume $\lambda_i \neq 0$ for all $i = 1, \dots, n$ since otherwise the result would follow immediately from the induction hypothesis. In particular it is sufficient to show that there exists some index k with $\lambda_k = 0$. Since $n \geq 3$ there exist two distinct indices i, j such that $m_i/m_j \notin \mathbb{F}_q^\times$. From the $n = 3$ case it follows that $K_{i,j} := K(\sqrt{m_i}, \sqrt{m_j})$ is a Galois extension of degree 4 over K. Thus there exists $\sigma \in \mathrm{Gal}(K_{i,j}/K)$ such that $\sigma(\sqrt{m_i}) = -\sqrt{m_i}$ and $\sigma(\sqrt{m_j}) = \sqrt{m_j}$. We may lift this to an element $\widetilde{\sigma} \in \mathrm{Gal}(K^s/K)$ where K^s is the separable closure of K. Then we find

$$0 = \widetilde{\sigma}\left(\sum_{k=1}^{n} \lambda_k \sqrt{m_k}\right) + \sum_{k=1}^{n} \lambda_k \sqrt{m_k} = 2\lambda_j \sqrt{m_j} + \sum_{k \neq i,j} \widetilde{\lambda}_k \sqrt{m_k},$$

where $\widetilde{\lambda}_k \in \{0, 2\lambda_k\}$. From the induction hypothesis we get $\lambda_j = 0$, which yields the desired result as remarked above. □

Proof of Lemma 5.5.1. First note that $F^*(\boldsymbol{c}) = 0$ if and only if

$$(F_1^{-1}c_1^3)^{1/2} + \cdots + (F_n^{-1}c_n^3)^{1/2} = 0, \tag{5.5.1}$$

for a suitable choice of square roots. Let $m_k \in \mathcal{O}$ be a square-free polynomial, which is either monic or has leading coefficient ζ. Say $i \in \mathcal{I}(k)$ if there exists some $d_i \in \mathcal{O}$ such that $F_i c_i^3 = m_k d_i^2$. From Lemma 5.5.2 we find that (5.5.1) implies

$$\sum_{i \in \mathcal{I}(k)} F_i^{-1} d_i = 0.$$

We have $c_i^2 \mid m_k d_i^2$ and consequently $c_i \mid d_i$ since m_k is square-free. Thus there exists $e_i \in \mathcal{O}$ such that $d_i = c_i e_i$. Substituting this into the relation $F_i c_i^3 = m_k d_i^2$ we find $c_i = m_k F_i^{-1} e_i^2$ and hence $d_i = c_i e_i = m_k F_i^{-1} e_i^3$. Therefore $F_i^{-1} d_i = m_k F_i \left(\frac{e_i}{F_i}\right)^3$ and the preceding display gives

$$\sum_{i \in \mathcal{I}(k)} F_i \left(\frac{e_i}{F_i}\right)^3 = 0. \tag{5.5.2}$$

We will now estimate the number of solutions $\boldsymbol{e}$ to (5.5.2) such that $|\boldsymbol{e}| \leq \widehat{E} = \sqrt{\widehat{C}/|m_k|}$. This will then enable us to estimate the number of solutions of (5.5.1). Via Hölder's inequality and Hua's Lemma in this context (cf. [35, Lemma 5.12]) we find

$$\#\left\{|\boldsymbol{e}| \leq \widehat{E} \colon \sum_{i \in \mathcal{I}(k)} F_i \left(\frac{e_i}{F_i}\right)^3 = 0\right\} \ll \begin{cases} 1 & \text{if } \#\mathcal{I}(k) = 1, \\ \widehat{E}^{2+\varepsilon} & \text{if } 2 \leq \#\mathcal{I}(k) \leq 4, \\ \widehat{E}^{\#\mathcal{I}(k)-2+\varepsilon} & \text{if } 5 \leq \#\mathcal{I}(k) \leq 6. \end{cases}$$

Note that at this point it is crucial to assume $\mathrm{char}(K) > 3$, because the Weyl differencing argument in the proof of Hua's lemma breaks down otherwise. Therefore for a fixed partition $\bigsqcup_j \mathcal{I}(k_j) = \{1, \ldots, n\}$ corresponding to $\{m_{k_j}\}$ the number of $|\boldsymbol{c}| \leq \widehat{C}$ satisfying (5.5.1) is bounded above by

$$\prod_j \left(\frac{\widehat{C}}{|m_{k_j}|}\right)^{e_{k_j}/2+\varepsilon},$$

where

$$e_{k_j} = \begin{cases} 0, & \text{if } \#\mathcal{I}(k_j) = 1 \\ 2, & \text{if } 2 \leq \#\mathcal{I}(k_j) \leq 4 \\ 3, & \text{if } \#\mathcal{I}(k_j) = 5 \\ 4, & \text{if } \#\mathcal{I}(k_j) = 6. \end{cases}$$

By considering all possible square-free elements $|m_{k_j}| \ll \widehat{C}$, we see that the total number of solutions of (5.5.1) corresponding to a fixed partition is bounded above by

$$\sum_{|m_{k_j}| \leq \widehat{C}} \prod_j \left(\frac{\widehat{C}}{|m_{k_j}|} \right)^{e_{k_j}/2+\varepsilon} \ll \prod_j \widehat{C}^{e_{k_j}/2+\varepsilon}.$$

It is easily checked that for any possible partition this is bounded above by $O(\widehat{C}^{3+\varepsilon})$ if $n = 6$. Therefore the total number of solutions to $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| \leq \widehat{C}$ has the same upper bound. In the case $n = 4$ one can similarly obtain $O(\widehat{C}^{1+\varepsilon})$ for the number of solutions corresponding to any partition, except in the case where $\#\mathcal{I}(k_1) = \#\mathcal{I}(k_2) = 2$. But solutions arising from such partitions are precisely the special solutions. This finishes the proof of the lemma. □

5.6 Circle method

As explained in the introduction, we are considering a diagonal cubic form $F \in \mathcal{O}[x_1, \ldots, x_n]$ of the shape

$$F(\boldsymbol{x}) = \sum_{i=1}^{n} F_i x_i^3, \quad F_i \in \mathcal{O} \setminus \{0\}.$$

Recall from (5.2.1) that the associated counting function can be written as

$$N(w, P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \int_{|\theta| < |r|^{-1}\widehat{Q}^{-1}} \sum_{\boldsymbol{c} \in \mathcal{O}^n} S_r(\boldsymbol{c}) I_r(\theta, \boldsymbol{c}) \mathrm{d}\theta.$$

Throughout the parameter Q is chosen in such a way that

$$|P|^{3/2} \leq \widehat{Q} \leq q|P|^{3/2} \tag{5.6.1}$$

ensuring that the measure of the set $\{|\theta| < |r|^{-1}\widehat{Q}^{-1}\}$ is $O(|P|^{-3})$ when $|r| = \widehat{Q}$. It follows from Lemma 5.3.1 that $I_r(\theta, \boldsymbol{c})$ vanishes unless $|\boldsymbol{c}| < |r||P|^{-1} \max\{q, H_F|P|^3\theta\}$. Since $H_F|P|^3|\theta| \leq H_F|P|^3\widehat{Q}^{-1}|r|^{-1}$ and $|P|^3\widehat{Q}^{-1}|r|^{-1} \gg 1$, we can truncate the sum over $\boldsymbol{c}$ in (5.2.1) at $|\boldsymbol{c}| \ll \widehat{C}$, where $\widehat{C} := |P|^2\widehat{Q}^{-1}$.

We now split up $N(w, P)$ according to the quality of our available estimates into

$$N(w, P) = N_0(P) + E_1(P) + E_2(P),$$

where

$$N_0(P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \int_{|\theta| < |r|^{-1}\widehat{Q}^{-1}} S_r(\boldsymbol{0}) I_r(\theta, \boldsymbol{0}) \mathrm{d}\theta, \tag{5.6.2}$$

$$E_1(P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \int_{|\theta| < |r|^{-1}\widehat{Q}^{-1}} \sum_{\substack{\boldsymbol{c} \in \mathcal{O}^n \\ F^*(\boldsymbol{c}) \neq 0}} S_r(\boldsymbol{c}) I_r(\theta, \boldsymbol{c}) \mathrm{d}\theta, \tag{5.6.3}$$

$$E_2(P) = |P|^n \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-n} \int_{|\theta| < |r|^{-1}\widehat{Q}^{-1}} \sum_{\substack{\boldsymbol{c} \in \mathcal{O}^n \setminus \{\boldsymbol{0}\} \\ F^*(\boldsymbol{c}) = 0}} S_r(\boldsymbol{c}) I_r(\theta, \boldsymbol{c}) \mathrm{d}\theta. \tag{5.6.4}$$

For $n = 4$ we will later divide the term $E_2(P)$ into special and ordinary solutions of $F^*(\boldsymbol{c}) = 0$ as defined in Section 5.5. Usually one expects that the main term in an asymptotic formula for

$N(w,P)$ should come from $N_0(P)$. As we are only interested in an upper bound for $N(w,P)$, the contribution from $N_0(P)$ will be rather straightforward to deal with. Handling the terms $E_1(P)$, $E_2(P)$ turns out to be a more challenging task and will occupy most of the remainder of our work. For $E_1(P)$ we can make use of the full power of our exponential sum estimates, in particular we gain an extra saving when averaging $S_r(\boldsymbol{c})$ over r. This is not possible for $E_2(P)$, but we shall benefit from the sparsity of $\boldsymbol{c}$'s such that $F^*(\boldsymbol{c})=0$, at least for ordinary solutions when $n=4$.

5.6.1 Contribution from $N_0(P)$

For this we write again $r=r_1r_2$, where r_1 is cube-free and r_2 is cube-full. It thus follows from (5.4.6) and Lemma 5.4.3 with $m=0$ that

$$S_r(\boldsymbol{c}) \ll |r_1|^{1+n/2+\varepsilon}|r_2|^{1+2n/3+\varepsilon}.$$

From Lemma 5.3.7 we obtain the estimate $I_r(\mathbf{0}) \ll |P|^{-3}$. We thus get

$$\begin{aligned} N_0(P) &\ll |P|^{n-3} \sum_{|r_1|\le \widehat{Q}} |r_1|^{-n} S_{r_1}(\boldsymbol{c}) \sum_{|r_2|\le \widehat{Q}/|r_1|} |r_2|^{-n} S_{r_2}(\boldsymbol{c}) \\ &\ll |P|^{n-3} \sum_{|r_1|\le \widehat{Q}} |r_1|^{1-n/2+\varepsilon} \sum_{|r_2|\le \widehat{Q}/|r_1|} |r_2|^{1-n/3+\varepsilon} \\ &\ll |P|^{n-3+\varepsilon}, \end{aligned}$$

since there are $O(\widehat{Y}^{1/3})$ cube-full r_2 with $|r_2|=\widehat{Y}$ and $n\ge 4$.

5.6.2 Contribution from $E_1(P)$

We begin with some preparations for the term $E_1(P)$. Let $0\le Y\le Q$ and fix the absolute value of r to be $\widehat{Y}$. As in Section 5.4.3, we will also fix a set of indices $\mathcal{T}\subset\{1,\dots,n\}$ of cardinality t, as well as a tuple $\boldsymbol{C}=(C_i)_{i\in\mathcal{T}}$, where $1\le C_i\le C$ and denote by $\mathcal{R}(\boldsymbol{C})$ the set of vectors $\boldsymbol{c}=(c_1,\dots,c_n)\in\mathcal{O}^n$ such that $|c_i|=\widehat{C}_i$ if $i\in\mathcal{T}$ and $c_j=0$ if $j\notin\mathcal{T}$. Let us put $\mathcal{C}=\max_{i\in\mathcal{T}}C_i$, so that $|\boldsymbol{c}|=\widehat{\mathcal{C}}$ whenever $\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})$. We then define $E_1(\mathcal{R}(\boldsymbol{C}),\widehat{Y})$ to be the contribution coming from $\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})$ and $|r|=\widehat{Y}$ in the definition of $E_1(P)$ given in (5.6.3). Explicitly, this means

$$E_1(\mathcal{R}(\boldsymbol{C}),\widehat{Y}) = \frac{|P|^n}{\widehat{Y}^n} \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} \sum_{\substack{r \text{ monic}\\ |r|=\widehat{Y}}} S_r(\boldsymbol{c}) I_{t^Y}(\boldsymbol{c}), \tag{5.6.5}$$

where we recall that

$$I_{t^Y}(\boldsymbol{c}) = \int_{|\theta|<\widehat{Y}^{-1}\widehat{Q}^{-1}} I_{t^Y}(\theta,\boldsymbol{c})\mathrm{d}\theta.$$

Note that here we used Lemma 5.3.5, which shows that the value of the double integral in the definition of $I_r(\boldsymbol{c})$ only depends on the absolute value of r for $\boldsymbol{c}$ fixed.

Note that there are $Q+1\ll |P|^{\varepsilon}$ possibilities for Y and $O(C^n)=O(|P|^{\varepsilon})$ choices for $\boldsymbol{C}$. In particular, if we can show that $E_1(\mathcal{R}(\boldsymbol{C}),\widehat{Y}) \ll |P|^{3n/4-3/2+\varepsilon}$ holds, then the same estimate will be true for $E_1(P)$ with a new value of $\varepsilon>0$. Next we tansform $E_1(P)$ in such a way that Lemma 5.4.2 and Lemma 5.4.4 are applicable. For this we write $r=b_1'b_1r_2$, where r_2 is the square-full part of r and $b_1'b_1$ is the square-free part of r. Moreover, if we let S be the set of prime divisors of $\Delta_F F^*(\boldsymbol{c})$, then we further require that $(b_1, S) = 1$ and each prime $\varpi \mid b_1'$ satisfies $\varpi \in S$. It then follows from (5.4.4) that $E_1(\mathcal{R}(\boldsymbol{C}), \widehat{Y})$ is given by

$$\frac{|P|^n}{\widehat{Y}^{(n-1)/2}} \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} I_{t^Y}(\boldsymbol{c}) \sum_{|r_2|\leq \widehat{Y}} \frac{S_{r_2}(\boldsymbol{c})}{|r_2|^{(n+1)/2}} \sum_{|b_1'|\leq \frac{\widehat{Y}}{|r_2|}} \frac{S_{b_1'}(\boldsymbol{c})}{|b_1'|^{(n+1)/2}} \sum_{\substack{|b_1|=\frac{\widehat{Y}}{|r_2 b_1'|}\\ (b_1,S)=1}} \frac{S_{b_1}(\boldsymbol{c})}{|b_1|^{(n+1)/2}}. \tag{5.6.6}$$

We can now apply Lemma 5.4.2 to the innermost sum to obtain

$$\sum_{\substack{|b_1|=\frac{\widehat{Y}}{|r_2 b_1'|}\\ (b_1,S)=1}} \frac{S_{b_1}(\boldsymbol{c})}{|b_1|^{(n+1)/2}} \ll \widehat{C}^{\varepsilon}(\widehat{Y}|r_2 b_1'|^{-1})^{1/2+\varepsilon}. \tag{5.6.7}$$

Moreover, by (5.4.6) and (5.4.4) we also have

$$\sum_{|b_1'|\leq \frac{\widehat{Y}}{|r_2|}} \frac{|S_{b_1'}(\boldsymbol{c})|}{|b_1'|^{n/2+1}} \ll |P|^{\varepsilon} \sum_{|b_1'|\leq \widehat{Y}/|r_2|} \frac{|(b_1', \nabla F^*(\boldsymbol{c}))|^{1/2}}{|b_1'|^{1/2}} \ll |P|^{\varepsilon}, \tag{5.6.8}$$

where we used that there at most $O((\widehat{Y}|r_2|^{-1}|F^*(\boldsymbol{c})|)^{\varepsilon}) = O(|P|^{\varepsilon})$ possibilities for square-free b_1' whose prime divisors are restricted to S with $|b_1'| \leq \widehat{Y}|r_2|^{-1}$. After inserting (5.6.7) and (5.6.8) into (5.6.6), we see that

$$E_1(\mathcal{R}(\boldsymbol{C}), \widehat{Y}) \ll \frac{|P|^{n+\varepsilon}}{\widehat{Y}^{n/2-1}} \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} |I_{t^Y}(\boldsymbol{c})| \sum_{|r_2|\leq \widehat{Y}} \frac{|S_{r_2}(\boldsymbol{c})|}{|r_2|^{n/2+1}}.$$

We can now estimate $I_{t^Y}(\boldsymbol{c})$ with Lemma 5.3.6:

$$\begin{aligned} I_{t^Y}(\boldsymbol{c}) &\ll \widehat{Y}^{-1}\widehat{Q}^{-1} \prod_{i=1}^{n} \min\left\{\left(\frac{|P||\boldsymbol{c}|}{\widehat{Y}}\right)^{-1/4}, \left(\frac{|P||c_i|}{\widehat{Y}}\right)^{-1/2}\right\} \\ &= \widehat{Y}^{-1}\widehat{Q}^{-1}\left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{(n-t)/4} \prod_{i\in\mathcal{T}} \min\left\{\left(\frac{|P|\widehat{\mathcal{C}}}{\widehat{Y}}\right)^{-1/4}, \left(\frac{|P|\widehat{C}_i}{\widehat{Y}}\right)^{-1/2}\right\}, \end{aligned}$$

where we used that $\min\left\{\left(\frac{|P|\widehat{\mathcal{C}}}{\widehat{Y}}\right)^{-1/4}, \left(\frac{|P||c_i|}{\widehat{Y}}\right)^{-1/2}\right\} = (|P|\widehat{\mathcal{C}}\widehat{Y}^{-1})^{-1/4}$ if $i \notin \mathcal{T}$. Denote the last product above by Π. Then after dividing r_2 into q-adic ranges, Lemma 5.4.4 implies

$$\begin{aligned} E_1(\mathcal{R}(\boldsymbol{C}), \widehat{Y}) &\ll \frac{|P|^{n+\varepsilon}}{\widehat{Y}^{n/2}\widehat{Q}}\left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{(n-t)/4} \Pi \sum_{\substack{\boldsymbol{c}\in\mathcal{R}(\boldsymbol{C})\\ F^*(\boldsymbol{c})\neq 0}} \sum_{|r_2|\leq \widehat{Y}} \frac{|S_{r_2}(\boldsymbol{c})|}{|r_2|^{n/2+1}} \\ &\ll \frac{|P|^{n+\varepsilon}}{\widehat{Y}^{n/2}\widehat{Q}}\left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{(n-t)/4} \widehat{Y}^{(n-t)/6}\Pi\#\mathcal{R}(\boldsymbol{C}). \end{aligned}$$

From the fact that $\#\mathcal{R}(\boldsymbol{C}) \ll \prod_{i\in\mathcal{T}} \widehat{C}_i$ we deduce that

$$\begin{aligned} \#\mathcal{R}(\boldsymbol{C})\Pi &\ll \prod_{i\in\mathcal{T}} \min\left\{\widehat{C}_i\left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{1/4}, \left(\frac{\widehat{C}_i\widehat{Y}}{|P|}\right)^{1/2}\right\} \\ &\ll \widehat{\mathcal{C}}^{\,t}\left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{t/4} \min\left\{1, \frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right\}^{t/4}, \end{aligned}$$

where we used that $\widehat{C_i} \leq \widehat{\mathcal{C}}$. Recalling (5.6.1), we therefore have

$$E_1(\mathcal{R}(\boldsymbol{C}), \widehat{Y}) \ll \frac{|P|^{n-3/2+\varepsilon}}{\widehat{Y}^{n/2}} \left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{n/4} \widehat{Y}^{(n-t)/6}\widehat{\mathcal{C}}^{\,t} \min\left\{1, \frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right\}^{t/4}.$$

One easily sees that the expression above is maximal either at $t = 0$ or $t = n$. For $t = 0$, we get

$$\begin{aligned} \frac{|P|^{n-3/2+\varepsilon}}{\widehat{Y}^{n/2}} \left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{n/4} \widehat{Y}^{n/6} &= |P|^{3n/4-3/2+\varepsilon}\widehat{Y}^{-n/12}\widehat{\mathcal{C}}^{-n/4} \\ &\ll |P|^{3n/4-3/2+\varepsilon} \end{aligned}$$

as desired. For $t = n$, we have

$$\begin{aligned} \frac{|P|^{n-3/2+\varepsilon}}{\widehat{Y}^{n/2}} \left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{n/4} \widehat{\mathcal{C}}^{n} \min\left\{1, \frac{\widehat{Y}}{\widehat{\mathcal{C}}|P|}\right\}^{n/4} &\ll |P|^{n/2-3/2+\varepsilon}\widehat{\mathcal{C}}^{n/2} \\ &\ll |P|^{3n/4-3/2+\varepsilon} \end{aligned}$$

since $\widehat{\mathcal{C}} \leq \widehat{C} \ll |P|^{1/2}$. This finishes our treatment of $E_1(P)$.

5.6.3 Contribution from $E_2(P)$ for ordinary solutions

Now we turn our attention to the term $E_2(P)$. For $n = 4$ we further divide it into $E_2(P) = E_2^{\mathsf{ord}}(P) + E_2^{\mathsf{spec}}(P)$, where $E_2^{\mathsf{spec}}(P)$ is restricted to special solutions of $F^*(\boldsymbol{c}) = 0$ in the sense of Section 5.5 and $E_2^{\mathsf{ord}}(P)$ to ordinary solutions of $F^*(\boldsymbol{c}) = 0$. In this section we deal with $E_2(P)$ for $n = 6$ and $E_2^{\mathsf{ord}}(P)$ for $n = 4$.

We shall again fix the absolute value of r to be $\widehat{Y}$ for some $0 \leq Y \leq Q$ and the absolute value of $\boldsymbol{c}$ to be $\widehat{\mathcal{C}}$ for some $0 < \mathcal{C} \leq C$. We will then consider the sum

$$E_2(Y, \mathcal{C}) := \frac{|P|^n}{\widehat{Y}^n} \sum_{\substack{|\boldsymbol{c}|=\widehat{\mathcal{C}} \\ F^*(\boldsymbol{c})=0}} \sum_{\substack{r \text{ monic} \\ |r|=\widehat{Y}}} S_r(\boldsymbol{c}) I_{t^Y}(\boldsymbol{c}),$$

where the sum over $\boldsymbol{c}$ is restricted to ordinary solutions of $F^*(\boldsymbol{c}) = 0$ for $n = 4$. Once we have shown $E_2(Y, \mathcal{C}) \ll |P|^{3n/4-3/2+\varepsilon}$ the same estimate will follow for $E_2(P)$ for $n = 6$ and for $E_2^{\mathsf{ord}}(P)$ for $n = 4$, because there are only $O(|P|^{\varepsilon})$ possible pairs of Y's and $\mathcal{C}$'s.

Lemma 5.6.1. *Let F be a non-singular cubic form in 4 or 6 variables, and let F^* be its dual form. Suppose there exists some $\eta > 0$ such that for any $\widehat{\mathcal{C}} \geq 1$ the following bound holds*

$$\#\{\boldsymbol{x} \in \mathcal{O}^n \colon \boldsymbol{x} \text{ is an ordinary solution to } F^*(\boldsymbol{x}) = 0, |\boldsymbol{x}| \leq \widehat{\mathcal{C}}\} \ll \widehat{\mathcal{C}}^{n-3+\eta}.$$

Then we have

$$E_2(P) \ll |P|^{3n/4-3/2+\eta/2+\varepsilon}.$$

Proof. If $D = \deg F^*$, then we see from (5.4.2) and Lemma 5.4.1 that F^* has non-zero monomials of the form $G_i x_i^D$ for every $i = 1, \ldots, n$. In particular, if $|\boldsymbol{c}| = \widehat{\mathcal{C}}$ and $F^*(\boldsymbol{c}) = 0$, then there must be at least two indices $i \neq j$ such that $\widehat{\mathcal{C}} \ll |c_i| \ll |c_j| \ll \widehat{\mathcal{C}}$. Therefore, from Lemma 5.3.6 we deduce

$$I_{t^Y}(\boldsymbol{c}) \ll \frac{\widehat{\mathcal{C}}}{|P|^2\widehat{Y}} \prod_{i=1}^{n} \min\left\{ \left(\frac{\widehat{Y}}{|P||c_i|}\right)^{1/2}, \left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{1/4} \right\} \ll \left(\frac{\widehat{Y}}{|P|\widehat{\mathcal{C}}}\right)^{(n-2)/4} |P|^{-3}. \tag{5.6.9}$$

Next we deal with the sum $S_r(\boldsymbol{c})$. Write $r = r_1r_2r_3$ into coprime monic factors r_i, where r_1 is cube-free, r_2 is cube-full and each prime divisor of r_3 divides $\prod F_i$.

Let us begin with $S_{r_2}(\boldsymbol{c})$. Suppose $\varpi^k \parallel r_2$ and write $H_\varpi = (\varpi^k, \boldsymbol{c})$. It follows that $\boldsymbol{c} = H_\varpi \boldsymbol{c}'$ for some $\boldsymbol{c}' \in \mathcal{O}^n$ with $(\varpi, \boldsymbol{c}') = 1$. It is again easy to see that any prime divisor of the coefficients G_i of the top-degree monomials x_i^D of F^* divides $\prod F_i$. In particular, if $H_\varpi \neq \varpi^k$, then $F^*(\boldsymbol{c}') = 0$ implies that at least two entries of $\boldsymbol{c}'$ are coprime to ϖ. On the other hand, if $H_\varpi = \varpi^k$, then $(\varpi^k, c_i) = \varpi^k$ for every $i = 1, \ldots, n$, so that in any case there are always least two distinct indices $i \neq j$ such that $(\varpi^k, c_i) = (\varpi^k, c_j) = H_\varpi$. Consequently it follows from Lemma 5.4.3 with $m = 2$ that

$$S_{r_2}(\boldsymbol{c}) \ll |r_2|^{2/3+2n/3+\varepsilon}|H|^{1/2},$$

where $H = \prod_{\varpi | r_2} H_\varpi$ divides each entry of $\boldsymbol{c}$.

In addition, (5.4.6) and Lemma 5.4.3 give us $S_{r_1}(\boldsymbol{c}) \ll |r_1|^{1+n/2+\varepsilon}$ and (5.4.8) tells us that $S_{r_3}(\boldsymbol{c}) \ll |r_3|^{1+2n/3+\varepsilon}$. To sum up, we have

$$S_r(\boldsymbol{c}) \ll |r|^\varepsilon |r_1|^{1+n/2}|r_2|^{2/3+2n/3}|r_3|^{1+2n/3}|H|^{1/2}.$$

Let us fix $|r_i| = \widehat{Y}_i$, where $0 \leq Y_i \leq Y$ and $Y_1 + Y_2 + Y_3 = Y$. We want to give an upper bound for

$$\mathcal{S} := \sum_{|r_i| = \widehat{Y}_i, i=1,2,3} \sum_{\substack{|\boldsymbol{c}| = \widehat{C} \\ F^*(\boldsymbol{c})=0}} |S_r(\boldsymbol{c})|.$$

Taking into account that the number of available r_1 and r_3 is $O(\widehat{Y}_1)$ and $O(|P|^\varepsilon)$ respectively, we see that

$$\begin{aligned}
\mathcal{S} &\ll |P|^\varepsilon \widehat{Y}_1^{2+n/2}\widehat{Y}_2^{2/3+2n/3}\widehat{Y}_3^{1+2n/3} \sum_{|r_2|=\widehat{Y}_2} \sum_{H|r_2} |H|^{1/2} \sum_{\substack{|\boldsymbol{c}|=\widehat{C}/|H| \\ F^*(\boldsymbol{c})=0}} 1 \\
&\ll |P|^\varepsilon \widehat{C}^{n-3+\eta}\widehat{Y}_1^{2+n/2}\widehat{Y}_2^{2/3+2n/3}\widehat{Y}_3^{1+2n/3} \sum_{|r_2|=\widehat{Y}_2} \sum_{H|r_2} |H|^{7/2-n-\eta},
\end{aligned}$$

where we used the main assumption of the lemma in order to bound the number of ordinary solutions of $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| = \widehat{C}/|H|$ for the second inequality. Since $n \geq 4$ clearly $7/2 - n - \eta \leq 0$ holds and since the number of available r_2 is $O(\widehat{Y}_2^{1/3})$, it follows that

$$\mathcal{S} \ll |P|^\varepsilon \widehat{C}^{n-3+\eta}\widehat{Y}_1^{2+n/2}\widehat{Y}_2^{1+2n/3}\widehat{Y}_3^{1+2n/3} \ll |P|^\varepsilon \widehat{C}^{n-3+\eta}\widehat{Y}^{2+n/2}, \tag{5.6.10}$$

because $2 + n/2 \geq 1 + 2n/3$ for $n \leq 6$. As there are only $O(|P|^\varepsilon)$ possibilities for permissible triples (Y_1, Y_2, Y_3), we deduce from (5.6.9) and (5.6.10) that

$$E_2(Y, \widehat{C}) \ll |P|^{3n/4-5/2+\varepsilon}\widehat{Y}^{3/2-n/4}\widehat{\mathcal{C}}^{3n/4-5/2+\eta}.$$

In particular, since $\widehat{\mathcal{C}} \ll |P|^{1/2}$ and $\widehat{Y} \ll |P|^{3/2}$, we thus obtain

$$\begin{aligned}
E_2(Y, \mathcal{C}) &\ll |P|^{3n/4-5/2+\varepsilon}|P|^{9/4-3n/8}|P|^{3n/8-5/4+\eta/2} \\
&\ll |P|^{3n/4-3/2+\eta/2+\varepsilon},
\end{aligned}$$

which completes the proof. □

At this point our treatment of $E_2(P)$ differs depending on the characteristic of K.

If $\mathrm{char}(K) > 3$, then by virtue of Lemma 5.5.1 we know that the number of ordinary solutions of the dual form $F^*(\boldsymbol{c}) = 0$ such that $|\boldsymbol{c}| \leq \widehat{\mathcal{C}}$ is bounded by $O(\widehat{\mathcal{C}}^{n-3+\varepsilon})$. Therefore Lemma 5.6.1 with $\eta = 0$ implies

$$E_2^{\mathrm{ord}}(P) \ll |P|^{3n/4-3/2+\varepsilon} \quad \text{and} \quad E_2(P) \ll |P|^{3n/4-3/2+\varepsilon},$$

for $n = 4$ and $n = 6$, respectively. This finishes our treatment of $E_2(P)$ in this case.

If $\mathrm{char}(K) = 2$, then we need to argue differently. We begin by considering the case when $n = 6$. According to Lemma 5.4.1 the dual form takes the shape of a non-singular diagonal cubic form. In particular, we can trivially bound the number of solutions to $F^*(\boldsymbol{c}) = 0$ such that $|\boldsymbol{c}| \leq \widehat{\mathcal{C}}$ by $O(\widehat{\mathcal{C}}^6) = O(\widehat{\mathcal{C}}^{n-3+\eta})$, where $\eta = 3$. Therefore, Lemma 5.6.1 gives

$$E_2(P) \ll |P|^{3n/4-3/2+\eta/2+\varepsilon} = |P|^{n-3+\eta/2+\varepsilon}.$$

This, together with our bounds for $N_0(P)$ and $E_1(P)$ established earlier in this section, shows that

$$N(P) \ll |P|^{n-3+\eta/2+\varepsilon}.$$

This holds for any non-singular, diagonal cubic form over K when $\mathrm{char}(K) = 2$. In particular, as a result we can bound the number of solutions to $F^*(\boldsymbol{c}) = 0$ with $|\boldsymbol{c}| \leq \widehat{\mathcal{C}}$ by $O(\widehat{\mathcal{C}}^{n-3+\eta/2+\varepsilon})$. Another application of Lemma 5.6.1 yields

$$E_2(P) \ll |P|^{3n/4-3/2+\eta/4+\varepsilon}$$

and we may argue as above to deduce

$$N(P) \ll |P|^{n-3+\eta/4+\varepsilon}.$$

If we repeat this process k-times, where $2^{-k+1} \leq \varepsilon$ we find

$$E_2(P) \ll |P|^{3n/4-3/2+2\varepsilon},$$

which concludes our treatment for $E_2(P)$ in this case.

On the other hand, if $n = 4$ we can trivially estimate the number of solutions to $F^*(\boldsymbol{c}) = 0$ of bounded height $\widehat{\mathcal{C}}$ by $O(\widehat{\mathcal{C}}^4) = O(\widehat{\mathcal{C}}^{n-3+\eta})$, where $\eta = 3$. Lemma 5.6.1 then yields

$$E_2(P) \ll |P|^{3n/4-3/2+\eta/2+\varepsilon} = |P|^{n-3+1/2+\eta/2+\varepsilon},$$

which in turn implies

$$N(P) \ll |P|^{n-3+1/2+\eta/2+\varepsilon}.$$

Repeating this process k-times, where $k > 1/\varepsilon$ we thus find

$$E_2(P) \ll |P|^{3n/4-3/2+1/2+2\varepsilon} = |P|^{2+2\varepsilon}.$$

5.7 Waring's problem and weak approximation

Having completed our task for $n = 6$, we will now apply it to Waring's problem and weak approximation for diagonal cubic hypersurfaces of dimension at least 5.

5.7.1 Waring's problem for $n \geq 7$

Recall that $\mathbb{J}_q^3[t]$ is the additive closure of all cubes in $\mathcal{O}$. Given $P \in \mathbb{J}_q^3[t]$, we define $B := \left\lceil \frac{\deg(P)}{3} \right\rceil + 1$ and the counting function

$$R_n(P) := \#\{\boldsymbol{x} \in \mathcal{O}^n \colon |\boldsymbol{x}| < \widehat{B}, x_1^3 + \cdots + x_n^3 = P\}.$$

Our next goal is to deduce Theorem 5.1.4 from our findings. We shall accomplish this goal with a classical version of the circle method. For $\alpha \in \mathbb{T}$, we define

$$T(\alpha) := \sum_{\substack{x \in \mathcal{O} \\ |x| < \widehat{B}}} \psi(\alpha x^3).$$

It then follows from Lemma 3.0.2 that we can write our counting function as

$$R_n(P) = \int_{\mathbb{T}} T(\alpha)^n \psi(-\alpha P) \mathrm{d}\alpha.$$

We then define our set of major arcs to be

$$\mathfrak{M} := \bigcup_{\substack{|r| \leq \widehat{B} \\ r \text{ monic}}} \bigcup_{\substack{|a| < |r| \\ (a,r)=1}} \{\alpha \in \mathbb{T} \colon |r\alpha - a| < \widehat{B}^{-2}\}$$

and $\mathfrak{m} := \mathbb{T} \setminus \mathfrak{M}$ constitutes our set of minor arcs. The following lemma is a consequence of [127, Theorem 30].

Lemma 5.7.1. *Suppose* $\mathrm{char}(K) \nmid 3$ *and* $n \geq 7$. *Then there exists* $\delta > 0$ *such that for all* $P \in \mathbb{J}_q^3[t]$ *we have*

$$\int_{\mathfrak{M}} T(\alpha)^n \psi(-\alpha P) \mathrm{d}\alpha = \mathfrak{S}(P) \sigma_\infty(P) \widehat{B}^{n-3} + O\left(\widehat{B}^{n-3-\delta}\right),$$

where $\mathfrak{S}(P)$ *and* $\sigma_\infty(P)$ *are the singular series and singular integral associated to* P. *Furthermore, they satisfy*

$$1 \ll \mathfrak{S}(P) \sigma_\infty(P) \ll 1.$$

Remark. In fact, Kubota states Lemma 5.7.1 only for $n \geq 10$. However, as explained by Liu–Wooley in [136, Lemma 5.2], this is a result of an oversight and Kubota's argument already works for $n \geq 7$.

We now have

$$\left| \int_{\mathfrak{m}} T(\alpha)^n \psi(-\alpha P) \mathrm{d}\alpha \right| \leq \sup_{\alpha \in \mathfrak{m}} |T(\alpha)|^{n-6} \int_{\mathbb{T}} |T(\alpha)|^6 \mathrm{d}\alpha. \tag{5.7.1}$$

If $\alpha \in \mathfrak{m}$, then Proposition 4.2.1 with $\widehat{Q} = \widehat{B}$ implies the existence of $a, r \in \mathcal{O}$ with r monic such that $|a| < |r| \leq \widehat{B}$, $(a, r) = 1$ and $|r\alpha - a| < \widehat{B}^{-1}$. As $\alpha \in \mathfrak{m}$, we must have $|\alpha - a/r| \geq \widehat{B}^{-2}|r|^{-1}$. Under these circumstances Weyl's inequality, see [35, Lemma 5.10] for $\mathrm{char}(K) > 3$ and [50, Proposition IV.4] for $\mathrm{char}(K) = 2$, guarantees the existence of $\delta > 0$ such that

$$\sup_{\alpha \in \mathfrak{m}} |T(\alpha)|^{n-6} \ll \widehat{B}^{(n-6)(1-\delta)}. \tag{5.7.2}$$

Since

$$\int_{\mathbb{T}} |T(\alpha)|^6 \mathrm{d}\alpha = \#\{\boldsymbol{x} \in \mathcal{O}^6 \colon |\boldsymbol{x}| < \widehat{B}, x_1^3 + x_2^3 + x_3^3 = x_4^3 + x_5^3 + x_6^3\},$$

Theorem 5.1.1 implies

$$\int_{\mathbb{T}} |T(\alpha)|^6 \mathrm{d}\alpha \ll \widehat{B}^{3+\varepsilon}. \tag{5.7.3}$$

Plugging (5.7.2) and (5.7.3) into (5.7.1) yields

$$\begin{aligned}\int_{\mathfrak{m}} T(\alpha)^n \psi(-\alpha P)\mathrm{d}\alpha &\ll \widehat{B}^{(n-6)(1-\delta)+3+\varepsilon}\\ &= \widehat{B}^{n-3-\delta(n-6)+\varepsilon}.\end{aligned}$$

After choosing $\varepsilon = \delta(n-6)/2$, we see that the contribution of the minor arcs is

$$\int_{\mathfrak{m}} T(\alpha)^n \psi(-\alpha P)\mathrm{d}\alpha \ll \widehat{B}^{n-3-\delta(n-6)/2}.$$

Since $n \geq 7$, combining this with Lemma 5.7.1 therefore completes the proof of Theorem 5.1.4.

5.7.2 Weak approximation for cubic diagonal hypersurfaces

We will show that weak approximation holds for the diagonal cubic hypersurface defined by $F(\boldsymbol{x}) = \sum_{i=1}^n F_i x_i^3$ if $n \geq 7$. Fix $\boldsymbol{x}_0 \in \mathbb{T}^n$ such that $F(\boldsymbol{x}_0) = 0$, $M \in \mathcal{O}$, $\boldsymbol{b} \in \mathcal{O}^n$ and $N \in \mathbb{Z}_{\geq 0}$ such that $|\boldsymbol{b}| < |M|$ and such that N is bounded in terms of M. Define the weight function $\widetilde{w}\colon K_\infty^n \to \mathbb{R}$ via

$$\widetilde{w}(\boldsymbol{x}) = \begin{cases} 1 & \text{if } |\boldsymbol{x} - \boldsymbol{x}_0| < \widehat{N}^{-1},\\ 0 & \text{otherwise.}\end{cases}$$

Further for $P \in \mathcal{O}$ we introduce the counting function

$$N(P, \widetilde{w}) := \sum_{\substack{\boldsymbol{x}\in\mathcal{O}^n\\ F(M\boldsymbol{x}+\boldsymbol{b})=0}} \widetilde{w}\left(\frac{M\boldsymbol{x}+\boldsymbol{b}}{P}\right).$$

As usual, we can write this as an integral over an exponential sum

$$N(P, \widetilde{w}) = \int_{\mathbb{T}} \widetilde{S}(\alpha)\mathrm{d}\alpha,$$

where

$$\widetilde{S}(\alpha) = \sum_{\boldsymbol{x}\in\mathcal{O}^n} \psi\left(\alpha F(M\boldsymbol{x}+\boldsymbol{b})\right)\widetilde{w}\left(\frac{M\boldsymbol{x}+\boldsymbol{b}}{P}\right).$$

Since F is diagonal we may factorise $\widetilde{S}(\alpha)$ as

$$\widetilde{S}(\alpha) = \prod_{i=1}^n \widetilde{T}_i(\alpha),$$

where

$$\widetilde{T}_i(\alpha) = \sum_{\substack{x\in\mathcal{O}\\ |Mx+b_i-x_{0,i}|<|P|\widehat{N}^{-1}}} \psi(\alpha F_i(Mx+b_i)^3).$$

Note that our counting function $N(P,\widetilde{w})$ agrees with the function $\rho_{M,\boldsymbol{b}}(n)$ and $\widetilde{S}(\alpha)$ agrees with $T(\alpha)$ in [131, Chapter 4]. In order to show weak approximation for the variety $X = V(F) \subset \mathbb{P}^{n-1}$, by the same argument as the one provided in Section 4.9 of [131], it is enough to show the following result.

Theorem 5.7.2. *Suppose* $\mathrm{char}(K) > 3$. *Then there exists some* $\delta > 0$ *such that*

$$N(P, \widetilde{w}) = |M|^{-3}\mathfrak{S}\mathfrak{J}|P|^{n-3} + O(|P|^{n-3-\delta}),$$

where $\mathfrak{S}$ *and* $\mathfrak{J}$ *are the singular series and the singular integral respectively as defined in* (5.7.6) *and* (5.7.8).

We tackle this using a traditional circle method argument.

We define the major arcs to be the set $\mathfrak{M} \subset \mathbb{T}$ given by

$$\mathfrak{M} = \bigcup_{\substack{r \in \mathcal{O} \\ |r| < |P|^{1/2} \\ r \text{ monic}}} \bigcup_{\substack{a \in \mathcal{O} \\ |a| < |r| \\ (a,q)=1}} \left\{ \alpha \in \mathbb{T} \colon |r\alpha - a| < H_F^{-1}|M|^{-3}|r||P|^{-5/2} \right\},$$

and we take the minor arcs to be the complement $\mathfrak{m} = \mathbb{T} \setminus \mathfrak{M}$.

In this context, provided $\mathrm{char}(K) > 3$, Weyl's inequality, as recorded in [131, Lemma 4.3.6] or Corollary 4.1.2, tells us that

$$|\widetilde{T}_i(\alpha)| \ll |P|^{1+\varepsilon} \left(\frac{|P| + |r| + |P|^3|r\alpha - a|}{|P|^3} + \frac{1}{|r| + |P|^3|r\alpha - a|} \right)^{1/4}$$

for $i = 1, \ldots, n$ if $a, r \in \mathcal{O}$ are such that $|a| < |r|$, r monic and $(a, r) = 1$. Using Proposition 4.2.1 and the definition of the minor arcs, a similar argument that handed us (5.7.2) gives

$$\sup_{\alpha \in \mathfrak{m}} \left| \widetilde{T}_i(\alpha) \right| \ll |P|^{7/8+\varepsilon}, \tag{5.7.4}$$

for any $\varepsilon > 0$. We are now ready to finish our treatment of the minor arcs. If $n \geq 7$ we obtain

$$\int_{\mathfrak{m}} |\widetilde{S}(\alpha)| \mathrm{d}\alpha = \int_{\mathfrak{m}} \left| \prod_{i=1}^{n} \widetilde{T}_i(\alpha) \right| \mathrm{d}\alpha \ll \sup_{\alpha \in \mathfrak{m}} \left| \widetilde{T}_7(\alpha) \cdots \widetilde{T}_n(\alpha) \right| \int_{\mathbb{T}} \left| \prod_{i=1}^{6} \widetilde{T}_i(\alpha) \right| \mathrm{d}\alpha.$$

The integral can be dealt with as follows. By Hölder's inequality we find

$$\int_{\mathfrak{m}} \left| \prod_{i=1}^{6} \widetilde{T}_i(\alpha) \right| \mathrm{d}\alpha \leq \prod_{i=1}^{6} \left(\int_{\mathbb{T}} |\widetilde{T}_i(\alpha)|^6 \mathrm{d}\alpha \right)^{1/6}.$$

Now the last quantity is equal to

$$\prod_{i=1}^{6} \# \left\{ \boldsymbol{x} \in \mathcal{O}^6 \colon x_j \equiv b_i \bmod M, |x_j/P - x_{0,i}| < \widehat{N}^{-1}, \text{ for all } j, \sum_{j=1}^{3} x_j^3 = \sum_{j=4}^{6} x_j^3 \right\}^{1/6},$$

which in turn is bounded by

$$\prod_{i=1}^{6} \#\{\boldsymbol{x} \in \mathcal{O}^6 \colon |\boldsymbol{x}| < |\boldsymbol{x}_0||P|, \, x_1^3 + x_2^3 + x_3^3 = x_4^3 + x_5^3 + x_6^3\}^{1/6},$$

if $|P|$ is sufficiently large. An application of Theorem 5.1.1 therefore yields

$$\int_{\mathbb{T}} \left| \prod_{i=1}^{6} \widetilde{T}_i(\alpha) \right| \mathrm{d}\alpha \ll |P|^{3+\varepsilon}.$$

Once combined with (5.7.4) we thus obtain

$$\int_{\mathfrak{m}} |\widetilde{S}(\alpha)| \mathrm{d}\alpha \ll |P|^{n-3-(n-7)/8+\varepsilon}$$

for any $\varepsilon > 0$, which is satisfactory if $n \geq 7$. We now turn to the major arcs. Given $a, r \in \mathcal{O}$ write

$$\widetilde{S}_r(a) := \sum_{|\boldsymbol{x}|<|r|} \psi\left(\frac{aF(M\boldsymbol{x}+\boldsymbol{b})}{r}\right).$$

For any $Y \in \mathbb{R}$ we define the truncated singular series

$$\mathfrak{S}(\widehat{Y}) := \sum_{\substack{|r|<\widehat{Y} \\ r \text{ monic}}} \sum_{\substack{|a|<|r| \\ (a,r)=1}} |r|^{-n} \widetilde{S}_r(a),$$

and the truncated singular integral to be

$$\mathfrak{I}(\widehat{Y}) = \int_{|\gamma|<H_F^{-1}\widehat{Y}} I(\gamma) \mathrm{d}\gamma,$$

where

$$I(\gamma) = \int_{\mathbb{T}^n} \psi(\gamma F(\boldsymbol{x})) \widetilde{w}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Then from (4.6.30) in [131] it follows that we have

$$\int_{\mathfrak{M}} \widetilde{S}(\alpha) \mathrm{d}\alpha = |M|^{-3} \mathfrak{S}(|P|^{1/2}) \mathfrak{I}(|P|^{1/2}) |P|^{n-3}.$$

It remains to study the convergence of the singular integral and singular series. In order to handle the singular series we will need upper bounds for $\widetilde{S}_r(a)$. First, we record the following multiplicative property, which is shown in [131, Lemma 4.7.2]. If $r_1, r_2 \in \mathcal{O}$ are coprime then

$$\widetilde{S}_{r_1 r_2}(a) = \widetilde{S}_{r_1}(a_1) \widetilde{S}_{r_2}(a_2),$$

where $a_i \in \mathcal{O}$ are such that $a_1 \equiv a\widetilde{r}_2 \mod r_1$ and $a_2 \equiv a\widetilde{r}_1 \mod r_2$, where $\widetilde{r}_1, \widetilde{r}_2$ denote the multiplicative inverses modulo r_2, r_1, respectively. Thus, from (5.4.8) in combination with the divisor estimate, it follows that we have

$$\widetilde{S}_r(a) \ll |r|^{2n/3+\varepsilon}, \tag{5.7.5}$$

where the constant may depend on M, b and ε.

Using this we see that

$$\sum_{\substack{|r|=\widehat{Y} \\ r \text{ monic}}} \sum_{\substack{|a|<|r| \\ (a,r)=1}} |r|^{-n} \left|\widetilde{S}_r(a)\right| \ll \widehat{Y}^{(2-n/3+\varepsilon)}.$$

Since $n \geq 7$ we deduce absolute convergence of the series

$$\mathfrak{S} = \sum_{r \text{ monic}} \sum_{\substack{|a|<|r| \\ (a,r)=1}} |r|^{-n} \widetilde{S}_r(a), \tag{5.7.6}$$

which is the singular series. Moreover choosing positive $\varepsilon < (n-6)/6$ we find

$$\mathfrak{S} - \mathfrak{S}(|P|^{1/2}) \ll |P|^{1-n/6+\varepsilon}, \tag{5.7.7}$$

if $n \geq 7$ upon redefining ε. We turn to the singular integral. Assuming N to be sufficiently large, in [34] it is shown in Lemma 7.5 and the paragraphs preceding it that

$$\mathfrak{I}(\widehat{Y}) = \mathfrak{I}(\widehat{N}/|\nabla F(\boldsymbol{x}_0)|) = \frac{1}{|\nabla F(\boldsymbol{x}_0)|\widehat{N}^{n-1}}$$

whenever $\widehat{Y} \geq \widehat{N}/|\nabla F(\boldsymbol{x}_0)|$. Thus clearly $\lim_{\widehat{Y}\to\infty} \mathfrak{I}(\widehat{Y})$ exists and is equal to

$$\mathfrak{I} := \lim_{\widehat{Y}\to\infty} \mathfrak{I}(\widehat{Y}) = \frac{1}{|\nabla F(\boldsymbol{x}_0)|\widehat{N}^{n-1}}. \tag{5.7.8}$$

We conclude that

$$N(P, \widetilde{w}) = |M|^{-3}\mathfrak{S}\mathfrak{I}|P|^{n-3} + O(|P|^{n-3-1/8+\varepsilon}),$$

as desired.

5.8 Special solutions and the case $n = 4$

In this section we will concern ourselves with understanding how the special solutions of $F^*(\boldsymbol{c}) = 0$ in the case $n = 4$ relate to the solutions of $F(\boldsymbol{x}) = 0$ on rational lines. The goal of this section is to prove the following lemma, from which Theorem 5.1.2 immediately follows.

Lemma 5.8.1. *For any $\varepsilon > 0$ the following holds*

$$|P|^4 \sum_{\substack{r \text{ monic} \\ |r|\leq \widehat{Q}}} |r|^{-4} \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \sum_{\boldsymbol{c}}^{\text{spec}} S_r(\boldsymbol{c}) I_r(\theta, \boldsymbol{c}) \mathrm{d}\theta = \sum_{\boldsymbol{x}}^{\text{line}} w(P^{-1}\boldsymbol{x}) + O(|P|^{3/2+\varepsilon}), \tag{5.8.1}$$

where $\sum_{\boldsymbol{c}}^{\text{spec}}$ denotes the sum over the special solutions $\boldsymbol{c} \in \mathcal{O}^4 \setminus \{\boldsymbol{0}\}$ *of $F^*(\boldsymbol{c}) = 0$ such that*

$$(F_1^{-1}c_1^3)^{1/2} \pm (F_2^{-1}c_2^3)^{1/2} = (F_3^{-1}c_3^3)^{1/2} \pm (F_4^{-1}c_4^3)^{1/2} = 0 \tag{5.8.2}$$

and $\sum_{\boldsymbol{x}}^{\text{line}}$ denotes the sum over points $\boldsymbol{x} \in \mathcal{O}^4$ satisfying

$$F_1x_1^3 + F_2x_2^3 = F_3x_3^3 + F_4x_4^3 = 0. \tag{5.8.3}$$

For notational convenience, this lemma only considers the case of lines such that $(i, j, k, l) = (1, 2, 3, 4)$ in the language of Theorem 5.1.2. By the symmetry of the situation at hand it is clear that the result follows for any permutation of indices.

5.8.1 Analysis of special solutions

We begin by noting that with an error of $O(|P|^{3/2+\varepsilon})$ we may include tuples $\boldsymbol{c} \in \mathcal{O}^4 \setminus \{\boldsymbol{0}\}$ satisfying (5.8.2) such that $c_i = 0$ for at least one i in the sum appearing in the left hand side of (5.8.1). Write $\widetilde{\sum}_{\boldsymbol{c}}^{\text{spec}}$ for the sum over such tuples $\boldsymbol{c}$. Note for such $\boldsymbol{c}$ Lemma 5.3.6 gives

$$I_r(\boldsymbol{c}) \ll |P|^{-5/2}|\boldsymbol{c}|^{-1},$$

for any $r \in \mathcal{O}$. Also note that $I_r(\theta, \boldsymbol{c}) = 0$ if $|\boldsymbol{c}| \gg |P|^{1/2}$. From (5.4.6) and Lemma 5.4.4, where we apply the second part with $m = 0$, we obtain

$$S_r(\boldsymbol{c}) \ll |r|^{\varepsilon}|r_1|^3|r_2|^{4-1/3},$$

where r_1 denotes the cube-free and r_2 the cube-full part of r. Hence

$$\sum_{\substack{r \text{ monic} \\ |r|\leq \widehat{Q}}} |r|^{-4} S_r(\boldsymbol{c}) \ll |P|^{\varepsilon} \left(\sum_{|r_1|\leq \widehat{Q}} |r_1|^{-1} \right) \left(\sum_{|r_2|\leq \widehat{Q}} |r_2|^{-1/3} \right) \ll |P|^{\varepsilon},$$

since the number of cube-full r_2 of a fixed absolute value of $\widehat{Y}$, say, is at most $P(\widehat{Y}^{1/3})$. To summarise, we found that the contribution to the left hand side of (5.8.1) is at most

$$|P|^4 \sum_{\substack{r \text{ monic} \\ |r|\leq \widehat{Q}}} |r|^{-4} \sum_{\boldsymbol{c}}^{\widetilde{\text{spec}}} S_r(\boldsymbol{c}) I_r(\boldsymbol{c}) \ll |P|^{3/2+\varepsilon} \sum_{0<|\boldsymbol{c}|\leq |P|^{1/2}}^{\widetilde{\text{spec}}} |\boldsymbol{c}|^{-1} \ll |P|^{3/2+\varepsilon},$$

where the last estimate follows since there are only $O(\widehat{C})$ vectors $\boldsymbol{c}$ of absolute value $\widehat{C}$, say, appearing in $\sum_{\boldsymbol{c}}^{\widetilde{\text{spec}}}$.

We may assume that both F_1/F_2 and F_3/F_4 are cubes in K. Otherwise the conclusion of the lemma is easily seen to be true, since there are no special solutions and $O(|P|)$ points $\boldsymbol{x}$ satisfying (5.8.3). Therefore there exist at most $O(1)$ many different possible $\rho_i \in \mathcal{O}$ with $(\rho_1, \rho_2) = (\rho_3, \rho_4) = 1$ and $\lambda, \mu \in \mathcal{O}$ such that

$$F_1 = \lambda \rho_1^3, \quad F_2 = \lambda \rho_2^3, \quad F_3 = \mu \rho_3^3, \quad F_4 = \mu \rho_4^3.$$

The different possibilites for ρ_i come from the potential existence of non-trivial third roots of unity in K. For a choice of $\rho_i \in \mathcal{O}$ if we write

$$c_1 = \rho_1 d_1, \quad c_2 = \rho_2 d_1, \quad c_3 = \rho_3 d_2, \quad c_4 = \rho_4 d_2,$$

then as we run through the possible choices of ρ_i and as $\boldsymbol{d}$ runs through $\mathcal{O}^2$, then $\boldsymbol{c}$ runs through solutions of $F^*(\boldsymbol{c}) = 0$ satisfying (5.8.2) . Given a choice of ρ_i there exist $\rho_i' \in \mathcal{O}$ such that

$$\rho_1 \rho_2' - \rho_2 \rho_1' = \rho_3 \rho_4' - \rho_4 \rho_3' = 1.$$

Then the change of variables $(x_1, x_2, x_3, x_4) \mapsto (y_1, y_2, z_1, z_2)$ given by

$$\begin{pmatrix} y_1 \\ z_1 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_2 & 0 & 0 \\ \rho_1' & \rho_2' & 0 & 0 \\ 0 & 0 & \rho_3 & \rho_4 \\ 0 & 0 & \rho_3' & \rho_4' \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

is unimodular. Moreover the inverse of this is easily seen to be

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \rho_2' & -\rho_2 & 0 & 0 \\ -\rho_1' & \rho_1 & 0 & 0 \\ 0 & 0 & \rho_4' & -\rho_4 \\ 0 & 0 & -\rho_3' & \rho_3 \end{pmatrix} \begin{pmatrix} y_1 \\ z_1 \\ y_2 \\ z_2 \end{pmatrix}.$$

We will write $\boldsymbol{x}(\boldsymbol{y}, \boldsymbol{z})$ for $\boldsymbol{x}$ arising from this linear transformation. An easy calculation reveals

$$F(\boldsymbol{x}(\boldsymbol{y}, \boldsymbol{z})) = y_1 Q_1(y_1, z_1) + y_2 Q_2(y_2, z_2) =: \widetilde{F}(\boldsymbol{y}, \boldsymbol{z}),$$

where Q_i are the quadratic forms given by

$$Q_1(y, z) = \frac{\lambda}{4} \left(y^2 + 3\{2\rho_1\rho_2 z - (\rho_1\rho_2' + \rho_1'\rho_2) y\}^2 \right),$$

and

$$Q_2(y,z) = \frac{\mu}{4}\left(y^2 + 3\{2\rho_3\rho_4 z - (\rho_3\rho_4' + \rho_3'\rho_4)y\}^2\right).$$

With this notation we then find

$$S_r(\boldsymbol{c}) = \sum_{|a|<|r|}{}' \sum_{|\boldsymbol{g}|,|\boldsymbol{h}|<|r|} \psi\left(\frac{a\widetilde{F}(\boldsymbol{g},\boldsymbol{h}) + \boldsymbol{g}\cdot\boldsymbol{d}}{r}\right),$$

and

$$I_r(\theta,\boldsymbol{c}) = \int_{K_\infty^2}\int_{K_\infty^2} w(\boldsymbol{x}(\boldsymbol{y},\boldsymbol{z}))\psi\left(\theta P^3\widetilde{F}(\boldsymbol{y},\boldsymbol{z}) + P\frac{\boldsymbol{y}\cdot\boldsymbol{d}}{r}\right)\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}.$$

We make the change of variables $\boldsymbol{y} = P^{-1}(\boldsymbol{g} + r\boldsymbol{v})$ in the integral to obtain

$$I_r(\theta,\boldsymbol{c}) = |r|^2|P|^{-2}\int_{K_\infty^2}\int_{K_\infty^2} w(\boldsymbol{x}(P^{-1}(\boldsymbol{g}+r\boldsymbol{v}),\boldsymbol{z})) \times \psi\left(\theta P^3\widetilde{F}(P^{-1}(\boldsymbol{g}+r\boldsymbol{v}),\boldsymbol{z}) + \frac{\boldsymbol{g}\cdot\boldsymbol{d}}{r}\right)\psi(\boldsymbol{v}\cdot\boldsymbol{d})\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{z}.$$

Hence we find

$$\sum_{\boldsymbol{c}}{}^{\mathrm{spec}} S_r(\boldsymbol{c})I_r(\theta,\boldsymbol{c}) = |r|^2|P|^{-2}\sum_{\rho_i}\sum_{|\boldsymbol{g}|<|r|}\int_{K_\infty^2}\sum_{\boldsymbol{d}\in\mathcal{O}^2}\int_{K_\infty^2} f_{\boldsymbol{g},\boldsymbol{z}}(\theta,\boldsymbol{v})\psi(\boldsymbol{v}\cdot\boldsymbol{d})\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{z},$$

where $\sum_{\rho_i}$ sums over the finitely many possible choices for $\rho_i \in \mathcal{O}$ as above and where

$$f_{\boldsymbol{g},\boldsymbol{z}}(\theta,\boldsymbol{v}) = \sum_{|a|<|r|}{}' \sum_{|\boldsymbol{h}|<|r|} w(\boldsymbol{x}(P^{-1}(\boldsymbol{g}+r\boldsymbol{v}),\boldsymbol{z}))\psi\left(\theta P^3\widetilde{F}(P^{-1}(\boldsymbol{g}+r\boldsymbol{v}),\boldsymbol{z}) + \frac{a\widetilde{F}(\boldsymbol{g},\boldsymbol{h})}{r}\right).$$

Poisson summation as in Proposition 3.0.5 yields

$$\sum_{\boldsymbol{d}\in\mathcal{O}^2}\int_{K_\infty^2} f_{\boldsymbol{g},\boldsymbol{z}}(\theta,\boldsymbol{v})\psi(\boldsymbol{v}\cdot\boldsymbol{d})\mathrm{d}\boldsymbol{v} = \sum_{\boldsymbol{s}\in\mathcal{O}^2} f_{\boldsymbol{g},\boldsymbol{z}}(\theta,\boldsymbol{s}).$$

We make the change of variables $\boldsymbol{j} = \boldsymbol{g} + r\boldsymbol{s}$ and the substitution $\boldsymbol{z} = P^{-1}\boldsymbol{t}$ in order to obtain

$$\sum_{\boldsymbol{c}}{}^{\mathrm{spec}} S_r(\boldsymbol{c})I_r(\boldsymbol{c}) = |r|^2|P|^{-4}\sum_{\rho_i}\sum_{\boldsymbol{j}\in\mathcal{O}^2} T_r(\boldsymbol{j})J_r(\boldsymbol{j},\theta),$$

where

$$T_r(\boldsymbol{j}) = \sum_{|a|<|r|}{}' \sum_{|\boldsymbol{h}|<|r|}\psi\left(\frac{a\widetilde{F}(\boldsymbol{j},\boldsymbol{h})}{r}\right),$$

and

$$J_r(\boldsymbol{j},\theta) = \int_{K_\infty^2} w(P^{-1}\boldsymbol{x}(\boldsymbol{j},\boldsymbol{t}))\psi(\theta\widetilde{F}(\boldsymbol{j},\boldsymbol{t}))\mathrm{d}\boldsymbol{t}.$$

Further we will write

$$J_r(\boldsymbol{j}) := \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} J_r(\boldsymbol{j},\theta)\mathrm{d}\theta.$$

We can summarise our findings until now as follows.

Lemma 5.8.2. *We have*

$$|P|^4\sum_{\substack{r \text{ monic}\\ |r|\le\widehat{Q}}}|r|^{-4}\sum_{\boldsymbol{c}}{}^{\mathrm{spec}} S_r(\boldsymbol{c})I_r(\boldsymbol{c}) = \sum_{\rho_i}\sum_{\substack{r \text{ monic}\\ |r|\le\widehat{Q}}}|r|^{-2}\sum_{\boldsymbol{j}\in\mathcal{O}^2} T_r(\boldsymbol{j})J_r(\boldsymbol{j}) + O(|P|^{3/2+\varepsilon}). \quad (5.8.4)$$

We now follow a strategy that is very similar to the usual delta method. The main term will come from $\boldsymbol{j} = \boldsymbol{0}$ and it then remains to estimate $T_r(\boldsymbol{j})$ and $J_r(\boldsymbol{j},\theta)$ for $\boldsymbol{j} \neq \boldsymbol{0}$.

5.8.2 The main term

Lemma 5.8.3. *For all $P \in \mathcal{O} \setminus \{0\}$ we have*

$$\sum_{\rho_i} \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} |r|^{-2} T_r(\mathbf{0}) J_r(\mathbf{0}) = \sum_{\boldsymbol{x}}^{\text{line}} w(P^{-1}\boldsymbol{x}) + O(1).$$

Proof. Since $\widetilde{F}(\mathbf{0}, \boldsymbol{z}) = 0$ for all $\boldsymbol{z} \in K_\infty^2$ we have

$$T_r(\mathbf{0}) = \sum_{|a|<|r|}{}' |r|^2,$$

and

$$J_r(\mathbf{0}, \theta) = \int_{K_\infty^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{t})) \mathrm{d}\boldsymbol{t}.$$

Therefore, the term arising from $\boldsymbol{j} = \mathbf{0}$ on the right hand side of (5.8.4) is equal to

$$\sum_{\rho_i} \int_{K_\infty^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{t})) \mathrm{d}\boldsymbol{t} \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \sum_{|a|<|r|}{}' \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \mathrm{d}\theta.$$

But from Proposition 4.2.1 we deduce that

$$\sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \sum_{|a|<|r|}{}' \int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \mathrm{d}\theta = \operatorname{vol}(\mathbb{T}) = 1.$$

Further, it is easily seen that

$$\sum_{\boldsymbol{x}}^{\text{line}} w(P^{-1}\boldsymbol{x}) = \sum_{\rho_i} \sum_{\boldsymbol{z} \in \mathcal{O}^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{z})).$$

But since $K_\infty^2 = \bigsqcup_{\boldsymbol{z} \in \mathcal{O}^2} (\boldsymbol{z} + \mathbb{T})$ we have

$$\int_{K_\infty^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{t})) \mathrm{d}\boldsymbol{t} = \sum_{\boldsymbol{z} \in \mathcal{O}^2} \int_{\mathbb{T}^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{z} + \boldsymbol{\alpha})) \mathrm{d}\boldsymbol{\alpha}.$$

If $\boldsymbol{z} \in \mathcal{O} \setminus \{\mathbf{0}\}$ then $|\boldsymbol{x}(\mathbf{0}, \boldsymbol{z} + \boldsymbol{\alpha})| = |\boldsymbol{x}(\mathbf{0}, \boldsymbol{z})|$ for all $\boldsymbol{\alpha} \in \mathbb{T}^2$ and so

$$\int_{\mathbb{T}^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{z} + \boldsymbol{\alpha})) \mathrm{d}\boldsymbol{\alpha} = w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{z}))$$

for such $\boldsymbol{z}$. We also clearly have $\int_{\mathbb{T}^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{\alpha})) \mathrm{d}\boldsymbol{\alpha} \ll 1$ and so

$$\int_{K_\infty^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{t})) \mathrm{d}\boldsymbol{t} = \sum_{\boldsymbol{z} \in \mathcal{O}^2} w(P^{-1}\boldsymbol{x}(\mathbf{0}, \boldsymbol{z})) + O(1),$$

whence the Lemma follows. □

5.8.3 Estimating the error term

In this section we make a choice of $\rho_1, \ldots, \rho_4$ and bound the contribution made from terms such that $\boldsymbol{j} \neq \boldsymbol{0}$. Once we showed the desired bound for a particular choice, Lemma 5.8.1 will follow since there are only $O(1)$ different possibilities for ρ_i.

We begin by bounding $J_r(\boldsymbol{j})$ in the case where $\boldsymbol{j} \neq \boldsymbol{0}$. Note first that $w(P^{-1}(\boldsymbol{x}(\boldsymbol{j}, \boldsymbol{t}))) = 0$ if $\boldsymbol{j} \gg |P|$ and so $J_r(\boldsymbol{j}) = 0$ if $\boldsymbol{j} \gg |P|$. Further this allows us to exchange the integral over θ with the sum over $\boldsymbol{j}$ in (5.8.4). Note that Lemma 3.0.2 implies

$$\int_{|\theta|<|r|^{-1}\widehat{Q}^{-1}} \psi(\theta \widetilde{F}(\boldsymbol{j}, \boldsymbol{t}))\mathrm{d}\theta = \begin{cases} |r|^{-1}\widehat{Q}^{-1}, & \text{if } |\widetilde{F}(\boldsymbol{j}, \boldsymbol{t})| < |r|\widehat{Q} \\ 0, & \text{otherwise.} \end{cases}$$

Thus we find

$$J_r(\boldsymbol{j}) \ll \mu(\boldsymbol{j}, r)|r|^{-1}\widehat{Q}^{-1},$$

where

$$\mu(\boldsymbol{j}, r) = \mathrm{vol}\left(\left\{\boldsymbol{t} \in K_\infty^2 \colon |\boldsymbol{t}| \ll |P|,\ |\widetilde{F}(\boldsymbol{j}, \boldsymbol{t})| < |r|\widehat{Q}\right\}\right).$$

To estimate this measure we simplify the expressions involved by making the substitution

$$u_1 = 2\rho_1\rho_2 t_1 - (\rho_1\rho_2' + \rho_1'\rho_2)j_1, \quad u_2 = 2\rho_3\rho_4 t_2 - (\rho_3\rho_4' + \rho_3'\rho_4)j_1.$$

After this linear change of variables $\widetilde{F}$ takes the form

$$\widetilde{G}(\boldsymbol{j}, \boldsymbol{u}) = \lambda j_1(3u_1^2 + j_1^2) + \mu j_2(3u_2^2 + j_2^2).$$

Since the change of variables is linear of constant, non-vanishing Jacobian it is sufficient to consider

$$\mu_{\widetilde{G}}(\boldsymbol{j}, r) \coloneqq \mathrm{vol}\left(\left\{\boldsymbol{u} \in K_\infty^2 \colon |\boldsymbol{u}| \ll |P|,\ |\widetilde{G}(\boldsymbol{j}, \boldsymbol{u})| < |r|\widehat{Q}\right\}\right).$$

If $j_2 = 0$ then using Lemma 5.3.4 it is easily seen that

$$\mu_{\widetilde{G}}(\boldsymbol{j}, r) \ll |P|\left(\frac{|r|\widehat{Q}}{|j_1|}\right)^{1/2},$$

and similarly if $j_1 = 0$. So assume $j_1 j_2 \neq 0$. In this case, note that we have

$$\mu_{\widetilde{G}}(\boldsymbol{u}, r) \ll \sum_{k,m=-\infty}^{\log_q |P|} \sum_{\substack{U_1 = q^k \\ U_2 = q^m}} \mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2),$$

where

$$\mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2) = \mathrm{vol}\left(\left\{\boldsymbol{u} \in K_\infty^2 \colon |u_1| = U_1,\ |u_2| = U_2,\ \left|\widetilde{G}(\boldsymbol{j}, \boldsymbol{u})\right| < |r|\widehat{Q}\right\}\right).$$

In the case where U_1 or $U_2 < |P|^{-1}$ we can use the trivial bound $O(U_1U_2)$ for $\mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2)$ to deduce that the total contribution arising from such U_1, U_2 is bounded by $O(1)$. For the remaining contribution note if $\boldsymbol{u}$ satisfies $\widetilde{G}(\boldsymbol{j}, \boldsymbol{u}) = 0$ then $u_1^2 = A + O(|r|\widehat{Q}/|j_1|)$ for some function $A(j_1, j_2, u_2)$ and thus u_1 lies in a subset of measure $O(|r|\widehat{Q}/(U_1|j_1|))$. Therefore $\mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2) \ll U_2|r|\widehat{Q}/(U_1|j_1|)$. Similarly, $\mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2) \ll U_1|r|\widehat{Q}/(U_2|j_2|)$. Putting this together yields

$$\mu_{\widetilde{G}}(\boldsymbol{j}, r, U_1, U_2) \ll |r|\widehat{Q}|j_1 j_2|^{-1/2}.$$

Since there are $|P|^\varepsilon$ pairs U_1, U_2 such that $|P|^{-1} \leq U_1, U_2 \leq |P|$ we deduce

$$\mu(\boldsymbol{j}, r) \ll 1 + |P|^\varepsilon |r|\widehat{Q}|j_1 j_2|^{-1/2}.$$

We summarise our observations in the following lemma.

Lemma 5.8.4. *Let $\boldsymbol{j} \in \mathcal{O}^2 \setminus \{\boldsymbol{0}\}$ be such that $|\boldsymbol{j}| \ll |P|$. If $j_1 j_2 \neq 0$, then we have*

$$J_r(\boldsymbol{j}) \ll |P|^{\varepsilon} |j_1 j_2|^{-1/2}. \tag{5.8.5}$$

If $j_2 = 0$, then we have

$$J_r(\boldsymbol{j}) \ll \frac{|P|^{1/4}}{(|j_1||r|)^{1/2}}. \tag{5.8.6}$$

Next, we turn to estimating the exponential sums $T_r(\boldsymbol{j})$. Via the Chinese remainder theorem we have for all $r_1, r_2 \in \mathcal{O}$ such that $(r_1, r_2) = 1$ that

$$T_{r_1 r_2}(\boldsymbol{j}) = T_{r_1}(\boldsymbol{j}) T_{r_2}(\boldsymbol{j}). \tag{5.8.7}$$

Thus we may put our focus on $T_r(\boldsymbol{j})$ where $r = \varpi^k$ for irreducible $\varpi \in \mathcal{O}$. Note that

$$\left| \sum_{|\boldsymbol{h}|<|r|} \psi\left(\frac{a\widetilde{F}(\boldsymbol{j},\boldsymbol{h})}{r}\right) \right| \leq \left| \sum_{|h_1|<|r|} \psi\left(\frac{aj_1 Q_1(j_1,h_1)}{r}\right) \right| \left| \sum_{|h_1|<|r|} \psi\left(\frac{aj_2 Q_2(j_2,h_2)}{r}\right) \right|.$$

A simple Weyl differencing type of argument further yields

$$\begin{aligned}
\left| \sum_{|h_1|<|r|} \psi\left(\frac{aj_1 Q_1(j_1,h_1)}{r}\right) \right|^2 &= \sum_{|h|,|h_1|<|r|} \psi\left(\frac{aj_1(Q_1(j_1,h+h_1) - Q_1(j_1,h_1))}{r}\right) \\
&\ll \sum_{|h|<|r|} \left| \sum_{|h_1|<|r|} \psi\left(\frac{6a\lambda j_1 \rho_1^2 \rho_2^2 j_1 h_1 h}{r}\right) \right| \\
&= |r| \, \#\{h \in \mathcal{O} \colon |h| < |r|, r \mid 6a\lambda j_1 \rho_1^2 \rho_2^2 j_1 h\} \\
&\ll |r| \, |(r, 6a\lambda j_1 \rho_1^2 \rho_2^2 j_1 h)| \\
&\ll |r| \, |(r, j_1)|.
\end{aligned}$$

We can find a similar estimate for the sum over h_2, which gives

$$T_r(\boldsymbol{j}) \ll |r|^2 |(r, j_1)|^{1/2} |(r, j_2)|^{1/2}.$$

This will be sufficient for our purposes if r is cube-full. However, for $r = \varpi$ or $r = \varpi^2$ we can do better. We begin by considering the case when $r = \varpi$ and we will further assume $\varpi \nmid (j_1, j_2)$. Note first that

$$\sideset{}{'}\sum_{|a|<|\varpi|} \psi\left(\frac{a\widetilde{F}(\boldsymbol{j},\boldsymbol{h})}{\varpi}\right) = \sum_{\substack{|a|<|\varpi| \\ a \neq 0}} \psi\left(\frac{a\widetilde{F}(\boldsymbol{j},\boldsymbol{h})}{\varpi}\right) = \begin{cases} |\varpi| - 1, & \text{if } \varpi \mid \widetilde{F}(\boldsymbol{j},\boldsymbol{h}), \\ -1, & \text{otherwise.} \end{cases}$$

Therefore we get

$$\begin{aligned}
T_\varpi(\boldsymbol{j}) &= (|\varpi| - 1)\# \left\{ |\boldsymbol{h}| < |\varpi| \colon \varpi \mid \widetilde{F}(\boldsymbol{j},\boldsymbol{h}) \right\} - \# \left\{ |\boldsymbol{h}| < |\varpi| \colon \varpi \nmid \widetilde{F}(\boldsymbol{j},\boldsymbol{h}) \right\} \\
&= |\varpi| \# \left\{ |\boldsymbol{h}| < |\varpi| \colon \varpi \mid \widetilde{F}(\boldsymbol{j},\boldsymbol{h}) \right\} - |\varpi|^2.
\end{aligned}$$

The equation $\widetilde{F}(\boldsymbol{j},\boldsymbol{h}) \equiv 0 \mod \varpi$ may be regarded as $Q(h_1, h_2, 1)$ for a ternary quadratic form $Q(x, y, z)$. The quadratic form Q is non-singular in $\mathcal{O}/\varpi$ if $\varpi \nmid j_1 j_2 F_0(\boldsymbol{j})$, where

$F_0(\boldsymbol{j}) = \lambda j_1^3 + \mu j_2^3$. Since ϖ is irreducible we have $\mathcal{O}/\varpi \cong \mathbb{F}_{|\varpi|}$ and so if $\varpi \nmid j_1 j_2 F_0(\boldsymbol{j})$ then Theorem 6.26 in [135] gives

$$\#\left\{|\boldsymbol{h}| < |\varpi| \colon \varpi \mid \widetilde{F}(\boldsymbol{j}, \boldsymbol{h})\right\} = |\varpi| + O(1).$$

We deduce $T_\varpi(\boldsymbol{j}) \ll |\varpi|$ in this case. Since $\varpi \nmid (j_1, j_2)$ the form Q does not vanish identically in $\mathcal{O}/\varpi$ and so we have

$$\#\left\{|\boldsymbol{h}| < |\varpi| \colon \varpi \mid \widetilde{F}(\boldsymbol{j}, \boldsymbol{h})\right\} \ll |\varpi|,$$

whence $T_\varpi(\boldsymbol{j}) \ll |\varpi|^2$ if $\varpi \mid j_1 j_2 F_0(\boldsymbol{j})$.

We now turn to analysing $T_{\varpi^2}(\boldsymbol{j})$. We assume $\varpi \nmid \lambda\mu \prod_{i=1}^5 \rho_i$. This condition affects only finitely many primes ϖ and so the estimates that we obtain under this condition hold in general by adjusting the resulting constant. Put

$$k_1 = 2\rho_1\rho_2 h_1 - (\rho_1\rho_2' + \rho_1'\rho_2) j_1, \quad \text{and} \quad k_2 = 2\rho_3\rho_4 h_2 - (\rho_3\rho_4' + \rho_3'\rho_4) j_2,$$

so that after this change of variables we have

$$\widetilde{F}(\boldsymbol{j}, \boldsymbol{k}(\boldsymbol{h})) = \frac{1}{4} F_0(\boldsymbol{j}) + \frac{3}{4}(\lambda j_1 k_1^2 + \mu j_2 k_2^2).$$

By our assumption on ϖ, as $\boldsymbol{h}$ ranges through values $|\boldsymbol{h}| < |\varpi^2|$ we also have that $\boldsymbol{k}$ ranges through $|\boldsymbol{k}| < |\varpi^2|$ under this change of variables. Hence we obtain

$$T_{\varpi^2}(\boldsymbol{j}) = \sideset{}{'}\sum_{|a| < |\varpi|^2} \psi\left(\frac{aF_0(\boldsymbol{j})}{4\varpi^2}\right) \sum_{|\boldsymbol{k}| < |\varpi|^2} \psi\left(\frac{3a(\lambda j_1 k_1^2 + \mu j_2 k_2^2)}{4\varpi^2}\right).$$

We can write $\boldsymbol{k} = \boldsymbol{u} + \varpi\boldsymbol{v}$ where $|\boldsymbol{u}|, |\boldsymbol{v}| < |\varpi|$. Then

$$\begin{aligned}
\sum_{|k_i| < |\varpi|^2} \psi\left(\frac{3a\lambda j_i k_i^2}{4\varpi^2}\right) &= \sum_{|u_i| < |\varpi|} \psi\left(\frac{3a\lambda j_i u_i^2}{4\varpi^2}\right) \sum_{|v_i| < |\varpi|} \psi\left(\frac{3a\lambda j_i u_i v_i}{4\varpi^2}\right) \\
&= |\varpi| \sum_{\substack{|u_i| < |\varpi| \\ \varpi \mid j_i u_i}} \psi\left(\frac{3a\lambda j_i u_i^2}{4\varpi^2}\right),
\end{aligned}$$

for $i = 1, 2$ since $\varpi \nmid a\lambda$. If $\varpi \nmid j_1 j_2$ the above expression is just $|\varpi|$ and so we get in this case

$$T_{\varpi^2}(\boldsymbol{j}) = |\varpi|^2 \sideset{}{'}\sum_{|a| < |\varpi|^2} \psi\left(\frac{aF_0(\boldsymbol{j})}{4\varpi^2}\right) = \begin{cases} 0, & \text{if } \varpi \nmid F_0(\boldsymbol{j}), \\ -|\varpi|^3 & \text{if } \varpi \parallel F_0(\boldsymbol{j}), \\ |\varpi|^4 - |\varpi|^3 & \text{if } \varpi^2 \mid F_0(\boldsymbol{j}), \end{cases}$$

and so in particular

$$T_{\varpi^2}(\boldsymbol{j}) \ll |\varpi|^2 |(\varpi^2, F_0(\boldsymbol{j}))|.$$

If, on the other hand, $\varpi \mid j_1$ we claim that $T_{\varpi^2}(\boldsymbol{j}) = 0$. Due to the standing assumption $\varpi \nmid (j_1, j_2)$ it follows that $\varpi \nmid j_2$ and thus the above gives

$$T_{\varpi^2}(\boldsymbol{j}) = |\varpi|^2 \sum_{|u_1| < |\varpi|} \sideset{}{'}\sum_{|a| < |\varpi|^2} \psi\left(\frac{a(F_0(\boldsymbol{j}) + 3\lambda j_1 u_1^2)}{4\varpi^2}\right).$$

This vanishes unless $\varpi \mid F_0(\boldsymbol{j}) + 3\lambda j_1 u_1^2$. But since $\varpi \mid j_1$ this would imply $\varpi \mid \mu j_2^3$ and hence $\varpi \mid j_2$. As we excluded this case by assumption the claim follows. We summarise our analysis of $T_r(\boldsymbol{j})$ in a lemma.

Lemma 5.8.5. *Let $\boldsymbol{j} \in \mathcal{O}^2 \setminus \{\boldsymbol{0}\}$. Then we have*

$$T_r(\boldsymbol{j}) \ll |r|^2 |(r, j_1)|^{1/2} |(r, j_2)|^{1/2}$$

for any $r \in \mathcal{O} \setminus \{0\}$. Further, if $r = \varpi$ or $r = \varpi^2$ for some irreducible $\varpi \in \mathcal{O}$ and if $\varpi \nmid (j_1, j_2)$ then we get

$$T_r(\boldsymbol{j}) \ll |r||(r, j_1 j_2 F_0(\boldsymbol{j}))|.$$

We are now finally in a position to give a sufficiently good upper bound for the right hand side of (5.8.4) and thus complete the proof of Theorem 5.1.2. For this we fix a choice of ρ_i and estimate the sum

$$\mathcal{S} := \sum_{\substack{r \text{ monic} \\ |r| \le \widehat{Q}}} |r|^{-2} \sum_{\substack{\boldsymbol{j} \in \mathcal{O}^2 \\ |\boldsymbol{j}| \ll |P|}} T_r(\boldsymbol{j}) J_r(\boldsymbol{j}).$$

Since there are $O(1)$ possibilities for the ρ_i's, this will be enough to show $\mathcal{S} \ll |P|^{3/2+\varepsilon}$.

We begin with the case when $j_1 j_2 F_0(\boldsymbol{j}) \neq 0$. In this situation Lemma 5.8.4 yields

$$\mathcal{S} \ll |P|^\varepsilon \sum_{\boldsymbol{j}} |j_1 j_2|^{-1/2} \sum_{\substack{r \text{ monic} \\ |r| \le \widehat{Q}}} |r|^{-2} |T_r(\boldsymbol{j})|. \tag{5.8.8}$$

Next we write $r = r_1 r_2$ where r_1, r_2 monic are coprime, and where r_1 is cube-free and $\varpi \mid r_1$ implies $\varpi \nmid (j_1, j_2)$. We can then factor $T_r(\boldsymbol{j})$ by (5.8.7) to obtain

$$\begin{aligned} \mathcal{S} &\ll |P|^\varepsilon \sum_{\boldsymbol{j}} |j_1 j_2|^{-1/2} \sum_{r_2} |r_2|^{-2} |T_{r_2}(\boldsymbol{j})| \sum_{r_1} |r_1|^{-2} |T_{r_1}(\boldsymbol{j})| \\ &\ll |P|^\varepsilon \sum_{\boldsymbol{j}} |j_1 j_2|^{-1/2} \sum_{r_2} |r_2|^{-2} |T_{r_2}(\boldsymbol{j})| \sum_{r_1} \frac{|(r_1, j_1 j_2 F_0(\boldsymbol{j}))|}{|r_1|}, \end{aligned}$$

where we used Lemma 5.8.5 to estimate $T_{r_1}(\boldsymbol{j})$. For the inner sum we have

$$\sum_{r_1} \frac{|(r_1, j_1 j_2 F_0(\boldsymbol{j}))|}{|r_1|} \ll |P|^\varepsilon |j_1 j_2 F_0(\boldsymbol{j})|^\varepsilon \ll |P|^{2\varepsilon},$$

since we assume $j_1 j_2 F_0(\boldsymbol{j}) \neq 0$ and in general it holds $\widehat{Y}^{-1} \sum_{|r| = \widehat{Y}} |(G, r)| \ll (|G| \widehat{Y})^\varepsilon$ for any $Y \in \mathbb{Z}_{\ge 0}$ and $G \in \mathcal{O}$.
Note that if $\varpi \parallel r_2$ or $\varpi^2 \parallel r_2$, then $\varpi \mid (j_1, j_2)$. In particular, if we put $\eta(r_2) = \prod \varpi$, where the product is over all $\varpi \mid r_2$ such that $\varpi \parallel r_2$ or $\varpi^2 \parallel r_2$, then we have $\boldsymbol{j} = \eta(r_2) \boldsymbol{k}$ for some $|\boldsymbol{k}| \ll |P| / |\eta(r_2)|$. It follows that

$$\begin{aligned} \mathcal{S} &\ll |P|^\varepsilon \sum_{\substack{r \text{ monic} \\ |r| \le \widehat{Q}}} |\eta(r)|^{-1} \sum_{\substack{|\boldsymbol{k}| \ll |P|/|\eta(r)| \\ k_1 k_2 \neq 0}} \frac{|(r, \eta(r) k_1)|^{1/2} |(r, \eta(r) k_2)|^{1/2}}{|k_1 k_2|^{1/2}} \\ &\ll |P|^\varepsilon \sum_{\substack{r \text{ monic} \\ |r| \le \widehat{Q}}} \sum_{\substack{|\boldsymbol{k}| \ll |P|/|\eta(r)| \\ k_1 k_2 \neq 0}} \frac{|(r, k_1)|^{1/2} |(r, k_2)|^{1/2}}{|k_1 k_2|^{1/2}}. \end{aligned}$$

The sum over $\boldsymbol{k}$ above factors into $(\sum_k |(r, k)|^{1/2} |k|^{-1/2})^2$, which we can estimate as

$$\begin{aligned} \sum_{\substack{|k| \ll |P|/|\eta(r)| \\ k \neq 0}} \frac{|(r, k)|^{1/2}}{|k|^{1/2}} &\ll \sum_{d \mid r} |d|^{1/2} \sum_{\substack{|k'| \ll |P|/|\eta(r) d| \\ (r, k') = 1}} |k' d|^{-1/2} \\ &\ll \sum_{d \mid r} |P|^{1/2} |\eta(r)|^{-1/2}. \end{aligned}$$

Since $\sum_{d|r} 1 \ll |r|^{\varepsilon} \ll |P|^{\varepsilon}$, we thus arrive at

$$\mathcal{S} \ll |P|^{1+\varepsilon} \sum_{|r| \leq \widehat{Q}} |\eta(r)|^{-1}.$$

Next we write $r = st_1^2 t_3$, where s, t_1, t_3 are pairwise coprime and monic, t_3 is cube-full and s is square-free. With this notation we clearly have $\eta(r) = st_1$ and there are at most $O(\widehat{Q}^{1/3}) = O(|P|^{1/2})$ available t_3, so that

$$\begin{aligned} \mathcal{S} &\ll |P|^{3/2+\varepsilon} \sum_{|s| \leq \widehat{Q}} |s|^{-1} \sum_{|t_1| \leq (\widehat{Q}/|s|)^{1/2}} |t_1|^{-1} \\ &\ll |P|^{3/2+\varepsilon} \sum_{|s| \leq \widehat{Q}} |s|^{-1} (\widehat{Q}/|s|)^{\varepsilon/2} \\ &\ll |P|^{3/2+\varepsilon} \widehat{Q}^{3\varepsilon/2}. \end{aligned}$$

With a new choice of ε this estimate suffices for our purpose.

Next we consider the case when $j_1 j_2 F_0(\boldsymbol{j}) = 0$. If $j_1 j_2 \neq 0$ but $F_0(\boldsymbol{j}) = 0$, then there exist some $j, \nu_i \in \mathcal{O}$ such that $j_i = \nu_i j$. The number of possible ν_i can be estimated by $O(1)$. In this case Lemma 5.8.4 and Lemma 5.8.5 yield

$$J_r(\boldsymbol{j}) \ll |P|^{\varepsilon} |j|^{-1}, \quad \text{and} \quad T_r(\boldsymbol{j}) \ll |r|^2 |(r, j)|.$$

The total contribution to $\mathcal{S}$ of such $\boldsymbol{j}$ is therefore bounded by

$$|P|^{\varepsilon} \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \sum_{\substack{j \ll P \\ j \neq 0}} |j|^{-1} |(r, j)| \ll |P|^{3/2+\varepsilon},$$

which is sufficient.

Finally we need to consider the case when one of $j_i = 0$. We may assume $j_2 = 0$ since the other case is analogous. Write $j_1 = j$, then the second part of Lemma 5.8.4 gives

$$J_r(\boldsymbol{j}) \ll \frac{|P|^{1/4}}{(|j||r|)^{1/2}}.$$

Combining the estimates in Lemma 5.8.5 also gives

$$T_r(\boldsymbol{j}) \ll |r|^{5/2+\varepsilon} |(j, r)| m(r)^{-1/2},$$

where $m(r) = \prod_{\varpi \| r} \varpi$. The contribution to $\mathcal{S}$ of $\boldsymbol{j}$ under consideration is therefore bounded by

$$|P|^{1/4} \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \sum_{\substack{j \ll P \\ j \neq 0}} |(j, r)| |j|^{-1/2} m(r)^{-1/2}.$$

Since $\sum_{0 < j \ll P} |(j, r)| |j|^{-1/2} \ll q^{\varepsilon} |P|^{1/2+\varepsilon}$ we get an overall bound

$$|P|^{3/4+\varepsilon} \sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} m(r)^{-1/2}.$$

Write $r = r_1 r_2$ where r_1 is square-free and r_2 is square-full. Note that then $m(r) = r_1$ and there are at most $O\left(\left(\widehat{Q}/|r_1|\right)^{1/2}\right)$ available r_2. Hence

$$\sum_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} m(r)^{-1/2} \ll \widehat{Q}^{1/2} \sum_{\substack{r_1 \text{ monic} \\ |r_1| \leq \widehat{Q}}} |r_1|^{-1} \ll |P|^{3/4+\varepsilon},$$

and so the desired bound of $O(|P|^{3/2+\varepsilon})$ contributed from $\boldsymbol{j}$'s such that either $j_1 = 0$ or $j_2 = 0$ follows. Altogether, we have shown

$$\mathcal{S} \ll |P|^{3/2+\varepsilon},$$

which completes the proof of Lemma 5.8.1.

CHAPTER 6

Rational points on complete intersections of cubic and quadratic hypersurfaces over $\mathbb{F}_q(t)$

This chapter is based on [85].

6.1 Introduction

If $X \subset \mathbb{P}^{n-1}$ is a variety over a global field K, then one key aspect entails understanding the counting function

$$N_X(B) := \#\{\boldsymbol{x} \in X(K) \colon H(\boldsymbol{x}) < B\}$$

for a suitable height function $H \colon X(K) \to \mathbb{R}_{\geq 0}$. In this chapter we shall focus on the case when $X = V(F_1, F_2) \subset \mathbb{P}^{n-1}$ is a non-singular complete intersection of a cubic and a quadric hypersurface over $K = \mathbb{F}_q(t)$. To state our main result, recall that $|\boldsymbol{x}| = \max q^{\deg x_i}$ for $\boldsymbol{x} \in \mathbb{F}_q[t]^n$. We take a smooth weight function $w \colon \mathbb{F}_q((t^{-1}))^n \to \mathbb{R}_{\geq 0}$ that is supported around a suitable point $\boldsymbol{x}_0 \in \mathbb{F}_q((t^{-1}))^n$ with $F_1(\boldsymbol{x}_0) = F_2(\boldsymbol{x}_0) = 0$. For non-zero $P \in \mathbb{F}_q[t]$, we consider the counting function

$$N(P) := \sum_{\substack{\boldsymbol{x} \in \mathbb{F}_q[t]^n \\ F_1(\boldsymbol{x}) = F_2(\boldsymbol{x}) = 0}} w\left(\frac{\boldsymbol{x}}{P}\right) \quad \text{as } |P| \to \infty.$$

Theorem 6.1.1. *Let $X \subset \mathbb{P}^{n-1}$ be a non-singular complete intersection of a cubic and a quadric hypersurface over $\mathbb{F}_q(t)$. If $n \geq 26$ and $\mathrm{char}(\mathbb{F}_q) > 3$, then there exists $\delta > 0$ such that*

$$N(P) = c|P|^{n-5} + O\left(|P|^{n-5-\delta}\right),$$

for some $c > 0$.

To put our result into context, Browning–Dietmann–Heath-Brown [37] proved the analogue of Theorem 6.1.1 over $K = \mathbb{Q}$ for $n \geq 29$. Another result in this direction when both F_1 and F_2 are diagonal is due to Wooley [209, 210]. Again working over $\mathbb{Q}$, he restricted the range of possible integer solutions to those having only small prime factors. Appealing to the theory of smooth Weyl sums this allowed him to provide an asymptotic formula for the number of such restricted solutions for $n \geq 13$ whenever F_1 and F_2 have at least 7 and 5 non-zero coefficients respectively.

A smooth complete intersection of a cubic and a quadratic hypersurface inside $\mathbb{P}^{n-1}$ is Fano as soon as $n \geq 6$. In particular, Theorem 6.1.1 falls under the realm of Conjecture 2.3.3 that was put forward by Manin and his collaborators. The order of growth established in Theorem 6.1.1 agrees with this prediction. For an overview of results for more general complete intersections, we refer the reader to Section 4.5.

As explained in Section 4.5, the Hasse principle for complete intersections defined by forms of degree $d_1, \dots, d_R$ inside $\mathbb{P}^{n-1}$ holds trivially as soon as $n > d_1^2 + \cdots + d_R^2$ by the theory of Lang–Tsen fields. However, establishing weak approximation remains a substantial challenge. Currently there are no results available regarding weak approximation for complete intersections of cubic and quadric hypersurfaces over $\mathbb{F}_q(t)$. Our second main result remedies this deficiency.

Theorem 6.1.2. *Let $X \subset \mathbb{P}^{n-1}$ be a non-singular complete intersection of a cubic and a quadric hypersurface over $\mathbb{F}_q(t)$. If $n \geq 26$ and $\mathrm{char}(\mathbb{F}_q) > 3$, then X satisfies weak approximation.*

The restriction on the characteristic in both Theorems 6.1.1 and 6.1.2 arises naturally in applications of the circle method. Typically, it comes from applications of Weyl differencing, which renders any estimates trivial when the characteristic is smaller than the degrees of the equations. In our situation we have to study both quadratic and cubic exponential sums that we can only bound satisfactorily when $\mathrm{char}(K) > 3$.

Browning and Vishe [44] have found a way to use the circle method over $\mathbb{F}_q(t)$ to obtain crude geometric information about the space of rational curves of fixed degree inside a hypersurface in sufficiently many variables compared to its degree. If one is willing to make all the estimates uniform in q, then our work is likely to give access to the analogous properties when X is the intersection of a cubic and a quadric hypersurface.

Using a geometric approach, Tian [195] has verified the Hasse principle for non-singular cubic hypersurfaces $X \subset \mathbb{P}^{n-1}$ when $\mathrm{char}(K) > 5$ and $n \geq 6$ and weak approximation for non-singular intersections of quadratic forms when $\mathrm{char}(K) > 2$ and $n \geq 6$. It would be interesting to see whether his methods carry over to say something useful about intersections of cubic and quadric hypersurfaces.

When the degree of a form F is small, the delta method developed by Duke, Friedlander and Iwaniec [69] and further refined by Heath-Brown [100] is capable of dealing with significantly fewer variables than the classical circle method. In particular, it has been successfully applied to quadratic forms [100] and cubic forms [116]. Over $\mathbb{F}_q(t)$ an identity analogous to the delta method turns out to be much simpler thanks to the non-archimedean nature and was successfully incorporated by Browning and Vishe [34]. However, until recently it was unclear how to construct an analogue of the delta method for systems of equations. Vishe [202] made substantial progress by developing a 2-dimensional analogue of the delta method over $K = \mathbb{F}_q(t)$ that enabled him to produce an asymptotic formula for the number of rational points of bounded height on non-singular intersections of two quadratic forms in $n \geq 9$ variables when $\mathrm{char}(K) > 2$. His innovation serves as the main input for our work and we shall now proceed to outline the main steps of our proof.

Outline of proof

In [202] the main new input is the development of a two-dimensional version of a Farey dissection over $\mathbb{F}_q(t)$. While Vishe's version only allows one to put squares around the approximating rationals, our application requires lopsided boxes in order to take into account the different degrees of the forms F_1 and F_2. In Section 6.2 we shall modify his development to accommodate our needs. We expect that the argument would carry over inductively to higher dimensions. In the case of hypersurfaces the delta method is particularly useful when the degree is at most 3. Unless one appeals to a similar strategy as devised by Marmon and Vishe [139] to deal with quartic hypersurfaces, it seems that when one considers intersections of two hypersurfaces, our situation is just at the barrier. That is, when the sum of the degrees of the individual hypersurfaces exceeds 5, it does not seem to give any improvements compared to the classical circle method.

Once we have achieved the Farey dissection, a standard application of the Poisson summation formula leads us to study certain oscillatory integrals and exponential sums. We provide upper bounds for the oscillatory integrals in Section 6.5 and for the exponential sums and averages thereof in Sections 6.6 and 6.7. When the modulus is square-free, we estimate the exponential sums by appealing to work of Katz [119], that ultimately relies on Deligne's resolution of the Riemann hypothesis over finite fields [67] and obtain cancellations when summing over the numerators. This is usually referred to as a "Kloosterman refinement". As in [202] and [34] it would have been desirable to obtain a double Kloosterman refinement, in which we extract cancellations when summing over both the numerators and denominators. In [202] and [34] the corresponding exponential sums are multiplicative and their averages over the denominator can be studied via the associated L-functions that satisfy a suitable version of the Riemann hypothesis. In our setting, we consider exponential sums associated to linear combinations of the cubic and quadratic form. That these are not homogeneous only allows for a "twisted" form of multiplicativity and it is not clear how to associate an L-function to study their averages. There remains the substantial task of providing estimates for exponential sums when the modulus is not square-free. We are unable to give upper bounds directly, but rather study averages of them over the dual variable in Section 6.7. The underlying arguments go back to work of Heath-Brown [98], but are significantly more complicated in our situation.

Any implied constant in this chapter is allowed to depend on q. Any further dependencies will be indicated by a subscript, unless mentioned otherwise.

6.2 Farey dissection

Vishe's strategy is to find a suitable family of lines in the unit square so that when we consider rational points on these lines they cover the whole square and at the same time stay sufficiently far away from each other to ensure an exact partition. Before reviewing his results in more detail, we need some notation. Let $K = \mathbb{F}_q(t)$ and $\mathcal{O} = \mathbb{F}_q[t]$ be its ring of integers. We will again use the notation introduced in Chapter 3.

If $\underline{c} = (c_1, c_2) \in \mathcal{O}^2$, then we say that $\underline{c}$ is *primitive* if $(c_1, c_2) = 1$ and either c_1 is monic or $c_1 = 0$ and c_2 is monic. For $d, k \in \mathcal{O}$ with $(d, k) = 1$ and $\underline{c} \in \mathcal{O}^2$ primitive we define the affine line

$$L_1(d\underline{c}, k) := \{\underline{x} \in K_\infty^2 : d\underline{c} \cdot \underline{x} = k\} \tag{6.2.1}$$

and the *generalised line*

$$L(d\underline{c}) := \{\underline{a}/r \in \mathbb{T}^2 \cap K^2 \colon (\underline{a}, r) = 1, \underline{a}/r \in L_1(d\underline{c}, k) \text{ for some } k \in \mathcal{O} \text{ with } (k,d) = 1\}. \tag{6.2.2}$$

Note that since $(k,d) = 1$, we must have $d \mid r$ if $\underline{a}/r \in L(d\underline{c})$ with $(\underline{a}, r) = 1$. We refer to $|d\underline{c}|$ as the *height* of $L(d\underline{c})$. For $\underline{x} \in \mathbb{T}^2$ and $R \in \mathbb{Z}$, we let $B(\underline{x}, \widehat{R}) := \{\underline{\theta} \in \mathbb{T}^2 \colon |\underline{\theta} - \underline{x}| < \widehat{R}\}$ be the ball of radius $\widehat{R}$ centered around $\underline{x}$. Similarly, for $\underline{\widehat{R}} = (\widehat{R}_1, \widehat{R}_2) \in \mathbb{Z}$ we set

$$R(\underline{x}, \underline{\widehat{R}}) := \{\underline{\theta} \in \mathbb{T}^2 \colon |\theta_i - x_i| < \widehat{R}_i \text{ for } i = 1, 2\}$$

to be a rectangle of sidelengths $\widehat{R}_1$ and $\widehat{R}_2$ centered around $\underline{x}$. In addition, $\underline{\widehat{R}}^{-1}$ denotes the vector $(\widehat{R}_1^{-1}, \widehat{R}_2^{-1})$. We are now in a position to state Vishe's partition of the unit square.

Theorem 6.2.1. *Let $Q \geq 1$. Then*

$$\mathbb{T}^2 = \bigsqcup_{\substack{r \text{ monic} \\ |r| \leq \widehat{Q}}} \bigsqcup_{\substack{d|r \text{ monic} \\ \underline{c} \in \mathcal{O}^2 \text{ primitive} \\ |r|\widehat{Q}^{-1/2} \leq |d\underline{c}| \leq |r|^{1/2} \\ |dc_2| < |r|^{1/2}}} \sideset{}{'}\bigsqcup_{\substack{|\underline{a}| < |r| \\ \underline{a}/r \in L(d\underline{c})}} B(\underline{a}/r, |r|^{-1}\widehat{Q}^{-1/2}),$$

where $\bigsqcup'$ indicates that we only consider vectors $\underline{a}$ such that $(\underline{a}, r) = 1$.

A few remarks are in order to explain the conditions on the lines in the theorem. First, from a standard application of Dirichlet's approximation theorem [131, Lemma 4.5.1] one obtains

$$\mathbb{T}^k = \bigcup_{\substack{|r| \leq \widehat{Q} \\ r \text{ monic}}} \sideset{}{'}\bigcup_{|\underline{a}| < |r|} B(\underline{a}/r, |r|^{-1}\widehat{Q}^{-1/k}) \tag{6.2.3}$$

for any $Q > 0$ and $k \geq 1$. Furthermore, using the pigeon-hole principle one can show that that any $\underline{a}/r \in K^2$ with $|\underline{a}| < |r|$ and $(\underline{a}, r) = 1$ lies on a line $L(d\underline{c})$ with $|d\underline{c}| \leq |r|^{1/2}$ and $|dc_2| < |r|^{1/2}$. It also clear from the definition (6.2.2) that we must have $d \mid r$. So the key condition in Theorem 6.2.1 is $|r|\widehat{Q}^{-1/2} \leq |d\underline{c}|$. This guarantees that

(i) rational points on an individual line stay sufficiently far away from each other,

(ii) rational points on distinct lines stay sufficiently far away from each other,

(iii) distinct lines don't intersect at rationals with small denominator.

With (i)–(iii) at hand it only remains to show that we can still cover $\mathbb{T}^2$ with balls centered on rationals $\underline{a}/r$ on lines $L(d\underline{c})$ such that $|r|\widehat{Q}^{-1/2} \leq |d\underline{c}|$. This is a consequence of one-dimensional Diophantine approximation, where in fact (6.2.3) already provides an exact partition of the unit interval.

We will follow this blueprint closely to obtain an analogue of Theorem 6.2.1 that allows for lopsided boxes. This requires us to go through most of Vishe's steps again, since we have to modify some of the proofs to get control over the distance between the individual coordinates of rational vectors. We begin with a two-dimensional version of Dirichlet's approximation theorem with rectangles.

Lemma 6.2.2. *let $R_1 \geq R_2 \geq 1$ be integers. Then*

$$\mathbb{T}^2 = \bigcup_{\substack{|r| \leq \widehat{R}_1 \widehat{R}_2 \\ r \text{ monic}}} \bigcup_{|\underline{a}|<|r|}' R(\underline{a}/r, |r|^{-1} \underline{\widehat{R}}^{-1}).$$

Proof. For any $\underline{x} \in \mathbb{T}^2$ the rectangle $R(\underline{x}, \underline{\widehat{R}}^{-1})$ has volume $(\widehat{R}_1 \widehat{R}_2)^{-1}$. We can therefore write

$$\mathbb{T}^2 = \bigsqcup_{i=1}^{\widehat{R}_1 \widehat{R}_2} R(\underline{x}_i, \underline{\widehat{R}}^{-1}) \tag{6.2.4}$$

for some $\underline{x}_i \in \mathbb{T}^2$. Observe that there are $\widehat{R}_1 \widehat{R}_2 q$ polynomials $r \in \mathcal{O}$ with $|r| \leq \widehat{R}_1 \widehat{R}_2$. In particular, if $\underline{x} \in \mathbb{T}^2$ and r runs through all $r \in \mathcal{O}$ with $|r| \leq \widehat{R}_1 \widehat{R}_2$, then two of them, say $r_1 \neq r_2$, must satisfy

$$\{r_i \underline{x}\} \in R(\underline{x}_j, \underline{\widehat{R}}^{-1}),$$

for $i = 1, 2$ and some $1 \leq j \leq \widehat{R}_1 \widehat{R}_2$, where $\{\cdot\}$ denotes the fractional part. If we let $r = r_1 - r_2$, then this implies

$$r\underline{x} - \underline{a} \in R(0, (\widehat{R}_1^{-1}, \widehat{R}_2^{-2})),$$

where $\underline{a}$ is the integer part of $(r_1 - r_2)\underline{x}$. We can divide through by $(\underline{a}, r)$ to ensure that $(\underline{a}, r) = 1$ and also multiply by a unit if necessary to guarantee that r is monic. □

Next we show that every rational lies on a generalised line of suitable height. This is the analogue of [202, Lemma 3.1], where the difference is that we allow the extra parameter T.

Lemma 6.2.3. *Let $T \geq 1$. Then for any $\underline{a}/r \in \mathbb{T}^2$ with $|\underline{a}| < |r|$ and $(\underline{a}, r) = 1$, there exists $d \mid r$ monic and $\underline{c} \in \mathcal{O}^2$ primitive such that $|dc_1| \leq \widehat{T}|r|^{1/2}$, $|dc_2| < \widehat{T}^{-1}|r|^{1/2}$ and $\underline{a}/r \in L(d\underline{c})$.*

Proof. The set

$$\{\underline{c} \in \mathcal{O}^2 \colon |c_1| \leq \widehat{T}|r|^{1/2}, |c_2| < \widehat{T}^{-1}|r|^{1/2}\}$$

has cardinality strictly bigger than $|r|$. In particular, there exists two distinct vectors $\underline{c}_1, \underline{c}_2$ in this set such that $\underline{c}_1 \cdot \underline{a} \equiv \underline{c}_2 \cdot \underline{a} \mod r$. Let $\underline{c}' = \underline{c}_1 - \underline{c}_2$. It then follows that $\underline{c}' \cdot \underline{a} = k'r$ for some $k' \in \mathcal{O}$. Now let $d' = \gcd(c_1, c_2)$, $d = d'/\gcd(d', k')$, $\underline{c} = \underline{c}'/d'$ and $k = k'/\gcd(d', k')$. We then have $d\underline{c} \cdot \underline{a} = kr$. Moreover, by construction $(d, k) = 1$, which also implies $d \mid r$. We can further guarantee that $\underline{c}$ is primitive and d monic by multiplying with a unit and changing k if necessary. □

For any $\underline{c} = (c_1, c_2) \in \mathcal{O}^2$ we let $\underline{c}^\perp = (-c_1, c_1)$. We also need the following result, which is [202, Lemma 3.5], about the distribution of rational points on an individual line.

Lemma 6.2.4. *Let $d \in \mathcal{O}$ be monic and $\underline{c} \in \mathcal{O}^2$ be primitive. Then for every $\underline{a}/r \in L(\underline{c})$ there exists a unique $a \in \mathcal{O}$ with $|a| < |r|$, $(a, r) = 1$ and a unique $\underline{d} \in \mathcal{O}^2$ with $|\underline{d}| < |\underline{c}|$ such that $\underline{a}/r = \frac{a}{r}\underline{c}^\perp + \underline{d}$. Moreover, for every $\underline{a}/r \in L(d\underline{c})$ there exists a unique $\underline{a}'/(r/d) \in L(\underline{c})$ with $|\underline{a}'| < |r/d|$ and a unique $\underline{d}' \in \mathcal{O}^2$ with $|\underline{d}|' < |d|$ such that $\underline{a}/r = \underline{a}'/r + \underline{d}'/d$.*

Next we turn to studying the distance between rational points on lines of the form $L(\underline{c})$. The following result is the analogue of [202, Lemma 3.6].

Lemma 6.2.5. *Let $\underline{c} \in \mathcal{O}^2$ be primitive and $\underline{a}/r \neq \underline{a}'/r' \in L(\underline{c})$. Then*

$$\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right| \geq \frac{|c_i^\perp|}{|rr'|} \text{ for both } i = 1,2 \quad \text{or} \quad \max_{i=1,2}\left\{|c_i|\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right|\right\} \geq 1.$$

The first case happens if $\underline{a}/r, \underline{a}'/r' \in L_1(\underline{c}, k)$ for some $k \in \mathcal{O}$. Moreover, if $\underline{a}/r$ is an element of $L(\underline{c}) \cap L_1(\underline{c}, k)$ with $|\underline{c}| \leq |r|^2$, there exists $\underline{b}/r_1 \in L(\underline{c}) \cap L_1(\underline{c}, k)$ such that $|r_1| = q|r|$ and $|a_i/r - b_i/r_1| = |c_i^\perp|/|rr_1|$.

Proof. We begin with the first part of the lemma. Since $\underline{a}/r, \underline{a}'/r' \in L(\underline{c})$, it follows from the definition of $L(\underline{c})$ that there exist $k, k' \in \mathcal{O}$ such that $(\underline{a}/r - \underline{a}'/r') \cdot \underline{c} = k - k'$. If $k \neq k'$, then by the ultrametric property we have $\max_{i=1,2} |a_i/r - a_i'/r'||c_i| \geq 1$, which is sufficient. Thanks to Lemma 6.2.4 we can also write $\underline{a}/r = \frac{a}{r}\underline{c}^\perp + \underline{d}$ and $\underline{a}'/r' = \frac{a'}{r'}\underline{c}^\perp + \underline{d}'$. If $k = k'$, then the conditions $|\underline{d}|, |\underline{d}'| < |\underline{c}|$ and the fact that $\underline{c}$ is primitive imply that $\underline{d} = \underline{d}'$. Therefore, we have $|a_i/r - a_i'/r'| = |a/r - a'/r||c_i^\perp| \geq |c_i^\perp|/|rr'|$. Furthermore, we have $k = k'$ if and only if $\underline{a}/r, \underline{a}'/r' \in L_1(\underline{c}, k)$.

For the second part of the lemma, one can check that Vishe's proof of [202, Lemma 3.6] in fact gives control over the distance of both coordinates of $\underline{a}/r - \underline{b}/r_1$. Moreover, he requires $|\underline{c}|^2 \leq |r|$, but his proof shows that $|\underline{c}| \leq |r|^2$ is in fact sufficient. □

We can also extend this result to arbitrary lines.

Lemma 6.2.6. *Let $d \in \mathcal{O}$ be monic and $\underline{c} \in \mathcal{O}^2$ be primitive. Then for $\underline{a}/r \neq \underline{a}'/r' \in L(d\underline{c})$, at least one of the following must hold:*

(i) $\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right| \geq \frac{|dc_i^\perp|}{|rr'|}$ *for both $i = 1, 2$,*

(ii) $\left|\frac{\underline{a}}{r} - \frac{\underline{a}'}{r'}\right| \geq \max\{|r|^{-1}, |r'|^{-1}\}$,

(iii) $\max_{i=1,2}\left\{|dc_i|\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right|\right\} \geq 1.$

Moreover, if $\underline{a}/r \in L_1(d\underline{c}, k)$, then there exists $\underline{a}/r \neq \underline{b}/r_2 \in L_1(d\underline{c}, k)$ such that

$$|a_i/r - b_i/r_2| \leq |dc_i^\perp|/|rr_2|$$

for both $i = 1, 2$.

Proof. We begin with the first part of the statement. Recall from Lemma 6.2.4 that we can write $\underline{a}/r = \underline{a}_1/r + \underline{d}/d$ and $\underline{a}'/r' = \underline{a}_2/r' + \underline{d}'/d$ where $\underline{a}_1/(r/d), \underline{a}_2/(r'/d) \in L(\underline{c})$ and $\underline{d}, \underline{d}' \in \mathcal{O}^2$. We thus have

$$\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right| = \left|\frac{a_{1,i}}{r} - \frac{a_{2,i}}{r'} + \frac{d_i - d_i'}{d}\right|$$

for $i = 1, 2$. If $\underline{d} \neq \underline{d}'$, then this is clearly at least $1/|d| \geq \max\{|r|^{-1}, |r'|^{-1}\}$ for one of $i = 1, 2$, since $d \mid r, r'$. One the other hand, if $\underline{d} = \underline{d}'$, we can use Lemma 6.2.5: In its first case we obtain $|a_{1,i}/(r/d) - a_{2,i}/(r'/d)| \geq |d^2||c_i^\perp|/|rr'|$ for both $i = 1, 2$, which implies

$$|a_i/r - a_i'/r'| \geq |dc_i^\perp|/|rr'| \quad \text{for} \quad i = 1, 2,$$

whereas in the second case $\max_{i=1,2}\{|c_i||a_{1,i}/(r/d) - a_{2,i}/(r/d)|\} \geq 1$, which implies

$$\max_{i=1,2} |dc_i||a_i/r - a_i'/r'| \geq 1.$$

For the second part of the lemma, we use Lemma 6.2.4 to write $\underline{a}/r = \underline{b}'/r + \underline{d}'/d$, where $\underline{b}'/(r/d) \in L_1(\underline{c}, k)$ for some $k \in \mathcal{O}$. It follows from the second part of Lemma 6.2.5 that there exists $\underline{b}''/r_1 \in L_1(\underline{c}, k)$ with $|r_1| = q|r/d|$ and $|b_i'/(r/d) - b_i''/r_1| = |dc_i^\perp|/|rr_1|$ for $i = 1, 2$. Now set $\underline{b}/r_2 = \underline{b}''/(r_1 d) + \underline{d}'/d$. We then have $d\underline{c} \cdot \underline{b}/r_2 = k + \underline{c} \cdot \underline{d}' = d\underline{c} \cdot \underline{a}/r$, so that $\underline{b}/r_2 \in L_1(d\underline{c}, k)$. Moreover, we also have $|b_i/r_2 - a_i/r| = |b_i''/r_1 - b_i'/(r/d)|/|d| = |c_i^\perp|/|rr_1| \leq |dc_i^\perp|/|rr_2|$, where we used that $|r_2| \leq |r_1 d|$. □

We also need the following lemma, which is the second part of [202, Lemma 3.9].

Lemma 6.2.7. *Let $\underline{a}/r \in L(d\underline{c}) \cap L(d'\underline{c}')$, where $d, d' \in \mathcal{O}$ are monic and $\underline{c}, \underline{c}' \in \mathcal{O}^2$ are primitive. If $|c_1 c_2'|, |c_2 c_1'| < |r/(dd')|$, then $d\underline{c} = d'\underline{c}'$.*

Note that in [202, Lemma 3.9] the extra condition $|d\underline{c}|^2, |d'\underline{c}'|^2 \leq |r|$ is required, but this is in fact not used in the proof. The next lemma is concerned with the distance between rational points that lie on distinct lines.

Lemma 6.2.8. *Let $\underline{a}/r \in L(d\underline{c})$ and $\underline{a}'/r' \in L(d'\underline{c}')$ with $d\underline{c} \neq d'\underline{c}'$, $|dd'c_1c_2'| < |rr'|^{1/2}$ and $|dd'c_2c_1'| < |rr'|^{1/2}$. Then we have $|a_i/r - a_i'/r'| \geq |dc_i r'|^{-1}$ for one of $i = 1, 2$ and $|a_i/r - a_i'/r'| \geq |d'c_i' r|^{-1}$ for one of $i = 1, 2$.*

Proof. First note that $\underline{a}/r$ and $\underline{a}'/r'$ must be distinct by Lemma 6.2.7 . By the second part of Lemma 6.2.6 there exists $\underline{b}/r_1 \in L(d\underline{c})$ such that $|a_i/r - b_i/r_1| \leq |dc_i^\perp|/|rr_1|$. Let

$$C = \begin{pmatrix} \underline{a}/r - \underline{b}/r_1 \\ \underline{a}/r - \underline{a}'/r' \end{pmatrix}.$$

Since $\underline{a}'/r' \notin L(d\underline{c})$, it follows that $\det(C) \neq 0$. It is therefore clear that $|\det(C)| \geq |rr'r_1|$. This implies that $|a_i/r - a_i'/r'| \geq 1/|dc_i r'|$ for one of $i = 1, 2$. Finally we can replace the role of r and r' to obtain the second inequality of the statement. □

We now have all ingredients at hand to prove the main result of this section.

Theorem 6.2.9. *Let $R_1 \geq R_2 \geq 1$ be integers. If we set $T = (R_1 - R_2)/2$, then*

$$\mathbb{T}^2 = \bigsqcup_{\substack{|r| \leq \widehat{R}_1 \widehat{R}_2 \\ r \text{ monic}}} \quad \bigsqcup_{\substack{d|r \text{ monic} \\ \underline{c} \in \mathcal{O}^2 \text{ primitive} \\ |dc_1| \leq \widehat{T}|r|^{1/2}, |dc_2| < \widehat{T}^{-1}|r|^{1/2} \\ \max\{\widehat{R}_i |dc_i^\perp|\} \geq |r|}} \quad {\bigsqcup_{\substack{|\underline{a}| < |r| \\ \underline{a}/r \in L(d\underline{c})}}}' R(\underline{a}/r, |r|^{-1}\underline{\widehat{R}}^{-1}). \tag{6.2.5}$$

Proof. We first show the union on the right hand side of (6.2.5) is disjoint. Let $\underline{a}/r \neq \underline{a}'/r'$ appear on the right hand side of (6.2.5) and suppose $|r'| \geq |r|$. We now have to distinguish a few cases. First, if $\underline{a}/r, \underline{a}'/r' \in L(d\underline{c})$ for some $L(d\underline{c})$ appearing on the right hand side of (6.2.5), then in case (i) of Lemma 6.2.6 we have

$$\left|\frac{a_i}{r} - \frac{a_i'}{r'}\right| \geq \frac{|dc_i^\perp|}{|rr'|} \geq \frac{1}{\widehat{R}_i |r|}$$

for one of $i = 1, 2$, where we used that $|dc_i^{\perp}| \geq \widehat{R}_i^{-1}|r'|$ for one of $i = 1, 2$. This is clearly sufficient to show $R(\underline{a}/r, |r|^{-1}\underline{\widehat{R}}^{-1})$ is disjoint from $R(\underline{a}'/r', |r'|^{-1}\underline{\widehat{R}}^{-1})$. On the other hand, case (ii) of Lemma 6.2.6 yields $|\underline{a}/r - \underline{a}'/r'| \geq \max\{|r|^{-1}, |r'|^{-1}\}$. However, then the rectangles around $\underline{a}/r$ and $\underline{a}'/r'$ must be disjoint since $\widehat{R}_1 \geq \widehat{R}_2 \geq 1$. If case (iii) in Lemma 6.2.6 holds for $i = 1$, then $dc_1 \neq 0$ and

$$\left|\frac{a_1}{r} - \frac{a_1'}{r'}\right| \geq \frac{1}{|dc_1|} \geq \frac{\widehat{R}_2|dc_1|}{\widehat{R}_1|r|} \geq \frac{1}{\widehat{R}_1|r|}.$$

A similar calculation shows that if the inequality holds for $i = 2$, then we obtain $|a_2/r - a_2'/r'| \geq \widehat{R}_2^{-1}|r|^{-1}$. This finishes the case $\underline{a}/r, \underline{a}'/r' \in L(d\underline{c})$. Next we are concerned about the case $\underline{a}/r \in L(d\underline{c})$, $\underline{a}'/r' \in L(d'\underline{c}')$ with $d\underline{c} \neq d'\underline{c}'$. Our constraints on $d\underline{c}, d'\underline{c}'$ guarantee that the requirements in Lemma 6.2.8 are met, and so we get $|a_i/r - a_i/r'| \geq |dc_i r|^{-1}$ for one of $i = 1, 2$. Note that $|dc_1| \leq (\widehat{R}_1/\widehat{R}_2)^{1/2}|r|^{1/2} \leq \widehat{R}_1$, since $|r| \leq \widehat{R}_1\widehat{R}_2$. Similarly we get $|dc_2| \leq \widehat{R}_2$, so that $|a_i/r - a_i/r'| \geq |dc_i r|^{-1} \geq \widehat{R}_i^{-1}|r|^{-1}$, which is sufficient. Finally, it remains to show that every rational $\underline{a}/r$ in the right hand side of (6.2.5) appears precisely once. This is a consequence of Lemma 6.2.7. Therefore, we have established that the right hand side of (6.2.5) is a disjoint union.

Now we show that if $\underline{a}/r \in \mathbb{T}^2$ is a rational with $|\underline{a}| < |r| \leq \widehat{R}_1\widehat{R}_2$ and $(\underline{a}, r) = 1$, then there exists a rational $\underline{a}'/r'$ appearing on the right hand side of (6.2.5) such that

$$R(\underline{a}/r, |r|^{-1}\underline{\widehat{R}}^{-1}) \subset R(\underline{a}'/r', |r'|^{-1}\underline{\widehat{R}}^{-1}).$$

By Lemma 6.2.3 this is enough to show the equality of sets in (6.2.5). It follows from Lemma 6.2.3 that there exist $d \in \mathcal{O}$ monic and $\underline{c} \in \mathcal{O}^2$ primitive such that $d \mid r$, $\underline{a}/r \in L(d\underline{c})$ and $|dc_1| \leq T|r|^{1/2}$, $|dc_2| < T^{-1}|r|^{1/2}$. If $\max\{\widehat{R}_i|dc_i^{\perp}|\} \geq |r|$ we are done. Otherwise, let $M = \max\{\widehat{R}_i|c_i^{\perp}|\}$ so that $M < |r|$. Lemma 6.2.4 allows to write $\underline{a}/r = \frac{a}{r}\underline{c}^{\perp} + \underline{d}/d$ for some $a \in \mathcal{O}$ and $\underline{d} \in \mathcal{O}^2$, where $a/(r/d) \in L(\underline{c})$. The one dimensional Dirichlet approximation theorem implies the existence of a rational a'/r_1 such that $|a'| < |r_1| \leq M|d|^{-1}$, $(a', r_1) = 1$ and $|a/(r/d) - a'/r_1| < |r_1|^{-1}|d|M^{-1}$. Now set $\underline{a}'/r' = \frac{a'}{r_1 d}\underline{c}^{\perp} + \underline{d}/d$. We then have $d\underline{c} \cdot \underline{a}'/r' = \underline{c} \cdot \underline{d} = d\underline{c} \cdot \underline{a}/r$, which implies $\underline{a}'/r' \in L(d\underline{c})$ and $d \mid r'$. Moreover, by construction we have $|r'| \leq |dr_1| \leq M$. We also have

$$|a_i/r - a_i'/r'| = |d|^{-1}|a_i/(r/d) - a'/r_1| < |r_1|^{-1}M^{-1} \leq |r'|^{-1}\widehat{R}_i^{-1}$$

for both $i = 1, 2$. This completes the proof of Theorem 6.2.9. □

Remark. Note that in the particular case $R_1 = R_2$ we recover Theorem 6.2.1 with $Q = 2R_1$ from Theorem 6.2.9.

The following corollary will be useful when evaluating the main contribution to our asymptotic formula.

Corollary 6.2.10. *Let $R_1 \geq R_2 \geq 1$ be integers. Then*

$$\bigsqcup_{\substack{|r| \leq \widehat{R}_2 \\ r \text{ monic}}} \bigsqcup_{\substack{d \mid r \text{ monic} \\ \underline{c} \in \mathcal{O}^2 \text{ primitive} \\ |dc_1| \leq \widehat{T}|r|^{1/2} \\ |dc_2| < \widehat{T}^{-1}|r|^{1/2}}} \bigsqcup_{\substack{|\underline{a}| < |r| \\ \underline{a}/r \in L(d\underline{c})}}{}' R(\underline{a}/r, |r|^{-1}\underline{\widehat{R}}^{-1}) = \bigsqcup_{\substack{|r| \leq \widehat{R}_2 \\ r \text{ monic}}} \bigsqcup_{|\underline{a}| < |r|}{}' R(\underline{a}/r, |r|^{-1}\underline{\widehat{R}}^{-1})$$

Proof. Clearly the left hand side of the claimed equality is contained in the right hand side. Moreover, by Lemma 6.2.3 the right hand side is contained in the left hand side. Theorem 6.2.9 implies that the left hand side is disjoint, since the condition $\max\{\widehat{R}_i|dc_i^{\perp}|\} \geq |r|$ is vacuously true for $|r| \leq \widehat{R}_2$. It remains to prove that the right hand side is disjoint. Suppose that $\underline{\alpha} \in R(\underline{a}_1/r_1, |r_1|^{-1}\underline{\widehat{R}}^{-1}) \cap R(\underline{a}_2/r_2, |r_2|^{-1}\underline{\widehat{R}}^{-1})$ with $\underline{a}_1/r_1 \neq \underline{a}_2/r_2$. We then have

$$\frac{1}{|r_1r_2|} \leq \left|\frac{\underline{a}_1}{r_1} - \frac{\underline{a}_2}{r_2}\right| = \left|\left(\frac{\underline{a}_1}{r_1} - \underline{\alpha}\right) + \left(\underline{\alpha} - \frac{\underline{a}_2}{r_2}\right)\right| < \max\left\{\frac{1}{|r_1|\widehat{R}_2}, \frac{1}{|r_2|\widehat{R}_2}\right\},$$

which is impossible since $|r_1|, |r_2| \leq \widehat{R}_2$. □

6.3 Geometry

If the complete intersection $X \subset \mathbb{P}^{n-1}$ is defined by a cubic form F_1 and a quadratic form F_2, then it is also defined by $F_1 + LF_2$ and F_2 for any linear form $L \in \mathcal{O}[x_1, \dots, x_n]$. In particular, given the degree of freedom we have, it is reasonable to expect that we can define X as the intersection of a non-singular cubic hypersurface and a quadratic hypersurface. This is indeed the case as demonstrated in Lemmas 3.2–3.3 of [37], whose proof we adjust to cope with positive characteristic.

Lemma 6.3.1. *Let $X \subset \mathbb{P}^{n-1}$ be a non-singular complete intersection of a cubic and a quadratic hypersurface over K. Then $X = V(F_1, F_2)$, where $F_1 \in \mathcal{O}[x_1, \dots, x_n]$ is a non-singular cubic form and F_2 is a quadratic form of rank at least $n-1$.*

Proof. Suppose $X = V(G_1, G_2)$, where G_1 is a cubic and G_2 a quadratic form respectively. For $U = \mathbb{P}^{n-1} \setminus V(G_2)$ define the morphism

$$\varphi\colon U \to \mathbb{P}^n, \quad (x_1\colon \cdots : x_n) \mapsto (G_1(\boldsymbol{x})\colon x_1G_2(\boldsymbol{x}) \cdots : x_nG_2(\boldsymbol{x})).$$

Assume for a moment that there exists a hyperplane $H \subset \mathbb{P}^n$ defined over K such that $\varphi^{-1}(H)$ is smooth. This means that there exist $\lambda_0, \dots, \lambda_n \in K$ such that

$$F_1 = \lambda_0G_1 + \lambda_1x_1G_2 + \cdots + \lambda_nx_nG_2$$

satisfies $U \cap \{F_1 = 0\}$ is smooth. However, $U \cap \{F_1 = 0\} = \{F_1 = 0\} \setminus \{G_2 = 0\}$, from which it follows that F_1 is non-singular since X is non-singular.

To prove the existence of the claimed λ_i's, Browning–Dietmann–Heath-Brown appeal to Bertini's Theorem, which does not hold in general in positive characteristic. However, it follows from work of Spreafico [189, Corollary 4.3] that the fiber above a general hyperplane $H \subset \mathbb{P}^n$ is smooth provided the induced extension of residues fields $\kappa(x)/\kappa(\varphi(x))$ is separable for any $x \in U$. Let $Y = V(G_1 - x_0G_2) \subset \mathbb{P}^n$. Then φ factors into

$$U \to Y \setminus V(G_2) \to Y \to \mathbb{P}^n,$$

where the first arrow is an isomorphism, the second an open embedding and the third a closed immersion. Indeed, if $G_2(\boldsymbol{x}) \neq 0$ with $\boldsymbol{x} = (x_1, \dots, x_n)$ and $(x_0\colon \cdots : x_n)$ lies on Y, then $x_0 = G_1(\boldsymbol{x})/G_2(\boldsymbol{x})$ and hence

$$(x_0\colon \cdots : x_n) = (G_1(\boldsymbol{x})\colon x_1G_2(\boldsymbol{x})\colon \cdots : x_nG_2(\boldsymbol{x})) = \varphi(\boldsymbol{x}).$$

Moreover, $Y \setminus V(G_2)$ is an open subset of Y and Y a closed subset of $\mathbb{P}^n$. Both open embeddings and closed immersions are unramified, as is the composition of unramified morphisms. It follows that φ is unramified and hence all the residue field extensions are separable.

It remains to show that G_2 has rank at least $n-1$. Aiming at a contradiction, suppose the opposite holds. This implies that G_2 is singular along a line. However, this line will intersect $\{F_1 = 0\}$ in a point that will then be a singular point of X, which is impossible. □

To deal with the exponential integrals appearing in our work, we will have to concentrate our weight function near a point such that linear combinations of its Hessian associated with F_1 and the matrix underlying the quadratic form F_2 always have large rank. This is only possible when $\mathrm{char}(K) > 3$ and so we assume this holds for the rest of our work. Let us now fix a symmetric matrix $M \in \mathcal{O}^{n\times n}$ such that $F_2(\boldsymbol{x}) = \boldsymbol{x}^t M \boldsymbol{x}$. Moreover, for $\boldsymbol{x} \in K_\infty^n$ we shall denote by $H(\boldsymbol{x}) = (\frac{\partial F_1}{\partial x_i \partial x_j})_{1\leq i,j\leq n}$ the Hessian of F_1 evaluated at $\boldsymbol{x}$.

Lemma 6.3.2. *For $1 \leq k \leq n-1$, let V_k be the Zariski closure of*

$$V_k' := \{\boldsymbol{x} \in \mathbb{A}^n \colon F_1(\boldsymbol{x}) = 0 \textit{ and } \mathrm{rk}(t_1 H(\boldsymbol{x}) + t_2 M) \leq k \textit{ for some } (t_1,t_2) \in \mathbb{A}^2 \setminus \{\boldsymbol{0}\}\}$$

inside $\mathbb{A}^n$. Then $\dim V_k \leq k$.

Proof. For $1 \leq k \leq n-1$ consider the incidence correspondence

$$I := \{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{A}^n \times \mathbb{A}^n \colon F_1(\boldsymbol{x}) = 0 \text{ and } \mathrm{rk}(H(\boldsymbol{x})\boldsymbol{y}, M\boldsymbol{y}) \leq 1\}.$$

Let V be an irreducible component of V_k. Since $V \times \{\boldsymbol{0}\} \subset I$ and $V \times \{\boldsymbol{0}\}$ is irreducible, there exists an irreducible component W of I containing $V \times \{\boldsymbol{0}\}$. Then the projection of I onto the first factor restricts to a surjective morphism $\psi \colon W \to V$ of irreducible varieties. We can therefore apply Chevalley's theorem [90, Proposition 14.109] to deduce the existence of an open dense subset $U \subset V$ such that $\dim \psi^{-1}(\boldsymbol{x}) = \dim W - \dim V$ for all $\boldsymbol{x} \in V$. Note that since V_k' is dense in V_k and V is an irreducible component of V_k, we must have $U \cap V_k' \neq \emptyset$. In addition, by the definition of V_k' we have $\dim \psi^{-1}(\boldsymbol{x}) \geq n-k$ for all $\boldsymbol{x} \in V_k'$, so that we must also have $\dim \psi^{-1}(\boldsymbol{x}) \geq n-k$ for all $\boldsymbol{x} \in U$.

Our next task is to bound the dimension of I. For this, let $\Delta \colon \mathbb{A}^n \to \mathbb{A}^n \times \mathbb{A}^n$ the diagonal embedding, that is $\boldsymbol{x} \mapsto (\boldsymbol{x},\boldsymbol{x})$, and let $S = \Delta(\mathbb{A}^n) \cap I$. If $\boldsymbol{x} \in S$, then there exists $(t_1,t_2) \in \mathbb{A}^2 \setminus \{\boldsymbol{0}\}$ such that $t_1 \nabla F_1(\boldsymbol{x}) + t_2 \nabla F_2(\boldsymbol{x}) = \boldsymbol{0}$. One the one hand, if $t_2 = 0$, then we get $\nabla F_1(\boldsymbol{x}) = \boldsymbol{0}$, which implies $\boldsymbol{x} = \boldsymbol{0}$ since F_1 is non-singular. On the other hand, if $t_2 \neq 0$, then after taking the inner product with $\boldsymbol{x}$, we get $F_2(\boldsymbol{x}) = 0$, so that $\boldsymbol{x}$ is a singular point on the affine cone of the non-singular complete intersection of F_1 and F_2, which implies $\boldsymbol{x} = \boldsymbol{0}$. Altogether we obtain $\dim S = 0$.

Having established an upper bound for $\dim S$, we are now in a position to get control over $\dim I$. From what we have just shown, it follows that

$$0 = \dim I \cap \Delta(\mathbb{A}^n) \geq \dim I + \dim \mathbb{A}^n - 2n = \dim I - n$$

and thus $\dim I \leq n$, from which we immediately deduce $\dim W \leq n$. Combining this with the information about the dimension of the fibers of ψ, we obtain for any $\boldsymbol{x} \in U$ the inequality $n - k \leq \dim \psi^{-1}(\boldsymbol{x}) = \dim W - \dim V_k$ and therefore $\dim V_k \leq k$ as claimed. □

Corollary 6.3.3. *Let $n \geq 14$. There exists $\boldsymbol{x}_0 \in K_\infty^n$ such that $F_1(\boldsymbol{x}_0) = F_2(\boldsymbol{x}_0) = 0$, rk$(H(\boldsymbol{x}_0)) \geq n-2$ and rk$(t_1 H(\boldsymbol{x}) + t_2 M) \geq n-2$ for all $(t_1, t_2) \in K_\infty^2 \setminus \{\mathbf{0}\}$.*

Proof. Let $X' \subset \mathbb{A}^n$ be the affine cone of the non-singular complete intersection $X = V(F_1, F_2) \subset \mathbb{P}^{n-1}$. It follows from Lemma 6.3.2 that V_{n-3} is a Zariski closed subset in $\mathbb{A}^n$ of dimension at most $n-3$. As $n \geq 13$, Lang–Tsen theory [91, Theorem 3.6] implies that $X(K_\infty) \neq \emptyset$ and since X is non-singular, it follows that $X'(K_\infty)$ is Zariski dense in X'. In particular, the fact that $\dim X' = n-2$ implies that $X'(K_\infty) \setminus V_{n-3}$ is non-empty and any point contained therein satisfies the conditions required in the statement of the corollary. □

We also need strong upper bounds for the number of integral points on an affine hypersurface, which are a special case of [155, Theorem 1.10] in the $\mathbb{F}_q[t]$ setting.

Theorem 6.3.4. *Let $G \in \mathcal{O}[x_1, \dots, x_n]$ be a polynomial of degree $d \geq 5$ whose degree d part is absolutely irreducible and let $B \geq 1$. Then there exists a constant $C > 0$ depending only on d, n and q such that*

$$\#\{\boldsymbol{x} \in \mathcal{O}^n : |\boldsymbol{x}| < \widehat{B}, G(\boldsymbol{x}) = 0\} \leq C\widehat{B}^{n-2}.$$

6.4 Activation of the circle method

In this section we collect the remaining facts needed to get the circle method started. Recall from Lemma 6.3.1 that we can assume $X = V(F_1, F_2)$ with $F_1 \in \mathcal{O}[x_1, \dots, x_n]$ a non-singular cubic form and $F_2 \in \mathcal{O}[x_1, \dots, x_n]$ a quadratic form of rank at least $n-1$. We shall fix such a choice of F_1 and F_2 once and for all and write $\underline{F} = (F_1, F_2)$. Moreover, we assume $M \in \mathrm{Mat}_{n\times n}(\mathcal{O})$ is a symmetric matrix such that $F_2(\boldsymbol{x}) = \boldsymbol{x}^t M \boldsymbol{x}$. In what follows, for $F \in K_\infty[x_1, \dots, x_n]$ we refer to the maximum of the absolute values of the coefficients of F as the height of F and denote it by H_F. We extend this definition to pairs of polynomials by $H_{\underline{G}} := \max\{H_{G_1}, H_{G_2}\}$.

Corollary 6.3.3 implies that there exists $\boldsymbol{x}_0 \in K_\infty^n$ such that

$$\begin{aligned} &F_1(\boldsymbol{x}_0) = F_2(\boldsymbol{x}_0) = 0, \\ &\mathrm{rk}(H(\boldsymbol{x}_0)) \geq n-2 \text{ and} \\ &\mathrm{rk}(\gamma_1 H(\boldsymbol{x}_0) + \gamma_2 M) \geq n-2 \text{ for all } (\gamma_1, \gamma_2) \in K_\infty^2 \setminus \{\mathbf{0}\}, \end{aligned} \tag{6.4.1}$$

where $H(\boldsymbol{x}_0)$ denotes the Hessian of the cubic form F_1 evaluated at $\boldsymbol{x}_0$. These properties are clearly invariant under scaling and so we may additionally assume $|\boldsymbol{x}_0| < H_{\underline{F}}^{-1}$. We will then work with the weight function $w\colon K_\infty^n \to \mathbb{R}_{\geq 0}$ defined by

$$w(\boldsymbol{x}) := \begin{cases} 1 & \text{if } |\boldsymbol{x} - \boldsymbol{x}_0| < \widehat{L}^{-1}, \\ 0 & \text{else,} \end{cases}, \tag{6.4.2}$$

where L is a large but fixed integer, whose exact value will be determined throughout our work. The non-archimedean nature of K_∞ ensures that rk$(H(\boldsymbol{x})) \geq n-2$ and $|\boldsymbol{x}| < 1/H_{\underline{F}}$ whenever $w(\boldsymbol{x}) \neq 0$ and L is sufficiently large. Moreover, we have seen in the proof of Corollary 6.3.3 that the set of points $\boldsymbol{x} \in K_\infty^n$ satisfying $\mathrm{rk}(\gamma_1 H(\boldsymbol{x}) + \gamma_2 M) \geq n-2$ for all $(\gamma_1, \gamma_2) \in K_\infty^2 \setminus \{\mathbf{0}\}$ is Zariski dense in K_∞^n. In particular, if L is large enough, we can guarantee that any $\boldsymbol{x} \in \operatorname{supp}(w)$ satisfies the third property in (6.4.1). Let us now fix $\boldsymbol{b} \in \mathcal{O}^n$ and $N \in \mathcal{O}$ such that $N \mid F_1(\boldsymbol{b}), F_2(\boldsymbol{b})$. The counting function we consider is now given by

$$N(P) := \sum_{\substack{\boldsymbol{x}\in\mathcal{O}^n \\ F_1(\boldsymbol{x})=F_2(\boldsymbol{x})=0 \\ \boldsymbol{x}\equiv\boldsymbol{b}\ (N)}} w\left(\frac{\boldsymbol{x}}{P}\right).$$

Combining the orthogonality relation from Lemma 3.0.2 with Theorem 6.2.9 immediately implies that $N(P)$ is given by

$$\sum_{\substack{|r|\leq\widehat{R}_1\widehat{R}_2 \\ r \text{ monic}}} \sum_{\substack{d|r \text{ monic} \\ \underline{c}\in\mathcal{O}^2 \text{ primitive} \\ |dc_1|\leq\widehat{T}|r|^{1/2}, |dc_2|<\widehat{T}^{-1}|r|^{1/2} \\ \max\{\widehat{R}_i|dc_i^{\perp}|\}\geq|r|}} \sideset{}{'}\sum_{\substack{|\underline{a}|<|r| \\ \underline{a}/r\in L(d\underline{c})}} \int_{|\theta_1|<|r|^{-1}\widehat{R}_1^{-1}} \int_{|\theta_2|<|r|^{-1}\widehat{R}_2^{-1}} S\left(\frac{\underline{a}}{r}+\underline{\theta}\right) \mathrm{d}\theta_2\mathrm{d}\theta_1, \tag{6.4.3}$$

where

$$S(\underline{\alpha}) := \sum_{\substack{\boldsymbol{x}\in\mathcal{O}^n \\ \boldsymbol{x}\equiv\boldsymbol{b}\ (N)}} \psi(\alpha_1F_1(\boldsymbol{x})+\alpha_2F_2(\boldsymbol{x}))w(\boldsymbol{x}/P), \tag{6.4.4}$$

for $\underline{\alpha} \in \mathbb{T}^2$ and $\sum'$ indicates the condition $(\underline{a}, r) = 1$. After splitting $\boldsymbol{x}$ into residue classes modulo r_N, where $r_N = rN/(r,N)$, it is a standard argument, see [37, Lemma 4.4] for example, to use the Poisson summation formula from Proposition 3.0.5 to evaluate $S(\underline{\theta}+\underline{a}/r)$ and transform (6.4.3) into

$$N(P) = |P|^n \sum_{\substack{|r|\leq\widehat{R}_1\widehat{R}_2 \\ r \text{ monic}}} |r_N|^{-n} \sum_{\substack{d|r \text{ monic} \\ \underline{c}\in\mathcal{O}^2 \text{ primitive} \\ |dc_1|\leq\widehat{T}|r|^{1/2}, |dc_2|<\widehat{T}^{-1}|r|^{1/2} \\ \max\{\widehat{R}_i|dc_i^{\perp}|\}\geq|r|}} \int_{D(|r|\underline{\widehat{R}})} \sum_{\boldsymbol{v}\in\mathcal{O}^n} S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) I_{r_N}(\underline{\theta},\boldsymbol{v})\mathrm{d}\underline{\theta}, \tag{6.4.5}$$

where $D(|r|\underline{\widehat{R}}) = \{\underline{\theta} \in \mathbb{T}^2 \colon |r\theta_i| < \widehat{R}_i^{-1} \text{ for } i = 1,2\}$,

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) := \sideset{}{'}\sum_{\underline{a}/r\in L(d\underline{c})} \sum_{\substack{|\boldsymbol{x}|<|r_M| \\ \boldsymbol{x}\equiv\boldsymbol{b}\ (N)}} \psi\left(\frac{a_1F_1(\boldsymbol{x})+a_2F_2(\boldsymbol{x})}{r}\right)\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{r_N}\right), \tag{6.4.6}$$

and

$$I_s(\underline{\theta},\boldsymbol{v}) := \int_{K_\infty^n} w(\boldsymbol{x})\psi\left(P^3\theta_1F_1(\boldsymbol{x})+P^2\theta_2F_2(\boldsymbol{x})+\frac{P\boldsymbol{v}\cdot\boldsymbol{x}}{s}\right)\mathrm{d}\boldsymbol{x} \tag{6.4.7}$$

for $s \in \mathcal{O}\setminus\{0\}$.

In our work we will assume throughout that R_1 and R_2 are chosen in such a way that

$$\widehat{R}_1 \asymp |P|^{4/3} \quad \text{and} \quad \widehat{R}_2 \asymp |P|^{1/3}, \tag{6.4.8}$$

so that $\widehat{T} \asymp |P|^{1/2}$. This ensures that $\operatorname{vol}(D(|r|\underline{\widehat{R}})) \asymp |P|^{-5}$ and $|I_r(\underline{\theta},\boldsymbol{v})| \asymp 1$ when $|r| = \widehat{R}_1\widehat{R}_2$. Let us now separate the terms from (6.4.5) that will go into the error term. For this, we write

$$N(P) = M(P) + E_1(P) + E_2(P), \tag{6.4.9}$$

where $E_2(P)$ consists of the contribution in (6.4.3) for which

1. $\boldsymbol{v} \neq \boldsymbol{0}$ with $|\theta_1| < |P|^{-9}\widehat{R}_2$ or $|\theta_2| < |P|^{-9}\widehat{R}_1$ or

2. $c_2 = 0$ with $|\theta_1| > |r|^{-1}|P|^{-\delta}$, where $\delta = 8(n-16)/(3n-24)$, and $|r| \leq |P|^{1-\eta}$ with $\eta = 2/n$,

holds. Observe that the set of $\underline{\theta}'$ for which (2) holds is non-empty when $n > 24$, because then $\delta > 4/3$. The terms $M(P)$ and $E_1(P)$ comprise the contribution from all r's and $\underline{\theta}$'s for which neither (1) nor (2) holds with the additional constraint that $\boldsymbol{v} = \mathbf{0}$ for $M(P)$ and $\boldsymbol{v} \neq \mathbf{0}$ for $E_1(P)$ respectively.

We begin with estimating the contribution to $E_2(P)$ defined by (1). Note that for any $r \in \mathcal{O}$ the measure of $\underline{\theta} \in \mathbb{T}^2$ for which (1) holds is $O(|P|^{-9}|r|^{-1})$. Estimating trivially $S(\underline{a}/r + \underline{\theta}) \leq |P|^n$ and using Lemma 6.6.3 to deduce that the number of $\underline{a}$ with $|\underline{a}| < |r|$ such that $\underline{a}/r \in L(d\underline{c})$ is at most $|dr|$, we see that the contribution from (1) in (6.4.3) is

$$\ll |P|^{n-9} \sum_{|r| \leq \widehat{R}_1\widehat{R}_2} \sum_{\substack{|dc_1| \leq \widehat{T}|r|^{1/2} \\ |dc_2| < \widehat{T}^{-1}|r|^{1/2} \\ d|r}} |d| \ll |P|^{n-9} \sum_{|r| \leq \widehat{R}_1\widehat{R}_2} |r|^{1+\varepsilon} \ll |P|^{n-17/3+\varepsilon},$$

upon recalling (6.4.8) for our choice of R_1 and R_2.

The reason for separating the contribution coming from (2) is that the integral estimates we provide in Section 6.5 are insufficient when $|r|$ is small and $|\underline{\theta}|$ is large. We eliminate this shortfall by dealing with this contribution in a manner akin to the treatment of the minor arcs in a classical application of the circle method. To begin with, let us fix the absolute values of r and θ_1 in the definition of $E_2(P)$ to be $|r| = \widehat{Y}$ and $|\theta_1| = \widehat{\Theta}_1$ with $-Y - \delta P \leq \Theta_1 \leq -Y - 4P/3$. The main tool to deal with this contribution is Weyl's inequality, whose function field analogue is provided by Lemma 4.3.6 in Lee's PhD thesis [131] and that we reproved in Corollary 4.1.2.

Lemma 6.4.1. *Let $\underline{\alpha} \in \mathbb{T}^2$ and $a_1/r \in K \cap \mathbb{T}$ be such that $(a_1, r) = 1$ and $\alpha_1 = a_1/r + \theta_1$. Then for $S(\underline{\alpha})$ given by* (6.4.4) *we have*

$$S(\underline{\alpha}) \ll_{\boldsymbol{b},N,F_1} |P|^{n+\varepsilon} \left(\frac{|P| + |r| + |P|^3|r_1\theta_1|}{|P|^3} + \frac{1}{|r_1| + |P|^3|r_1\theta_1|} \right)^{n/8}.$$

Remark. Lee states Lemma 4.3.6 without the appearance of the quadratic form that features in the definition of $S(\underline{\alpha})$. However, as we saw in the proof of Corollary 4.1.2 the process of Weyl differencing eliminates the effect of the quadratic form, so that the estimate in Lemma 6.4.1 only depends on F_1.

We now wish to apply Lemma 6.4.1. The problem is that if $\underline{a}/r \in L(d\underline{c})$, we do not necessarily have $(a_1, r) = 1$. However, recall from (6.2.2) that $\underline{a}/r \in L(d\underline{c})$ if and only if $d\underline{a} \cdot \underline{c} = kr$ for some $k \in \mathcal{O}$ with $(k, d) = 1$. In particular, if $c_2 = 0$, then $c_1 = 1$ and so $\underline{a}/r \in L(d(1, 0))$ if and only if $a_1/r = k/d$ with $(k, d) = 1$. It is now easily checked that the constraints coming from (2) together with Lemma 6.4.1 yield

$$S(\underline{a}/r + \underline{\theta}) \ll |P|^{n+\varepsilon}(|P|^{-2} + |d\theta_1| + |P|^{-3}|d\theta_1|^{-1})^{n/8}.$$

Since $c_1 \neq 0$, the condition $\max\{\widehat{R}_i|dc_i^{\perp}|\} \geq |r|$ implies $|d| \gg |r||P|^{-1/3}$. We deduce that the contribution from $|r| = \widehat{Y}$ with $\widehat{Y} \leq |P|^{1-\eta}$ and $|\theta_1| = \widehat{\Theta}_1$ to $E_2(P)$ via (6.4.3) is

$$\begin{aligned}
&\ll \sum_{|r|=\widehat{Y}} \sum_{\substack{|d|\leq \widehat{Y} \\ |d|\gg \widehat{Y}|P|^{-1/3} \\ d|r}} \sideset{}{'}\sum_{\substack{|\underline{a}|<|r| \\ \underline{a}/r\in L(d(1,0))}} \int_{|\theta_1|=\widehat{\Theta}_1} \int_{|\theta_2|\ll |P|^{-1/3}\widehat{Y}^{-1}} S(\underline{a}/r+\underline{\theta})\mathrm{d}\underline{\theta} \\
&\ll |P|^{n-1/3+\varepsilon}\widehat{\Theta}_1 \sum_{|r|=\widehat{Y}} \sum_{\substack{|d|\leq \widehat{Y} \\ |d|\gg \widehat{Y}|P|^{-1/3} \\ d|r}} |d| \left(|P|^{-2} + |d|\widehat{\Theta}_1 + |P|^{-3}|d|^{-1}\widehat{\Theta}_1^{-1}\right)^{n/8} \\
&\ll |P|^{3n/4-5/3+\varepsilon}\widehat{Y} + |P|^{n-1/3+\varepsilon}\widehat{Y}^{2+n/8}\widehat{\Theta}_1^{1+n/8} + |P|^{2n/3-2/3+\varepsilon}\widehat{Y}(\widehat{\Theta}_1\widehat{Y})^{1-n/8} \\
&\ll |P|^{3n/4-2/3-\eta+\varepsilon} + |P|^{5n/6-2/3-\eta+\varepsilon} + |P|^{2n/3-2/3-\delta(1-n/8)+\varepsilon}\widehat{Y}
\end{aligned}$$

where we used again Lemma 6.6.3 to bound the number of $\underline{a}$'s, that $\widehat{\Theta}_1 \ll \widehat{Y}^{-1}|P|^{-4/3}$ and (2) to estimate $\widehat{Y}$ and $\widehat{\Theta}_1$. We have $3n/4-2/3 < n-5$ for $n \geq 18$, so that he first term is sufficiently small. Moreover, $5n/6-2/3-\eta < n-5$ as soon as $n \geq 26$, which is also satisfactory. Lastly, the third term above is

$$|P|^{2n/3-2/3-\delta(1-n/8)+\varepsilon}\widehat{Y} \ll |P|^{2n/3+1/3-\eta-\delta(1-n/8)+\varepsilon} = |P|^{n-5-\eta+\varepsilon},$$

where of course δ was chosen in such a way as to simplify the exponent above. Therefore, this contribution is also satisfactory. Since there are $O(|P|^{\varepsilon})$ choices for Y and Θ_1, we have thus shown that

$$E_2(P) \ll |P|^{n-5-\kappa} \tag{6.4.10}$$

for some $\kappa > 0$ if $n \geq 26$.

The goal for the remainder of this chapter is to establish the following result.

Proposition 6.4.2. *If $n \geq 26$ and $\mathrm{char}(K) > 3$, then*

$$N(P) = c|P|^{n-5} + O\left(|P|^{n-5-\delta'}\right)$$

for some $\delta' > 0$, where $c > 0$ if for every prime ϖ there exists $\boldsymbol{x}_\varpi \in \mathcal{O}_\varpi^n$ such that $F_1(\boldsymbol{x}_\varpi) = F_2(\boldsymbol{x}_\varpi) = 0$ and $|\boldsymbol{b} - \boldsymbol{x}_\varpi|_\varpi < |N|_\varpi$ and where the implied constant depends on F_1, F_2, $\boldsymbol{b}$ and N.

Once we have established Proposition 6.4.2, there is no difficulty in deducing the weak approximation property for X. The exact details do not merit repetition here and can for example be found in Section 7.1 of [34] in the case of cubic hypersurfaces. Theorem 6.1.1 is the special case $N = 1$, in which case a non-zero solution $\boldsymbol{x}_\varpi \in \mathcal{O}_\varpi^n$ to $F_1(\boldsymbol{x}_\varpi) = F_2(\boldsymbol{x}_\varpi) = 0$ for every prime ϖ is guaranteed by the Lang–Tsen theory [91, Theorem 3.6] and the homogeneity of F_1 and F_2 under the weaker assumption $n \geq 14$. Therefore, Theorems 6.1.1 and 6.1.2 are a consequence of Proposition 6.4.2. In the light of (6.4.9) and (6.4.10) it will be enough to show that

$$M(P) = c|P|^{n-5} + O\left(|P|^{n-5-\kappa'}\right) \quad \text{and} \quad E_1(P) \ll |P|^{n-5-\kappa''},$$

for some $\kappa', \kappa'' > 0$, where c satisfies the properties claimed in Proposition 6.4.2. This goal will ultimately be achieved in Section 6.8 and requires a thorough analysis of the exponential sums and oscillatory integrals that appear in (6.4.5). We carry out this investigation in the subsequent three sections.

6.5 Exponential integrals

To get control over $I_s(\underline{\theta}, \boldsymbol{v})$, we consider for $\boldsymbol{w} \in K_\infty^n$, $\underline{\gamma} \in K_\infty^2$ and $G_1, G_2 \in K_\infty[x_1, \dots, x_n]$ the following oscillatory integral

$$J_{\underline{G}}(\underline{\gamma}, \boldsymbol{w}) := \int_{\mathbb{T}^n} \psi\left(\underline{\gamma} \cdot \underline{G}(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x},$$

where we henceforth adopt the notation $\underline{\gamma} \cdot \underline{G}(\boldsymbol{x}) = \gamma_1 G_1(\boldsymbol{x}) + \gamma_2 G_2(\boldsymbol{x})$. The main ingredients to deal with the exponential integrals appearing in our work are [202, Lemma 2.1–2.2], which we recall here for our convenience.

Lemma 6.5.1. *We have $J_{\underline{G}}(\underline{\gamma}, \boldsymbol{w}) = 0$ if $|\boldsymbol{w}| > \max\{1, |\gamma_1| H_{G_1}, |\gamma_2| H_{G_2}\}$.*

Lemma 6.5.2. *Let $\Omega = \{\boldsymbol{x} \in \mathbb{T}^n \colon |\gamma_1 \nabla G_1(\boldsymbol{x}) + \gamma_2 \nabla G_2(\boldsymbol{x}) + \boldsymbol{w}| \leq H_{\underline{G}} \max\{1, |\underline{\gamma}|^{1/2}\}\}$. Then*

$$J_{\underline{G}}(\underline{\gamma}, \boldsymbol{w}) = \int_{\Omega} \psi\left(\underline{\gamma} \cdot \underline{G}(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x}.$$

To begin our treatment of $I_s(\underline{\theta}, \boldsymbol{v})$, note that by the definition of w in (6.4.2) we have

$$\begin{aligned} I_s(\underline{\theta}, \boldsymbol{v}) &= \widehat{L}^{-n} \psi\left(\frac{P\boldsymbol{v} \cdot \boldsymbol{x}_0}{s}\right) \int_{\mathbb{T}^n} \psi\left(P^3 \theta_1 G_1(\boldsymbol{y}) + P^2 \theta_2 G_2(\boldsymbol{y}) + \frac{P\boldsymbol{v} \cdot \boldsymbol{y}}{t^L s}\right) \mathrm{d}\underline{\theta} \\ &= \widehat{L}^{-n} \psi\left(\frac{P\boldsymbol{v} \cdot \boldsymbol{x}_0}{s}\right) J_{\underline{G}}(P^3\theta_1, P^2\theta_2, (P\boldsymbol{v} t^{-L}/s)), \end{aligned} \tag{6.5.1}$$

where $G_i(\boldsymbol{y}) = F_i(\boldsymbol{x}_0 + t^{-L}\boldsymbol{y})$ for $i = 1, 2$ and we applied the change of variables $\boldsymbol{y} = t^L(\boldsymbol{x} - \boldsymbol{x}_0)$. It is clear that G_i is a polynomial with coefficients in K_∞ and $H_{\underline{G}} \leq H_{\underline{F}}$. Therefore, it follows from Lemma 6.5.1 that

$$I_s(\underline{\theta}, \boldsymbol{v}) = 0 \quad \text{if} \quad |\boldsymbol{v}| > \widehat{L} H_{\underline{F}} |s| \frac{\max\{1, |P|^3|\theta_1|, |P|^2|\theta_2|\}}{|P|}. \tag{6.5.2}$$

Before we can derive upper bounds for $I_s(\underline{\theta}, v)$ from Lemma 6.5.2, we need a preliminary step.

Lemma 6.5.3. *Let $C \subset K_\infty^2$ be compact and bounded away from $\mathbf{0}$. If we define $A(\underline{\gamma}, \boldsymbol{x})$ to be the maximum of the absolute values of the $(n-2) \times (n-2)$-minors of $\gamma_1 H(\boldsymbol{x}) + \gamma_2 M$, then*

$$A(\underline{\gamma}, \boldsymbol{x}) \gg_{C, w, \underline{F}} 1$$

for $\underline{\gamma} \in C$ and $\boldsymbol{x} \in \operatorname{supp}(w)$.

Proof. Suppose by contradiction that the statement of the lemma is false. Then there exists a sequence $(\underline{\gamma}_k, \boldsymbol{x}_k) \in C \times \operatorname{supp}(w)$ such that $A(\underline{\gamma}_k, \boldsymbol{x}_k) \leq 1/k$ for all $k \geq 1$. Since $C \times \operatorname{supp}(w)$ is compact, we can pass to a convergent subsequence with limit $(\underline{\gamma}', \boldsymbol{x}') \in C \times \operatorname{supp}(w)$. However, since the map $(\underline{\gamma}, \boldsymbol{x}) \mapsto A(\underline{\gamma}, \boldsymbol{x})$ is continuous, this implies that every $(n-2) \times (n-2)$ minor of $\gamma_1' H(\boldsymbol{x}') + \gamma_2' M$ vanishes. Therefore, rk$(\gamma_1' H(\boldsymbol{x}') + \gamma_2' M) \leq n - 3$, which is a contradiction since any $\boldsymbol{x} \in \operatorname{supp}(w)$ satisfies the third condition in (6.4.1). □

When $\underline{\Theta} = (\Theta_1, \Theta_2) \in \mathbb{Z}^2$, then we shall henceforth adopt the convention that $|\underline{\gamma}| = \widehat{\underline{\Theta}}$ means $|\theta_1| = \widehat{\Theta}_1$ and $|\theta_2| = \widehat{\Theta}_2$. We finally have all the ingredients at hand to provide an upper bound for an average of $I_s(\underline{\theta}, \boldsymbol{v})$ over $\underline{\theta}$.

Proposition 6.5.4. *Let $\underline{\Theta} \in \mathbb{Z}^2$, $\boldsymbol{v} \in \mathcal{O}^n$, $s \in \mathcal{O}$ be monic and define the quantity $\widehat{Z} = \max\{1, |P|^3\widehat{\Theta}_1, |P|^2\widehat{\Theta}_2\}$. Then if $|\theta_2| \gg_{F_1,F_2,w} |P||\theta_1|$, we have*

$$\int_{|\underline{\theta}|=\widehat{\underline{\Theta}}} I_s(\underline{\theta}, \boldsymbol{v})\mathrm{d}\underline{\gamma} \ll_{F_1,F_2,w} \widehat{\Theta}_1\widehat{\Theta}_2\widehat{Z}^{-(n-1)/2}$$

while if $|\theta_2| \ll_{F_1,F_2,w} |P||\theta_1|$, then

$$\int_{|\underline{\theta}|=\widehat{\underline{\Theta}}} I_s(\underline{\theta}, \boldsymbol{v})\mathrm{d}\underline{\gamma} \ll_{F_1,F_2,w} \widehat{\Theta}_1\widehat{\Theta}_2\widehat{Z}^{-(n-2)/2}.$$

Proof. For the ease of notation, let us write $\boldsymbol{w} = P\boldsymbol{v}(st^L)^{-1}$ and $\gamma_i = t^{(4-i)P}\theta_i$ for $i = 1, 2$. If $\widehat{Z} = 1$, then we use the trivial estimate $I_s(\underline{\theta}, \boldsymbol{v}) \leq \widehat{L}^{-n}$ that is an immediate consequence of (6.5.1). We shall therefore assume $\widehat{Z} > 1$ from now on. It then follows from (6.5.1) and Lemma 6.5.2 after an obvious change of variables that

$$\begin{aligned}|I_s(\underline{\theta}, \boldsymbol{v})| &\leq \widehat{L}^{-n}\,\mathrm{vol}\{\boldsymbol{x} \in \mathbb{T}^n \colon |P^3\theta_1\nabla G_1(\boldsymbol{x}) + \theta_2P^2\nabla G_2(\boldsymbol{x}) + \boldsymbol{w}| < H_{\underline{G}}\widehat{Z}^{1/2}\} \\ &\leq \mathrm{vol}\{\boldsymbol{x} \in \mathbb{T}^n \colon |\boldsymbol{x} - \boldsymbol{x}_0| < \widehat{L}^{-1}, |\theta_1P^3\nabla F_1(\boldsymbol{x}) + \theta_2P^2\nabla F_2(\boldsymbol{x}) + \boldsymbol{w}| < H_{\underline{F}}\widehat{Z}^{1/2}\}.\end{aligned}$$

Now let us denote the last set whose measure we want to estimate by Ω and suppose $\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{x}' \in \Omega$. By definition of Ω, we must then have

$$|\gamma_1(\nabla F_1(\boldsymbol{x} + \boldsymbol{x}') - \nabla F_1(\boldsymbol{x})) + \gamma_2\nabla F_2(\boldsymbol{x}')| < H_{\underline{F}}\widehat{Z}^{1/2}. \tag{6.5.3}$$

We now distinguish between the relative sizes of γ_1 and γ_2. Firstly, suppose $|\gamma_2| \gg |\gamma_1|$, so that $\widehat{Z} = |\gamma_2|$. Since $\mathrm{rk}(M) \geq n-1$, there exists indices $1 \leq i, j \leq n$ such that the submatrix M' obtained from M by deleting the ith row and jth column has rank $n-1$. Let us now fix $a \in \mathbb{T}^n$ and consider the set Ω_a of $\boldsymbol{x} \in \Omega$ whose jth entry is a. Assume Ω_a is non-empty and $\boldsymbol{x}', \boldsymbol{x}' + \boldsymbol{x}$ are both in Ω_a. We shall now write $\boldsymbol{x}'_{\widehat{j}}$ for the vector obtained from $\boldsymbol{x}'$ by deleting the jth entry and similarly for $\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{x}'$. Note that the jth entry of $\boldsymbol{x}$ must be 0. In addition, H' denotes the submatrix of H after deleting the ith row and jth column. It then follows from (6.5.3) that

$$|(\gamma_1H'(\boldsymbol{x} + \boldsymbol{x}') + \gamma_2M')\boldsymbol{x}_{\widehat{j}}| \leq |(\gamma_1H(\boldsymbol{x} + \boldsymbol{x}') + \gamma_2M)\boldsymbol{x}| \ll \widehat{Z}^{1/2}. \tag{6.5.4}$$

Since $\mathrm{rk}\, M' = n-1$, we have $M'\boldsymbol{x}_{\widehat{j}} \gg |\boldsymbol{x}_{\widehat{j}}|$. In particular, the trivial estimate $H'(\boldsymbol{x} + \boldsymbol{x}')\boldsymbol{x}_{\widehat{j}} \ll |\boldsymbol{x}_{\widehat{j}}|$ together with the assumption $|\gamma_1| \ll |\gamma_2|$ implies that (6.5.4) can only hold if

$$|\gamma_2M'\boldsymbol{x}_{\widehat{j}}| \ll \widehat{Z}^{1/2}.$$

We can now multiply the left hand side by M'^{-1}, whose entries have absolute value $O(1)$, to deduce that $|\boldsymbol{x}_{\widehat{j}}| \ll \widehat{Z}^{-1/2}$ and thus

$$\int_{|\underline{\gamma}|=\widehat{\underline{\Theta}}} I_s(\underline{\theta}, \boldsymbol{v})\mathrm{d}\underline{\gamma} \ll \int_{|\underline{\gamma}|=\widehat{\underline{\Theta}}}\int_{\mathbb{T}^n} \mathrm{vol}(\Omega_a)\mathrm{d}a\mathrm{d}\underline{\gamma} \ll \widehat{\Theta}_1\widehat{\Theta}_2\widehat{Z}^{-(n-1)/2},$$

which is satisfactory.

We now treat the more complicated case when $|\gamma_1| \gg |\gamma_2|$, so that $\widehat{Z} = |\gamma_1|$. For $\bar{i} = \{i_1, i_2\}$, $\bar{j} = \{j_1, j_2\} \subset \{1, \ldots, n\}$ and a matrix $B \in \mathsf{Mat}_{n\times n}(K_\infty)$, we write $B_{\bar{i},\bar{j}}$

for the matrix obtained from B by deleting the i_1th and i_2th rows as well as the j_1th and j_2th columns. It follows from Lemma 6.5.3 that

$$A := \max_{\substack{\bar{i},\bar{j}\subset\{1,\dots,n\}\\ |\bar{i}|=|\bar{j}|=2}} |\det((\gamma_1 H(\boldsymbol{x}) + \gamma_2 M)_{\bar{i},\bar{j}}| \gg |\gamma_1|$$

for $\boldsymbol{x} \in \operatorname{supp}(w)$. Next we divide Ω into at most $(n(n-1)/2)^2$ subsets according to the indices at which the maximum above occurs, that is for $\bar{i}, \bar{j} \subset \{1, \dots, n\}$ we set

$$\Omega_{\bar{i},\bar{j}} := \{\boldsymbol{x} \in \Omega\colon A = |\det((\gamma_1 H(\boldsymbol{x}) + \gamma_2 M)_{\bar{i},\bar{j}}|\}.$$

Moreover, to estimate the measure of $\Omega_{\bar{i},\bar{j}}$ we shall again fix the j_1th and j_2th entries of $\boldsymbol{x}$ and denote by $\boldsymbol{x}_{\bar{j}}$ the vector obtained from $\boldsymbol{x}$ by deleting the j_1th and j_2th entries, so that

$$\operatorname{vol}(\Omega_{\bar{i},\bar{j}}) \leq \int_{\mathbb{T}^2} \operatorname{vol}\{\boldsymbol{x} \in \Omega_{\bar{i},\bar{j}}\colon x_{j_k} = a_k \text{ for } k = 1, 2\}\mathrm{d}\underline{a}.$$

If $\boldsymbol{x}', \boldsymbol{x} + \boldsymbol{x}'$ are both in $\Omega_{\bar{i},\bar{j}}$ and $x'_{j_k}, x_{j_k} + x'_{j_k} = a_k$ for $k = 1, 2$, then (6.5.3) implies that

$$|(H(\boldsymbol{x} + \boldsymbol{x}') + \gamma_2\gamma_1^{-1}M)_{\bar{i},\bar{j}}\boldsymbol{x}_{\bar{j}}| \leq |(H(\boldsymbol{x} + \boldsymbol{x}') + \gamma_2\gamma_1^{-1}M)\boldsymbol{x}| \ll \widehat{Z}^{-1/2}.$$

Since $\boldsymbol{x} + \boldsymbol{x}' \in \Omega_{\bar{i},\bar{j}}$, the entries of the inverse of $(H(\boldsymbol{x} + \boldsymbol{x}') + \gamma_2\gamma_1^{-1}M)_{\bar{i},\bar{j}}$ have absolute value $O(1)$. In particular, after multiplying the last equation above with it from the left we get that $|\boldsymbol{x}_{\bar{i},\bar{j}}| \ll \widehat{Z}^{-1/2}$. From what we have shown so far, it thus follows that

$$\begin{aligned}\int_{|\underline{\gamma}|=\widehat{\underline{\Theta}}} I_s(\underline{\theta}, \boldsymbol{v})\mathrm{d}\underline{\gamma} &\ll \sum_{\substack{\bar{i},\bar{j}\subset\{1,\dots,n\}\\ |\bar{i}|=|\bar{j}|=2}} \int_{|\underline{\gamma}|=\widehat{\underline{\Theta}}} \operatorname{vol}(\Omega_{\bar{i},\bar{j}})\mathrm{d}\underline{\gamma} \\ &\ll \widehat{Z}^{-(n-2)/2}\widehat{\Theta}_1\widehat{\Theta}_2,\end{aligned}$$

which completes the proof. □

6.6 Exponential sums: pointwise estimates

The aim of this section is to collect estimates for the complete exponential sums $S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})$ defined in (6.4.6). These sums enjoy a twisted multiplicativity property, which essentially reduces the task of estimating them to the case of prime power moduli. For $r, R \in \mathcal{O}$, we adopt the notation

$$r \mid R^\infty$$

to mean that every prime divisor of r also divides R.

Lemma 6.6.1. *Suppose $d \mid r$ and $r = r_1r_2$ with $(r_1, r_2) = 1$. If we write $N = N_1N_2N_3$, where $N_i \mid r_i^\infty$ for $i = 1, 2$ and $(r, N_3) = 1$, and let $s_i = r_iN_i/(r_i, N_i)$ for $i = 1, 2$, then there exist $\boldsymbol{b}' \in (\mathcal{O}/N_3\mathcal{O})^n$ and $t_i \in (\mathcal{O}/s_i\mathcal{O})^\times$ for $i = 1, 2$ such that*

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) = S_{d_1\underline{c},r_1,\boldsymbol{b},N_1}(t_1\boldsymbol{v})S_{d_2\underline{c},r_2,\boldsymbol{b},N_2}(t_2\boldsymbol{v})\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{b}'}{N_3}\right).$$

where $d = d_1d_2$ with $d_i \mid r_i$ for $i = 1, 2$.

Proof. By construction, s_1, s_2 and N_3 are pairwise coprime so that $r_N = s_1 s_2 N_3$. In particular, if $\boldsymbol{y}_i$ runs through a complete sets of residues modulo s_i for $i = 1, 2$ and $\boldsymbol{y}_3$ modulo N_3, then

$$\boldsymbol{x} = s_2 N_3 \boldsymbol{y}_1 + s_1 N_3 \boldsymbol{y}_2 + s_1 s_2 \boldsymbol{y}_3$$

constitutes a complete set of residues modulo r. Next, for $\underline{a}/r \in L(d\underline{c})$, we write $\underline{a} = r_2 \underline{a}_1 + r_1 \underline{a}_2$, where $|\underline{a}_i| < |r_i|$ and $(\underline{a}_i, r_i) = 1$. It is then clear that

$$\psi\left(\frac{\underline{a}\cdot \underline{F}(\boldsymbol{x})}{r}\right) = \psi\left(\frac{\underline{a}_1 \cdot \underline{F}(s_2 N_3 \boldsymbol{y}_1)}{r_1}\right)\psi\left(\frac{\underline{a}_2 \cdot \underline{F}(s_1 N_3 \boldsymbol{y}_2)}{r_2}\right)$$

and

$$\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{r_N}\right) = \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{y}_1}{s_1}\right)\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{y}_2}{s_2}\right)\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{y}_3}{N_3}\right).$$

Moreover, it is demonstrated in the proof of Lemma 5.2 in [202] that $\underline{a}/r \in L(d\underline{c})$ if and only if $\underline{a}_i/r_i \in L(d_i\underline{c})$ for both $i = 1, 2$. The result now follows after the change of variables $\boldsymbol{x}_1 = s_2 N_3 \boldsymbol{y}_1$ and $\boldsymbol{x}_2 = s_1 N_3 \boldsymbol{y}_2$ and taking $t_1 \equiv (s_2 N_3)^{-1}\ (s_1)$, $t_2 \equiv (s_1 N_3)^{-1}\ (s_2)$ and $\boldsymbol{b}' \equiv (s_1 s_2)^{-1}\boldsymbol{b}\ (s_3)$. □

In some cases we will obtain estimates for the sums $S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})$ by considering their relatives

$$T(\underline{a}, r, \boldsymbol{v}) := \sum_{|\boldsymbol{x}|<|r|} \psi\left(\frac{a_1 G_1(\boldsymbol{x}) + a_2 G_2(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x}}{r}\right) \tag{6.6.1}$$

for appropriate polynomials $G_1, G_2 \in \mathcal{O}[x_1, \dots, x_n]$. These sums satisfy the following twisted multiplicativity property.

Lemma 6.6.2. *Let $r = r_1 r_2$ with $(r_1, r_2) = 1$. Then*

$$T(\underline{a}, r, \boldsymbol{v}) = T(\underline{a}_{r_2}, r_1, \boldsymbol{v}) T(\underline{a}_{r_1}, r_2, \boldsymbol{v}),$$

where $\underline{a}_s := (s^2 a_1, s a_2)$ for $s \in \mathcal{O}$.

Proof. As $\boldsymbol{x}_i$ runs over a full set of residues mod r_i, $\boldsymbol{x} = r_2\boldsymbol{x}_1 + r_1\boldsymbol{x}_2$ runs over a full set of residues mod r. Moreover, using Taylor's formula it is easy to see that

$$\psi\left(\frac{a_i F_i(\boldsymbol{x})}{r}\right) = \psi\left(\frac{a_i r_2^{4-i} F_i(\boldsymbol{x}_1)}{r_1}\right)\psi\left(\frac{a_i r_1^{4-i} F_i(\boldsymbol{x}_2)}{r_2}\right)$$

for $i = 1, 2$ and

$$\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{r}\right) = \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}_1}{r_1}\right)\psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}_2}{r_2}\right),$$

from which the statement of the lemma follows. □

For our investigation we shall also need a good understanding of the distribution of rational points $\underline{a}/r$ on an individual line $L(d\underline{c})$ when r is fixed. By Lemma 6.6.1 it suffices to consider the case $r = \varpi^k$ and $d = \varpi^m$ with $m \le k$. The following lemma summarises the content of equations (6.9)–(6.11) of [202].

Lemma 6.6.3. *If $1 \leq m < k$, then modulo ϖ^k we have the following equality of sets*

$$\begin{aligned}\{\underline{a}\colon \underline{a}/\varpi^k \in L(\varpi^m \underline{c})\} =&\{a\underline{c}^\perp + \varpi^{k-m}\underline{d}\colon |a| < |\varpi|^{k-m}, (a,\varpi)=1, |\underline{d}| < |\varpi|^m\}\setminus \\ &\{a\underline{c}^\perp + \varpi^{k-m+1}\underline{d}\colon |a| < |\varpi|^{k-m+1}, |\underline{d}| < |\varpi|^{m-1}, (a,\varpi)=1\}\end{aligned}$$

and for $k = m$ we have

$$\begin{aligned}\{\underline{a}\colon \underline{a}/\varpi^k \in L(\varpi^k \underline{c})\} =&\{\underline{d}\colon (\underline{d},\varpi)=1, |\underline{d}| < |\varpi|^k\}\setminus \\ &\{a\underline{c}^\perp + \varpi\underline{d}\colon (a,\varpi)=1, |a| < |\varpi|, |\underline{d}| < |\varpi|^{k-1}\}.\end{aligned}$$

Moreover, when $m = 0$, then

$$\{\underline{a}\colon \underline{a}/\varpi^k \in L(\underline{c})\} = \{a\underline{c}^\perp\colon (a,\varpi)=1, |a| < |\varpi|^k\}.$$

In particular, we have $\#\{\underline{a}\colon \underline{a}/r \in L(d\underline{c})\} \leq |d||r|$.

6.6.1 Square-free moduli

We will now deal with $S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})$ when r is square-free. A key player in our estimates is the dual form $F_1^* \in \mathcal{O}[x_1,\dots,x_n]$, whose zero locus parameterises hyperplanes that have singular intersection with the projective hypersurface defined by F. It is well known [79, Example 4.4.3] that F_1^* is absolutely irreducible and of degree $3 \times 2^{n-2}$, providing $\mathrm{char}(K) > 3$. We begin our treatment by assuming that $d = 1$. In this case Lemma 6.6.3 tells us that $S_{\underline{c},\varpi,\boldsymbol{0},1}(\boldsymbol{v})$ equals the familiar exponential sum

$$S_\varpi(\boldsymbol{v}) := \sideset{}{'}\sum_{a\,(\varpi)} \sum_{\boldsymbol{x}\,(\varpi)} \psi\left(\frac{aF_{\underline{c}}(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right),$$

where $F_{\underline{c}}(\boldsymbol{x}) = -c_2F_1(\boldsymbol{x}) + c_1F_2(\boldsymbol{x})$. Let $\mathbb{F}_\varpi = \mathcal{O}/\varpi\mathcal{O}$ be the residue field of ϖ. Our main ingredient is the following special case of a result due to Katz [119, Theorem 4].

Theorem 6.6.4. *Let $X \subset \mathbb{P}^n_{\mathbb{F}_\varpi}$ be a complete intersection of dimension r defined by forms of degrees $d_1,\dots,d_{n-r}$ and let $L, H \in H^0(\mathbb{P}^n, O_{\mathbb{P}^n}(1))$. If*

(i) $X \cap L \cap H$ has dimension $r-2$,

(ii) the singular locus of $X \cap L$ has dimension ε and

(iii) the singular locus of $X \cap L \cap H$ has dimension $\delta \geq \varepsilon$,

then there exists a constant $C > 0$ depending only on $n, d_1,\dots,d_{n-r}$ such that for $f = H/L$ it holds that

$$\left|\sum_{\boldsymbol{x}\in X[1/L]} \psi\left(\frac{f(\boldsymbol{x})}{\varpi}\right)\right| \leq C|\varpi|^{(r+1+\delta)/2},$$

where $X[1/L]$ is the affine variety defined as the complement of the hyperplane cut out by L in X.

Remark. Katz states Theorem 4 for arbitrary closed subvarieties of projective space that are geometrically integral or equidimensional and Cohen-Macauly. However, our assumption that X is a complete intersection implies that X is Cohen-Macauly and equidimensional, thereby allowing us to state the simplified version above.

Suppose that $\varpi \nmid \boldsymbol{v}c_2$. Using orthogonality of characters we see that

$$S_\varpi(\boldsymbol{v}) = |\varpi| \sum_{\substack{\boldsymbol{x}\in\mathbb{F}_\varpi^n \\ F_{\underline{c}}(\boldsymbol{x})=0}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right).$$

Let $F(x_0, \boldsymbol{x}) \in \mathcal{O}[x_0, x_1, \dots, x_n]$ be the homogenization of $F_{\underline{c}}$, that is

$$F(x_0, \boldsymbol{x}) = -c_2F_1(\boldsymbol{x}) + x_0c_1F_2(\boldsymbol{x}),$$

and define $X = V(F) \subset \mathbb{P}^n_{\mathbb{F}_\varpi}$ to be the projective variety cut out by the reduction of F modulo ϖ. Note that the point $(1, 0, \dots, 0)$ will always be a singularity of X.

Moreover, we also set $L(x_0, \boldsymbol{x}) = x_0$ and $H(\boldsymbol{x}) = -\boldsymbol{v}\cdot\boldsymbol{x}$. In our situation we thus have $X \cap L = V(F_1)$ and $X \cap L \cap H = V(F_1, \boldsymbol{v}\cdot\boldsymbol{x})$. In particular, $\delta = \varepsilon = -1$ provided $\varpi \nmid \Delta_{F_1}F_1^*(\boldsymbol{v})$, where Δ_{F_1} is the discriminant of F_1. Indeed, the condition $\varpi \nmid \Delta_{F_1}$ guarantees that the reduction of F_1 modulo ϖ is non-singular and $\varpi \nmid F_1^*(\boldsymbol{v})$ implies that $V(F_1, \boldsymbol{v}\cdot\boldsymbol{x}) \subset \mathbb{P}^{n-1}_{\mathbb{F}_\varpi}$ is non-singular.

We can thus apply Theorem 6.6.4 with $\delta = -1$ and $r = n-1$ to deduce that

$$|\varpi|^{-1}|S_\varpi(\boldsymbol{v})| \leq C|\varpi|^{(n-1)/2},$$

where C is a constant that only depends on the degrees of F_1 and F_2 and n. Absorbing the primes $\varpi \mid \Delta_{F_1}$ into the constant and invoking Lemma 6.6.1, we have thus established the following result.

Lemma 6.6.5. *Suppose that $r \in \mathcal{O}$ is square-free with $(r, F_1^*(\boldsymbol{v})c_2) = 1$. There exists a constant $C > 0$ depending only on $\Delta_{F_1}, \deg F_1, \deg F_2$ and n such that*

$$|S_{\underline{c},r,\mathbf{0},1}(\boldsymbol{v})| \leq C^{\omega(r)}|r|^{(n+1)/2},$$

where $\omega(r)$ denotes the number of prime divisors of r

Let us now turn to the case $d \neq 1$. We shall use the following estimate of Deligne [66, Théorème 8.4], which states that for a polynomial $F \in \mathbb{F}_\varpi[x_1, \dots, x_n]$ of degree d with $\mathrm{char}(\mathbb{F}_\varpi) \nmid d$ such that the highest degree part cuts out a smooth projective hypersurface in $\mathbb{P}^{n-1}_{\mathbb{F}_\varpi}$, one has

$$\left|\sum_{\boldsymbol{x}\in\mathbb{F}_\varpi^n} \psi\left(\frac{F(\boldsymbol{x})}{\varpi}\right)\right| \leq (d-1)^n|\varpi|^{n/2}. \tag{6.6.2}$$

Recalling the definition of the sum T in (6.6.1) with $G_i = F_i$ for $i = 1, 2$, the estimate (6.6.2) implies

$$|T(\underline{a}, \varpi, \boldsymbol{v})| \leq 2^n|\varpi|^{n/2}$$

whenever $\varpi \nmid a_1\Delta_{F_1}$. On the other hand, if $\varpi \mid a_1$, then $\varpi \nmid a_2$ and then

$$|T(\underline{a}, \varpi, \boldsymbol{v})| \leq |\varpi|^{(n+1)/2}, \tag{6.6.3}$$

provided F_2 is a quadratic form of rank at least $n-1$ modulo ϖ, as for example follows from [202, Lemma 3.5]. Now let us assume $r \in \mathcal{O}$ is square-free and write $r = r_1r_2$, with $(r_1, r_2) = 1$ and $r_2 \mid N^\infty$. If $d = d_1d_2$ with $d_1 \mid r_1$ and $d_2 \mid r_2$, then by Lemma 6.6.1 we have

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) = S_{d_1\underline{c},r_1,\mathbf{0},1}(t_1\boldsymbol{v})S_{d_2\underline{c},r_2,\boldsymbol{b},N}(t_2\boldsymbol{v})$$

for some $t_i \in (\mathcal{O}/r_i\mathcal{O})^\times$. After absorbing the primes $\varpi \mid \Delta_{F_1}$ or for which the reduction of F_2 has rank strictly less than $n-1$ into the constant, it now follows from the estimates we just recorded and Lemma 6.6.3 that

$$S_{d_1\underline{c},r_1,\mathbf{0},1}(t_1\boldsymbol{v}) = \sum_{\underline{a}/r \in L(d_1\underline{c})} T(\underline{a}, r_1, t_1\boldsymbol{v}) \leq C^{\omega(r_1)}|d_1||r_1|^{(n+3)/2},$$

for some constant $C > 0$.

We can now estimate $S_{d_2\underline{c},r_2,\boldsymbol{b},M}(t_2\boldsymbol{v})$ trivially to arrive at the following result.

Lemma 6.6.6. *Suppose r is square-free and $d \mid r$. Then there exists a constant $C > 0$ depending only on $\deg F_1, F_2, \Delta_{F_1}$ and N such that*

$$|S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})| \leq C^{\omega(r)}|d||r|^{(n+3)/2}.$$

Moreover, if $(c_2, r) = 1$, then

$$|S_{\underline{c},r,\mathbf{0},1}(\boldsymbol{v})| \leq C^{\omega(r)}|r|^{n/2+1}.$$

6.6.2 Square-full moduli

To satisfactorily deal with square-full moduli, we begin with the case $r = \varpi^2$. Our main ingredient is the following result due to Heath-Brown [99, p. 395]. It should be noted that Heath-Brown's proves his result solely over the integers. However, it is a routine exercise and the required adaptations are minor to check that his argument holds over $\mathbb{F}_q[t]$ as well. Let $F \in \mathcal{O}[x_1, \ldots, x_n]$ be a polynomial of degree d with $\mathrm{char}(\mathbb{F}_q) > d$ and suppose the reduction of the top degree part of F defines a smooth projective hypersurface modulo ϖ. Then it holds that

$$\left| \sum_{\boldsymbol{x}\ (\varpi^2)} \psi\left(\frac{F(\boldsymbol{x})}{\varpi^2}\right) \right| \leq (d-1)^n |\varpi|^n. \tag{6.6.4}$$

After absorbing the contribution from the primes dividing Δ_{F_1} into the constant and employing Lemma 6.6.3, we arrive at the following estimate.

Lemma 6.6.7. *If $\varpi \nmid a_1$, then*

$$T(\underline{a}, \varpi^2, \boldsymbol{v}) \ll_{\Delta_{F_1}} |\varpi|^n.$$

In particular, we also have

$$S_{\underline{c},\varpi^2,\mathbf{0},1}(\boldsymbol{v}) \ll_{\Delta_{F_1}} |\varpi|^{2+n}$$

provided $\varpi \nmid c_2$.

Since F_2 is a quadratic form, there exists a matrix $G \in \mathrm{GL}_n(K)$ with entries in $\mathcal{O}$ such that after the change of variables $\boldsymbol{y} = G\boldsymbol{x}$ one has

$$F_2(\boldsymbol{y}) = \sum_{i=1}^n b_i y_i^2 \quad \text{with } b_2 \ldots b_n \neq 0.$$

If $\varpi \mid \det(G)$, we may still locally diagonalise F_2 with a matrix $G_\varpi \in \mathrm{GL}_n(K_\varpi)$ that has coefficients in $\mathcal{O}_\varpi$, so that after the change of variables $\boldsymbol{y} = G_\varpi\boldsymbol{x}$ we have

$$F_2(\boldsymbol{y}) = \sum_{i=1}^n b_{\varpi,i} y_i^2 \quad \text{with } b_{\varpi,2} \cdots b_{\varpi,n} \neq 0.$$

Let us define

$$\Delta_{F_2} = \begin{cases} b_1 \cdots b_n \prod_{\varpi | \det(G)} \varpi^{\nu_\varpi(b_{\varpi,1}\cdots b_{\varpi,n})} & \text{if } \mathrm{rk}(M) = n, \\ b_2 \cdots b_n \prod_{\varpi | \det(G)} \varpi^{\nu_\varpi(b_{\varpi,2}\cdots b_{\varpi,n})} & \text{if } \mathrm{rk}(M) = n-1. \end{cases}$$

and

$$v_\varpi = \nu_\varpi(\Delta_{F_2}). \tag{6.6.5}$$

We then have the following result [37, Lemma 6.2].

Lemma 6.6.8. *Let $k \geq 2$ and suppose $\varpi^{1+v_\varpi} \mid a_1$. Then for any $\boldsymbol{v} \in \mathcal{O}^n$ we have*

$$T(\underline{a}, \varpi^k, \boldsymbol{v}) \ll_{F_2} |\varpi|^{(n+1)/2}.$$

Lemma 6.2 in [37] is again only proved for the analogous sum over the integers. Moreover, they assume that F_2 is diagonal from the beginning, a difference we take care of by the diagonalisation process above. The proof goes through verbatim in our setting, and so we shall not repeat it here.

Corollary 6.6.9. *Let $r \in \mathcal{O}$ be such that $\varpi \mid r$ implies $\varpi^2 \parallel r$. Then there exists a constant $C > 0$ depending on F_1, F_2 and N such that*

$$|S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})| \leq C^{\omega(r)}|d||r|^{(n+3)/2}.$$

Proof. Write $r = r_1 r_2$ with coprime $r_1, r_2 \in \mathcal{O}$ such that $(r_1, N) = 1$ and $r_2 \mid N^\infty$. As in the proof of Lemma 6.6.6 it suffices to obtain and upper bound for the sum $S_{d_1\underline{c},r_1,\boldsymbol{0},1}(t_1\boldsymbol{v})$, where $t_1 \in (\mathcal{O}/r_1\mathcal{O})^\times$, and estimate the sum corresponding to r_2 trivially. By definition we then have

$$S_{d_1\underline{c},r_1,\boldsymbol{0},1}(t_1\boldsymbol{v}) = \sum_{\underline{a}/r \in L(d_1\underline{c})} T(\underline{a}, r_1, t_1\boldsymbol{v}).$$

Using the mulitplicativity property recorded in Lemma 6.6.2, we can now invoke Lemma 6.6.7 in conjunction with Lemma 6.6.8 to obtain $|T(\underline{a}, r_1, t_1\boldsymbol{v})| \leq C^{\omega(r_1)}|r_1|^{(n+1)/2}$. Lemma 6.6.3 provides us with an upper bound for the number of $\underline{a}$'s such that $\underline{a}/r_1 \in L(d_1\underline{c})$ that completes the proof. □

6.6.3 The case $c_2 = 0$.

We now consider separately the case $c_2 = 0$. This can only occur if $c_1 = 1$ and so to ease notation, we write $\underline{c}_0 := (1, 0)$. By Lemma 6.6.1 we can reduce to the case when $r = \varpi^k$ and $d = \varpi^m$ with $m \leq k$ and we again begin our treatment assuming that $r = \varpi$. When $d = 1$, Lemma 6.6.6 already provides sufficiently good upper bounds. However, when $d = \varpi$, we have to do better and establishing an estimate that is superior to Lemma 6.6.6 is our first goal.

For $k \geq 1$, let us define

$\rho_1(\varpi^k) = \#\{\boldsymbol{x}\ (\varpi) \colon F_1(\boldsymbol{x}) \equiv F_2(\boldsymbol{x}) \equiv 0\ (\varpi^k)\}$ and $\rho_2(\varpi^k) = \#\{\boldsymbol{x}\ (\varpi) \colon F_2(\boldsymbol{x}) \equiv 0\ (\varpi^k)\}$.

By Lemma 6.6.3 we have

$$\begin{aligned} S_{\varpi\underline{c}_0,\varpi,\boldsymbol{0},1}(\boldsymbol{v}) &= \sum_{a_1\ (\varpi)}{}' \sum_{a_2\ (\varpi)} \sum_{\boldsymbol{x}\ (\varpi)} \psi\left(\frac{a_1F_1(\boldsymbol{x}) + a_2F_2(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right) \\ &= |\varpi|\left(|\varpi| \sum_{\substack{\boldsymbol{x}\ (\varpi) \\ F_1(\boldsymbol{x}) \equiv F_2(\boldsymbol{x}) \equiv 0\ (\varpi)}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right) - \sum_{\substack{\boldsymbol{x}\ (\varpi) \\ F_2(\boldsymbol{x}) \equiv 0\ (\varpi)}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right)\right). \end{aligned}$$

If $\varpi \mid \boldsymbol{v}$, then the expression above simplifies to

$$S_{\varpi\underline{c}_0,\varpi,\mathbf{0},1}(\boldsymbol{v}) = |\varpi|(|\varpi|\rho_1(\varpi) - \rho_2(\varpi)).$$

Since the reduction of X modulo ϖ is non-singular for $|\varpi|$ sufficiently large, we have

$$\rho_1(\varpi) = |\varpi|^{n-2} + O\left(|\varpi|^{(n-1)/2}\right),$$

as follows for example from Equation (3.12) in [34]. Moreover, because F_2 has rank at least $n-1$, it holds that

$$\rho_2(\varpi) = |\varpi|^{n-1} + O\left(|\varpi|^{(n+1)/2}\right).$$

So that in total we have

$$S_{\varpi\underline{c}_0,\varpi,\mathbf{0},1}(\boldsymbol{v}) \ll |\varpi|^{(n+3)/2} \tag{6.6.6}$$

whenever $\varpi \mid \boldsymbol{v}$, where the implied constant only depends on F_1 and F_2. Let us now deal with the opposite case $\varpi \nmid \boldsymbol{v}$. We want to apply Theorem 6.6.4 to our situation. For this, we define $X'_\varpi = V(F_1, F_2) \subset \mathbb{P}^n = \mathrm{Proj}(\mathbb{F}_\varpi[x_0, x_1, \dots, x_n])$. In addition, we set $L(x_0, \boldsymbol{x}) = x_0$ and $H(x_0, \boldsymbol{x}) = -\boldsymbol{v}\cdot\boldsymbol{x}$. Provided $|\varpi|$ is sufficiently large, X'_ϖ is a complete intersection of dimension $n-2$ with the only singularity at $(1:0:\dots:0) \in \mathbb{P}^n$. Moreover, we have $X'_\varpi \cap L = V(F_1, F_2) \subset \mathbb{P}^{n-1}$, which is non-singular. It follows from a result of Zak and Fulton–Lazarsfeld [80, Remark 7.5] that $X'_\varpi \cap H \cap L = V(F_1, F_1) \cap H \subset \mathbb{P}^{n-1}$ has at worst isolated singularities, so that in the notation of Theorem 6.6.4 we have $\varepsilon = -1$ and $\delta \leq 0$. In particular,

$$\sum_{\substack{\boldsymbol{x}\ (\varpi) \\ F_1(\boldsymbol{x}) \equiv F_2(\boldsymbol{x}) \equiv 0\ (\varpi)}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right) \ll |\varpi|^{(n-1)/2}.$$

Combining this with (6.6.3), we infer

$$\begin{aligned} S_{\varpi\underline{c}_0,\varpi,\mathbf{0},1}(\varpi) &= |\varpi|^2 \sum_{\substack{\boldsymbol{x}\ (\varpi) \\ F_1(\boldsymbol{x}) \equiv F_2(\boldsymbol{x}) \equiv 0\ (\varpi)}} \psi\left(\frac{-\boldsymbol{v}\cdot x}{\varpi}\right) - \sum_{a_2\ (\varpi)} \sum_{\boldsymbol{x}\ (\varpi)} \psi\left(\frac{a_2 F_2(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x}}{\varpi}\right) \\ &\ll |\varpi|^{(n+3)/2} + |\varpi|^{(n+3)/2}. \end{aligned}$$

Estimating the contribution from the primes $\varpi \mid N$ trivially, and using Lemma 6.6.1, it thus follows from (6.6.6) that

$$|S_{d\underline{c}_0,d,\boldsymbol{b},N}(\boldsymbol{v})| \leq C^{\omega(d)}|d|^{(n+3)/2} \tag{6.6.7}$$

for some constant $C > 0$ that only depends on F_1, F_2 and N.

We also require strong upper bounds for the sums

$$S_1 = S_{\varpi\underline{c}_0,\varpi^2,\mathbf{0},1}(\boldsymbol{v}) \quad \text{and} \quad S_2 = S_{\varpi^2\underline{c}_0,\varpi^2,\mathbf{0},1}(\boldsymbol{v}).$$

Let us begin with the former. In this case Lemma 6.6.3 implies

$$\begin{aligned}
S_1 &= \sideset{}{'}\sum_{|a|<|\varpi|} \sum_{|\underline{d}|<|\varpi|} \sum_{|\boldsymbol{x}|<|\varpi|^2} \psi\left(\frac{(a+\varpi d_2)F_2(\boldsymbol{x})+\varpi d_1 F_1(\boldsymbol{x})-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2}\right)\\
&= |\varpi| \sideset{}{'}\sum_{|a|<|\varpi|^2} \sum_{\substack{|\boldsymbol{x}|<|\varpi|^2\\ F_1(\boldsymbol{x})\equiv 0\,(\varpi)}} \psi\left(\frac{aF_2(\boldsymbol{x})-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2}\right)\\
&= |\varpi|^3\left(\sum_{\substack{|\boldsymbol{x}|<|\varpi|^2\\ F_1(\boldsymbol{x})\equiv 0\,(\varpi)\\ F_2(\boldsymbol{x})\equiv 0\,(\varpi^2)}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2}\right) - |\varpi|^{-1} \sum_{\substack{|\boldsymbol{x}|<|\varpi|^2\\ F_1(\boldsymbol{x})\equiv 0\,(\varpi)\\ F_2(\boldsymbol{x})\equiv 0\,(\varpi)}} \psi\left(\frac{-\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2}\right)\right)\\
&= |\varpi|^3\left(\Sigma_1 - |\varpi|^{-1}\Sigma_2\right)
\end{aligned}$$

say. The conditions $F_1(\boldsymbol{x})\equiv 0\,(\varpi)$ and $F_2(\boldsymbol{x})\equiv 0\,(\varpi^2)$ are invariant under scaling $\boldsymbol{x}$ by any b with $(b,\varpi)=1$ and so we deduce that that

$$\begin{aligned}
\Sigma_1 &= \frac{1}{|\varpi|^2(1-|\varpi|^{-1})} \sum_{\substack{|\boldsymbol{x}|<|\varpi|^2\\ F_1(\boldsymbol{x})\equiv 0\,(\varpi)\\ F_2(\boldsymbol{x})\equiv 0\,(\varpi^2)}} \sideset{}{'}\sum_{|b|<|\varpi|^2} \psi\left(\frac{b\boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2}\right)\\
&= \frac{1}{1-|\varpi|^{-1}} \sum_{\substack{\boldsymbol{x}\,(\varpi)\\ F_1(\boldsymbol{x})\equiv F_2(\boldsymbol{x}\equiv 0\,(\varpi)\\ \boldsymbol{v}\cdot\boldsymbol{x}\equiv 0\,(\varpi)}} (\rho_1(\boldsymbol{x})-|\varpi|^{-1}\rho_2(\boldsymbol{x})),
\end{aligned}$$

where

$$\rho_1(\boldsymbol{x}) := \#\{\boldsymbol{y}\,(\varpi^2)\colon \boldsymbol{y}\equiv\boldsymbol{x}\,(\varpi), F_2(\boldsymbol{x})\equiv\boldsymbol{v}\cdot\boldsymbol{x}\equiv 0\,(\varpi^2)\}$$

and

$$\rho_2(\boldsymbol{x}) := \#\{\boldsymbol{y}\,(\varpi^2)\colon \boldsymbol{y}\equiv\boldsymbol{x}\,(\varpi), F_2(\boldsymbol{x})\equiv 0\,(\varpi^2)\}.$$

Running the exact argument again yields

$$\Sigma_2 = \frac{1}{1-|\varpi|^{-1}} \sum_{\substack{\boldsymbol{x}\,(\varpi)\\ F_1(\boldsymbol{x})\equiv F_2(\boldsymbol{x})\equiv 0\,(\varpi)\\ \boldsymbol{v}\cdot\boldsymbol{x}\equiv 0\,(\varpi)}} (\rho_1'(\boldsymbol{x})-|\varpi|^{n-1}),$$

where

$$\rho_1'(\boldsymbol{x}) := \#\{\boldsymbol{y}\,(\varpi^2)\colon \boldsymbol{y}\equiv\boldsymbol{x}\,(\varpi), \boldsymbol{v}\cdot\boldsymbol{x}\equiv 0\,(\varpi^2)\}.$$

Suppose that $\boldsymbol{y}=\boldsymbol{x}+\varpi\boldsymbol{z}$. Then $\boldsymbol{y}$ is counted by $\rho_1(\boldsymbol{x})$ if and only if $\varpi \mid (\boldsymbol{x}^t M\boldsymbol{y}+\varpi^{-1}F_2(\boldsymbol{x}))$ and $\varpi \mid (\boldsymbol{v}\cdot\boldsymbol{z}+\varpi^{-1}\boldsymbol{v}\cdot\boldsymbol{x})$. Similarly, $\boldsymbol{y}$ is counted by $\rho_2(\boldsymbol{x})$ if and only if ϖ divides $(\boldsymbol{x}^t M\boldsymbol{y}+\varpi^{-1}F_2(\boldsymbol{x}))$. In particular, we see that $\rho_1(\boldsymbol{x})-|\varpi|^{-1}\rho_2(\boldsymbol{x})=0$ unless

$$\operatorname{rk}\begin{pmatrix}\boldsymbol{v}\\ M\boldsymbol{x}\end{pmatrix} = \operatorname{rk}(M\boldsymbol{x}) \mod \varpi. \tag{6.6.8}$$

Note that for $\boldsymbol{x}\not\equiv 0\,(\varpi)$ and $|\varpi|$ sufficiently large this can only happen if $\boldsymbol{v}$ and $M\boldsymbol{x}$ are proportional, since then the non-singularity of X implies that $M\boldsymbol{x}\not\equiv 0\,(\varpi)$. In particular, if (6.6.8) holds and $\boldsymbol{x}\not\equiv 0\,(\varpi)$, then $\rho_1(\boldsymbol{x})-|\varpi|^{-1}\rho_2(\boldsymbol{x})=|\varpi|^{n-1}-|\varpi|^{n-2}$, while if $\boldsymbol{v}\equiv 0\,(\varpi)$, then $\rho_1(\boldsymbol{0})-|\varpi|^{-1}\rho_2(\boldsymbol{0})=|\varpi|^n-|\varpi|^{n-1}$. Moreover, we have $\rho_1'(\boldsymbol{x})=|\varpi|^{n-1}$ unless $\varpi\mid\boldsymbol{v}$, in which case $\rho_1(\boldsymbol{x})=|\varpi|^n$. In total, we thus have

$$S_1 = |\varpi|^3 \left(\frac{|\varpi|^{n-1}}{(1-|\varpi|^{-1})^2}\mathcal{N}_1 - \frac{|\varpi|^{n-1}}{(1-|\varpi|^{-1})^2}\mathcal{N}_2 \right) + O\left(|\varpi|^{n+3}\right).$$

The error term takes care of the contribution from $\boldsymbol{x} \equiv \mathbf{0}\ (\varpi)$ in Σ_1 and Σ_2, and where we have defined

$$\mathcal{N}_1 := \#\{\in \mathbb{F}_\varpi^n \setminus \{\mathbf{0}\} \colon \boldsymbol{v}\cdot\boldsymbol{x} = F_1(\boldsymbol{x}) = F_2(\boldsymbol{x}) = 0, \text{(6.6.8) holds}\}$$

and

$$\mathcal{N}_2 := \#\{\boldsymbol{x} \in \mathbb{F}_\varpi^n \setminus \{\mathbf{0}\} \colon F_1(\boldsymbol{x}) = F_2(\boldsymbol{x}) = 0, \boldsymbol{v} \equiv \mathbf{0}\ (\varpi)\}.$$

Let us first deal with the case $\varpi \mid \boldsymbol{v}$. It follows that $\mathcal{N}_1 = \mathcal{N}_2$ and thus

$$S_1 = O(|\varpi|^{n+3}).$$

Suppose next that $\varpi \nmid \boldsymbol{v}$. In this case $\mathcal{N}_2 = 0$ and (6.6.8) holds if and only if $M\boldsymbol{x}$ and $\boldsymbol{v}$ are proportional. Since M has rank at least $n-1$, this can happen for at most $O(|\varpi|^2)$ choices of $\boldsymbol{x}$, so that

$$S_1 \ll |\varpi|^{n+2}\mathcal{N}_1 + |\varpi|^{n+3} \ll |\varpi|^{n+4}.$$

In total we have therefore established that

$$S_1 \ll |\varpi|^{n+4}. \tag{6.6.9}$$

Let us now turn to the sum S_2. It follows from Lemma 6.6.3 that

$$S_2 \leq \sideset{}{'}\sum_{a_1\ (\varpi^2)} \sum_{a_2\ (\varpi^2)} \left| \sum_{\boldsymbol{x}\ (\varpi^2)} \psi\left(\frac{a_1F_1(\boldsymbol{x}) + a_2F_2(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x}}{\varpi^2} \right) \right|.$$

Since $(a_1, \varpi) = 1$, we can apply Lemma 6.6.7 to deduce that the sum over $\boldsymbol{x}$ is $O(|\varpi|^n)$ and hence

$$S_2 \ll |\varpi|^{n+4}, \tag{6.6.10}$$

which completes our treatment of S_1 and S_2.

Finally, when $\varpi \nmid \Delta_{F_2}$ and $0 \leq m < k$, we can invoke Lemma 6.6.8 to deduce that

$$S_{\varpi^m \underline{c}_0, \varpi^k, \mathbf{0}, 1}(\boldsymbol{v}) \ll |\varpi|^{m+k(n+3)/2}. \tag{6.6.11}$$

Using Lemma 6.6.1 and estimating the contribution from N trivially, we see that the following result summarises the content of (6.6.7) and (6.6.9)–(6.6.11).

Proposition 6.6.10. *Let $d_1, d_2, d_3 \in \mathcal{O}$ be square-free. Then there exists a constant $C > 0$ depending on F_1, F_2 and N such that*

$$|S_{d_1d_2d_3^2\underline{c}_0, d_1d_2^2d_3^2, \boldsymbol{b}, N}(\boldsymbol{v})| \leq C^{\omega(d_1d_2d_3)}|d_1|^{(n+3)/2}|d_2d_3|^{n+4}.$$

In addition, let $d, r \in \mathcal{O}$ be both monic such that $d \mid r$ and $(r, \Delta_{F_2}) = 1$. If $\nu_\varpi(d) < \nu_\varpi(r)$ for all $\varpi \mid d$, then

$$|S_{d\underline{c}_0, r, \mathbf{0}, 1}(\boldsymbol{v})| \leq C^{\omega(r)}|d||r|^{(n+3)/2}.$$

6.7 Exponential sums: averages

We also need to deal with certain averages over exponential sums. Our first ingredient is the following result, which we apply to our situation in the corollary directly afterwards.

Lemma 6.7.1. *Let $\boldsymbol{v}_0 \in K_\infty^n$ and $V \geq 1$. If r is cube-full and $\underline{a} \in \mathcal{O}^2$ is such that $|(a_1, r)| \ll 1$, then*

$$\sum_{\substack{\boldsymbol{v} \in \mathcal{O}^n \\ |\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}}} |T(\underline{a}, r, \boldsymbol{v})| \ll_{F_1} |r|^{n/2+\varepsilon} \left(\widehat{V}^n + |r|^{n/3}\right).$$

The Lemma we just stated follows from equation (6.9) in [34]. There only the case when $(a_1, r) = 1$ is considered. However, as explained in the paragraph after Lemma 6.4 in [37], the argument leading to the estimate continues to hold when $|(a_1, r)| \ll 1$ after employing some minor modifications.

Corollary 6.7.2. *Let $\boldsymbol{v}_0 \in K_\infty^n$ and $V \geq 1$. Suppose r is cube-full such that $d \mid r$ and define*

$$P_1(r) := \{\varpi^k \parallel r \colon (\varpi, \Delta_{F_2} c_2) = 1, \varpi^k \nmid d\} \text{ and } P_2(r) = \{\varpi^k \parallel r \colon (\varpi, \Delta_{F_2}) = 1, \varpi^k \mid d\}.$$

If we write $r = r_1 r_2$, where

$$r_1 = \begin{cases} \prod_{\varpi^k \in P_1(r)} \varpi^k & \text{if } c_2 \neq 0, \\ \prod_{\varpi^k \in P_2(r)} \varpi^k & \text{else,} \end{cases} \tag{6.7.1}$$

and $r_2 = r/r_1$, then

$$\sum_{\substack{\boldsymbol{v} \in \mathcal{O}^n \\ |\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}}} |S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})| \ll_{F_1,F_2,N} |d||r|^{n/2+1+\varepsilon}|r_2|^{1/2} \left(\widehat{V}^n + |r|^{n/3}\right).$$

Note that since $d \mid r$, the condition $\varpi^k \nmid d$ in the definition of $P_1(r)$ means that every prime factor of d that divides r_1 in fact properly divides r_1 when $c_2 \neq 0$.

Proof. Denote the sum to be estimated by S. After making the change of variables $\boldsymbol{x} = \boldsymbol{y}N + \boldsymbol{b}$, we obtain the identity

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) = \psi\left(\frac{-\boldsymbol{b}\cdot\boldsymbol{v}}{r_N}\right) \sum_{\underline{a}/r \in L(d\underline{c})} T(\underline{a}, r/(r,N), \boldsymbol{v})$$

with underlying polynomials $G_i(\boldsymbol{y}) = (r,N)^{-1}F_i(N\boldsymbol{y}+\boldsymbol{b})$ for $i = 1, 2$ in the definition (6.6.1). Since $N \mid F_i(\boldsymbol{b})$, it follows that G_i has coefficients in $\mathcal{O}$. Moreover, the cubic part of G_1 is given by the non-singular polynomial $g_0(\boldsymbol{y}) = (r,N)^{-1}N^3F_1(\boldsymbol{y})$. We now factor $r/(r,N)$ into its cube-free part t and its cube-full part s. Since r is cube-full, we must have $|t| \leq |N|$. Using Lemma 6.6.2 and estimating the contribution from the sum corresponding to t trivially, we see that

$$|S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})| \leq |N|^n \sum_{\underline{a}/r \in L(d\underline{c})} |T(\underline{a}_t, s, \boldsymbol{v})|. \tag{6.7.2}$$

Next we write $s = s_1 s_2$, where

$$s_1 = \prod_{\substack{\varpi^k \| s \\ \varpi^{1+v_\varpi} \nmid a_1}} \varpi^k \quad \text{and} \quad s_2 = \prod_{\substack{\varpi^k \| s \\ \varpi^{1+v_\varpi} | a_1}} \varpi^k,$$

with v_ϖ defined in (6.6.5). It then follows from Lemma 6.6.8 that $T(\underline{a}_{ts_1}, s_2, \boldsymbol{v}) \ll |s_2|^{(n+1)/2+\varepsilon}$. Therefore, after applying Lemma 6.6.2 and Lemma 6.7.1 together with the identity (6.7.2), we obtain

$$\begin{aligned} S &\ll |r|^\varepsilon \sum_{\underline{a}/r \in L(d\underline{c})} |s_1|^{n/2} |s_2|^{(n+1)/2} \left(\widehat{V}^n + |s_1|^{n/3} \right) \\ &\ll |r|^{n/2+\varepsilon} \sum_{\underline{a}/r \in L(d\underline{c})} |s_2|^{1/2} \left(\widehat{V}^n + |r|^{n/3} \right). \end{aligned}$$

Let us write $r = r_1 r_2$ as in the statement of the lemma and $d = d_1 d_2$ with $d_i \mid r_i^\infty$ for $i = 1, 2$. The explicit description of $L(d\underline{c})$ in Lemma 6.6.3 implies that if $\underline{b}/r_1 \in L(d_1\underline{c})$, then $(b_1, r_1) = 1$. It is shown in the proof of [202, Lemma 5.2] that if $|\underline{a}| < |r|$ and $\underline{a} = r_2 \underline{b} + r_1 \underline{b}'$ with $|\underline{b}| < |r_1|, |\underline{b}'| < |r_2|$, then $\underline{a}/r \in L(d\underline{c})$ if and only if $\underline{b}/r_1 \in L(d_1\underline{c})$ and $\underline{b}'/r_2 \in L(d_2\underline{c})$. In particular, we must have $|s_2| \le |r_2|$. Thus it follows from Lemma 6.6.3 that

$$S \ll |d||r|^{n/2+1+\varepsilon} |r_2|^{1/2} \left(\widehat{V}^n + |r|^{n/3} \right)$$

as desired. □

Our next result is concerned about averages of $S_{d\underline{c},s,\boldsymbol{b},N}(\boldsymbol{v})$ over a sparse set of $\boldsymbol{v} \in \mathcal{O}^n$. Let $V \ge 1$, $C_1 \ge C_2 \ge 1$ and $\boldsymbol{v}_0 \in K_\infty^n$. For $d, s \in \mathcal{O}$ with $d \mid s$ and s cube-full, we proceed to consider the average

$$S(V, C_1, C_2) := \sum_{\substack{\underline{c} \in \mathcal{O}^2_{\text{prim}} \\ |c_i| \le \widehat{C}_i}} \sum_{\substack{|\boldsymbol{v} - \boldsymbol{v}_0| < \widehat{V} \\ F_1^*(\boldsymbol{v}) = 0}} |S_{d\underline{c},s,\boldsymbol{b},N}(\boldsymbol{v})|, \tag{6.7.3}$$

where F_1^* is the dual form of F_1 that we already met in Section 6. Note that upon replacing $\boldsymbol{v}_0$ with the nearest integer vector, we can assume without loss of generality that $\boldsymbol{v}_0 \in \mathcal{O}^n$.

Our basic strategy is to relate $S_{d\underline{c},s,\boldsymbol{b},N}(\boldsymbol{v})$ to a point-counting problem and gain savings when summing this problem over $\boldsymbol{v}$ and $\underline{c}$ first. For this let us write $s = r'\widetilde{s}$ into coprime $r', \widetilde{s} \in \mathcal{O}$ with

$$r' := \prod_{\nu_\varpi(s) \ge \nu_\varpi(N)+3} \varpi^{\nu_\varpi(s)}.$$

Note that r' is cube-full and $\varpi \mid \widetilde{s}$ implies $\nu_\varpi(N) \ge 1$ since s is cube-full. In particular, we have $|\widetilde{s}| \le |N|^3$ and thus by Lemma 6.6.1 that

$$|S_{d\underline{c},s,\boldsymbol{b},N}(\boldsymbol{v})| \le |N|^{3(n+2)} |S_{d'\underline{c},r',\boldsymbol{b},N'}(t\boldsymbol{v})| \tag{6.7.4}$$

for some $N' \mid r'$, $d' \mid r'$ and $t \in (\mathcal{O}/(r'N'/(r', N'))\mathcal{O})^\times$. Next we write $r' = r(r', N')$ and make the change of variables $\boldsymbol{x} = \boldsymbol{y}N' + \boldsymbol{b}$, so that

$$S_{d'\underline{c},r',\boldsymbol{b},N'}(t\boldsymbol{v}) = \psi\left(\frac{-\boldsymbol{b} \cdot \boldsymbol{v}}{r_{N'}} \right) \sum_{\underline{a}/r' \in L(d'\underline{c})} T(\underline{a}, r, t\boldsymbol{v}). \tag{6.7.5}$$

Let us now further write $r = e^2 f$, where $f \mid e$ and

$$f := \prod_{2 \nmid \nu_\varpi(r)} \varpi. \tag{6.7.6}$$

Our first step is to deduce to a congruence condition for $\boldsymbol{v}$ from the sum $T(\underline{a}, r, t\boldsymbol{v})$. This will be achieved in the next lemma.

Lemma 6.7.3. *Let $r \in \mathcal{O}$ be cube-full, $\underline{a} \in \mathcal{O}$, $\boldsymbol{v} \in \mathcal{O}^n$ and $r = e^2 f$ with $f \mid e$ and f given by* (6.7.6)*. Then*

$$T(\underline{a}, r, \boldsymbol{v}) = |e|^n \sum_{\substack{|\boldsymbol{y}|<|ef| \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e)}} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y}) - \boldsymbol{v}\cdot\boldsymbol{y}}{r}\right).$$

Proof. Let us write $\boldsymbol{x} = \boldsymbol{y} + ef\boldsymbol{z}$ with $|\boldsymbol{y}| < |ef|$ and $|\boldsymbol{z}| < |e|$. Then Taylor's formula implies

$$\begin{aligned} T(\underline{a}, r, \boldsymbol{v}) &= \sum_{|y|<|ef|} \sum_{|\boldsymbol{z}|<|e|} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y}+ef\boldsymbol{z}) - \boldsymbol{v}(\boldsymbol{y}+ef\boldsymbol{z})}{r}\right) \\ &= \sum_{|\boldsymbol{y}|<|ef|} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y}) - \boldsymbol{v}\cdot\boldsymbol{y}}{r}\right) \sum_{|\boldsymbol{z}|<|e|} \psi\left(\frac{\boldsymbol{z}\cdot(\nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y}) - \boldsymbol{v})}{e}\right) \\ &= |e|^n \sum_{\substack{|\boldsymbol{y}|<|ef| \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e)}} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y}) - \boldsymbol{v}\cdot\boldsymbol{y}}{r}\right). \end{aligned}$$

□

Next we want to establish extra congruence conditions for $F_1(\boldsymbol{y})$ and $F_2(\boldsymbol{y})$ by considering the sum over $\underline{a}/r \in L(d\underline{c})$. This step underpins the first substantial deviation from the treatment of the averages of exponential sums in [34] and results in a significant complication of the argument. The reason for this extra difficulty is that in the setting of one polynomial the underlying exponential sum is a Ramanujan sum, whose behaviour is well understood, while in our case the orthogonality relations we obtain stem from the more involved structure of rational points on the lines $L(d\underline{c})$.

Before we begin our treatment, we make the following convention to ease notation. Whenever we have a sum of the form $\sum'_{|\underline{a}_i|<|g_i|}$, we understand $'$ to mean that $(\underline{a}_i, r_i') = 1$. It then follows from (6.7.4) and (6.7.5) combined with Lemma 6.7.3 that

$$S(V, C_1, C_2) \le |N|^{3(n+2)} |e|^n \sum_{\substack{\underline{c}\in\mathcal{O}^2_{\text{prim}} \\ |c_i|\le\widehat{C}_i}} \sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V} \\ F_1^*(\boldsymbol{v})=0}} \left| \sum_{|\boldsymbol{y}|<|ef|} \psi\left(\frac{-t\boldsymbol{v}\cdot\boldsymbol{y}}{r}\right) \sum_{\substack{\underline{a}/r'\in L(d'\underline{c}) \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})\equiv t\boldsymbol{v}\ (e)}} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y})}{r}\right) \right|. \tag{6.7.7}$$

Our goal is now to investigate the sum

$$\Gamma(\boldsymbol{v}, \boldsymbol{y}) := \sum_{\substack{\underline{a}/r'\in L(d\underline{c}) \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})\equiv \boldsymbol{v}\ (e)}} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y})}{r}\right).$$

for $r, r' \in \mathcal{O}$ with $r \mid r'$. To do so, let us write $(r', N) = kk'$, where $(k, k') = (r, k') = 1$. Then we factor $r' = r_1' r_2' r_3'$ with pairwise coprime r_i's, $(r_1', d) = 1$ and

$$r_2' := \prod_{\substack{\nu_\varpi(ek') \geq \nu_\varpi(r') - \nu_\varpi(d) + 1 \\ \nu_\varpi(d) > 0}} \varpi^{\nu_\varpi(r')}. \tag{6.7.8}$$

Accordingly we shall also write $d = d_2 d_3$, $e = e_1 e_2 e_3$, $f = f_1 f_2 f_3$, $k = k_1 k_2 k_3$, $k' = k_1' k_2' k_3'$ and $r = r_1 r_2 r_3$ with $d_i, e_i, f_i, k_i, k_i', r_i \mid r_i'$, so that $r_i = e_i^2 f_i$ for $i = 1, 2, 3$. Moreover, we let d_3' be the maximal divisor of d_3 that divides r_3. In particular, we have $|d_3'| \asymp |d_3|$. The definition of $L(d\underline{c})$ implies that $\underline{a}/r' \in L(d\underline{c})$ if and only if $\varpi^{\nu_\varpi(r') - \nu_\varpi(d)} \parallel \underline{a} \cdot \underline{c}$ when $\nu_\varpi(d) \geq 1$ and $\varpi^{\nu_\varpi(r')} \mid \underline{a} \cdot \underline{c}$ when $\nu_\varpi(d) = 0$ for all $\varpi \mid r'$.

Since $r_i \mid r_i'$, it is therefore clear that the sum we are investigating is multiplicative and accordingly we shall denote the sum corresponding to r_i' by S_i for $i = 1, 2, 3$, so that $\Gamma(\boldsymbol{v}, \boldsymbol{y}) = S_1 S_2 S_3$.

We now treat each sum individually and start with S_1. When $|\underline{a}| < |r_1'|$, then by Lemma 6.2.4 we have $\underline{a}/r_1' \in L(\underline{c})$ if and only if $\underline{a} \equiv a\underline{c}^\perp \ (r_1')$ with $a \in (\mathcal{O}/r_1'\mathcal{O})^\times$. Since $r_1' = r_1(r_1', N)$, we can write $a = a_1 + e_1 k_1' a_2$ with $|a_1| < |e_1 k_1'|$ and $|a_2| < |r_1'||e_1 k_1'|^{-1} = |e_1 f_1 k_1|$. From the definition of e_1 and k_1' it is clear that $(a, r_1') = 1$ if and only if $(a_1, r_1') = 1$. Therefore, after splitting a_2 into residue classes modulo $e_1 f_1$ and using the fact that $(k_1', e_1 f_1) = 1$, we have

$$\begin{aligned} S_1 &= |k_1| \sideset{}{'}\sum_{\substack{|a_1| < |e_1 k_1'| \\ a_1 \nabla F_{\underline{c}}(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_1)}} \psi\left(\frac{a_1 F_{\underline{c}}(\boldsymbol{y})}{r_1}\right) \sum_{a_2\ (e_1 f_1)} \psi\left(\frac{a_2 k_1' F_{\underline{c}}(\boldsymbol{y})}{e_1 f_1}\right) \\ &= |k_1 e_1 f_1| \sideset{}{'}\sum_{\substack{|a_1| < |e_1 k_1'| \\ a_1 \nabla F_{\underline{c}}(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_1) \\ F_{\underline{c}}(\boldsymbol{y}) \equiv 0\ (e_1 f_1)}} \psi\left(\frac{a_1 F_{\underline{c}}(\boldsymbol{y})}{r_1}\right). \end{aligned} \tag{6.7.9}$$

Next we deal with the sum S_2. As before we make the change of variables $\underline{a} = \underline{a}_1 + k_2' e_2 \underline{a}_2$ with $|\underline{a}_1| < |e_2 k_2'|$ and $|\underline{a}_2| < |e_2 f_2 k_2|$, so that $(\underline{a}, r_2') = 1$ if and only if $(\underline{a}_1, r_2') = 1$. Moreover, it follows from the definition of r_2 that $\nu_\varpi(r_2') - \nu_\varpi(d_2) = \nu_\varpi(\underline{a} \cdot \underline{c})$ if and only if $\nu_\varpi(r_2') - \nu_\varpi(d_2) = \nu_\varpi(\underline{a}_1 \cdot \underline{c})$, so that $\underline{a}/r_2' \in L(d_2\underline{c})$ if and only if $\underline{a}_1/r_2' \in L(d_2\underline{c})$. We can again divide $\underline{a}_2$ into residue classes modulo $e_2 f_2$ to obtain

$$\begin{aligned} S_2 &= |k_2'|^2 \sum_{\substack{|\underline{a}_1| < |e_2 k_2'| \\ \underline{a}_1/r_2' \in L(d_2\underline{c}) \\ \nabla(\underline{a}_1 \cdot \underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_2)}} \psi\left(\frac{\underline{a}_1 \cdot \underline{F}(\boldsymbol{y})}{r_2}\right) \sum_{\underline{a}_2\ (e_2 f_2)} \psi\left(\frac{k_2' \underline{a}_2 \cdot \underline{F}(\boldsymbol{y})}{e_2 f_2}\right) \\ &= |e_2 f_2 k_2'|^2 \sum_{\substack{|\underline{a}_1| < |e_2 k_2'| \\ \underline{a}_1/r_2' \in L(d_2\underline{c}) \\ \nabla(\underline{a}_1 \cdot \underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_2) \\ F_1(\boldsymbol{y}) \equiv F_2(\boldsymbol{y}) \equiv 0\ (e_2 f_2)}} \psi\left(\frac{\underline{a}_1 \cdot \underline{F}(\boldsymbol{y})}{r_2}\right), \end{aligned} \tag{6.7.10}$$

where we used that $(k_2', e_2 f_2) = 1$.

Finally, we begin our treatment of the sum S_3, which is slightly more involved. First, we introduce character sums to detect the condition $\nu_\varpi(r_3') - \nu_\varpi(d_3) = \nu_\varpi(\underline{a}\cdot\underline{c})$:

$$S_3 = |r_3'^{-1} d_3| \sideset{}{'}\sum_{\substack{\underline{a}\ (r_3') \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})\equiv \boldsymbol{v}\ (e_3)}} \psi\left(\frac{\underline{a}\cdot \underline{F}(\boldsymbol{y})}{r_3}\right)$$
$$\times \prod_{\substack{\varpi^k \| r_3' \\ \varpi^m \| d_3}} \left(\sum_{b_0\ (\varpi^{k-m})} \psi\left(\frac{b_0\underline{a}\cdot\underline{c}}{\varpi^{k-m}}\right) - |\varpi|^{-1} \sum_{b_1\ (\varpi^{k-m+1})} \psi\left(\frac{b_1\underline{a}\cdot\underline{c}}{\varpi^{k-m+1}}\right)\right).$$

Then we make the change of variables $\underline{a} = \underline{a}_1 + e_3 k_3' \underline{a}_2$, with $|\underline{a}_1| < |e_3 k_3'|$ and $|\underline{a}_2| < |k_3 e_3 f_3|$. It follows from the definition of e_3 and k_3 that $(\underline{a}, r_3') = 1$ if and only if $(\underline{a}_1, r_3') = 1$. Moreover, r_3' was defined in such a way that $\nu_\varpi(e_3 k_3') \leq k - m$ for all $\varpi \mid r_3'$. Slightly abusing notation, note that we have

$$\sum_{|\underline{a}_2| < |e_3 f_3 k_3|} \psi\left(\frac{k_3'\underline{a}_2\cdot \underline{F}(\boldsymbol{y})}{e_3 f_3}\right) \prod_{\substack{\varpi^k \| r_3' \\ \varpi^m \| d_3}} \left(\sum_{b_0\ (\varpi^{k-m})} \psi\left(\frac{b_0\underline{a}\cdot\underline{c}}{\varpi^{k-m}}\right) - |\varpi|^{-1} \sum_{b_1\ (\varpi^{k-m+1})} \psi\left(\frac{b_1\underline{a}\cdot\underline{c}}{\varpi^{k-m+1}}\right)\right)$$
$$= \prod_{\substack{\varpi^k \| r_3' \\ \varpi^m \| d_3}} \left(\sum_{b_0\ (\varpi^{k-m})} \psi\left(\frac{b_0\underline{a}_1\cdot\underline{c}}{\varpi^{k-m}}\right) S_0(\varpi) - |\varpi|^{-1} \sum_{b_1\ (\varpi^{k-m+1})} \psi\left(\frac{b_1\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+1}}\right) S_1(\varpi)\right),$$

where

$$S_i(\varpi) := \sum_{\underline{a}_2\ (\varpi^{k-l})} \psi\left(\frac{k_3'\underline{a}_2\cdot(\underline{F}(\boldsymbol{y}) + \varpi^{m-i} b_i \underline{c})}{\varpi^{k'-l}}\right)$$
$$= |\varpi|^{2(k-l)} \delta_{-k_3' F_j(\boldsymbol{y}) \equiv \varpi^{m-i} b_i c_j\ (\varpi^{k'-l})}$$

for $i = 0, 1$ and where we temporarily wrote $l = \nu_\varpi(e_3 k_3')$ and $k' = \nu_\varpi(r_3)$. Observe that for $\varpi \mid k_3'$ the sums $S_i(\varpi)$ are independent of b_i. For $\varpi \mid r_3$, it follows upon making the change of variables $b_i = b_i' + \varpi^{k'-m-l+i} b_i''$ with $|b_i'| < |\varpi|^{k'-m-l+i}$ and $|b_i''| < |\varpi|^{k-k'+l}$ that

$$\sum_{b_i\ (\varpi^{k-m+i})} \psi\left(\frac{b_i\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+i}}\right) S_i(\varpi) = |\varpi|^{2(k-l)} \sum_{\substack{b_i\ (\varpi^{k-m+i}) \\ -k_3' F_j(\boldsymbol{y}) \equiv \varpi^{m-i} b_i c_j\ (\varpi^{k'-l})}} \psi\left(\frac{b_i\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+i}}\right)$$
$$= |\varpi|^{2(k-l)} \sum_{\substack{|b_i'| < |\varpi|^{k'-l-(m-i)} \\ -k_3' F_j(\boldsymbol{y}) \equiv \varpi^{m-i} b_i c_j\ (\varpi^{k'-l})}} \psi\left(\frac{b_i'\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+i}}\right)$$
$$\times \sum_{b_i''\ (\varpi^{k-k'+l})} \psi\left(\frac{b_i''\underline{a}_1\cdot\underline{c}}{\varpi^{k-k'+l}}\right)$$
$$= \frac{|\varpi|^{3k}}{|\varpi|^{l+k'}} \delta_{\varpi^{k-k'+l} | \underline{a}_1\cdot\underline{c}} \sum_{\substack{|b_i'| < |\varpi|^{k'-l-(m-i)} \\ -k_3' F_j(\boldsymbol{y}) \equiv \varpi^{m-i} b_i' c_j\ (\varpi^{k'-l})}} \psi\left(\frac{b_i'\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+i}}\right)$$

for $i = 0, 1$. Note that since $(c_1, c_2) = 1$, we have $(c_i, \varpi) = 1$ for $i = 1$ or $i = 2$. In particular, there is a unique b_i' with $|b_i'| < |\varpi|^{k-l-m+i}$ and $F_j(\boldsymbol{y}) \equiv \varpi^{m-i} b_i' k_3' c_j\ (\varpi^{k-l})$ for $j = 1, 2$. In addition, the latter equation implies $F_{\underline{c}}(\boldsymbol{y}) \equiv 0\ (\varpi^{k'-l})$ and $F_j(\boldsymbol{y}) \equiv 0\ (\varpi^m)$ for $j=1,2$. Using that $k'-l = \nu_\varpi(e_3f_3)$ for $\varpi \mid r_3$ and $k-k' = \nu_\varpi(k_3)$, we arrive at the identity

$$S_3 = |d_3e_3f_3k_3^2||k_3'|^{-1}\delta_{F_1(\boldsymbol{y})\equiv F_2(\boldsymbol{y})\equiv 0\ (d_3')}\delta_{F_{\underline{c}}(\boldsymbol{y})\equiv 0\ (e_3f_3)} \sideset{}{'}\sum_{\substack{|\underline{a}_1|<|e_3k_3'| \\ \nabla(\underline{a}_1\cdot\underline{F})(\boldsymbol{y})\equiv\boldsymbol{v}\ (e_3) \\ \varpi^{\nu_\varpi(k_3e_3)}\|\underline{a}_1\cdot\underline{c}}} \psi\left(\frac{\underline{a}_1\cdot\underline{c}}{r_3}\right)\Pi(\underline{a}_1), \tag{6.7.11}$$

where d_3' is the maximal divisor of d_3 dividing r_3 and

$$\Pi(\underline{a}_1) := \prod_{\substack{\varpi^k\|r_3' \\ \varpi^m\|d_3 \\ \varpi|r_3}} \left(\sum_{\substack{|b_0'|<|\varpi|^{k'-l-m} \\ F_i(\boldsymbol{y})\equiv\varpi^m b_1' c_i\ (\varpi^{k'-l})}} \psi\left(\frac{b_0'\underline{a}_1\cdot\underline{c}}{\varpi^{k-m}}\right) - |\varpi|^{-1} \sum_{\substack{|b_1'|<|\varpi|^{k'-l-(m-1)} \\ F_i(\boldsymbol{y})\equiv\varpi^{m-1}b_2'c_i\ (\varpi^{k'-l})}} \psi\left(\frac{b_1'\underline{a}_1\cdot\underline{c}}{\varpi^{k-m}}\right)\right)$$
$$\times \prod_{\substack{\varpi^k\|r_3' \\ \varpi^m\|d_3 \\ \varpi\nmid r_3}} \left(\sum_{b_0\ (\varpi^{k-m})} \psi\left(\frac{b_0\underline{a}_1\cdot\underline{c}}{\varpi^{k-m}}\right) S_0(\varpi) - |\varpi|^{-1} \sum_{b_1\ (\varpi^{k-m+1})} \psi\left(\frac{b_1\underline{a}_1\cdot\underline{c}}{\varpi^{k-m+1}}\right) S_1(\varpi)\right).$$

Moreover, since there is a unique b_i' with $|b_i'| < |\varpi|^{k-l-m-i}$ and $F_i(\boldsymbol{y}) \equiv \varpi^{m-i}b_i'k_3'c_i\ (\varpi^{k-l})$ it is easy to see that $\Pi(\underline{a}_1) \le |k_3'|$. After relabelling the variables, (6.7.9), (6.7.10) and (6.7.11) show that we have established the following result.

Lemma 6.7.4. *Let $r' \in \mathcal{O}$ be cube-full, $d \in \mathcal{O}$ with $d \mid r$, $\boldsymbol{v} \in \mathcal{O}^n$. Define $r = r'/(r', M)$ and $r = e^2 f$ with $f \mid e$ and f given by* (6.7.6)*. Then with the notation introduced in* (6.7.8) *for $\boldsymbol{y}, \boldsymbol{v} \in \mathcal{O}^n$, we have*

$$\sum_{\substack{\underline{a}/r'\in L(d\underline{c}) \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})\equiv\boldsymbol{v}\ (e)}} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y})}{r}\right) = |k_1e_1f_1||e_2f_2k_2'|^2|d_3e_3f_3k_3^2||k_3'|^{-1}\delta_{F_1(\boldsymbol{y})\equiv F_2(\boldsymbol{y})\equiv 0\ (e_2f_2d_3')}$$
$$\times\ \delta_{F_{\underline{c}}(\boldsymbol{y})\equiv 0\ (e_1f_1e_3f_3)} \sum\nolimits^{(1)} \psi\left(\frac{a_1F_{\underline{c}}(\boldsymbol{y})}{r_1}\right) \sum\nolimits^{(2)} \psi\left(\frac{\underline{a}_2\cdot\underline{F}(\boldsymbol{y})}{r_2}\right) \sum\nolimits^{(3)} \psi\left(\frac{\underline{a}_3\cdot\underline{c}}{r_3}\right)\Pi(\underline{a}_3),$$

where (1) *indicates that we are summing over $|a_1| < |e_1k_1'|$ subject to $(a_1, r_1') = 1$, $a_1\nabla F_{\underline{c}}(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_1)$;* (2) *that $|\underline{a}_2| < |e_2k_2'|$ with $(\underline{a}_2, r_2') = 1$, $\underline{a}_2/r_2' \in L(d_2\underline{c})$ and $\nabla(\underline{a}_2\cdot\underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_2)$; and* (3) *that $|\underline{a}_3| < |e_3k_3'|$ with $(\underline{a}_3, r_3') = 1$, $\nabla(\underline{a}_3\cdot\underline{F})(\boldsymbol{y}) \equiv \boldsymbol{v}\ (e_3)$ and $\varpi^{\nu_\varpi(k_3e_3)} \parallel \underline{a}_3\cdot\underline{c}$. In addition, $|\Pi(\underline{a}_3)| \le |k_3'|$.*

Recall that N, k_i, k_i' are all $O(1)$. In particular, once we combine Lemmas 6.7.3 and 6.7.4 with the Chinese remainder theorem, we obtain from (6.7.7) that

$$|S_{d\underline{c},s,\boldsymbol{b},N}(\boldsymbol{v})| \ll |e|^{n+1}|fe_2f_2d_3| \sideset{}{'}\sum_{|a_1|<|e_1k_1'|} \sideset{}{'}\sum_{\substack{|\underline{a}_2|<|e_2k_2'| \\ \underline{a}_2/r_2'\in L(d_2\underline{c})}} \sideset{}{'}\sum_{\substack{|\underline{a}_3|<|e_3k_3'| \\ \varpi^{\nu_\varpi(k_3e_3)}\|\underline{a}_3\cdot\underline{c}}} \left| \sum_{|\boldsymbol{y}|<|ef|}^{(4)} \psi\left(\frac{\underline{a}\cdot\underline{F}(\boldsymbol{y}) - t\boldsymbol{v}\cdot\boldsymbol{y}}{r}\right)\right|, \tag{6.7.12}$$

where $\underline{a} = a_1\underline{c}^\perp r_2r_3 + \underline{a}_2r_1r_3 + \underline{a}_3r_1r_2$ and (4) denotes the conditions $\nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y}) \equiv t\boldsymbol{v}\ (e)$, $F_1(\boldsymbol{y}) \equiv F_2(\boldsymbol{y}) \equiv 0\ (e_2f_2d_3')$ and $F_{\underline{c}}(\boldsymbol{y}) \equiv 0\ (e_1f_1e_3f_3)$.

Recall that $r = e^2 f$. Next we write $\boldsymbol{y} = \boldsymbol{y}_1 + e\boldsymbol{y}_2$ with $|\boldsymbol{y}_1| < |e|$ and $|\boldsymbol{y}_2| < |f|$. Note that the definition of r_3 implies that $\nu_\varpi(d_3') \leq \nu_\varpi(e_3 f_3)$. We shall therefore write

$$d_3' = e_3' f_3', \quad \text{where} \quad f_3' = \prod_{\nu_\varpi(d_3') = \nu_\varpi(e_3)+1} \varpi, \tag{6.7.13}$$

so that $e_3' \mid e_3$ and $f_3' \mid f_3$. Hence $F_i(\boldsymbol{y}) \equiv 0 \; (e_2 f_2 d_3')$ if and only if $F_i(\boldsymbol{y}_1) = e_2 e_3' m_i$ say and $f_2 f_3' \mid (m_i + e/(e_2 e_3') \boldsymbol{y}_2 \cdot \nabla F_i(\boldsymbol{y}_1))$ for $i = 1, 2$. Similarly, if $F_{\underline{c}}(\boldsymbol{y}_1) = e_1 e_3 n$, then it must hold that $f_1 f_3 \mid (n + e_2 \boldsymbol{y}_2 \cdot \nabla F_{\underline{c}}(\boldsymbol{y}_1))$. In addition, if $\nabla(\underline{a} \cdot \underline{F})(\boldsymbol{y}_1) = t\boldsymbol{v} + e\boldsymbol{k}$, then upon writing $\underline{a} = (a_1, a_2)$ we have

$$\underline{a} \cdot \underline{F}(\boldsymbol{y}) - t\boldsymbol{v} \cdot \boldsymbol{y} \equiv \underline{a} \cdot \underline{F}(\boldsymbol{y}_1) - t\boldsymbol{v} \cdot \boldsymbol{y}_1 + e^2 \left(a_1 \boldsymbol{y}_1 \cdot \nabla F_1(\boldsymbol{y}_2) + a_2 F_2(\boldsymbol{y}_2) + \boldsymbol{y}_2 \cdot \boldsymbol{k} \right) \; (r).$$

It thus follows that

$$\begin{aligned} &\left| \sum_{|\boldsymbol{y}| < |ef|}^{(4)} \psi\left(\frac{\underline{a} \cdot \underline{F}(\boldsymbol{y}) - t\boldsymbol{v} \cdot \boldsymbol{y}}{r} \right) \right| \\ &\qquad \leq \sum_{\substack{|\boldsymbol{y}_1| < |e| \\ e_2 e_3' | F_i(\boldsymbol{y}_1),\, i=1,2 \\ e_1 e_3 | F_{\underline{c}}(\boldsymbol{y}_1) \\ \nabla(\underline{a} \cdot \underline{F})(\boldsymbol{y}_1) \equiv t\boldsymbol{v} \; (e)}} \max_{\boldsymbol{k}} \left| \sum_{\boldsymbol{y}_2 \; (f)}^{(5)} \psi\left(\frac{a_1 \boldsymbol{y}_1 \cdot \nabla F_1(\boldsymbol{y}_2) + a_2 F_2(\boldsymbol{y}_2) + \boldsymbol{y}_2 \cdot \boldsymbol{k}}{f} \right) \right|, \end{aligned} \tag{6.7.14}$$

where (5) denotes the conditions

$$f_2 f_3' \mid (m_i + e/(e_2 e_3') \boldsymbol{y}_2 \cdot \nabla F_i(\boldsymbol{y}_1)) \quad \text{and} \quad f_1 f_3 \mid (n + e_2 \boldsymbol{y}_2 \cdot \nabla F_{\underline{c}}(\boldsymbol{y}_1)).$$

By abuse of notation we denote the sum over $\boldsymbol{y}_2$ by $\Sigma^{(5)}$. We can then use orthogonality of characters to detect the congruence conditions in (5). After employing the triangle inequality, a standard squaring and differencing argument delivers

$$\begin{aligned} \sum^{(5)} &\leq |(f_2 f_3')^2 f_1 f_3|^{-1} \sum_{b_0 \; (f_1 f_3)} \sum_{\underline{b}_2 \; (f_2 f_3')} \left| \sum_{\boldsymbol{y}_2 \; (f)} \psi\left(\frac{a_1 \boldsymbol{y}_1 \cdot \nabla F_1(\boldsymbol{y}_2) + a_2 F_2(\boldsymbol{y}_2) + \boldsymbol{y}_2 \cdot \boldsymbol{k}'}{f} \right) \right| \\ &\leq |f|^{n/2} N_f(\underline{a}, \boldsymbol{y}_1)^{1/2}, \end{aligned} \tag{6.7.15}$$

where $\boldsymbol{k}'$ is a term that depends at most on $m_1, m_2, \boldsymbol{y}_1$ and the e_i's, and

$$N_f(\underline{a}, \boldsymbol{y}) := \#\{\boldsymbol{z} \; (e) \colon (a_1 H(\boldsymbol{y}) + a_2 M)\boldsymbol{z} \equiv 0 \; (f)\}.$$

We now pause for a moment and collect what we have achieved so far. Inserting (6.7.14) and (6.7.15) into (6.7.12), we get

$$\begin{aligned} |S_{d\underline{c}, s, \boldsymbol{b}, N}(\boldsymbol{v})| &\ll |e|^{n+1} |f|^{n/2+1} |e_2 f_2 d_3| \sideset{}{'}\sum_{|a_1| < |e_1 k_1'|} \\ &\times \sideset{}{'}\sum_{\substack{|\underline{a}_2| < |e_2 k_2'| \\ \underline{a}_2 / r_2' \in L(d_2 \underline{c})}} \sideset{}{'}\sum_{\substack{|\underline{a}_3| < |e_3 k_3'| \\ \varpi^{\nu_\varpi(k_3 e_3)} \| \underline{a}_3 \cdot \underline{c}}} \sum_{\substack{|\boldsymbol{y}_1| < |e| \\ e_2 e_3' | F_i(\boldsymbol{y}_1), i=1,2 \\ e_1 e_3 | F_{\underline{c}}(\boldsymbol{y}) \\ \nabla(\underline{a} \cdot \underline{F})(\boldsymbol{y}_1) \equiv t\boldsymbol{v} \; (e)}} N_f(\underline{a}, \boldsymbol{y})^{1/2}. \end{aligned} \tag{6.7.16}$$

The innermost sum is clearly multiplicative, and so our next step is to focus on the sums

$$\begin{aligned}
S_1(e_1,f_1) &\coloneqq \sum_{\substack{\boldsymbol{y}\,(e_1)\\ F_{\underline{c}}(\boldsymbol{y})\equiv 0\,(e_1)\\ F_1(\boldsymbol{y})F_2(\boldsymbol{y})\equiv 0\,(f_1)}} N_{f_1}(\underline{c}^{\perp},\boldsymbol{y})^{1/2},\\
S_1'(e_1,f_1) &\coloneqq \sum_{\substack{\boldsymbol{y}\,(e_1)\\ F_{\underline{c}}(\boldsymbol{y})\equiv 0\,(e_1)\\ F_1(\boldsymbol{y})F_2(\boldsymbol{y})\not\equiv 0\,(f_1)}} N_{f_1}(\underline{c}^{\perp},\boldsymbol{y})^{1/2},\\
S_2(e_2,f_2) &\coloneqq \sum_{\substack{\boldsymbol{y}\,(e_2)\\ F_1(\boldsymbol{y})\equiv F_2(\boldsymbol{y})\equiv 0\,(e_2)}} N_{f_2}(\underline{a},\boldsymbol{y})^{1/2}\\
S_3(e_3,f_3) &\coloneqq \sum_{\substack{\boldsymbol{y}\,(e_3)\\ F_1(\boldsymbol{y})\equiv F_2(\boldsymbol{y})\equiv 0\,(e_3')\\ F_{\underline{c}}(\boldsymbol{y})\equiv 0\,(e_3)}} N_{f_3}(\underline{a},\boldsymbol{y})^{1/2}.
\end{aligned}$$

We will establish sufficiently strong estimates for $S_i(e_i,f_i)$ when $i=1,2,3$, while we will obtain an additional saving by averaging $S_1'(e_1,f_1)$ over $\underline{c}$. Before we can provide upper bounds for them, we prove the following intermediate step.

Lemma 6.7.5. *Let $r',r\in\mathcal{O}$ with $r'\mid r$ and $n\geq 13$. Then we have*

$$\begin{aligned}
&N_1(r)\coloneqq \#\{\boldsymbol{x}\,(r)\colon F_{\underline{c}}(\boldsymbol{x})\equiv 0\,(r)\}\ll |r|^{n-1+\varepsilon},\\
&N_2(r)\coloneqq \#\{\boldsymbol{x}\,(r)\colon F_1(\boldsymbol{x})\equiv F_2(\boldsymbol{x})\equiv 0\,(r)\}\ll |r|^{n-2+\varepsilon} \text{ and}\\
&N_3(r)\coloneqq \#\{\boldsymbol{x}\,(r)\colon F_1(\boldsymbol{x})\equiv F_2(\boldsymbol{x})\equiv 0\,(r'), F_{\underline{c}}(\boldsymbol{x})\equiv 0\,(r)\}\ll |r|^{n-1+\varepsilon}|r'|^{-1}.
\end{aligned}$$

Proof. All the quantities are mutliplicative by the Chinese remainder theorem, and so we may assume that $r=\varpi^k$ and $r'=\varpi^m$ with $m\leq k$ during the proof. Let us begin with the treatment of $N_3(\varpi^k)$ by detecting the congruence condition with character sums:

$$|\varpi|^{2m+k}N_3(\varpi^k)=\sum_{\underline{a}\,(\varpi^m)}\sum_{b\,(\varpi^k)}\sum_{\boldsymbol{x}\,(\varpi^k)}\psi\left(\frac{(\varpi^{k-m}\underline{a}+b\underline{c}^{\perp})\cdot\underline{F}(\boldsymbol{x})}{\varpi^k}\right)$$

Suppose now that $0\leq l\leq k-1$ is such that $\varpi^l\parallel\varpi^{k-m}\underline{a}+b\underline{c}^{\perp}$. Then we claim that the sum over $\boldsymbol{x}$ above is

$$|\varpi|^{ln}\sum_{\boldsymbol{x}\,(\varpi^{k-l})}\psi\left(\frac{\varpi^{-l}(\varpi^{k-m}\underline{a}+b\underline{c}^{\perp})\cdot\underline{F}(\boldsymbol{x})}{\varpi^{k-l}}\right)\ll|\varpi|^{ln+5(k-l)n/6}.$$

Indeed, if $\varpi^{1+v_\varpi}\mid\varpi^{-l}(\varpi^{k-m}a_1-bc_2)$, then the sum is $O(|\varpi|^{ln+(k-l)(n+1)/2})$ by Lemma 6.6.8, while if $\varpi^{1+v_\varpi}\nmid\varpi^{-l}(\varpi^{k-m}a_1-bc_2)$, then we can apply Lemma 6.7.1 with $\boldsymbol{v_0}=\boldsymbol{0}$ and $\widehat{V}=1$ to obtain the claimed estimate.

For $0\leq l\leq k$ fixed, let us now determine the number of triples (a_1,a_2,b) such that $\varpi^{k-m}\underline{a}+b\underline{c}^{\perp}\equiv 0\,(\varpi^l)$. If $l\leq k-m$, then this holds if and only if $\varpi^l\parallel b$ since $\underline{c}$ is primitive, so that the number of available (a_1,a_2,b) is $O(|\varpi|^{2m+k-l})$. On the other hand, if $l>k-m$, then again because $\underline{c}$ is primitive, we can without loss of generality assume that $(c_1,\varpi)=1$. This implies

$$b\equiv\varpi^{k-m}a_2c_1^{-1}\,(\varpi^l),$$

which determines b uniquely modulo ϖ^l. We thus also have

$$\varpi^{k-m}a_1 \equiv -c_2 b \equiv \varpi^{k-m}a_2 c_2 c_1^{-1}\,(\varpi^l),$$

which determines a_1 uniquely modulo ϖ^{l-k+m} provided a_2 is given. In total we get that the number of such (a_1, a_2, b) is at most $|\varpi|^{m+k-l+m-(l-k+m)} = |\varpi|^{2(k-l)+m}$. We conclude that

$$|\varpi|^{2m+k}N_3(\varpi^k) \ll |\varpi|^{m+nk} + |\varpi|^{m+nk}\sum_{l=0}^{k-1}|\varpi|^{(n/6-2)(l-k)} \ll |\varpi|^{m+nk} + |\varpi|^{m+nk}$$

for $n/6 - 2 > 0$.

The proof of the statement for the quantities $N_1(r)$ and $N_2(r)$ runs along the same lines and is in fact less involved. The argument again relies on the estimates provided by Lemmas 6.6.8 and 6.7.1 and we do not provide details here. □

Before we continue with our study of the sums $S_i(e_i, f_i)$, we make some preliminary observations. First of all, if $(a_1, f) \ll 1$, then it follows from Equation (6.12) in [34] that

$$\sum_{\boldsymbol{y}\,(f)} N_f(\underline{a}, \boldsymbol{y}) \ll |f|^{n+\varepsilon}. \tag{6.7.17}$$

Moreover, if $f \mid a_1$, then $\mathrm{rk}(M) \geq n-1$ readily implies $N_f(\underline{a}, \boldsymbol{y}) \ll |f|$.

Lemma 6.7.6. *Let $e_i, f_i \in \mathcal{O}$ with $f_i \mid e_i$ and f_i square-free for $i = 1, 2, 3$. Then for $n \geq 13$, we have*

$$S_1(e_1, f_1) \ll |e_1|^{n-1+\varepsilon}|f_1|^{1/2},$$
$$S_2(e_2, f_2) \ll |e_2|^{n-2+\varepsilon}|f_2| \text{ and}$$
$$S_3(e_3, f_3) \ll |e_3|^{n-1+\varepsilon}|e_3'|^{-1}|f_3|.$$

Proof. All of the sums under consideration are multiplicative, and so we only have to prove the corresponding estimates when $e_i = \varpi^k$ and $f_i = 1, \varpi$. Moreover, we shall write $m = \nu_\varpi(e_3')$, so that $k \geq m \geq 1$.

Let X_ϖ be the reduction of $V(F_1, F_2)$ modulo ϖ. When $i = 2, 3$, we begin with the case when $f_i = 1$ or $a_1 \equiv 0\,(\varpi)$, while when $i = 1$ we assume $f_i = 1$ or $c_2 \equiv 0\,(\varpi)$. Since $N_\varpi((0, a_2), \boldsymbol{y}) \ll |\varpi|$, in our situation we thus see that

$$S_i(\varpi^k, f_i) \ll |f_i|^{1/2}N_i(\varpi^k).$$

Lemma 6.7.5 provides estimates for $N_i(\varpi^k)$ that are satisfactory for the statement of the lemma.

Moreover, when $c_1 \equiv 0\,(\varpi)$, then it follows from the second display after Equation (6.15) in [34] that $S_1(\varpi^k, f_1) \ll \varpi^{k(n-1)}|f_1|^{1/2}$, which is also sufficient.

We may therefore assume that $f_i = \varpi$ from now on. When $i = 1$, we are left with the case $(c_2, \varpi) = (c_1, \varpi) = 1$, while for $i = 2, 3$ we have to deal with the case when $(a_1, \varpi) = 1$. Let us first assume that X_ϖ, the reduction of $V(F_1, F_2)$ modulo ϖ, is singular. Since this can only happen for at most finitely ϖ's, it must hold that $N_\varpi(\underline{a}, \boldsymbol{y}) \ll 1$. In particular, we obtain

$$S_i(\varpi^k, f_i) \ll N_i(\varpi)$$

in this case, which is again satisfactory by Lemma 6.7.5.

So let us now assume that X_ϖ is non-singular. We first provide an upper bound for the contribution from $\boldsymbol{y} \not\equiv \boldsymbol{0}\,(\varpi^k)$ to $S_i(\varpi^k,\varpi)$. If $F_{\underline{c}}(\boldsymbol{y}) \equiv 0\,(\varpi^l)$ and $F_1(\boldsymbol{y})F_2(\boldsymbol{y}) \equiv 0\,(\varpi)$, then $F_{\underline{c}}(\boldsymbol{y}+\varpi^l\boldsymbol{z}) \equiv 0\;(\varpi^{l+1})$ if and only if $\varpi^{-l}F_{\underline{c}}(\boldsymbol{y}) \equiv -\boldsymbol{z}\cdot\nabla F_{\underline{c}}(\boldsymbol{y})\;(\varpi)$. Since $c_1c_2 \not\equiv 0\;(\varpi)$, the condition $F_1(\boldsymbol{y})F_2(\boldsymbol{y}) \equiv 0\;(\varpi)$ forces that $\nabla F_{\underline{c}}(\boldsymbol{y}) \not\equiv \boldsymbol{0}\;(\varpi)$ as otherwise $\boldsymbol{y}$ would be a singular point of X_ϖ. In particular for $\boldsymbol{y} \not\equiv \boldsymbol{0}\;(\varpi)$, inductively we obtain

$$M_1(\boldsymbol{y}) := \#\{\boldsymbol{z}\;(\varpi^k) \colon \boldsymbol{z} \equiv \boldsymbol{y}\;(\varpi), F_{\underline{c}}(\boldsymbol{z}) \equiv 0\;(\varpi^k)\} \ll |\varpi|^{(k-1)(n-1)}.$$

Similarly, if $F_1(\boldsymbol{y}) \equiv F_2(\boldsymbol{y}) \equiv 0\;(\varpi^l)$, then $F_1(\boldsymbol{y}+\varpi^l\boldsymbol{z}) \equiv F_1(\boldsymbol{y}+\varpi^l\boldsymbol{z}) \equiv 0\;(\varpi^{l+1})$ holds if and only if $\varpi^{-l}F_i(\boldsymbol{y}) \equiv -\boldsymbol{z}\cdot\nabla F_i(\boldsymbol{y})\;(\varpi)$. As X_ϖ is non-singular and $\boldsymbol{y} \not\equiv \boldsymbol{0}\;(\varpi)$, we must have $\mathrm{rk}(\nabla F_1(\boldsymbol{y}), \nabla F_2(\boldsymbol{y})) = 2$. Therefore, it follows by induction that

$$M_2(\boldsymbol{y}) := \#\{\boldsymbol{z}\;(\varpi^k) \colon \boldsymbol{z} \equiv \boldsymbol{y}\;(\varpi), F_1(\boldsymbol{z}) \equiv F_2(\boldsymbol{z}) \equiv 0\;(\varpi^k)\} \ll |\varpi|^{(k-1)(n-2)}.$$

Finally, if $F_1(\boldsymbol{y}) \equiv F_2(\boldsymbol{y}) \equiv 0\;(\varpi^m)$, then $F_{\underline{c}}(\boldsymbol{y}+\varpi^m\boldsymbol{z}) \equiv 0\;(\varpi^{m+1})$ if and only if $\varpi^{-m}F_{\underline{c}}(\boldsymbol{y}) \equiv -\boldsymbol{z}\cdot\nabla F_{\underline{c}}(\boldsymbol{y})\,(\varpi)$. As $F_1(\boldsymbol{y}) \equiv F_2(\boldsymbol{y}) \equiv 0\,(\varpi)$, we must have $\nabla F_c(\boldsymbol{y}) \not\equiv \boldsymbol{0}\,(\varpi)$, as otherwise $\boldsymbol{y}$ would be a singular point of X_ϖ. Combing this with the arguments that were used to estimate $M_1(\boldsymbol{y})$ and $M_2(\boldsymbol{y})$, we obtain

$$\begin{aligned} M_3(\boldsymbol{y}) &:= \{\boldsymbol{z}\;(\varpi^k) \colon \boldsymbol{z} \equiv \boldsymbol{y}, F_1(\boldsymbol{z}) \equiv F_2(\boldsymbol{z}) \equiv 0\;(\varpi^m), F_{\boldsymbol{c}}(\boldsymbol{z}) \equiv 0\;(\varpi^k)\} \\ &\ll |\varpi|^{(m-1)(n-2)+(k-m)(n-1)}. \end{aligned}$$

It now follows from an application of the Cauchy-Schwarz inequality that the contribution from $\boldsymbol{y} \not\equiv 0\;(\varpi)$ to $S_i(\varpi^k,\varpi)$ is at most

$$\max_{\boldsymbol{y}\not\equiv 0\;(\varpi)} M_i(\boldsymbol{y})N_i(\varpi)^{1/2}\Big(\sum_{\boldsymbol{y}\;(\varpi)} N_\varpi(\underline{a},\boldsymbol{y})\Big)^{1/2} \ll |\varpi|^{n/2} \max_{\boldsymbol{y}\not\equiv 0\;(\varpi)} M_i(\boldsymbol{y})N_i(\varpi)^{1/2}$$

by (6.7.17). Combining the estimates we just provided for $M_i(\boldsymbol{y})$ with Lemma 6.7.5 to bound $N_i(\varpi)$, we see that this contribution is sufficient for the conclusion of the lemma to hold.

We are thus left with estimating the contribution from $\boldsymbol{y} \equiv \boldsymbol{0}\;(\varpi)$ to $S_i(\varpi^k,\varpi)$. In this case we have $N_\varpi(\underline{a},\boldsymbol{y}) \ll |\varpi|$, so that

$$S_i(\varpi^k,\varpi) \ll |\varpi|^{1/2}N_i(\varpi^k),$$

which is again satisfactory by Lemma 6.7.5. □

We now return to the main task of this section: estimating the quantity $S(V,C_1,C_2)$ that was defined in (6.7.3). Equation (6.7.16) gives

$$\begin{aligned} S(V,C_1,C_2) \ll |e|^{n+1}|f|^{n/2+1}|e_2f_2d_3| &\sum_{\substack{\underline{c}\in\mathcal{O}^2_{\text{prim}} \\ |c_i|\le\widehat{C}_i}} \sideset{}{'}\sum_{|a_1|<|e_1k_1'|} \sideset{}{'}\sum_{\substack{|\underline{a}_1|<|e_2k'| \\ \underline{a}_1/r_2'\in L(d_2\underline{c})}} \sideset{}{'}\sum_{\substack{|\underline{a}_1'|<|e_3k_3'| \\ \varpi^{\nu_\varpi(k_3e_3)}\|\underline{a}_1\cdot\underline{c}}} \\ &\times \sum_{\boldsymbol{y}\;(e)} N_f(\underline{a},\boldsymbol{y})^{1/2} \sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0|\le\widehat{V} \\ F_1^*(\boldsymbol{v})=0 \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})\equiv t\boldsymbol{v}\;(e)}} 1. \end{aligned} \tag{6.7.18}$$

Let us now write $\boldsymbol{v} = \boldsymbol{v}_0+\boldsymbol{v}_1+\boldsymbol{v}_2e$, where $|\boldsymbol{v}_1| < |e|$, $t\boldsymbol{v}_1 \equiv \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y})(e)$ and $|\boldsymbol{v}_2| < \widehat{V}|e|^{-1}$. Observe that $(e,t) = 1$ implies that $\boldsymbol{v}_1$ is unique. Note that $G(\boldsymbol{v}_2) = F_1^*(\boldsymbol{v}_0+\boldsymbol{v}_1+\boldsymbol{v}_2e)$ is of degree $3 \times 2^{n-2}$ and its leading degree part is absolutely irreducible, so that we may invoke Lemma 6.3.4 to deduce that

$$\sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0| \leq \widehat{V} \\ F_1^*(\boldsymbol{v})=0 \\ \nabla(\underline{a}\cdot\underline{F})(\boldsymbol{y}) \equiv t\boldsymbol{v}\ (e)}} 1 \ll 1 + \left(\frac{\widehat{V}}{|e|}\right)^{n-2}. \tag{6.7.19}$$

Using the Chinese remainder theorem together with Lemma 6.7.6, we obtain

$$\begin{aligned}\sum_{\boldsymbol{y}\ (e)} N_f(\underline{a}, \boldsymbol{y})^{1/2} &= S_2(e_2, f_2) S_3(e_3, f_3) \left(S_1(e_1, f_1) + S_1'(e_1, f_1)\right) \\ &\ll |e_2|^{n-2+\varepsilon} |f_2| |e_3|^{n-1+\varepsilon} |e_3'|^{-1} |f_3| \left(|e_1|^{n-1+\varepsilon} |f_1|^{1/2} + S_1'(e_1, f_1)\right).\end{aligned} \tag{6.7.20}$$

Moreover, the conjunction of the conditions $F_1(\boldsymbol{y}) F_2(\boldsymbol{y}) \not\equiv 0\ (f_1)$ and $F_{\underline{c}}(\boldsymbol{y}) \equiv 0\ (f_1)$ can only hold if $(c_1, f_1) = (c_2, f_1) = 1$. Therefore, for $C_1 \geq C_2 \geq 1$ we have

$$\begin{aligned}\sum_{\substack{\underline{c} \in \mathcal{O}^2_{\text{prim}} \\ |c_i| \leq \widehat{C}_i}} S_1'(e_1, f_1) &\leq \max_{\substack{\underline{a}\ (f_1) \\ (a_1, f_1) = (a_2, f_1) = 1}} \sum_{\substack{\boldsymbol{y}\ (e_1) \\ F_1(\boldsymbol{y}) F_2(\boldsymbol{y}) \not\equiv 0 (f_1)}} N_{f_1}(\underline{a}, \boldsymbol{y})^{1/2} \sum_{\substack{\underline{c} \in \mathcal{O}^2_{\text{prim}} \\ |c_i| \leq \widehat{C}_i \\ F_{\underline{c}}(\boldsymbol{y}) \equiv 0\ (e_1)}} 1 \\ &\leq \widehat{C}_2 \left(1 + \frac{\widehat{C}_1}{|e_1|}\right) \max_{\substack{\underline{a}\ (f_1) \\ (a_1, f_1) = (a_2, f_1) = 1}} \left(\frac{|e_1|}{|f_1|}\right)^n \sum_{\boldsymbol{y}\ (f_1)} N_{f_1}(\underline{a}, \boldsymbol{y})^{1/2} \\ &\ll |e_1|^n \widehat{C}_2 \left(1 + \frac{\widehat{C}_1}{|e_1|}\right),\end{aligned} \tag{6.7.21}$$

where we used (6.7.17) together with the Cauchy-Schwarz inequality to arrive at the last estimate.

Next observe that $\underline{a}_2 / r_2' \in L(d_2 \underline{c})$ implies that $\varpi^{\nu_\varpi(r_2' d_2^{-1})} \mid \underline{a}_2 \cdot \underline{c}$, so that that

$$\#\{|\underline{a}_2| < |e_2 k'| \colon \underline{a}_2 / r_2' \in L(d_2 \underline{c})\} \leq \left(\frac{|e_2 k'|}{|r_2' d_2^{-1}|}\right)^2 |r_2' d_2^{-1}| = |e_2 k'|^2 |r_2'|^{-1} |d_2|. \tag{6.7.22}$$

A similar argument delivers

$$\#\{|\underline{a}_3| < |e_3 k_3'| \colon \varpi^{\nu_\varpi(e_3)} \mid \underline{a}_3 \cdot \underline{c}\} \leq |k_3'|^2 |e_3|. \tag{6.7.23}$$

Recall that $e = e_1 e_2 e_3$, $f = f_1 f_2 f_3$ and $d = d_2 d_3$. Then since k, k' are $O(1)$ and $|s| \asymp |r|$, we can combine (6.7.19) – (6.7.23) with (6.7.18) to obtain

$$\begin{aligned} S(V, C_1, C_2) \ll& |e|^{n+1} |f|^{n/2+1} |e_1 e_3|^2 |e_2| |f_1|^{1/2} |f_2|^2 |f_3| |d_2 d_3| |r_2'|^{-1} |e_3'|^{-1} \left(|e| + \widehat{V}\right)^{n-2} \\ &\times \widehat{C}_2 \left(\widehat{C}_1 + |e_1| |f_1|^{-1/2} + \widehat{C}_1 |f_1|^{-1/2}\right) \\ \ll& |d| |s|^{(n+3)/2} \frac{|f_2^3 f_3|^{1/2}}{|r_2' e_3'|} \widehat{C}_2 \left(|e| + \widehat{V}\right)^{n-2} \left(\widehat{C}_1 + |e_1| |f_1|^{-1/2}\right), \end{aligned}$$

where we used that $|e^2 f| \asymp |s|$ and $d_2 d_3 = d$. Since r_2 is cube-full and $|r_2'| \asymp |r_2|$, we have $f_2^3 \mid r_2$ and thus $|f_2|^{3/2} |r_2'|^{-1} \ll 1$. Moreover, since $\nu_\varpi(d_3') \geq 1$ for all $\varpi \mid e_3$, the definition of e_3' in (6.7.13) implies that $f_3 \mid e_3'$ and hence $|f_3|^{1/2} |e_3'| \ll 1$. We have thus established the following result.

Lemma 6.7.7. *Let $s \in \mathcal{O}$ be cube-full and $d \mid s$. Then with the notation of* (6.7.8)*, it holds that*

$$S(V, C_1, C_2) \ll_{F_1, F_2, N} |d| |s|^{(n+3)/2} \widehat{C}_2 \left(|s|^{n/2-1} + \widehat{V}^{n-2}\right) \left(\widehat{C}_1 + |e_1| f_1|^{-1/2}\right).$$

6.8 Return to the circle method

In this section we combine the estimates for the various exponential sums and integrals that we have produced so far to finish our treatment of $N(P)$. To ease of notation, for $d \in \mathcal{O}$ and $\underline{c} \in \mathcal{O}^2$ we abbreviate the properties d monic, $\underline{c}$ primitive, $|dc_1| \leq \min\{\widehat{T}|r|^{1/2}, |r|\}$, $|dc_2| < \widehat{T}^{-1}|r|^{1/2}$ and $\max\{\widehat{R}_i|dc_i^{\perp}|\} \geq |r|$ by $P(d, \underline{c})$. In addition, throughout this section we shall assume $n \geq 26$. We will not make the dependence of the implied constants explicit anymore, but allow them to depend at most on F_1, F_2, w and N as well as on ε if it appears in the inequality. Recall the decomposition of $N(P)$ in (6.4.9).

6.8.1 The main term

We begin to carry out the analysis of $M(P)$. Since we chose $\widehat{R}_2 \asymp |P|^{1/3}$ in (6.4.8), it follows from Corollary 6.2.10 that

$$M(P) = |P|^n \sum_{\substack{|r| \ll |P|^{1/3} \\ r \text{ monic}}} |r_N|^{-n} S_r K_r + |P|^n \sum_{\substack{|P|^{1/3} \ll |r| \leq \widehat{R}_1\widehat{R}_2 \\ r \text{ monic}}} |r_N|^{-1} \sum_{\substack{d\underline{c}\colon P(d,\underline{c}) \\ d|r}} S_{d\underline{c},r,\boldsymbol{b},N}(\mathbf{0}) K_r, \tag{6.8.1}$$

where

$$S_r = \sideset{}{'}\sum_{\underline{a}\ (r)} \sum_{\substack{|\boldsymbol{x}|<|r_N| \\ \boldsymbol{x} \equiv \boldsymbol{b}\ (N)}} \psi\left(\frac{\underline{a} \cdot \underline{F}(\boldsymbol{x})}{r}\right)$$

and

$$K_r = \int_{|\theta_1| < \widehat{R}_1^{-1}|r|^{-1}} \int_{|\theta_2| < \widehat{R}_2^{-1}|r|^{-1}} I_{r_N}(\underline{\theta}, \mathbf{0}) \mathrm{d}\underline{\theta}.$$

It is a consequence of Proposition 6.5.4 that

$$K_r \ll |P|^{-5+\varepsilon} \tag{6.8.2}$$

and from Corollary 6.7.2 with $\boldsymbol{v}_0 = \mathbf{0}$ and $\widehat{V} = 1$ we deduce

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) \ll |d||r|^{5n/6+3/2+\varepsilon}. \tag{6.8.3}$$

Note that if $c_2 \neq 0$, then since $\widehat{T} \asymp |P|^{1/2}$, the condition $|dc_2| < \widehat{T}^{-1}|r|^{1/2}$ can only hold if $|r| \gg |P|$. In particular, it is now easy to see that

$$\sum_{\substack{d\underline{c}\colon P(d,\underline{c}) \\ d|r}} |d| \ll |r|^{1+\varepsilon}. \tag{6.8.4}$$

It follows from (6.8.2)–(6.8.4) that the rightmost term in (6.8.1) is of order

$$\begin{aligned} |P|^{n-5+\varepsilon} \sum_{\substack{|P|^{1/3} \ll |r| \leq \widehat{R}_1\widehat{R}_2 \\ r \text{ monic}}} \frac{|r|^{5n/6+3/2}}{|r_N|^n} \sum_{\substack{d\underline{c}\ P(d,\underline{c}) \\ d|r}} |d| &\ll |P|^{n-5+\varepsilon} \sum_{\substack{|P|^{1/3} \ll |r| \leq \widehat{R}_1\widehat{R}_2 \\ r \text{ monic}}} |r|^{5/2-n/6} \\ &\ll |P|^{n-5-(7/2-n/6)/3+\varepsilon}, \end{aligned}$$

where we used that that $7/2 - n/6 < 0$ for $n > 21$. Consequently, the contribution from this term is negligible and it remains to investigate the first term on the right hand side of (6.8.1).

The first step we take is to analyse the integral K_r. Let $C > 0$ be a fixed positive integer, whose exact value will be determined in due course. We then split up the integral K_r into

$$K_r = \int_{|\theta_1|<\widehat{C}^{-1}|P|^{-3}} \int_{|\theta_2|<\widehat{C}^{-1}|P|^{-2}} I_{r_N}(\underline{\theta}, \mathbf{0})\mathrm{d}\underline{\theta} + \int_{\Xi} I_{r_N}(\underline{\theta}, \mathbf{0})\mathrm{d}\underline{\theta}, \tag{6.8.5}$$

where Ξ is defined by

$$\Xi := \{\underline{\theta} \in \mathbb{T}^2 \colon |\theta_i| < \widehat{R}_i^{-1}|r|^{-1} \text{ for } i = 1,2 \text{ and } \widehat{C}^{-1}|P|^{-3} \leq |\theta_1| \text{ or } \widehat{C}^{-1}|P|^{-2} \leq |\theta_2|\}.$$

Note that Ξ is non-empty only if $|r| < \max_{i=1,2}\{\widehat{C}\widehat{R}_i^{-1}|P|^{4-i}\}$. Since $\widehat{R}_i^{-1}|P|^{4-i} \asymp |P|^{5/3}$ by (6.4.8), this will certainly hold for $|r| \ll |P|^{1/3}$ and P sufficiently large. We will show that the second integral vanishes and produce a lower bound for the first one. Beginning with the former task, we have by (6.5.1) that the second integral is equal to

$$\widehat{L}^{-n} \int_{\Xi} \int_{\mathbb{T}^n} \psi\left(P^3\theta_1 G_1(\boldsymbol{x}) + P^2\theta_2 G_2(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\underline{\theta},$$

where $G_i(\boldsymbol{x}) = F_i(\boldsymbol{x}_0 + t^{-L}\boldsymbol{x})$ for $i = 1,2$. Let $\Gamma = (\{1\} \times \mathbb{T}) \cup (\mathbb{T} \times \{1\})$ and define

$$\lambda = \min_{\underline{\gamma}\in\Gamma} |\gamma_1 \nabla F_1(\boldsymbol{x}_0) + \gamma_2 \nabla F_2(\boldsymbol{x}_0)|.$$

Observe that $0 < \lambda < 1$ since Γ is compact and $\boldsymbol{x}_0$ a non-singular point of the variety defined by F_1 and F_2 with $|\boldsymbol{x}_0| < H_{\underline{F}}^{-1}$. To simplify notation, we write $\gamma_1 = P^3\theta_1$ and $\gamma_2 = P^2\theta_2$. Let now $\gamma \in \mathbb{T}$ be such that $|\gamma| = |\underline{\gamma}|$. Then by the ultrametric property we have

$$|\gamma_1 \nabla G_1(\boldsymbol{x}) + \gamma_2 \nabla G_2(\boldsymbol{x})| = \widehat{L}^{-1}|\gamma||\gamma_1 \nabla F_1(\boldsymbol{x}_0)/\gamma + \gamma_2 \nabla F_2(\boldsymbol{x}_0)/\gamma| \geq \widehat{L}^{-1}|\gamma|\lambda,$$

provided L is sufficiently large. Moreover, all higher partial derivatives of $\gamma_1 G_1(\boldsymbol{x}) + \gamma_2 G_2(\boldsymbol{x})$ are of order $O(\widehat{L}^{-2}|\gamma|)$. By the second derivative test [34, Lemma 2.5] we thus have $K_r = 0$ if $\widehat{L}^{-1}|\gamma|\lambda \geq 1$. Since $|\gamma| \geq \widehat{C}^{-1}$, this can be ensured if we make the choice $\widehat{C} = \lambda\widehat{L}^{-1}$, which we henceforth assume.

We proceed to investigate the first integral in (6.8.5). After making the change of variables $\gamma_i = t^{(4-i)P}\theta_i$, by (6.5.1) we have

$$K_r = \widehat{L}^{-n}|P|^{-5} \int_{|\underline{\gamma}|<\widehat{C}^{-1}} \int_{\mathbb{T}^n} \psi\left(\gamma_1 G_1(\boldsymbol{x}) + \gamma_2 G_2(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\underline{\gamma}.$$

It is now clear that K_r is in fact independent of r and to emphasise this, we define

$$\sigma_\infty := \widehat{L}^{-n} \int_{|\underline{\gamma}|<\widehat{C}^{-1}} \int_{\mathbb{T}^n} \psi\left(\gamma_1 G_1(\boldsymbol{x}) + \gamma_2 G_2(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\underline{\gamma}.$$

The integral σ_∞ is the singular integral associated to our counting problem and our next step is to show that $\sigma_\infty > 0$. To do so, we exchange the order of integration and apply Lemma 3.0.2 to deduce that

$$\sigma_\infty = \widehat{L}^{-n}\widehat{C}^{-2} \operatorname{vol}\left(\{\boldsymbol{x} \in \mathbb{T}^n \colon |G_i(\boldsymbol{x})| < \widehat{C} \text{ for } i = 1,2\}\right).$$

Using Taylor expansion and the fact that $F_i(\boldsymbol{x}_0) = 0$ for $i = 1,2$, it follows that

$$G_1(\boldsymbol{x}) = t^{-L}\boldsymbol{x} \cdot \nabla F_1(\boldsymbol{x}_0) + \frac{1}{2}t^{-2L}\boldsymbol{x}^t H(\boldsymbol{x}_0)\boldsymbol{x} + t^{-3L}F_1(\boldsymbol{x})$$

and
$$G_2(\boldsymbol{x}) = t^{-L}\boldsymbol{x}\cdot\nabla F_2(\boldsymbol{x}_0) + t^{-2L}F_2(\boldsymbol{x}).$$
Provided L is sufficiently large, we must then have
$$|G_1(\boldsymbol{x})| = \widehat{L}^{-1}|\boldsymbol{x}\cdot\nabla F_1(\boldsymbol{x}_0)| \quad \text{and} \quad |G_2(\boldsymbol{x})| = \widehat{L}^{-1}|\boldsymbol{x}\cdot\nabla F_2(\boldsymbol{x}_0)|.$$
However, if we recall that $\widehat{C} = \widehat{L}^{-1}\lambda$, then for $|\boldsymbol{x}| < \lambda$ it is clear that
$$\widehat{L}^{-1}|\boldsymbol{x}\cdot\nabla F_i(\boldsymbol{x}_0)| < \widehat{L}^{-1}\lambda = \widehat{C},$$
so that $\sigma_\infty > \widehat{L}^{-n}\widehat{C}^{-2}\lambda^n$. We summarise our investigation of the integral K_r in the following result.

Lemma 6.8.1. *Let $|r| \ll |P|^{1/3}$. Then*
$$K_r = \sigma_\infty |P|^{-5},$$
where $\sigma_\infty > 0$ depends only on the weight function w.

It follows from Lemma 6.8.1 and the upper bound provided after (6.8.1) that
$$M(P) = \sigma_\infty \mathfrak{S}_{\boldsymbol{b},N}(1/3)|P|^{n-5} + O\left(|P|^{n-5-\delta'}\right),$$
for some $\delta' > 0$, where for $\Delta > 0$ we have defined
$$\mathfrak{S}_{\boldsymbol{b},N}(\Delta) \coloneqq \sum_{\substack{|r|\ll|P|^\Delta \\ r \text{ monic}}} |r_N|^{-N} S_r$$
to be the truncated singular series associated to our counting problem. Let
$$\mathfrak{S}_{\boldsymbol{b},N} \coloneqq \sum_{r \text{ monic}} |r_N|^{-n} S_r$$
be the completed singular series. It follows from Lemma 6.7.1 and Lemma 6.6.8 that
$$S_r = \sideset{}{'}\sum_{\underline{a}\,(r)} T(\underline{a}, r, \mathbf{0}) \ll |r|^{2+5n/6+\varepsilon},$$
so that $\mathfrak{S}_{\boldsymbol{b},N}$ converges absolutes for $n > 18$ and
$$|\mathfrak{S}_{\boldsymbol{b},N}(\Delta) - \mathfrak{S}_{\boldsymbol{b},N}| \ll \sum_{|r|>|P|^\Delta} |r|^{2-n/6+\varepsilon} \ll |P|^{\Delta(3-n/6+\varepsilon)}.$$
It is a routine exercise to show that $\mathfrak{S}_{\boldsymbol{b},N} > 0$ provided there exists $\boldsymbol{x}_\varpi \in \mathcal{O}_\varpi^n$ such that $F_1(\boldsymbol{x}_\varpi) = F_2(\boldsymbol{x}_\varpi) = 0$ and $|\boldsymbol{b} - \boldsymbol{x}_\varpi|_\varpi < |N|_\varpi$ for all ϖ, see for example Corollary 4.4.7 of [131] for arbitrary complete intersections. In particular, we have established the following result.

Proposition 6.8.2. *For $n > 18$ we have*
$$M(P) = \sigma_\infty \mathfrak{S}_{\boldsymbol{b},N}|P|^{n-5} + O\left(|P|^{n-5-\delta''}\right)$$
for some $\delta'' > 0$ with $\sigma_\infty > 0$. Moreover, $\mathfrak{S}_{\boldsymbol{b},N} > 0$ if for every ϖ there exists $\boldsymbol{x}_\varpi \in \mathcal{O}_\varpi^n$ such that $F_1(\boldsymbol{x}_\varpi) = F_2(\boldsymbol{x}_\varpi) = 0$ and $|\boldsymbol{b} - \boldsymbol{x}_\varpi|_\varpi < |N|_\varpi$.

To prove Proposition 6.4.2 and hence also Theorems 6.1.1 and 6.1.2, it remains to give a satisfactory upper bound for the error term $E_1(P)$ defined in (6.4.9). This will occupy the remainder of our work and makes use of the various estimates we have provided for the oscillatory integrals and exponentials sums under consideration.

6.8.2 Preparation for the error terms

We continue our investigation of the error term $E_1(P)$. Before doing so, we take some preliminary steps. Firstly, we shall fix the absolute value of r to be $\widehat{Y}$ and of θ_i to be $\widehat{\Theta}_i$ for $i = 1, 2$ respectively, where

$$1 \leq Y \leq R_1 + R_2, \qquad -9P + R_2 \leq \Theta_1 < -Y - R_1 \quad \text{and} \quad -9P + R_1 \leq \Theta_2 < -Y - R_2. \tag{6.8.6}$$

and neither of the conditions (1) nor (2) recorded after (6.4.9) hold. Observe that by our choice of R_1 and R_2 in (6.4.8) the number of admissible triples (Y, Θ_1, Θ_2) is $O(|P|^{\varepsilon})$. Secondly, we treat separately the contribution from $c_2 \neq 0$ and $c_2 = 0$ and denote the contribution of such r's, $\underline{\theta}$'s and $\underline{c}$'s to $E_1(P)$ by $E_{1,a}(Y, \Theta_1, \Theta_2)$ and $E_{1,b}(Y, \Theta_1, \Theta_2)$ respectively, so that

$$E_{1,a}(Y, \Theta_1, \Theta_2) = |P|^n \sum_{\substack{|r| = \widehat{Y} \\ r \text{ monic}}} |r_N|^{-n} \sum_{\substack{d\underline{c}: P(d\underline{c}) \\ c_2 \neq 0 \\ d | r}} \int_{|\theta_i| = \widehat{\Theta}_i} \sum_{\boldsymbol{v} \in \mathcal{O}^n \setminus \{\boldsymbol{0}\}} S_{d\underline{c}, r, \boldsymbol{b}, N}(\boldsymbol{v}) I_{r_N}(\underline{\theta}, \boldsymbol{v}) \mathrm{d}\underline{\theta} \tag{6.8.7}$$

and

$$E_{1,b}(Y, \Theta_1, \Theta_2) = |P|^n \sum_{\substack{|r| = \widehat{Y} \\ r \text{ monic}}} |r_N|^{-n} \sum_{\substack{|d| \leq \widehat{Y}^{1/2} P^{1/2} \\ d | r}} \int_{|\theta_i| = \widehat{\Theta}_i} \sum_{\boldsymbol{v} \in \mathcal{O}^n \setminus \{\boldsymbol{0}\}} S_{d\underline{c}_0, r, \boldsymbol{b}, N}(\boldsymbol{v}) I_{r_N}(\underline{\theta}, \boldsymbol{v}) \mathrm{d}\underline{\theta}, \tag{6.8.8}$$

where $\underline{c}_0 = (1, 0)$ since by our convention $\underline{c}$ with $c_2 = 0$ can only be primitive when $c_1 = 1$. If we can show that $E_{1,i}(Y, \Theta_1, \Theta_2) \ll |P|^{n-5-\kappa}$ for $i = a, b$ and some $\kappa > 0$, then since the number of admissible triples (Y, Θ_1, Θ_2) is $O(|P|^{\varepsilon})$, the same estimate will hold with a new choice of κ for $E_1(P)$. Moreover, if we let $\widehat{Z} = \max\{1, |P|^3 \widehat{\Theta}_1, |P|^2 \widehat{\Theta}_2\}$, then (6.5.2) implies that the summation range of $\boldsymbol{v}$ in the definition of $E_{1,i}(Y, \Theta_1, \Theta_2)$ is empty unless

$$|\boldsymbol{v}| \ll \widehat{V}, \quad \text{where} \quad \widehat{V} = \frac{\widehat{Y} \widehat{Z}}{|P|}. \tag{6.8.9}$$

In particular, since $\boldsymbol{v} \neq \boldsymbol{0}$, it must also hold that

$$\widehat{Y} \widehat{Z} \gg |P|. \tag{6.8.10}$$

Finally, we use the convention that $\varepsilon > 0$ is an arbitrarily small real number whose exact value may change from one appearance to the next.

Treatment of $E_{1,a}(Y, \Theta_1, \Theta_2)$

Note that we must have

$$\widehat{Y} \gg |P|, \tag{6.8.11}$$

because as $c_2 \neq 0$ and $\widehat{T} \asymp |P|^{1/2}$, the inequality $|dc_2| < \widehat{T}^{-1} \widehat{Y}^{1/2}$ can only hold if $\widehat{Y} \gg |P|$. Since we assume that $c_2 \neq 0$, Lemma 6.6.5 gives strong upper bounds for $S_{\underline{c}, r, \boldsymbol{b}, N}(\boldsymbol{v})$ provided r is square-free and $(r, F_1^*(\boldsymbol{v})) = 1$. Accordingly, we shall further split up $E_{1,a}(Y, \Theta_1, \Theta_2)$ into the contribution from those $\boldsymbol{v}$ with $F_1^*(\boldsymbol{v}) \neq 0$ and $F_1^*(\boldsymbol{v}) = 0$ and denote it by E_1' and E_2' respectively. In the treatment of E_2' we compensate the worse exponential sum estimates compared to E_1' by exploiting the sparsity of vectors $\boldsymbol{v}$ such that $F_1^*(\boldsymbol{v}) = 0$.

Let us begin by dealing with the term E_1'. Applying (6.5.1) and Lemma 6.5.1 to the integral $I_{r_N}(\underline{\theta}, \boldsymbol{v})$ in (6.8.7), it follows that

$$E_1' \leq |N|^n \widehat{L}^{-n} \frac{|P|^n}{\widehat{Y}^n} \sum_{\substack{d\underline{c}\colon P(d\underline{c}) \\ c_2 \neq 0}} \int_{|\underline{\theta}|=\widehat{\Theta}} \int_{\mathbb{T}^n} \sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}\widehat{Z}^{-1/2} \\ F_1^*(\boldsymbol{v})\neq 0}} \sum_{\substack{|r|=\widehat{Y} \\ d|r}} |S_{d\underline{c},r,\boldsymbol{b},M}(\boldsymbol{v})| \mathrm{d}\boldsymbol{x} \mathrm{d}\underline{\theta}, \tag{6.8.12}$$

where $\boldsymbol{v}_0 = -r_N t^L (P^3 \nabla F_1(\boldsymbol{x}_0 + t^{-L}\boldsymbol{x}) + P^2 \nabla F_2(\boldsymbol{x}_0 + t^{-L}\boldsymbol{x}))$. Our next goal is to estimate the sum

$$S := \sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}\widehat{Z}^{-1/2} \\ F_1^*(\boldsymbol{v})\neq 0}} \sum_{\substack{|r|=\widehat{Y} \\ d|r}} |S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})|.$$

For this we write $r = b_1 b_1' b_2 b_2' r_3$ into pairwise coprime $b_1, b_1', b_2, b_2', r_3 \in \mathcal{O}$, where $b_1 b_1'$ is the square-free part of r satisfying $(b_1, dNF_1^*(\boldsymbol{v})c_2) = 1$; $b_2 b_2'$ is such that $\nu_\varpi(b_2 b_2') = 2$ for all $\varpi \mid b_2 b_2'$ and $(b_2, dNc_2) = 1$ and r_3 is the cube-full part of r. Accordingly we shall also write $d = d_1 d_2 d_3$ with $d_1 \mid b_1'$, $d_2 \mid b_2'$ and $d_3 \mid r_3$. We can then use the multiplicativity of $S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})$ recorded in Lemma 6.6.1 to deduce for appropriate $t_1, t_1', t_2, t_2', t_3, N_1, N_2, N_3 \in \mathcal{O}$ that

$$\begin{aligned} |S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})| &= |S_{b_1}(t_1\boldsymbol{v}) S_{d_1\underline{c},b_1',\boldsymbol{b},N_1}(t_1'\boldsymbol{v}) S_{b_2}(t_2\boldsymbol{v}) S_{d_2\underline{c},b_2',\boldsymbol{b},N_2}(t_2'\boldsymbol{v}) S_{d_3\underline{c},r_3,\boldsymbol{b},N_3}(t_3\boldsymbol{v})| \\ &\ll |r|^\varepsilon |b_1|^{(n+1)/2} |b_2|^{n/2+1} |d_1 d_2| |b_1' b_2'|^{(n+3)/2} |S_{d_3\underline{c},r_3,\boldsymbol{b},N_3}(t_3\boldsymbol{v})|, \end{aligned} \tag{6.8.13}$$

where we used Lemmas 6.6.5, 6.6.6 and 6.6.7 to estimate the sums corresponding to b_1, b_1' and b_2 respectively and Corollary 6.6.9 for the sum corresponding to b_2'. Moreover, by Corollary 6.7.2 we have

$$\sum_{\substack{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}\widehat{Z}^{-1/2} \\ F_1^*(\boldsymbol{v})\neq 0}} |S_{d_3\underline{c},r_3,\boldsymbol{b},N_3}(t_3\boldsymbol{v})| \ll |d_3||r_3|^{n/2+1+\varepsilon} |r_3''|^{1/2} \left(\widehat{V}^n \widehat{Z}^{-n/2} + |r_3|^{n/3}\right), \tag{6.8.14}$$

where $r_3 = r_3' r_3''$ with $(r_3', d) = 1$ are defined in (6.7.1). Consequently, plugging (6.8.13) and (6.8.14) into the definition of S yields

$$\begin{aligned} S &\ll \widehat{Y}^{(n+1)/2+\varepsilon} |d| \left(\widehat{V}^n \widehat{Z}^{-n/2} + \widehat{Y}^{n/3}\right) \sum_{\substack{|r|=\widehat{Y} \\ d|r}} |b_2 r_3'|^{1/2} |b_1' b_2' r_3''| \\ &= \widehat{Y}^{(n+3)/2+\varepsilon} |d| \left(\widehat{V}^n \widehat{Z}^{-n/2} + \widehat{Y}^{n/3}\right) \sum_{\substack{|r|=\widehat{Y} \\ d|r}} |b_1|^{-1} |b_2 r_3'|^{-1/2}, \end{aligned}$$

since $|b_1 b_1' b_2 b_2' r_3' r_3''| = \widehat{Y}$. By definition, we must have $b_1' b_2' r_3'' \mid (dNc_2 F_1^*(\boldsymbol{v}))^\infty$. The number of available b_1', b_2', r_3'' with $|b_1' b_2' r_3''| \leq \widehat{Y}$ is hence $O(\widehat{Y}^\varepsilon)$. Moreover, since $b_2 r_3'$ is square-full, the number of b_2 and r_3''s of fixed absolute value $\widehat{B}$ is $O(\widehat{B}^{1/2})$. After summing over q-adic intervals it thus follows that

$$S \ll \widehat{Y}^{(n+3)/2+\varepsilon} |d| \left(\widehat{V}^n \widehat{Z}^{-n/2} + \widehat{Y}^{n/3}\right)$$

and hence

$$\begin{aligned} E_1' &\ll |P|^n \widehat{Y}^{3/2-n/2+\varepsilon} \widehat{\Theta}_1 \widehat{\Theta}_2 \sum_{d\underline{c}\colon P(d\underline{c})} |d| \left(\widehat{V}^n \widehat{Z}^{-n/2} + \widehat{Y}^{n/3}\right) \\ &\ll |P|^{n-5} \widehat{Y}^{5/2-n/2+\varepsilon} \widehat{Z}^2 \left(\widehat{V}^n \widehat{Z}^{-n/2} + \widehat{Y}^{n/3}\right), \end{aligned}$$

where we used that $\widehat{\Theta}_1\widehat{\Theta}_2 \ll |P|^{-5}\widehat{Z}^2$ and $\sum_{d\underline{c}\colon P(d\underline{c})} |d| \ll \widehat{Y}^{1+\varepsilon}$. If the first term in the brackets dominates, we get

$$E_1' \ll |P|^{-5}\widehat{Y}^{5/2+n/2+\varepsilon}\widehat{Z}^{2+n/2} \ll |P|^{5n/6-5+10/3}\widehat{Y}^{1/2+\varepsilon} \ll |P|^{5n/6-5/3+5/6+\varepsilon},$$

because $\widehat{Z} \ll |P|^{5/3}\widehat{Y}^{-1}$ and $\widehat{Y} \ll |P|^{5/3}$. Thus, the contribution from E_1' in this case is $O(|P|^{n-5-\kappa})$ for some $\kappa > 0$ as soon as $n > 25$. If the second term dominates, then

$$E_1' \ll |P|^{n-5}\widehat{Y}^{5/2-n/6+\varepsilon}\widehat{Z}^2 \ll |P|^{n-5+10/3}\widehat{Y}^{1/2-n/6+\varepsilon} \ll |P|^{n-5+10/3+1/2-n/6+\varepsilon},$$

where we used that $\widehat{Y} \gg |P|$ by (6.8.11). This is satisfactory as soon as $n > 23$, which completes our treatment of E_1'.

Next we consider the contribution from E_2'. This time we apply Lemma 6.5.2 to the integral $I_{r_N}(\underline{\theta}, \boldsymbol{v})$ and obtain

$$E_2' \ll \frac{|P|^n}{\widehat{Y}^n}\widehat{Z}^{1-n/2}\widehat{\Theta}_1\widehat{\Theta}_2 \sum_{|d| \leq \widehat{T}^{-1/2}|\widehat{Y}|^{1/2}} \sum_{\substack{|c_i| < \widehat{C}_i \\ c_2 \neq 0}} \sum_{\substack{0 < |\boldsymbol{v}| \leq \widehat{V} \\ F_1^*(\boldsymbol{v}) = 0}} \sum_{\substack{|r| = \widehat{Y} \\ d \mid r}} |S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})|,$$

where $\widehat{C}_1 = \widehat{T}\widehat{Y}^{1/2}|d|^{-1}$ and $\widehat{C}_2 = \widehat{T}^{-1}\widehat{Y}^{1/2}|d|^{-1}$. We proceed to consider the sum

$$S' \coloneqq \sum_{\substack{|c_i| < \widehat{C}_i \\ c_2 \neq 0}} \sum_{\substack{0 < |\boldsymbol{v}| \leq \widehat{V} \\ F_1^*(\boldsymbol{v}) = 0}} \sum_{\substack{|r| = \widehat{Y} \\ d \mid r}} |S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})|.$$

For this we first factor $r = b_1b_2r_3$ into pairwise coprime $b_1, b_2, r_3 \in \mathcal{O}$ and $d = d_2d_3$ with $d_2 \mid b_2$ and $d_3 \mid r_3$, where b_1 is square-free, b_2 is cube-free, $(b_1, dNc_2) = 1$ and r_3 is the cube-full part of r. Parallel to our argument for E_1' we use Lemma 6.6.1 to factor $S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v})$ and invoke Lemmas 6.6.6 and 6.6.7 to bound the sum corresponding to b_1 as well as Lemma 6.6.6 and Corollary 6.6.9 to bound the sum corresponding to b_2 to obtain

$$S_{d\underline{c},r,\boldsymbol{b},N}(\boldsymbol{v}) \ll |r|^{\varepsilon}|d_2||b_1|^{n/2+1}|b_2|^{(n+3)/2}|S_{d_3\underline{c},r_3,\boldsymbol{b},N'}(t\boldsymbol{v})|$$

for appropriate $t, N' \in \mathcal{O}$. We wish to apply Lemma 6.7.7 to estimate the average

$$A \coloneqq \sum_{|c_i| < \widehat{C}_i} \sum_{\substack{0 < |\boldsymbol{v}| \leq \widehat{V} \\ F_1^*(\boldsymbol{v}) = 0}} |S_{d_3\underline{c},r_3,\boldsymbol{b},N'}(t\boldsymbol{v})|.$$

For this note that with the notation of Lemma 6.7.7 for P sufficiently large we have $|e_1| \leq \widehat{Y}^{1/2} \ll \widehat{Y}^{1/2}\widehat{T}|d|^{-1} = \widehat{C}_1$, since $\widehat{T} \asymp |P|^{1/2}$ and $|d| \leq \widehat{Y}^{1/2}\widehat{T}^{-1} \ll |P|^{1/3}$. In particular, Lemma 6.7.7 hands us

$$A \ll |d_3||r_3|^{n/2+3/2+\varepsilon}\widehat{C}_1\widehat{C}_2\left(|r_3|^{n/2-1} + \widehat{V}^{n-2}\right). \tag{6.8.15}$$

We may also forget about the condition $F_1^*(\boldsymbol{v}) = 0$ and use Corollary 6.7.2 to obtain the alternative estimate

$$A \ll \widehat{C}_1\widehat{C}_2|d_3||r_3|^{n/2+1+\varepsilon}|r_3''|^{1/2}\left(\widehat{V}^n + |r_3|^{n/3}\right), \tag{6.8.16}$$

where $r_3 = r_3'r_3''$ with r_3' given by (6.7.1). In particular, we have shown so far that

$$A \ll \widehat{C}_1\widehat{C}_2|d_3||r_3|^{n/2+1+\varepsilon}\min\left\{|r_3|^{1/2}\left(\widehat{V}^{n-2} + |r_3|^{n/2-1}\right), |r_3''|^{1/2}\left(\widehat{V}^n + |r_3|^{n/3}\right)\right\}. \tag{6.8.17}$$

We begin with the contribution from $\widehat{V} \geq |r_3|^{1/2}$ or $\widehat{V} \leq |r_3|^{1/3}$. Since $b_2 r_3'' \mid (Ndc_2)^\infty$, there are at most $O(|P|^\varepsilon)$ pairs (b_2, r_3''). Moreover, since r_3 is cube-full, there are $O(|r_3|^{1/3})$ available r_3 of fixed absolute. One now easily derives

$$\sum_{\substack{|r|=\widehat{Y} \\ d|r}} |b_2 r_3|^{1/2} \ll \widehat{Y}^{1+\varepsilon}.$$

In addition, we also have

$$\sum_{|d| \leq \widehat{T}^{-1}\widehat{Y}^{1/2}} \widehat{C}_1 \widehat{C}_2 |d| = \widehat{Y} \sum_{|d| \leq \widehat{T}^{-1}\widehat{Y}^{1/2}} |d|^{-1} \ll \widehat{Y}^{1+\varepsilon}. \tag{6.8.18}$$

After employing the estimates $\widehat{\Theta}_1 \widehat{\Theta}_2 \ll |P|^{-5}\widehat{Z}^2$ and $|r_3| \leq \widehat{Y}$, we have in the cases under consideration

$$E_2' \ll |P|^{n-5} \widehat{Y}^{3-n/2} \widehat{Z}^{3-n/2} \left(\widehat{V}^{n-2} + \widehat{Y}^{n/3} \right). \tag{6.8.19}$$

If the first term dominates, we obtain

$$E_2' \ll |P|^{-3} \widehat{Y}^{1+n/2+\varepsilon} \widehat{Z}^{1+n/2} \ll |P|^{5n/6-3+5/3+\varepsilon},$$

which is satisfactory if $n > 22$. On the other hand, if $\widehat{Y}^{n/3} \geq \widehat{V}^{n-2}$, then

$$E_2' \ll |P|^{n-5} \widehat{Y}^{3-n/6+\varepsilon} \widehat{Z}^{3-n/2} \ll |P|^{n-5+(3-n/6)+\varepsilon},$$

by (6.8.11) and because $\widehat{Z} \geq 1$. Thus this contribution is sufficiently small as soon as $n > 18$.

Finally, we have to deal with the contribution from $|r_3|^{1/3} < \widehat{V} < |r_3|^{1/2}$. For this, we use that for any real numbers $A, B > 0$ and $0 \leq \kappa \leq 1$ that $\min\{A, B\} \leq A^{1-\kappa} B^\kappa$ with $\kappa = 1/(n-2)$ to deduce that

$$\begin{aligned} S' &\ll \widehat{C}_1 \widehat{C}_2 |d| \widehat{V}^{n(1-\kappa)} \widehat{Y}^{n/2+1+\varepsilon} \sum_{|r|=\widehat{Y}} |b_2|^{1/2} |r_3''|^{(1-\kappa)/2} |r_3|^{\kappa(n-1)/2} \\ &\ll \widehat{C}_1 \widehat{C}_2 |d| \widehat{V}^{n(1-\kappa)} \widehat{Y}^{n/2+2+\varepsilon} \sum_{|b_2 r_3' r_3''| \leq \widehat{Y}} |b_2|^{-1/2} |r_3'|^{(n-1)/(2n-4)-1}. \end{aligned}$$

The number of available b_2 of fixed absolute value is $O(|b_2|^{1/2})$. Moreover, there are at most $O(\widehat{Y}^\varepsilon)$ possibilities for r_3''. Since $(n-1)/(2n-4) - 1 < -1/3$ for $n \geq 6$ and the number of $|r_3'|$ of fixed absolute value is $O(|r_3'|^{1/3})$, it follows that the sum above is $O(\widehat{Y}^\varepsilon)$. Therefore, we have

$$S' \ll \widehat{C}_1 \widehat{C}_2 |d| \widehat{V}^{n(1-\kappa)} \widehat{Y}^{n/2+2+\varepsilon}. \tag{6.8.20}$$

Therefore, the contribution to E_2' is at most

$$\begin{aligned} \frac{|P|^n}{\widehat{Y}^n} \widehat{Z}^{1-n/2} \widehat{\Theta}_1 \widehat{\Theta}_2 \sum_d S' &\ll |P|^{n-5/3+\varepsilon} \widehat{Y}^{1-n/2} \widehat{Z}^{1-n/2} \widehat{V}^{n(1-\kappa)} \\ &= |P|^{-5/3+n\kappa+\varepsilon} \widehat{Y}^{1+n/2-\kappa n} \widehat{Z}^{1+n/2-\kappa n} \\ &\ll |P|^{5n/6-2\kappa n/3+\varepsilon}, \end{aligned}$$

where we used $\widehat{\Theta}_1 \widehat{\Theta}_2 \ll |P|^{-5/3} \widehat{Y}^{-2}$ and (6.8.18) to estimate the sum over d. One can check that $5n/6 - 2n/(3n-6) < n - 5$ provided $n \geq 26$, which completes our treatment of E_2' and thus also of $E_{1,a}(Y, \Theta_1, \Theta_2)$.

Treatment of $E_{1,b}(Y,\Theta_1,\Theta_2)$

We differ our treatment according to the size of Y. When $\widehat{Y}\geq |P|^{1-\eta}$, where η is as in (2) after (6.4.9), it is more efficient to estimate the sum over $\boldsymbol{v}$ via the same trick that we used to arrive at (6.8.12), whereas when $\widehat{Y}\leq |P|^{1-\eta}$, we estimate the integral directly. Before doing so, for $W\geq 1$ we focus on the sum

$$\Sigma(W):=\sum_{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{W}}\sum_{|d|\leq \widehat{K}}\sum_{\substack{|r|=\widehat{Y}\\ d|r}}|S_{d\underline{c}_0,r,\boldsymbol{b},N}(\boldsymbol{v})|,$$

where $\widehat{K}=\min\{\widehat{Y},\widehat{Y}^{1/2}\widehat{T}\}$.

To begin with, let us write $r=sr_3$, where s is the cube-free part of r and r_3 is cube-full. Moreover, we write $d=ef_1f_2^2d_3$ into pairwise coprime $e,f_1,f_2,d_3\in\mathcal{O}$, where ef_1f_2 is square-free, e is the greatest common divisor of s and the square-free part of d, $ef_1^2f_2^2\mid s$ and $d_3\mid r_3$. The multiplicativity of $S_{d\underline{c}_0,r,\boldsymbol{b},N}(\boldsymbol{v})$ together with Proposition 6.6.10 then imply

$$S_{d\underline{c}_0,r,\boldsymbol{b},N}(\boldsymbol{v})\ll \widehat{Y}^{\varepsilon}|f_1f_2|^{1/2}|s|^{(n+3)/2}|S_{d_3\underline{c}_0,r_3,\boldsymbol{b},N'}(t\boldsymbol{v})|,$$

for some $N'\mid N$ and $t\in\mathcal{O}$ with $(t,r_3)=1$. Moreover, the number of available s is $O(\widehat{Y}|e_1f_1^2f_2^2r_3|^{-1})$, so that

$$S\ll \widehat{Y}^{(n+5)/2+\varepsilon}\sum_{d}\sum_{\boldsymbol{v}}\sum_{d=ef_1f_2^2d_2}|e_1f_1|^{-1}|f_2|^{-1/2}\sum_{\substack{|r_3|\leq\widehat{Y}\\ d_3|r_3}}\frac{|S_{d_3\underline{c}_0,r_3,\boldsymbol{b},N'}(t\boldsymbol{v})|}{|r_3|^{(n+5)/2}}.$$

Next we factor r_3 into $d_3's_1s_2$ into pairwise coprime d_3',s_1,s_2, where $d_3'\mid d_3$, $(s_2,N'\Delta_{F_2})=1$ and $\nu_\varpi(s_2)>\nu_\varpi(d_3)$, so that $s_1\mid(N'\Delta_{F_2})^\infty$. Accordingly we shall also write $d_3=d_3'g_1g_2$, so that $g_i\mid s_i$ for $i=1,2$. Let $u\in\mathcal{O}$ and suppose that

$$u=\prod b_i^i,\quad (b_i,b_j)=1 \text{ if } i\neq j,\quad b_i \text{ square-free}.$$

We then define the function

$$m(u):=b_1^3b_2^3\prod_{i\geq 3}b_i^{i+1}.$$

Note that since s_3 is cube-full and $\nu_\varpi(s_2)>\nu_\varpi(g_2)$, we must then have $m(g_2)\mid s_2$. By Lemma 6.6.1 we then have for some $t_1,t_2\in\mathcal{O}$ with $(t_1,d_3's_1)=(t_2,s_2)=1$ and $N''\mid N'$ that

$$S_{d_3\underline{c}_0,r_3,\boldsymbol{b},N'}(t\boldsymbol{v})=S_{d_3'g_1\underline{c}_0,d_3's_1,\boldsymbol{b},N''}(t_1\boldsymbol{v})S_{g_2\underline{c}_0,s_2,\boldsymbol{0},1}(t_2\boldsymbol{v}).$$

Therefore, it follows from Proposition 6.6.10 that

$$\begin{aligned}\sum_{\substack{|r_3|\leq\widehat{Y}\\ d_3|r_3}}\frac{|S_{d_3\underline{c}_0,r_3,\boldsymbol{b},N'}(t\boldsymbol{v})|}{|r_3|^{(n+5)/2}}&\ll\widehat{Y}^{\varepsilon}\sum_{d_3=d_3'g_1g_2}\sum_{\substack{|s_1|\leq\frac{\widehat{Y}}{|d_3'|}\\ g_1|s_1}}\frac{|S_{d_3'g_1\underline{c}_0,d_3's_1,\boldsymbol{b},N''}(t_1\boldsymbol{v})|}{|d_3's_1|^{(n+5)/2}}\sum_{\substack{|s_2|\leq\widehat{Y}\\ m(g_2)|s_2}}|g_2||s_2|^{-1}\\
&\ll\widehat{Y}^{\varepsilon}\sum_{d_3=d_3'g_1g_2}|g_2||m(g_2)|^{-1}\sum_{\substack{|s_1|\leq\frac{\widehat{Y}}{|d_3'|}\\ g_1|s_1}}\frac{|S_{d_3'g_1\underline{c}_0,d_3's_1,\boldsymbol{b},N''}(t_1\boldsymbol{v})|}{|d_3's_1|^{(n+5)/2}},\end{aligned}$$

where we used that there are at most $O(|s_2|^{1/3})$ available s_2 of fixed absolute value, because s_2 is cube-full. Next we change the order of summation and make the sum over $\boldsymbol{v}$ the innermost one. We then use Corollary 6.7.2 to deduce

$$\sum_{\substack{|s_1|\le \frac{\widehat{Y}}{|d_3'|}\\ g|s_1}} \sum_{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{W}} \frac{|S_{d_3'g_1\underline{c}_0,d_3's_1,\boldsymbol{b},N''}(t_1\boldsymbol{v})|}{|d_3's_1|^{(n+5)/2}} \ll |d_3'|^{-1/2}|g_1| \sum_{\substack{|s_1|\le \frac{\widehat{Y}}{|d_3'|}\\ g_1|s_1}} |s_1|^{-1}\left(\widehat{W}^n + |d_3's_1|^{n/3}\right)$$
$$\ll |d_3'|^{-1/2}\widehat{W}^n + |d_3'|^{-1/2}\widehat{Y}^{n/3+\varepsilon}.$$

From what we have shown so far, it follows that $\Sigma(W)$ is bounded from above by

$$\widehat{Y}^{(n+5)/2+\varepsilon} \sum_{|d|\le\widehat{K}} \sum_{d=e_1f_1f_2^2d_3} \sum_{d_3=d_3'g_1g_2} |e_1f_1|^{-1}|f_2|^{-1/2}|g_2||m(g_2)|^{-1}|d_3'|^{-1/2}\left(\widehat{W}+\widehat{Y}^{n/3}\right).$$

Let now $k\in\mathbb{Z}_{>0}$ be such that $1/k<\varepsilon$ and write

$$d = h'h_1h_2^2\cdots h_k^kh_{k+1}, \quad (h',h_i)=(h_i,h_j)=1, h' \mid (N\Delta_{F_2})^\infty \text{ for } i\neq j,$$

where h_i is square-free for $i=1,\dots,k$ and h_{k+1} is $(k+1)$th-powerful. Recalling the definition of $m(g_2)$ and that d_3' is cube-full, it is then not hard to see that

$$\sum_{|d|\le\widehat{K}} \sum_{d=e_1f_1f_2^2d_3} \sum_{d_3=d_3'g_1g_2} |e_1f_1|^{-1}|f_2|^{-1/2}|g_2||m(g_2)|^{-1}|d_3'|^{-1/2}$$
$$\le \sum_{|d|\le\widehat{K}} |h_1\cdots h_k|^{-1} \sum_{d=ef_1f_2^2d_2} \sum_{d_3=d_3'g_1g_2} 1$$
$$\ll \sum_{|d|\le\widehat{K}} |d|^{\varepsilon}|h_1\cdots h_k|^{-1}$$
$$\ll \widehat{K}^{1/k+\varepsilon}.$$

Therefore, the definition of S and our choice of k implies after redefining ε that

$$\Sigma(W) \ll \widehat{Y}^{(n+5)/2+\varepsilon}\left(\widehat{W}+\widehat{Y}^{n/3}\right). \tag{6.8.21}$$

We now apply (6.8.21) to estimate $E_{1,b}(Y,\Theta_1,\Theta_2)$ in two different ways according to the size of Y. Let us begin by assuming that $\widehat{Y}\ge|P|^{1-\eta}$. In this case, we deduce from (6.5.1) and Lemma 6.5.1 that

$$E_{1,b}(Y,\Theta_1,\Theta_2) \le |N|^n\widehat{L}^{-n}\frac{|P|^n}{\widehat{Y}^n}\int_{|\underline{\theta}|=\widehat{\underline{\Theta}}}\int_{\mathbb{T}^n} \sum_{|\boldsymbol{v}-\boldsymbol{v}_0|<\widehat{V}\widehat{Z}^{-1/2}} \sum_{|d|\le\widehat{K}} \sum_{\substack{|r|=\widehat{Y}\\ d|r}} |S_{d\underline{c}_0,r,\boldsymbol{b},N}(\boldsymbol{v})|\mathrm{d}\boldsymbol{x}\mathrm{d}\underline{\theta},$$

where $\boldsymbol{v}_0 = -r_Mt^L(P^3\nabla F_1(\boldsymbol{x}_0+t^{-L}\boldsymbol{x}) + P^2\nabla F_2(\boldsymbol{x}_0+t^{-L}\boldsymbol{x}))$. We can now use (6.8.21) with $\widehat{W}=\widehat{V}\widehat{Z}^{-1/2}$ to deduce that

$$E_{1,b}(Y,\Theta_1,\Theta_2) \ll |P|^n\widehat{Y}^{5/2-n/2+\varepsilon}\widehat{\Theta}_1\widehat{\Theta}_2\left(\widehat{V}^n\widehat{Z}^{-n/2}+\widehat{Y}^{n/3}\right)$$
$$= \widehat{Y}^{5/2+n/2+\varepsilon}\widehat{\Theta}_1\widehat{\Theta}_2\widehat{Z}^{n/2} + |P|^n\widehat{Y}^{5/2-n/6+\varepsilon}\widehat{\Theta}_1\widehat{\Theta}_2$$
$$\ll |P|^{5n/6-5/3}\widehat{Y}^{1/2+\varepsilon} + |P|^{n-5/3+(1-\eta)(1/2-n/6+\varepsilon)}.$$

The first term is $O(|P|^{5n/6-5/6+\varepsilon})$, which is satisfactory as soon as $n\ge 26$. Moreover, $n-5/3+(1/2-n/6)<n-5$ if $n\ge 24$. In particular, the second term is sufficiently small provided ε is small.

If $\widehat{Y} \leq |P|^{1-\eta}$ we instead estimate the integral $I_{r_N}(\underline{\theta}, \boldsymbol{v})$ directly via Lemma 6.5.2 to obtain

$$E_{1,b}(Y, \Theta_1, \Theta_2) \ll \frac{|P|^n}{\widehat{Y}^n} \widehat{\Theta}_1 \widehat{\Theta}_2 \widehat{Z}^{1-n/2-\mu} \Sigma(\widehat{V}),$$

where

$$\mu = \begin{cases} 1/2 & \text{if } |P|\widehat{\Theta}_1 \ll \widehat{\Theta}_2, \\ 0 & \text{else.} \end{cases}$$

We can now apply (6.8.21) with $\widehat{W} = \widehat{V}$ to deduce that

$$\begin{aligned} E_{1,b}(Y, \Theta_1, \Theta_2) &\ll |P|^n \widehat{Y}^{5/2-n/2+\varepsilon} \widehat{\Theta}_1 \widehat{\Theta}_2 \widehat{Z}^{1-n/2-\mu} \left(\widehat{V}^n + \widehat{Y}^{n/3} \right) \\ &= \widehat{Y}^{5/2+n/2+\varepsilon} \widehat{\Theta}_1 \widehat{\Theta}_2 \widehat{Z}^{1+n/2-\mu} + |P|^{n-5} \widehat{Y}^{5/2-n/6+\varepsilon} \widehat{Z}^{3-n/2-\mu}. \end{aligned}$$

The second term is satisfactory by (6.8.10) if $n > 15$, so we may assume the first one dominates. Recall that we already dealt with the case when $\widehat{Y}\widehat{\Theta}_1 \geq |P|^{-\delta}$, where $\delta = 8(n-16)/(3n-24)$, since we assume (2) after (6.4.9) does not hold. So we may suppose the contrary is true. There are now two cases: First, we assume that $\widehat{\Theta}_2 \ll |P|\widehat{\Theta}_1$. In this situation we have $\widehat{Z} \ll |P|^{3-\delta} \widehat{Y}^{-1}$ and $\mu = 0$, so that

$$E_{1,b}(Y, \Theta_1, \Theta_2) \ll |P|^{(3-\delta)(1+n/2)+1-2\delta} \widehat{Y}^{-1/2+\varepsilon}.$$

A rather involved computation or a check with a computer algebra program verifies that $(3-\delta)(1+n/2) + 1 - 2\delta < n - 5$ if $n \geq 25$, which is satisfactory.

The only case that remains is when $|P|\widehat{\Theta}_1 \ll \widehat{\Theta}_2$, in which case $\mu = 1/2$ and hence

$$E_{1,b}(Y, \Theta_1, \Theta_2) \ll |P|^{5n/6+5/6-5/3} \widehat{Y}^{\varepsilon},$$

which is satisfactory for $n > 25$.

CHAPTER 7

Rational points on del Pezzo surfaces of low degree over global fields

The content of this chapter is based on the joint work [89] with Hochfilzer.

7.1 Introduction

Recall that if X is a Fano variety over a global field and $U \subset X(K)$ is a suitable complement of a thin set, then as explained in Section 2.3 Manin's conjecture, predicts and asymptotic formula of the shape

$$N_U(B) \sim cB(\log B)^{\mathrm{rk}\,\mathrm{Pic}(X)-1},$$

where $N_U(B)$ counts rational points of X lying on U of anti-canonical height at most B. While Manin's conjecture for curves is well understood, the case of surfaces is already much more mysterious. A surface X that is Fano is called a *del Pezzo surface* and is classified by the degree $d = K_X^2$, which satisfies $1 \leq d \leq 9$. Moreover, it is generally believed that for $d \geq 2$ one can take U to be the complement of the exceptional curves in Manin's conjecture.

If the degree satisfies $6 \leq d \leq 9$ then any del Pezzo surface is a toric variety. Thanks to work of Batyrev and Tschinkel [6] for number fields and Bourqui [22, 21] in positive characteristic, Manin's conjecture is therefore known for all del Pezzo surfaces of degree $6 \leq d \leq 9$. If $d \leq 5$, much less is known. One of the notable exceptions is de la Bretèche's work [25], in which he proved Manin's conjecture for split del Pezzo surfaces of degree 5 over $\mathbb{Q}$. Recently, with a different approach Browning [36] obtained the same result with a better error term. In addition, de la Bretèche and Fouvry [26] verified Manin's conjecture for del Pezzo surfaces of degree 5 over $\mathbb{Q}$ that are the blow-up of a pair of points that are defined over $\mathbb{Q}$ and a pair of conjugate points over $\mathbb{Q}(i)$. When $d = 4$, the tour de force [24] of de la Bretèche and Browning provides us with the only example of a del Pezzo surface of degree 4 over $\mathbb{Q}$ for which we know Manin's conjecture. These results already reflect the guiding principle that the arithmetic of del Pezzo surfaces becomes harder to understand the smaller the degree is. In particular, for $1 \leq d \leq 3$ we do not know the truth of Manin's conjecture for any single example of a del Pezzo surface and in fact, if the ground field is not $\mathbb{Q}$, we do not even know it for any del Pezzo surface of degree $1 \leq d \leq 5$.

While an asymptotic formula remains largely elusive for small degrees, even providing upper bounds remains a substantial challenge in itself. For $2 \leq d \leq 5$, currently the best upper bounds are found in work of Salberger [169], in which he showed that $N_U(B) \ll_X B^{3/\sqrt{d}+\varepsilon}$ when the ground field is $\mathbb{Q}$. When X is a cubic surface over $\mathbb{Q}$, he [168] is able to improve his estimate to $N_U(B) \ll B^{12/7+\varepsilon}$. Again working over $\mathbb{Q}$, it follows from combining work of Bhargava et al. [12] on pointwise bound for the ranks of elliptic curves with work of Helfgott and Venkatesh [107] on integral points of elliptic curves that $N_U(B) \ll_X B^{2.87}$ when X is a del Pezzo surface of degree 1. In his thesis, Tschinkel [198] proved that $N_U(B) \ll B^{1+\varepsilon}$ for any split quintic del Pezzo surface over any number field. Moreover, when $K = \mathbb{Q}$ and $d = 3$ Heath Brown [103] succeeded in showing that $N_U(B) \ll_X B^{3/2+\varepsilon}$ and the underlying cubic forms is diagonal conditional on certain conjectures for Hasse-Weil L-functions associated to a family of cubic threefolds. The authors [88] showed that the same upper bounds holds unconditionally over $\mathbb{F}_q(t)$ when $\mathrm{char}(\mathbb{F}_q) > 3$.

Let us now consider the following conjecture and its consequences for Manin's conjecture for del Pezzo surfaces.

Conjecture 7.1.1. *Let E be an elliptic curve over a global field K with $\mathrm{char}(K) \neq 2, 3$ and $\mathscr{C}_E$ its conductor. Then*

$$\mathrm{rk}\, E = o(\log \mathrm{N}(\mathscr{C}_E)) \quad \textit{as } \mathrm{N}(\mathscr{C}_E) \to \infty,$$

where $\mathrm{N}(\mathscr{C}_E)$ denotes the norm of $\mathscr{C}_E$.

Mestre [144] showed that Conjecture 7.1.1 is implied by the Birch and Swinnerton-Dyer conjecture. Moreover, 2-descent shows that we always have $\mathrm{rk}\, E = O(\log \mathrm{N}(\mathscr{C}_E))$. In fact, when $\mathrm{char}(K) > 3$, the conjecture was proven by Brumer [46]. The relevance of this conjecture is that the pullback of a generic hyperplane under the rational map $X \to \mathbb{P}^d$ induced by the anti-canonical divisor is generically a smooth genus 1 curve. Assuming Conjecture 7.1.1, Heath-Brown [102] obtained *uniform* upper bounds for the number of rational points of bounded height on planar elliptic curves and showed that $N_U(B) \ll_X B^{4/3+\varepsilon}$ for any cubic surface over $K = \mathbb{Q}$. Our first main result extends this to any del Pezzo surface of degree at most 5 and to any global field whose characteristic exceeds 3 if it is positive.

Theorem 7.1.2. *Let X be a del Pezzo surface of degree $1 \leq d \leq 5$ over a global field K with $\mathrm{char}(K) \neq 2, 3$. Then*

$$N_U(B) \ll_X B^{1+1/d+\varepsilon},$$

unconditionally when $\mathrm{char}(K) > 3$ and assuming Conjecture 7.1.1 when $\mathrm{char}(K) = 0$. Moreover, when $d = 1$ the implied constant is independent of X.

Using an approach based on exponential sums, Bonolis and Browning [18] obtained the estimate $N_U(B) \ll B^{3-1/20}(\log B)^2$ for $d = 1$ over $\mathbb{Q}$. While this estimate is weaker than the one obtained by Bhargava et al. [12], it has the advantage that it is uniform with respect to the underlying surface. The upper bound from Theorem 7.1.2 for $d = 1$ shares the same uniformity.

Recall that if there is a dominant K-morphism $X \to \mathbb{P}^1$ such that all fibers are plane conics, we say that X admits a *conic bundle structure*. When a del Pezzo surface comes with such extra structure, one can use it to get better control over the number of rational points. In particular, Heath-Brown [101] has shown that $N_U(B) \ll_X B^{4/3+\varepsilon}$ for $d = 3$ over $\mathbb{Q}$ when X has three coplanar lines defined over $\mathbb{Q}$, which give rise to three inequivalent conic bundle structures. This result was later generalised to number fields by Broberg [27]. Moreover, Browning and Sofos [32] proved that

$$B(\log B)^{\mathrm{rk}\,\mathrm{Pic}(X)-1} \ll_X N_U(B) \ll_X B(\log B)^{\mathrm{rk}\,\mathrm{Pic}(X)-1}$$

when $d = 4$ for $K = \mathbb{Q}$ assuming that $X(K) \neq \emptyset$. Building on ideas of Salberger announced at the conference "Higher dimensional varieties and rational points" at Budapest in 2001, work of Browning an Swarbrick-Jones [43] gives $N_U(B) \ll_X B^{1+\varepsilon}$ when $d = 4$ and K is a number field. When $d = 2$ and $K = \mathbb{Q}$, Salberger announced at the conference "Géomètrie arithmétique et variétés rationnelles" at Luminy in 2007 the result that $N_U(B) \ll_X B^{11/6+\varepsilon}$ provided X is split. We are now ready to reveal our second main result.

Theorem 7.1.3. *Let X be a del Pezzo surface of degree 4 or 5 over a global field K of characteristic $\mathrm{char}(K) \neq 2$ with a conic bundle structure. Then $N_U(B) \ll_X B^{1+\varepsilon}$ for an effectively computable implied constant.*

This result is new when $d = 4$ and $\mathrm{char}(K) > 0$ and new for any global field when $d = 5$. We also note that the constants in Theorem 7.1.3 are all effectively computable, which is in contrast to the result due to Browning and Swarbrick-Jones [43]. This is because we avoid an application of the Thue–Siegel–Roth theorem, which we note is in general not true in positive characteristic (cf. [154]).

The results listed so far are by no means exhaustive. In particular, when one considers *singular* del Pezzo surfaces, one can even obtain an asymptotic formula when the degree is 2 or 3, but we restrict to the more difficult case of smooth surfaces in this work. Moreover, in [77] Frei, Loughran and Sofos studied lower bounds for del Pezzo surfaces with a conic bundle structure and showed that $N_U(B) \gg B(\log B)^{\mathrm{rk}\,\mathrm{Pic}(X)-1}$ for del Pezzo surfaces over number fields whose rank of the Picard group is sufficiently large with respect to d.

7.1.1 Outline.

The basic idea underlying the proof of Theorem 7.1.2 is simple: a generic hyperplane section of a del Pezzo surface is a non-singular genus 1 curve and the rank growth hypothesis allows us to obtain uniform upper bounds for the number of rational points of bounded height on elliptic curves. In fact, in Section 7.4 we establish the following result.

Proposition 7.1.4. *Let $E \subset \mathbb{P}^n$ be a non-singular genus 1 curve of degree d over a global field K with $\mathrm{char}(K) \neq 2, 3$. Assuming that Conjecture 7.1.1 holds when $\mathrm{char}(K) = 0$, we have*

$$\#\{x \in E(K) \colon H(x) < B\} \ll B^{\varepsilon},$$

where the implied constant only depends on d, n, K and ε and $H \colon \mathbb{P}^n(K) \to \mathbb{R}_{>0}$ is the usual height on projective space.

Heath-Brown [102] previously proved Proposition 7.1.4 for plane elliptic curves over $\mathbb{Q}$. However, his results still had a dependence on the height of the elliptic curve, which we were able to remove. Proposition 7.1.4 alone is not sufficient to prove Theorem 7.1.2. The hyperplane sections can also be singular and we need uniform upper bounds for the number of rational points of bounded height on curves. The determinant method developed by Heath-Brown [104] has been used by many authors over the last two decades to establish such estimates. Moreover, the recent work of Paredes and Sasyk [155] extended these results using the global determinant method due to Salberger [169] to any global field, which enables us to work over arbitrary global fields. These results will also prove vital when bounding the number of degenerate hyperplane sections, which correspond to rational points on the dual variety of X.

When $d = 4$ or $d = 5$, even the bounds coming from the determinant method are not strong enough when the hyperplane section contains an irreducible component of degree $d - 1$, as we need an additional saving with respect to the height of the hyperplane. To overcome this difficulty, we show that such curves are in fact rational and provide upper bounds by exhibiting a uniform parameterisation of the rational points.

Finally, to prove Theorem 7.1.3, we use uniform upper bounds for rational points on conics of Browning and Heath-Brown [28], that we transfer to the setting of positive characteristic. For quintic del Pezzo surface we combine them with techniques from the geometry of numbers in a similar fashion as Bonolis, Browning and Huang [19]. In particular, in Section 7.3.1 we prove a result regarding the number of lattice points in a box depending on the successive minima of the lattice for all global fields, which is likely to be useful for applications outside the context of this paper.

7.1.2 Conventions

In Sections 7.5 and 7.6 the letter A denotes an arbitrarily large constant, whose value can change from line to the next. In particular, we may write expressions like $B^{2A} \ll B^A$, which has the advantage of avoiding introducing notation like $A', A'', A''', \dots$. Finally, we shall use the notation $\alpha \sim R$ to indicate that α lies in the dyadic interval $(R, 2R]$.

7.2 Background

7.2.1 Geometry

Throughout this work a *variety* over a field K is a separated K-scheme of finite type. In this section K denotes an arbitrary field.

del Pezzo surfaces

In this subsection we review some of the basic geometric properties of del Pezzo surfaces, which can for example be found in the book by Manin [137]. A del Pezzo surface over a field K is a smooth, projective and geometrically integral surface over K whose anti-canonical divisor $-K_X$ is ample. Let $\overline{X} = X \times K^{\mathrm{sep}}$, where K^{sep} denotes a separable closure of K. The geometric Picard group $\mathrm{Pic}(\overline{X})$ is a finitely generated $\mathbb{Z}$-module and since X is smooth, we can identify $\mathrm{Pic}(\overline{X})$ with the class group of Weil divisors on X. Therefore, $\mathrm{Pic}(\overline{X})$ comes with a symmetric bilinear intersection pairing $(\,\cdot\,)\colon \mathrm{Pic}(\overline{X}) \times \mathrm{Pic}(\overline{X}) \to \mathbb{Z}$. If $C \in \mathrm{Pic}(\overline{X})$, then by abuse of notation we shall write C^2 for $C \cdot C$. The degree of X is defined to be K_X^2 and satisfies $1 \leq d \leq 9$.

Definition 7.2.1. Let $C \subset \overline{X}$ be an irreducible curve. Then we say that C is *exceptional* if $C^2 = C \cdot K_X = -1$.

By the adjunction formula, an exceptional curve has arithmetic genus 0 and hence is $\overline{K}$-isomorphic to $\mathbb{P}^1$. There are at most finitely many exceptional curves on a del Pezzo surface and their precise number is given in Table 7.1. The anti-canonical divisor induces a rational map $X \to \mathbb{P}^d$, which is in fact a morphism for $d \geq 2$. When $d \geq 3$ the map is an embedding and realises X as a non-degenerate surface of degree d in $\mathbb{P}^d$. If $C \subset \overline{X}$ is a geometrically connected curve of arithmetic genus 0 with $C \cdot K_X = -2$, we say that C is a *conic*.

Definition 7.2.2. We say that X admits a conic bundle structure over K if there is a dominant K-morphism $X \to \mathbb{P}^1$ all of whose fibers are plane conics.

It follows from Lemma 5.1 in [77] that X admits a conic bundle structure over K if and only if it contains a conic defined over K. In [77] the authors assume the ground field to be perfect in their statement; however, an inspection of the proof reveals that this assumption is not used.

Dual varieties

Let $X \subset \mathbb{P}^n$ be a variety and denote by $\widehat{\mathbb{P}}^n$ the dual projective space parameterising hyperplanes in $\mathbb{P}^n$. We define the *conormal variety* of X to be

$$Z(X) = \overline{\{(x, H) \in X_{sm} \times \widehat{\mathbb{P}}^n \colon T_x X \subset H\}},$$

equipped with the reduced scheme structure, where $T_x X \subset \mathbb{P}^n$ denotes the embedded tangent space of X at x and X_{sm} denotes the smooth locus of X. Let $\phi\colon Z(X) \to \widehat{\mathbb{P}}^n$ denote the projection onto the second factor. Then the image $\phi(Z(X)) \subset \widehat{\mathbb{P}}^n$ is a variety, which is called the *dual variety* of X and denoted by X^*. We always have the inequality $\dim(X^*) \leq n - 1$. We say that X is *reflexive* if $Z(X) = Z(X^*)$ under the natural identification $\mathbb{P}^n = \widehat{\widehat{\mathbb{P}}}^n$.

If X is smooth, then a hyperplane $H \in \widehat{\mathbb{P}}^n$ has singular intersection with X if and only if $H \in X^*$, while if X is not smooth and $X \cap H$ is singular, then $H \in X^*$ or H intersects the singular locus of X.

Our goal of this section is to compute $\deg(X^*)$ when X is a del Pezzo surface. To do so, we need to recall the properties of Chern classes as found in [79]. For a smooth variety X over K, let $CH(X)$ denote the Chow ring whose kth graded piece A^k is the group of cycles of codimension k modulo rational equivalence and multiplication is given by the intersection pairing. Associated to any vector bundle $\mathcal{F}$ over X are the Chern classes $c_i(\mathcal{F}) \in A^i$ satisfying

(1) $c_0(\mathcal{F}) = 1$,

(2) $c_i(\mathcal{F}) = 0$ if $i > \mathrm{rk}(\mathcal{F})$,

(3) if $\mathcal{F}$ is of rank r, then $c_1(\mathcal{F}) = c_1(\wedge^r \mathcal{F})$.

Moreover, if X is a smooth surface and $\mathcal{L}$ is a line bundle on X corresponding to a Weil divisor D, then $\deg(c_1(\mathcal{L}) \cap E) = D \cdot E$ for any Weil divisor E of X, where $\cdot$ is the usual intersection pairing on the class group of Weil divisors.

degree	9	8	7	6	5	4	3	2	1
	0	0 or 1	3	6	10	16	27	56	240

Table 7.1: Number of exceptional curves

Let $\mathcal{T}_X$ be the tangent bundle on a smooth variety $X \subset \mathbb{P}^n$ of dimension k. Then we define the *delta invariants* of X to be

$$\delta_i(X) := \sum_{j=i}^{k}(-1)^{k-j}\binom{j+1}{i+1}\deg(c_{k-j}(X)),$$

where $c_i(X) = c_i(\mathcal{T}_X)$ and $\deg(c_i(X)) = \deg(c_i(X) \cap H^{k-i})$ with H the class of a hyperplane section on X. We then have the following result due to Holme [110, Theorem 3.4].

Proposition 7.2.3. *Let $X \subset \mathbb{P}^n$ be a smooth variety and suppose that r_0 is such that $\delta_0(X) = \cdots = \delta_{r_0-1}(X) = 0$, but $\delta_{r_0}(X) \neq 0$. Then $\dim(X^*) = n-1-r_0$ and X is reflexive if and only if $\deg(X^*) = \delta_{r_0}(X)$.*

Remark. In fact Holme's result is even valid for singular varieties. However, in this case the delta invariants may differ from the definition above.

We now have everything at hand to compute the degree of the dual variety of a del Pezzo surface.

Proposition 7.2.4. *Let $X \subset \mathbb{P}^d$ be a del Pezzo surface of degree $d \geq 3$ embedded anti-canonically over a field K with $\mathrm{char}(K) \neq 2, 3$. Then $X^* \subset \widehat{\mathbb{P}}^d$ is a hypersurface with $\deg(X^*) = 12$.*

Proof. Let us first compute $\delta_0(X)$. Since X is embedded anti-canonically, the definition of $\delta_0(X)$ gives

$$\delta_0(X) = \deg(c_2(X)) - 2\deg(c_1(X) \cap (-K_X)) + 3(-K_X)^2. \tag{7.2.1}$$

We have by definition $(-K_X)^2 = d$. Moreover, as $\mathcal{T}_X$ has rank 2, we have $c_1(X) = c_1(\wedge^2\mathcal{T}_X) = c_1(-K_X)$ and hence $\deg(c_1(X) \cap (-K_X)) = d$ as well. It thus remains to compute $\deg(c_2(X))$. Let $\chi(X, \mathcal{O}_X)$ be the Euler characteristic of the structure sheaf on X. By Lemma 3.2.1 of Kollár [124], we have $\chi(X, \mathcal{O}_X) = 1$. Moreover, Noether's formula (see Example 15.2.2 of Fulton [79]) gives

$$\chi(X, \mathcal{O}_X) = ((-K_X)^2 + \deg(c_2(X))/12,$$

which implies $\deg(c_2(X)) = 12 - d$. Once combined with (7.2.1), we obtain $\delta_0(X) = 12$.

In the light of Proposition 7.2.3 it thus remains to show that X is reflexive. To do so, we make use of the Monge–Segre–Wallace criterion in the form given by Kleiman [122, (4) Theorem], which asserts that X is reflexive if and only if the map $\phi\colon Z(X) \to X^*$ is separable. The map ϕ is known to be finite and by definition its degree $\deg(\phi)$ is the degree of the induced extension of function fields $K(X^*)/K(Z(X))$. By the last equation on page 152 of Holme [110], we have $\deg(\phi)\deg(X^*) = \delta_0(X)$. In particular, $\deg(\phi) \mid 12$. As we assume that $\mathrm{char}(K) \neq 2, 3$, this automatically implies that $K(X^*)/K(Z(X))$ is separable and hence X is reflexive by the Monge–Segre–Wallace criterion. □

Remark. At least the condition $\mathrm{char}(K) \neq 2$ is necessary in the last proposition. Indeed, it follows for example from Lemma 4.1 of the authors' work [88] that if $K = \mathbb{F}_q(t)^{\mathsf{sep}}$ with $\mathrm{char}(\mathbb{F}_q) = 2$, then the dual variety of the Fermat cubic surface is again a Fermat cubic surface.

7.2.2 Algebraic number theory

We call K a *global field* if it is a finite extension of $\mathbb{Q}$ or the function field of a curve over a finite field. When K is a function field of positive characteristic p, suppose $\mathbb{F}_q$ is the field of constants of K. We can then always find an element $t \in K$ that is transcendental over $\mathbb{F}_q$ and such that $K/\mathbb{F}_q(t)$ is separable. Note that the choice of t is not unique, but this is irrelevant to us. We shall then write $k = \mathbb{F}_q(t)$ and if K is a number field, we write $k = \mathbb{Q}$ and in either case define $d_K = [K\colon k]$.

The ring of integers $\mathcal{O}_K$ of K is by definition the integral closure of $\mathbb{Z}$ and $\mathbb{F}_q[t]$ respectively. By definition, a non-archimedean place $\mathfrak{p}$ of K is a discrete valuation ring $\mathcal{O}_{(\mathfrak{p})} \subset K$ with field of fractions K with the additional constraint $\mathbb{F}_q \subset \mathcal{O}_{(\mathfrak{p})}$ when $\mathrm{char}(K) > 0$. A non-archimedean place corresponds to an embedding $K \to \mathbb{C}$ and only exists in characteristic 0. We then define $\Omega_{K,\infty}$ to be the set of places lying above the infinite place t^{-1} in $\mathbb{F}_q(t)$ when $\mathrm{char}(K) > 0$ and to be the set of archimedean places when $\mathrm{char}(K) = 0$. In either case Ω_K denotes the set of all places of K and $\Omega_{K,f} = \Omega_K \setminus \Omega_{K,\infty}$ the set of finite places. In addition, we let s_K be the cardinality of $\Omega_{K,\infty}$.

If ν is a place of K and μ the corresponding place of k lying below it, we let K_ν and k_μ be the completions of K and k with respect to ν and μ respectively and define the local degree $d_\nu := [K_\nu\colon k_\mu]$. If ν is non-archimedean, we let $\mathcal{O}_\nu \subset K_\nu$ be the ring of integers and $\mathfrak{m}_\nu$ its maximal ideal. For any place ν, we define an absolute value on K via

$$|x|_\nu := \begin{cases} |x|_\infty^{d_\nu} & \text{if } \nu \text{ is archimedean,} \\ \#(\mathcal{O}_\nu/\mathfrak{m}_\nu)^{-v_\nu(x)} & \text{if } \nu \text{ is non-archimedean,} \end{cases}$$

where $|\cdot|_\infty$ denotes the usual absolute value on $\mathbb{C}$ and v_ν the normalised valuation on K induced by ν. Via the embedding $K \to K_\nu$ corresponding to the place ν this gives rise to an absolute value on K.

Remark. Note that if ν is a complex place, then strictly speaking $|\cdot|_\nu$ is not an absolute value in the usual sense, as it does not satisfy the triangle inequality. However, it still satisfies $|x+y|_\nu \leq 4\max\{|x|_\nu, |y|_\nu\}$.

When $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$, we define its norm to be $\mathrm{N}(\mathfrak{p}) := \#(\mathcal{O}_K/\mathfrak{p})$ and extend it multiplicatively to all fractional ideals of $\mathcal{O}_K$. If $\alpha_1, \dots, \alpha_m \in K$, we let $\langle \alpha_1, \dots, \alpha_m \rangle$ be the fractional ideal generated by $\alpha_1, \dots, \alpha_m$. By abuse of notation, we shall then write $\mathrm{N}(\alpha_1, \dots, \alpha_m)$ instead of $\mathrm{N}(\langle \alpha_1, \dots, \alpha_m \rangle)$. Similarly, if ν is a non-archimedean place we also write $\mathrm{N}(\cdot)$ for the ideal norm on $\mathcal{O}_\nu$. When K is a function field, a divisor is by definition a formal sum $\mathfrak{a} = \sum_{\nu \in \Omega_K} e_\nu \nu$, where $e_\nu \in \mathbb{Z}$ is non-zero for at most finitely many ν. We then define $\deg(\mathfrak{a}) = \sum e_\nu \deg(\nu)$, where $\deg(\nu) = [(\mathcal{O}_\nu/\mathfrak{m}_\nu)\colon \mathbb{F}_q]$ is the degree of the residue field extension. We shall then also write $\mathrm{N}(\mathfrak{a}) = q^{\deg(\mathfrak{a})}$ and refer to it as the norm of $\mathfrak{a}$.

Further, given K_ν there exists a standard additive character $\psi_\nu\colon K_\nu \to \mathbb{C}^\times$ as defined in [161, Chapter 7]. In particular, these have the property that when ν is a finite place then $\mathcal{O}_\nu \subset K_\nu$ is the maximal subgroup on which ψ_ν acts trivially. As a result we obtain the following character orthogonality relation. The proof is standard, so we omit it here.

Lemma 7.2.5. *Let ν be a finite place of K, let π be a uniformizer of K_ν and let $r \geq 1$ be an integer. If $a \in \mathcal{O}_\nu$ we have*

$$\frac{1}{\mathrm{N}(\pi)^r} \sum_{x \in \mathcal{O}_\nu/(\pi^r)} \psi_\nu\left(\frac{ax}{\pi^r}\right) = \begin{cases} 1 & \text{if } a \in (\pi)^r, \\ 0 & \text{otherwise.} \end{cases}$$

7.2.3 Height functions

We can construct an exponential height on $\mathbb{P}^n(K)$ via

$$H(\boldsymbol{x}) = \prod_{\nu \in \Omega_K} \max_{0 \le i \le n} |x_i|_\nu$$

whenever $\boldsymbol{x} = [x_0, \dots, x_n] \in \mathbb{P}^n(K)$. Note that our normalisation of the absolute values implies that the product formula $\prod_{\nu \in \Omega_K} |x|_\nu = 1$ holds for any $x \in K$, so that the height on $\mathbb{P}^n(K)$ is indeed well defined. Observe that if $\boldsymbol{x} = [x_0, \dots, x_n] \in \mathbb{P}^n(K)$ with $(x_0, \dots, x_n) \in \mathcal{O}_K^{n+1}$, then

$$H(\boldsymbol{x}) = \frac{1}{\mathrm{N}(x_0, \dots, x_n)} \prod_{\nu \in \Omega_{K,\infty}} \max_{0 \le i \le n} |x_i|_\nu.$$

In addition, for $\boldsymbol{x} \in \mathcal{O}_K^{n+1}$ we define the norms

$$\|\boldsymbol{x}\| := \max_{0 \le i \le n} \max_{\nu \mid \infty} |x_i|_\nu \quad \text{and} \quad \|\boldsymbol{x}\|_\infty := \prod_{\nu \mid \infty} \max_{0 \le i \le n} |x_i|_\nu.$$

Given $\boldsymbol{R} = (R_\nu)_{\nu \mid \infty} \in \mathbb{R}_{>0}^{s_K}$ we define $|\boldsymbol{R}| = \prod_{\nu \mid \infty} R_\nu$ and

$$L(\boldsymbol{R}) = \{x \in \mathcal{O}_K \colon |x|_\nu \le R_\nu \text{ for all } \nu \mid \infty\}.$$

The following results are all standard over number fields and should be well known over function fields. Due to a lack of statements in the literature, we provide full proofs in the case of function fields.

Lemma 7.2.6. *Let $\boldsymbol{R} = (R_\nu)_{\nu \mid \infty} \in \mathbb{R}_{>0}^{s_K}$ and let $\mathfrak{a} \subset \mathcal{O}_K$ be an integral ideal. We have*

$$|\boldsymbol{R}| \mathrm{N}(\mathfrak{a})^{-1} \ll_K \#(L(\boldsymbol{R}) \cap \mathfrak{a}) \ll_K \max\{1, |\boldsymbol{R}| \mathrm{N}(\mathfrak{a})^{-1}\}.$$

Proof. If K is a number field, this is Theorem 0 in Chapter V§1 of Lang [128]. We may therefore assume that K is a function field. Without loss of generality, we may assume that $R_\nu = q^{r_\nu}$ for some $r_\nu \in \mathbb{Z}$, where $q^{r_\nu} \in |K_\nu|_\nu$. In particular, we may choose elements $a_\nu \in K_\nu$ such that $|a_\nu|_\nu = R_\nu$ for all $\nu \mid \infty$. The condition $x \in (L(\boldsymbol{R}) \cap \mathfrak{a})$ is then equivalent to

$$v_\nu(x) \ge -v_\nu(a_\nu) \text{ for all } \nu \mid \infty \quad \text{and} \quad v_\mathfrak{p}(x) \ge v_\mathfrak{p}(\mathfrak{a}) \text{ for all } \mathfrak{p} \mid \mathfrak{a}. \qquad (7.2.2)$$

Let C be the smooth projective curve associated to the function field K and let g be its genus. We then define the divisor $D \in \mathrm{Pic}(C)$ to be $D = \sum_{\nu \mid \infty} v_\nu(a_\nu)\nu - \sum_{\mathfrak{p} \nmid \infty} v_\mathfrak{p}(\mathfrak{a})\mathfrak{p}$. For $x \in K$, denote by $(x) = \sum_\nu v_\nu(x)\nu$ the associated divisor inside $\mathrm{Pic}(C)$ and define the Riemann–Roch space

$$L(D) = \{x \in K^* \colon (x) + D \ge 0\} \cup \{0\},$$

where the notation $(x) + D \ge 0$ means that the divisor is effective. It is then clear from (7.2.2) that $L(D) = L(\boldsymbol{R}) \cap \mathfrak{a}$. The Riemann-Roch space defines an $\mathbb{F}_q$-vector space and we denote its dimension by $\ell(D)$.

Let $\deg \colon \mathrm{Pic}(C) \to \mathbb{Z}$ be the degree map, which for the divisor D defined above satisfies $\deg(D) = \sum_{\nu \mid \infty} \log_q(R_\nu) - \log_q(\mathrm{N}(\mathfrak{a}))$. If $\deg(D) \le 2g - 2$, then as the degree map has finite fibers whose cardinality is given by the class number of K and $L(D)$ only depends on the class of D in $\mathrm{Pic}(C)$, there are only $O(1)$ possibilities for $\ell(D)$. We may therefore assume that $\deg(D) > 2g - 2$ from now on. In this case the Riemann-Roch Theorem in the form given by Rosen [164, Corollary 4] tells us that $\ell(D) = \deg(D) - g + 1$, so that

$$\#(L(\boldsymbol{R}) \cap \mathfrak{a}) = q^{\ell(D)} = q^{1-g} |\boldsymbol{R}| \mathrm{N}(\mathfrak{a})^{-1},$$

which completes the proof. □

Fix $\mathfrak{a}_1, \dots, \mathfrak{a}_h \subset \mathcal{O}_K$ to be a full set of representatives in the class group of K once and for all. Whenever we indicate that an implied constant depends on K, then it is implicitly allowed to also depend on our choice of the ideal class group representatives.

Unless K is a principal ideal domain one cannot longer ensure that an element in $\mathbb{P}^n(K)$ has a representative in $\mathcal{O}_K^{n+1}$ such that the coordinate entries generate $\mathcal{O}_K$. Instead in this more general setting the set of primitive vectors is replaced by the set

$$Z_n' \coloneqq \left\{ \boldsymbol{x} \in \mathcal{O}_K^{n+1} \setminus \{\boldsymbol{0}\} \colon (x_0, \dots, x_n) = \mathfrak{a}_i \text{ for some } i = 1, \dots, h \right\}.$$

Note that if $\boldsymbol{x} \in Z_n'$ is a representative for an element $x \in \mathbb{P}^n(K)$ then we have $H(x) \asymp \|\boldsymbol{x}\|_\infty$. Since $\|u\|_\infty = 1$ for any unit $u \in \mathcal{O}_K$ the set $\{\boldsymbol{x} \in Z_n' \colon \|\boldsymbol{x}\|_\infty \leq B\}$ is potentially infinite.

Lemma 7.2.7. *Let $\lambda_\nu \in \mathbb{R}_{>0}$ be given for $\nu \mid \infty$. Then there exists a unit $u \in \mathcal{O}_K^\times$ and $t \in \mathbb{R}$ such that*

$$\lambda_\nu \asymp_K \frac{|u|_\nu}{t}.$$

Proof. If K is a number field, this is proved in the Lemma on p. 187 of [182], so that we may assume that K is a function field from now on. Upon defining $\lambda_\nu' = \log_q \lambda_\nu$, the statement of the lemma is equivalent to

$$|\lambda_\nu' + \log_q(t) - \log_q(u)| \leq C,$$

for some fixed constant $C = C(K) > 0$. If we define

$$\phi \colon \mathcal{O}_K^\times \to \mathbb{R}^{s_K}, \quad u \mapsto (\log_q |u|_\nu)_{\nu | \infty},$$

then $\Gamma = \phi(\mathcal{O}_K^\times) \oplus \mathbf{1}\mathbb{Z}$ is a lattice of rank s_K, where $\mathbf{1} = (1, \dots, , 1)$. Thus the quotient $\mathbb{R}^{s_K}/\Gamma$ is compact and we can find a unit $u \in \mathcal{O}_K^\times$ and $t' \in \mathbb{Z}$ such that

$$|\lambda_\nu' + t' - \log_q(u)| \leq C,$$

where C only depends on K. The statement now follows upon setting $t = q^{t'}$. □

Lemma 7.2.8. *There exist constants $c_1, c_2 > 0$ depending on K such that every member of $\mathbb{P}^n(K)$ has a representative $\boldsymbol{x} \in Z_n'$ such that $c_1 \|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_\infty^{1/s_K} \leq c_2 \|\boldsymbol{x}\|$.*

Proof. Let $\boldsymbol{y} \in \mathbb{P}^n(K)$ and suppose $\boldsymbol{x}' = (x_0', \dots, x_n') \in Z_n'$ is a representative for $\boldsymbol{y}$. Define $\lambda_\nu = \max\{|x_0'|_\nu, \dots, |x_n'|_\nu\}$. Then by Lemma 7.2.7 we can find a unit $u \in \mathcal{O}_K^\times$ and $t \in \mathbb{R}$ such that $\lambda_\nu \asymp_K |u|_\nu / t$ for all $\nu \mid \infty$. Taking the product over all infinite places, we deduce that

$$1 \asymp \prod_{\nu | \infty} t \max\{|x_0'|_\nu, \dots, |x_n'|_\nu\} = t^{s_K} \|\boldsymbol{x}'\|_\infty,$$

which implies that $t^{-1} \asymp \|\boldsymbol{x}'\|_\infty^{1/s_K}$. Moreover, if we define $\boldsymbol{x} = u^{-1}\boldsymbol{x}'$, then we have

$$\max\{|x_0|_\nu, \dots, |x_n|_\nu\} \asymp t^{-1} \asymp \|\boldsymbol{x}\|_\infty^{1/s_K},$$

since $\|\boldsymbol{x}\|_\infty = \|\boldsymbol{x}'\|_\infty$. As $\boldsymbol{x}$ is also a representative for $\boldsymbol{y}$ in $\mathbb{P}^n(K)$, the result follows. □

Let c_1, c_2 be the constants from Lemma 7.2.8. We then define

$$Z_n \coloneqq \{\boldsymbol{x} \in Z'_n \colon c_1\|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_\infty^{1/s_K} \leq c_2\|\boldsymbol{x}\|\},$$

so that in particular every element in $\mathbb{P}^n(K)$ has a representative in Z_n.

For $x \in K$, we define the affine height

$$h(x) \coloneqq H(1, x).$$

Lemma 7.2.9. *Let K be a global field. Then*

$$\{u \in \mathcal{O}_K^\times \colon h(u) \leq B\} \ll_{d_K} (\log B)^{s_K}.$$

Proof. If K is a number field this is proved by Broberg [27, Proposition 4] and so we may assume that K is a function field. Let $\mathbb{F}_q$ be the field of constants of K. If we define

$$\phi \colon \mathcal{O}_K^\times \to \mathbb{Z}^{s_K}, \quad u \mapsto (\log_q |u|_\nu)_{\nu|\infty},$$

then $\phi(\mathcal{O}_K^\times)$ is a lattice of rank $s_K - 1$ inside $\mathbb{Z}^{s_K}$ and ϕ is a group homorphism with kernel $\mathbb{F}_q^\times$. Note that for any $u \in \mathcal{O}_K^\times$ we have $h(u) = h(u^{-1})$, so that if $h(u) \leq B$ holds, then we must have $\max_{\nu|\infty} |\log_q |u|_\nu| \leq \log_q B$. It follows that the number of units in question is $O((\log B)^{s_K})$ as desired. □

Lemma 7.2.10. *Let L/K be an extension of degree d and let $y \in \mathcal{O}_K \setminus \{0\}$. Then*

$$\#\{(y_1, y_2) \in \mathcal{O}_L^2 \colon y_1y_2 = y \text{ and } |y_i|_\nu \leq R \text{ for all } \nu \in \Omega_{K,\infty}, i = 1, 2\} \ll_{d,K} (R\mathrm{N}(y))^\varepsilon,$$

where $|\cdot|_\nu$ is extended uniquely to L.

Proof. Note that by the usual divisor bound, there are $O(\mathrm{N}(y)^\varepsilon)$ ideals in $\mathcal{O}_K$ that divide (y). Moreover, if $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$ and we have a factorisation $\mathfrak{p} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_1}$ into prime ideals of $\mathcal{O}_L$, then we must have $\sum_{i=1}^r e_i \leq d$. It follows that there $O(\mathrm{N}(y)^{\varepsilon O_d(1)}) = O_d(\mathrm{N}(y)^\varepsilon)$ divisors of the ideal (y) in $\mathcal{O}_L$.

Now suppose that y_i and z_i generate the same ideal and $y_1y_2 = z_1z_2$. This implies that $z_1 = uy_1$ and $z_2 = u^{-1}y_2$ for some unit $u \in \mathcal{O}_L$. Thus if $|z_i|_\nu \leq R$, we have $|u|_\nu \leq R/|y_i|_\nu$ for $i = 1, 2$. Moreover, if ω is a place of L lying above ν, then $|u|_\omega = |u|_\nu^{d_\omega}$, where d_ω is the degree of the extension K_ω/K_ν. In particular,

$$h(u) = \prod_{\omega\in\Omega_{L,\infty}} \max\{1, |u|_\omega\} \ll \prod_{\omega\in\Omega_{L,\infty}} R^{d_\omega}|y_i|_\omega^{-1} \ll R^d\mathrm{N}_L(y_i)^{-1} \leq R^d.$$

If $\mathrm{char}(K) > 0$ and $\mathbb{F}_q$ is the field of constants of K, then the field of constants of L is an extension of $\mathbb{F}_q$ of degree at most d. Therefore, $s_L \leq ds_K$ and Lemma 7.2.9 implies that there are $O_d((\log R)^{ds_K}) = O_d(R^\varepsilon)$ available u, which completes the proof. □

Suppose we are given a morphism $\phi \colon \mathbb{P}^n \to \mathbb{P}^m$ of the form $\phi(\boldsymbol{x}) = (\phi_0(\boldsymbol{x}), \dots, \phi_m(\boldsymbol{x}))$ where $\phi_0, \dots, \phi_m \in \mathcal{O}_K[x_0, \dots, x_n]$ are homogeneous forms of degree e without a common zero in $\overline{K}$. Then functoriality of heights implies that

$$H(\phi(\boldsymbol{x})) \asymp H(\boldsymbol{x})^e,$$

where the implied constant depends on n, m, K and ϕ.

Given $f \in K[x_1, \ldots, x_n]$ we define $\|f\| = \|\boldsymbol{f}\|$, where $\boldsymbol{f}$ is the coefficient vector of f. Similarly, we can extend it to vectors $F = (f_1, \ldots, f_r) \in K[x_1, \ldots, x_n]^r$ by setting $\|F\| = \max \|f_i\|$. We require a version of the functoriality of heights for morphisms $\mathbb{P}^1 \to \mathbb{P}^n$ with an explicit dependence on the height of the morphism.

Lemma 7.2.11. *Let $\psi \colon \mathbb{P}^1 \to \mathbb{P}^n$ be a morphism over K given by*

$$\psi([u, v]) = [\psi_0(u, v), \ldots, \psi_n(u, v)],$$

where $\psi_i \in \mathcal{O}_K[u, v]$ is homogeneous of degree d for $i = 0, \ldots, n$ and the ψ_i do not share a common non-constant factor. Then

$$\|\psi\|^{-B} H([u, v])^d \ll_{K,n,d} H(\psi([u, v])) \ll_{K,n,d} \|\psi\|^B H([u, v])^d,$$

where the implied constants and B only depend on K, n and d.

Proof. The upper bound is easy: for any $(u, v) \in Z_1$, the triangle inequality implies

$$H(\psi([u, v])) = \mathrm{N}(\psi(u, v))^{-1} \prod_{\nu \mid \infty} \max\{|\psi_i(u, v)|_\nu\} \ll \prod_{\nu \mid \infty} \max\{|\psi_{ij}|_\nu\} \max\{|u|_\nu, |v|_\nu\}^d,$$

from which the claim follows.

For the lower bound, a repeated application of the Euclidean algorithm gives

$$Cu^e = \sum_{i=0}^{n} f_i \psi_i(u, v) \quad \text{and} \quad Cv^e = \sum_{i=0}^{n} f_i' \psi_i(u, v)$$

for some homogeneous forms $f_i, f_i' \in \mathcal{O}_K[u, v]$ of degree $e - d$ and $C \in \mathcal{O}_K$. Moreover, an inspection of the algorithm reveals that e only depends on n; that the f_i and f_i' can be taken to satisfy $\|f_i\|, \|f_i'\| = O(\|\psi\|^B)$ and $\|C\| = O(\|\psi\|^{B'})$ for some absolute constants $B, B' > 0$. It follows that if $\mathfrak{a} \mid (\psi(u, v))$ for $(u, v) \in Z_1$, then $\mathfrak{a} \mid (C)(\mathfrak{a}_1 \cdots \mathfrak{a}_h)^e$, where $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ are a set of representatives for CL_K. Therefore, for any $\nu \mid \infty$,

$$\begin{aligned} |C|_\nu \max\{|u|_\nu^e, |v|_\nu^e\} &= \max\left\{\left|\sum f_i \psi(u, v)\right|_\nu, \left|\sum f_i' \psi_i(u, v)\right|_\nu\right\} \\ &\ll \|\psi\|^B \max\{|u|_\nu^{e-d}, |v|_\nu^{e-d}\} \max\{|\psi_i(u, v)|\}, \end{aligned}$$

which implies

$$\begin{aligned} H(\psi(u, v)) &= \mathrm{N}(\psi(u, v))^{-1} \prod_{\nu \mid \infty} \max\{|\psi_i(u, v)|_\nu\} \\ &\gg \mathrm{N}(C)^{-1} \|\psi\|^{-B} H([u, v])^d \\ &\gg \|\psi\|^{-B''} H([u, v])^d \end{aligned}$$

for some constant $B'' > 0$. □

7.2.4 Rational points on varieties

For $\boldsymbol{c} \in \mathbb{P}^n(K)$, we shall write $H_{\boldsymbol{c}} \subset \mathbb{P}^n$ for the hyperplane defined by $\boldsymbol{c} \cdot \boldsymbol{x} = 0$. We then have the following result, which follows for number fields from work of Bombieri and Vaaler [17] and for function fields from work of Thunder [194].

Lemma 7.2.12 (Siegel's Lemma)**.** *Let K be a global field and $\boldsymbol{x} \in \mathbb{P}^n(K)$ with $H(\boldsymbol{x}) \leq R$. Then there exists $\boldsymbol{c} \in \mathbb{P}^n(K)$ with $H(\boldsymbol{c}) \ll_{n,k} R^{1/n}$ and $\boldsymbol{x} \in H_{\boldsymbol{c}}$.*

The next result [155, Theorem 1.8] will be helpful when dealing with singular hyperplane sections.

Lemma 7.2.13. *Let $C \subset \mathbb{P}^n$ be an irreducible curve of degree $e \geq 1$. Then*

$$\#\{\boldsymbol{x} \in C(K) \colon H(\boldsymbol{x}) \leq B\} \ll_{e,n,K} B^{2/e}.$$

We will need strong upper bounds for the number of rational points on irreducible varieties of large degree to control the number of degenerate hyperplane sections. Let $X \subset \mathbb{P}^n$ be an irreducible variety of degree d. Then the estimate

$$\#\{\boldsymbol{x} \in X(K) \colon \mathbb{P}^n(K)\} \ll_{d,n,K} B^{\dim(X)+1} \tag{7.2.3}$$

is proved via a simple inductive process of taking hyperplane sections, see for example Lemma 5.6 of [35] for the case $K = \mathbb{Q}$. We omit a proof here. Using the determinant method, one can do significantly better. The principal input is the following result [155, Theorem 1.11] building on work of Salberger [169] and Heath-Brown [104], that demonstrates the truth of the dimension growth conjecture for all global fields.

Proposition 7.2.14. *Let $X \subset \mathbb{P}^n$ be an integral variety of degree $d \geq 5$. Then*

$$\#\{\boldsymbol{x} \in X(K) \colon H(\boldsymbol{x}) \leq B\} \ll_{d,n,K} B^{\dim(X)}.$$

We shall also make us of the following result about representations by binary quadratic forms.

Lemma 7.2.15. *Let $Q \in \mathcal{O}_K[s,t]$ be a binary quadratic form without repeated roots and let $\boldsymbol{R} \in \mathbb{R}^{s_K}_{>1}$. If $\operatorname{char}(K) \neq 2$, then for any $\gamma \in \mathcal{O}_K$ we have*

$$r_Q(\gamma, \boldsymbol{R}) \coloneqq \#\{(s,t) \in \mathcal{O}_K^2 \colon |(s,t)|_\nu \leq R_v \text{ for all } \nu \mid \infty, Q(s,t) = \gamma\} \ll_K (|\boldsymbol{R}| \|Q\| \mathrm{N}(\gamma))^{\varepsilon}.$$

Proof. As $\operatorname{char}(K) \neq 2$, we can always find $P \in \mathrm{GL}_2(K)$ such that $Q(P(s,t))$ is diagonal. Moreover, it is clear that P can be chosen in such a way that its entries p_{ij} satisfy $\|p_{ij}\| \ll \|Q\|^B$. Setting $(x,y) = P(s,t)$, the equation then becomes $ax^2 + by^2 = \gamma$, where $\|(a,b)\| \ll \|Q\|^2$ and $|(x,y)|_\nu \ll \|Q\|^B R_\nu$ for all $\nu \mid \infty$. Multiplying both sides by a suitable element in $\mathcal{O}_K$, we may assume that a is a square in $\mathcal{O}_K$ and $a, b \in \mathcal{O}_K$, so that after applying another change of variables, it transpires that

$$r_Q(k, \boldsymbol{R}) \leq \#\{(x,y) \in \mathcal{O}_K^2 \colon |(x,y)|_\nu \ll \|Q\|^B R_\nu,\ x^2 + dy^2 = \gamma'\},$$

where $\|d\| \ll \|Q\|^B$ and $|\gamma'|_\nu \ll \|Q\|^B |\gamma|_\nu$ for all $\nu \mid \infty$. Let $L = K(\sqrt{d})$ and note that in L we have the factorisation of integral ideals

$$(x - \sqrt{d}y)(x + \sqrt{d}y) = (\gamma').$$

Let $z_1 = x - \sqrt{d}y$ and $z_2 = x + \sqrt{d}y$. Since Q is square-free, the assignment $(x, y) \mapsto (z_1, z_2)$ is injective. Moreover, if we also denote by $|\cdot|_\nu$ the extension of $|\cdot|_\nu$ to L, then we have

$$|z_i|_\nu \ll |\sqrt{d}|_\nu \max\{|x|_\nu, |y|_\nu\} \ll \|Q\|^B R_\nu \leq \|Q\|^B |\boldsymbol{R}|$$

for $i = 1, 2$. Therefore, Lemma 7.2.10 implies that the number of available $(z_1, z_2) \in \mathcal{O}_L$ is $O((\|Q\|^B |\boldsymbol{R}| \mathrm{N}(\gamma'))^\varepsilon) = O((\|Q\||\boldsymbol{R}|\mathrm{N}(\gamma))^\varepsilon))$ as desired. □

Corollary 7.2.16. *Let $Q \in \mathcal{O}_K[s,t]$ be a square-free binary quadratic form. Suppose we are given $\boldsymbol{R} \in \mathbb{R}_{>1}^{s_K}$ and $S > 0$. If $\mathfrak{a} \subset \mathcal{O}_K$ is an ideal, then*

$$\#\{(s,t) \in Z_1 \colon |(s,t)|_\nu \leq R_\nu, \|Q(s,t)\|_\infty \leq S, Q(s,t) \in \mathfrak{a}\} \ll_K \max\left\{1, \frac{|S|}{\mathrm{N}(\mathfrak{a})}\right\} (|\boldsymbol{R}|S\|Q\|)^\varepsilon.$$

Proof. There are $O(1)$ possible $(s,t) \in Z_1$ for which $Q(s,t) = 0$, and so it suffices to bound the contribution from those with $Q(s,t) \neq 0$. Fix a place $\omega \mid \infty$ of K and let us split $|Q(s,t)|_\nu$ into dyadic intervals for $\nu \neq \omega$, say $|Q(s,t)|_\nu \sim w_\nu$. The quantity we want to estimate is then at most

$$\sum_{w_\nu} \#\Big\{(s,t) \in \mathcal{O}_K^2 \colon |(s,t)| \leq R_\nu, Q(s,t) \in \mathfrak{a}, |Q(s,t)|_\omega \ll S \prod_{\substack{\nu|\infty \\ \nu \neq \omega}} w_\nu^{-1}, |Q(s,t)|_\nu \leq w_\nu\Big\}.$$

In the notation of Lemma 7.2.15, each term in the sum is at most

$$\begin{aligned}\sum_{\gamma \in \mathfrak{a}} r_Q(\gamma, \boldsymbol{R}) &\ll \sum_{\gamma \in \mathfrak{a}} (\|Q\|\mathrm{N}(\gamma)|\boldsymbol{R}|)^\varepsilon \\ &\ll (\|Q\|S|\boldsymbol{R}|)^\varepsilon \max\left\{1, \frac{S}{\mathrm{N}(\mathfrak{a})}\right\}\end{aligned}$$

where the sum over γ runs over all $\gamma \in \mathfrak{a}$ such that $|\gamma|_\nu \leq w_\nu$ for $\nu \neq \omega$ and $|\gamma|_\omega \ll S \prod w_\nu^{-1}$ and we successively applied Lemmas 7.2.15 and 7.2.6. We clearly have $|Q(s,t)|_\nu \ll \|Q\| R_\nu^2$ and hence also $|Q(s,t)|_\nu \gg \|Q\|^{-s_K - 1} |\boldsymbol{R}|^{-2(s_K+1)}$, so that there are $O((\|Q\||\boldsymbol{R}|)^\varepsilon)$ possibilities for w_ν, which gives the result. □

We also need the following easy generalisation of a result due to Broberg [27, Lemma 9].

Lemma 7.2.17. *Let $G \in K[x_1, \dots, x_n]$ be a polynomial of degree d and $R, S \geq 1$. Then*

$$M(G, R, S) \coloneqq \#\{\boldsymbol{x} \in \mathcal{O}_K^n \colon \|\boldsymbol{x}\| \leq R^{1/s_K}, \mathrm{N}(G(\boldsymbol{x})) \leq S\} \ll_G R^{n-1+\varepsilon} S^{1/d}.$$

Proof. We first give a proof for the case $n = 1$, so that $G \in K[x]$. Upon dividing through by the leading coefficient, we may assume that G is monic. Over an algebraic closure $\overline{K}$ of K we then have the factorisation

$$G(x) = \prod_{i=1}^d (x - a_i),$$

for some $a_i \in \overline{K}$. As $L = K(a_1, \dots, a_n)$ is a finite degree extension of K, every place $\nu \in \Omega_{K,\infty}$ extends uniquely to L and by abuse of notation we shall also denote it by ν. We then have

$$\mathrm{N}(G(x)) = \prod_{i=1}^d \prod_{\nu|\infty} |x - a_i|_\nu.$$

In particular, if $\mathrm{N}(G(x)) \leq S$, then we must have

$$\prod_{\nu \mid \infty} |x - a_i|_\nu \leq S^{1/d}$$

for some $1 \leq i \leq d$. Note that $\|G\| + R^{1/s_K - d} \ll |x - a_i|_\nu \ll \|G\| + R^{1/s_K}$, where the upper bound follows from $|x|_\nu \leq R^{1/s_K}$ and the lower bound is a consequence of the upper bound and $1 \leq \mathrm{N}(G(x))$. Let us now fix a place $\omega \mid \infty$ of K and put $|x - a_i|_\nu$ into dyadic intervals for $\nu \neq \omega$. We thus have that $M(G, R, S)$ is bounded above by

$$\sum_{i=1}^{d} \sum_{w_\nu} \#\Big\{ x \in \mathcal{O}_K \colon |x|_\nu \leq R^{1/s_K}, |x - a_i|_\omega \ll S^{1/d} \prod_{\nu \neq \omega} w_\nu^{-1}, |x - a_i|_\nu \leq w_\nu \text{ for all } \nu \neq \omega \Big\},$$

where the sum over w_ν is over all tuples of dyadic powers $(w_\nu)_{\nu \neq \omega}$ such that $|w_\nu| \ll (R + \|G\|)$ and $1 \leq \prod_\nu w_\nu \leq S^{1/d}$. If $x, x' \in \mathcal{O}_K$ satisfy $|x - a_i|_\nu \leq w_\nu$ and $|x' - a_i|_\nu \leq w_\nu$, then $|x - x'|_\nu \ll w_\nu$. From Lemma 7.2.6 we thus obtain

$$M(G, R, S) \ll \sum_{w_\nu} S^{1/d} \ll (R\|G\|)^\varepsilon S^{1/d},$$

which is satisfactory.

Next we assume that $n > 1$. For $\boldsymbol{a} = (a_2, \ldots, a_n) \in \mathcal{O}_K^{n-1}$, define $\boldsymbol{x}' = (x_1, x_2 + a_2 x_1, \ldots, x_n + a_n x_1)$. Let G_d be the homogeneous degree d part of G. We now claim that we can take $\boldsymbol{a} \in \mathcal{O}_K^{n-1}$ such that $\deg_{x_1'}(G(\boldsymbol{x}')) = d$. Indeed, the assertion is equivalent to $G_d(1, a_2, \ldots, a_n) \neq 0$ and since G_d is not the zero polynomial, we can always find such an $\boldsymbol{a}$ with $\|\boldsymbol{a}\| \ll 1$. The new variables $\boldsymbol{x}'$ now satisfy $|\boldsymbol{x}'|_\nu \ll R^{1/s_K}$, where the implied constant depends on $\boldsymbol{a}$. Once $x_2', \ldots, x_n'$ are fixed, then by construction $g(x_1') = G(x_1', \ldots, x_n')$ is a degree d polynomial and since $\|x_i'\| \ll R^{1/s_K}$ for $i = 2, \ldots, n$, we have $\|g\| \ll R^{d/s_K}$. In particular, from the case $n = 1$ and Lemma 7.2.6, it follows that

$$M(G, R, S) \ll \sum_{\substack{x_2', \ldots, x_n' \in \mathcal{O}_K \\ |x_i'|_\nu \ll R^{1/s_K}}} \#\{x_1' \in \mathcal{O}_K \colon \mathrm{N}(g(x_1')) \leq S, |x_1'|_\nu \ll R^{1/s_K}\} \ll R^{n-1+\varepsilon} S^{1/d}.$$

□

For separable binary forms we can prove a stronger result, which we will require.

Lemma 7.2.18. *Let $f \in K[s,t]$ be a separable binary form of degree d and let $R, S \geq 1$. Then we have*

$$M(f, R, S) = \#\left\{(s,t) \in Z_1 \colon \|(s,t)\|_\infty \asymp R,\, \mathrm{N}(f(s,t)) \leq S\right\} \ll_f R^{1+\varepsilon}\left(1 + \frac{S}{R^{d-1}}\right)$$

Proof. First note that there are $O(1)$ many solutions in Z_1 to $f(s,t) = 0$. For the remainder we will therefore proceed to only count the contribution such that $f(s,t) \neq 0$. After rescaling f by a constant we may assume that it is of the form

$$f = \prod_{i=1}^{d} (s - a_i t),$$

where $a_i \in \overline{K}$. Since $K(a_1, \ldots, a_d)$ is a finite algebraic extension of K we may extend each place $\nu \mid \infty$. By an abuse of notation we will denote the extended place by ν. The elements

a_i are pairwise different since f was assumed to be separable. Therefore, if we take a constant $c > 0$ sufficiently small in terms of the polynomial f, then given $\nu \mid \infty$ there is at most one index $i_\nu \in \{1, \ldots, d\}$ such that

$$|s - a_{i_\nu} t|_\nu < c \max\{|s|_\nu, |t|_\nu\}.$$

Therefore we have

$$\mathrm{N}(f(s,t)) = \prod_{\nu|\infty} \prod_{i=1}^{d} |s - a_i t|_\nu \asymp R^{d-1} \prod_{\nu|\infty} |s - a_{i_\nu} t|_\nu.$$

Fix a place $\omega \mid \infty$ and divide the absolute values $|s - a_{i_\nu} t|_\nu$ for $\nu \neq \omega$ into dyadic intervals $(\gamma_\nu, 2\gamma_\nu]$, say. Then since $\mathrm{N}(f(s,t)) \leq S$ we have that

$$|s - a_{i_\omega} t|_\omega \ll \frac{S}{R^{d-1}} \prod_{\nu \neq \omega} \gamma_\nu^{-1}.$$

For fixed t these restrictions imply that there are at most $O(1 + S/R^{d-1})$ many s available. Since $\|t\| \ll R$, from Lemma 7.2.6 we see that each such dyadic decomposition contributes at most $O(R(1+S/R^{d-1}))$. Clearly we must only consider dyadic intervals such that $\gamma_\nu \ll R^{1/s_K}$. Since $\mathrm{N}(f(s,t)) \geq 1$ whenever $(s,t) \in \mathcal{O}_K^2$ is not a zero of f we have that

$$\gamma_\nu \gg R^{-d-s_K}.$$

It follows that there are $O(R^\varepsilon)$ relevant dyadic intervals to consider, which completes the proof of the lemma. □

7.3 Geometry of numbers and rational points on conics

In this section we will first recall basic properties of $\mathcal{O}_K$-lattices and establish analogues of classical results from the geometry of numbers. In the second part of this section, we use them to obtain uniform upper bounds for the number of rational points on conics.

7.3.1 Lattices

Let R be a Dedekind domain with field of fractions K. We call a finitely generated torsion-free module Λ an R-*lattice* and denote the corresponding vector space $K\Lambda$ by V. The dimension of V is called the *rank* of Λ. Suppose that $\Lambda' \subset V$ is another R-lattice of the same rank r as Λ. By the invariant factor theorem [162, Theorem 4.14], there exist elements $e_1, \ldots, e_r \in \Lambda$ and unique fractional ideals $\mathfrak{b}_1, \ldots, \mathfrak{b}_r, \mathfrak{c}_1, \ldots, \mathfrak{c}_r \subset K$ such that $\mathfrak{c}_1 \subset \cdots \subset \mathfrak{c}_r$ and

$$\Lambda = \bigoplus_{i=1}^{r} \mathfrak{b}_i e_i \quad \text{and} \quad \Lambda' = \bigoplus_{i=1}^{r} \mathfrak{c}_i \mathfrak{b}_i e_i.$$

Moreover, $\Lambda' \subset \Lambda$ if and only if $\mathfrak{c}_i \subset R$ for $i = 1, \ldots, r$. We then define the *index ideal* of Λ' in Λ to be the fractional ideal $(\Lambda \colon \Lambda') = \mathfrak{c}_1 \cdots \mathfrak{c}_r$. Note that if $\Lambda' \subset \Lambda$, then $\mathrm{N}((\Lambda \colon \Lambda')) = [\Lambda \colon \Lambda']$, where the right hand side denotes the ordinary index of abelian groups. In addition, if $a \in (\Lambda \colon \Lambda')$, then we have $\Lambda \subset a^{-1}\Lambda'$.

Let us now specialise to the case where $R = \mathcal{O}_K$ is the ring of integers of a global field. Suppose we are given an $\mathcal{O}_K$-lattice Λ. We can then form the $\mathcal{O}_\mathfrak{p}$-lattice $\Lambda_\mathfrak{p} := \Lambda \otimes \mathcal{O}_\mathfrak{p}$

for any finite prime $\mathfrak{p}$. If $\Lambda' \subset V$ is another lattice of full rank, then we have $\Lambda_\mathfrak{p} = \Lambda'_\mathfrak{p}$ for almost all $\mathfrak{p}$ and so it is clear that we have $(\Lambda\colon \Lambda') = \bigcap(\Lambda_\mathfrak{p}\colon \Lambda'_\mathfrak{p})$, where the intersection takes place in $\mathcal{O}_K$. If we are given sublattices $L_\mathfrak{p} \subset V_\mathfrak{p} = K_\mathfrak{p}V$ for all finite primes $\mathfrak{p}$ such that $L_\mathfrak{p} = \Lambda_\mathfrak{p}$ for all but finitely many $\mathfrak{p}$, then $\Lambda' \coloneqq \bigcap L_\mathfrak{p}$ defines an $\mathcal{O}_K$-lattice in V such that $\Lambda'_\mathfrak{p} = L_\mathfrak{p}$ for all $\mathfrak{p}$. It follows from the local description of the index ideal that if $A \in \mathrm{GL}(V)$, we have $(\Lambda\colon A\Lambda') = (\det(A))(\Lambda\colon \Lambda')$. For details we refer the reader to Sections 4 and 5 of Reiner [162].

If $V = K^n$ and $\Lambda \subset V$ is an $\mathcal{O}_K$-lattice of rank n, then the determinant $\det(\Lambda)$ is defined to be $\mathrm{N}((\mathcal{O}_K^n\colon \Lambda))$. For $\nu \mid \infty$ let S_ν be an open, symmetric, convex and bounded subset of K_ν^n such that for all complex places ν we have $S_\nu = \alpha S_\nu$ whenever $\alpha \in k_\nu$ and $|\alpha|_\nu = 1$. We shall consider

$$S = \prod_{\nu\mid\infty} S_\nu.$$

We may realize Λ via the diagonal embedding inside $\prod_{\nu\mid\infty} K_\nu^n$. Let k_∞ be the completion of k at the infinite place and $|\cdot|_\infty$ be the associated absolute value on k_∞. That is, if $k = \mathbb{Q}$ then $k_\infty = \mathbb{R}$ and $|\cdot|_\infty$ is the usual absolute value, whereas if $k = \mathbb{F}_q(t)$, then $k_\infty = \mathbb{F}_q((t^{-1}))$ and $|\cdot|_\infty$ corresponds to the absolute value induced by the degree on the $\mathbb{F}_q(t)$. We define the i-th successive minimum of Λ with respect to S to be

$$\lambda_{i,S} \coloneqq \inf\{|\lambda|_\infty\colon \lambda \in k_\infty \text{ and } \Lambda \cap \lambda S \text{ contains } i \text{ linearly independent vectors}\}.$$

It is clear that we have

$$\lambda_{1,S} \le \lambda_{2,S} \le \cdots \le \lambda_{n,S}.$$

Frequently we shall take

$$S_\nu = \{\boldsymbol{x} \in K_\nu^n\colon |x_i|_\nu < 1 \text{ for } i = 1,\dots,n\}, \tag{7.3.1}$$

in which case we will denote the successive minima with respect to S simply by λ_i. Recall that K_ν is a locally compact abelian group, so that in particular it can be endowed with a Haar measure $\mathrm{d}x_\nu$ that we normalise in such a way that $\mathrm{d}x_\nu$ is the usual Lebesgue measure when ν is real, 2 times the Lebesgue measure when ν is complex and $\int_{\mathcal{O}_\nu} \mathrm{d}x_\nu = |\mathfrak{D}_\nu|_\nu^{1/2}$, where $\mathfrak{D}_\nu$ is the local different of the non-archimidean place ν. We can extend the Haar measure to K_ν^n by $\mathrm{d}\boldsymbol{x}_\nu = \mathrm{d}x_{1,\nu}\cdots\mathrm{d}x_{n,\nu}$. If $S_\nu \subset K_\nu^n$ is measurable, we shall write $\mathrm{vol}(S_\nu) = \int_{S_\nu}\mathrm{d}\boldsymbol{x}_\nu$. We have the following version of Minkwoski's second theorem over global fields.

Lemma 7.3.1. *Let Λ and S be as above and let $\lambda_{1,S} \le \cdots \le \lambda_{n,S}$ the successive minima of Λ with respect to S. Then we have*

$$\det(\Lambda) \asymp_{K,n} \prod_{\nu\mid\infty} \mathrm{vol}(S_\nu)(\lambda_{1,S}\cdots\lambda_{n,S})^{d_K}.$$

This was proven independently by Bombieri–Vaaler [17, Theorems 3 and 6] and McFeat [142] for number fields and by McFeat [142] in the case of function fields. The formulation in the respective works looks slightly different, but one can arrive at our presentation of the result by the same considerations as in the proof of the Corollary to Lemma 5 by Broberg [27]. We remark that if we choose S_ν as in (7.3.1) then Lemma 7.3.1 states

$$\det(\Lambda) \asymp (\lambda_1\cdots\lambda_n)^{d_K}. \tag{7.3.2}$$

Note that if $\lambda \in k_\infty$, then $|\lambda|_\infty = |\lambda|_\nu^{d_\nu}$, where d_ν is the degree of the extension K_ν/k_∞. Therefore, by the discreteness of Λ and the definition of the successive minima, (7.3.2) implies that Λ contains an element $\boldsymbol{x}$ with $|\boldsymbol{x}|_\nu \leq \lambda_1^{d_\nu}$ with equality for at least one $\nu \mid \infty$, so that in particular

$$\|\boldsymbol{x}\|_\infty = \prod_{\nu\mid\infty} |\boldsymbol{x}|_\nu \leq \prod_{\nu\mid\infty} \lambda_1^{d_\nu} = \lambda_1^{d_K} \ll \det(\Lambda)^{1/n}.$$

A useful consequence of this is the following. Let $\mathfrak{a} \subset \mathcal{O}_K$ be an ideal. Then $\mathfrak{a}$ is clearly an $\mathcal{O}_K$-module with $\det(\mathfrak{a}) = \mathrm{N}(\mathfrak{a})$ and by (7.3.2) it contains an element $x \in \mathcal{O}_K$ such that $\mathrm{N}(x) = \|x\|_\infty \ll \mathrm{N}(\mathfrak{a})$. Moreover, because $(x) \subset \mathfrak{a}$, we also have $\mathrm{N}(\mathfrak{a}) \leq \mathrm{N}(x)$, so that in fact $\mathrm{N}(\mathfrak{a}) \asymp \mathrm{N}(x)$. We will make frequent use of this fact without further comment.

For the following lemma, we need to introduce some notation. For any place $\nu \mid \infty$, we let P_ν denote the maximal compact subgroup of $\mathrm{GL}_n(K_\nu)$. When $K_\nu = \mathbb{R}$, this is just $O_n(\mathbb{R})$, when $K_\nu = \mathbb{C}$ this is $U_n(\mathbb{C})$, while for non-archimedean places it is $\mathrm{GL}_n(\mathcal{O}_\nu)$, where $\mathcal{O}_\nu$ is the ring of integers of K_ν. By the Iwasawa decomposition, as presented by Weil [207, Chapter II, §2, Theorem 1], for any matrix $M \in \mathrm{GL}_n(K_\nu)$, there exists a matrix $A_\nu \in P_\nu$ such that $A_\nu M$ is upper triangular. Moreover, we have $|A_\nu \boldsymbol{x}|_\nu \asymp |\boldsymbol{x}|_\nu$ for any $\boldsymbol{x} \in K_\nu^n$ and $A_\nu \in P_\nu$.

Lemma 7.3.2. *Let $\Lambda \subset K^n$ be a lattice with successive minima $\lambda_1 \leq \cdots \leq \lambda_n$. Then there exists a free $\mathcal{O}_K$ lattice $\Lambda' \subset K^n$ with basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ such that*

(i) $\Lambda \subset \Lambda'$,

(ii) $\det(\Lambda) \asymp_{K,n} \det(\Lambda')$,

(iii) there exist matrices $A_\nu \in P_\nu$ such that $A_\nu \boldsymbol{b}_i = (b_{1,\nu}^{(1)}, \ldots, b_{i,\nu}^{(i)}, 0, \ldots, 0)$ satisfying $|b_{i,\nu}^{(i)}|_\nu \asymp_{K,n} \lambda_i^{d_\nu}$ for all $1 \leq i \leq n$ and $\nu \mid \infty$,

(iv) if $\boldsymbol{x} \in \Lambda$ is given by $\boldsymbol{x} = \sum_{i=1}^n y_i \boldsymbol{b}_i$ with $y_i \in \mathcal{O}_K$, then $|y_i|_\nu \ll_{K,n} |\boldsymbol{x}|_\nu \lambda_i^{-d_\nu}$ for all $\nu \mid \infty$.

Proof. By definition of the successive minima, we can find K-linearly independent vectors $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \in \Lambda$ such that $|\boldsymbol{c}_i|_\nu \leq \lambda_i^{d_\nu}$ for all $\nu \mid \infty$ with equality for at least one ν. Let now L be the free $\mathcal{O}_K$-lattice given by

$$L \coloneqq \bigoplus_{i=1}^n \mathcal{O}_K \boldsymbol{c}_i.$$

Note that we clearly have $L \subset \Lambda$. Our next goal is to show that $\det(L) \asymp \det(\Lambda)$.

To do so, observe that since L is free, we have $\det(L) = \mathrm{N}(\det(\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n))$, where $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$ is the $n \times n$ matrix having $\boldsymbol{c}_i$ as its ith column. For every place $\nu \mid \infty$, choose a matrix A_ν in the maximal compact subgroup of $\mathrm{GL}_n(K_\nu)$ such that

$$A_\nu \boldsymbol{c}_i = (c_{i,\nu}^{(1)}, \ldots, c_{i,\nu}^{(i)}, 0, \ldots, 0), \tag{7.3.3}$$

which is always possible by the Iwasawa decomposition. Since $|\det(A_\nu)|_\nu = 1$, we then have

$$\mathrm{N}(\det(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)) = \prod_{\nu\mid\infty} |c_{1,\nu}^{(1)} \cdots c_{n,\nu}^{(n)}|_\nu. \tag{7.3.4}$$

As A_ν is norm preserving, we also have $|c_{i,\nu}^{(i)}|_\nu \ll |\boldsymbol{c}_i|_\nu \leq \lambda_i^{d_\nu}$. Since $L \subset \Lambda$, we must have $\det(L) \geq \det(\Lambda)$, whence

$$\det(\Lambda) \leq \det(L) \ll (\lambda_1 \cdots \lambda_n)^{d_K}$$

by (7.3.4). However, by Lemma 7.3.1 we have $\det(\Lambda) \asymp (\lambda_1 \cdots \lambda_n)^{d_K}$ and thus

$$\det(L) \asymp (\lambda_1 \cdots \lambda_n)^{d_K} \asymp \det \Lambda$$

as claimed.

We now continue with the construction of the lattice Λ' whose existence we want to show. Let us choose $b \in (\Lambda \colon L)$, so that $\Lambda \subset b^{-1}L$. We have

$$1 \asymp \frac{\mathrm{N}(\mathcal{O}_K \colon L)}{\mathrm{N}(\mathcal{O}_K \colon \Lambda)} = \mathrm{N}(\Lambda \colon L),$$

so that upon multiplying b with a unit if necessary, we may assume that $|b|_\nu \asymp 1$ for all $\nu \mid \infty$. If we define the lattice $\Lambda' = b^{-1}L$, then by construction we have $\Lambda \subset \Lambda'$. In addition, it holds that

$$\det(\Lambda') = \mathrm{N}(\mathcal{O}_K \colon b^{-1}\Lambda) = \frac{\det(\Lambda)}{\mathrm{N}(b)^n} \asymp \det(\Lambda),$$

as claimed in (ii). It thus remains to verify the properties asserted in (iii) and (iv) of the lemma. If we define $\boldsymbol{b}_i = b^{-1}\boldsymbol{c}_i$ for $i = 1, \ldots, n$, then by construction Λ' is the free $\mathcal{O}_K$-lattice with basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Let $A_\nu \in P_\nu$ be the matrices in (7.3.3) and define $b_{i,\nu}^{(j)} \in K_\nu$ to be the jth entry of $A_\nu \boldsymbol{b}_i$. Note that this value is explicitly given by $b^{-1}c_{i,\nu}^{(j)}$. Returning to (7.3.4) and recalling that $\det(\Lambda) \asymp \det(L) = \mathrm{N}(\det(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n))$, we deduce from Lemma 7.3.1 that

$$\begin{aligned}(\lambda_1 \cdots \lambda_n)^{d_K} &\asymp \prod_{\nu \mid \infty} |c_{1,\nu}^{(1)} \cdots c_{n,\nu}^{(\nu)}|_\nu \\ &\ll (\lambda_1 \cdots \lambda_{i-1})^{d_K} \lambda_i^{d_K - d_\nu} |c_{i,\nu}^{(i)}|_\nu (\lambda_{i+1} \cdots \lambda_n)^{d_K},\end{aligned}$$

for any $1 \leq i \leq n$ and $\nu \mid \infty$, where we used that $|c_{i,\nu}^{(i)}|_\nu \ll |\boldsymbol{c}_i|_\nu \leq \lambda_i^{d_\nu}$ and that $\sum d_\nu = d_K$. It follows that $|c_{i,\nu}^{(i)}|_\nu \asymp \lambda_i^{d_\nu}$. As $|b|_\nu \asymp 1$ for all $\nu \mid \infty$ and $|b_{i,\nu}^{(i)}|_\nu = |b|_\nu^{-1} |c_{i,\nu}^{(i)}|_\nu$, this completes the verification of (iii).

Turning to (iv), suppose that $\boldsymbol{x} \in \Lambda$ is given by $\boldsymbol{x} = \sum_{i=1}^n y_i \boldsymbol{b}_i$ with $y_i \in \mathcal{O}_K$. Then we have

$$|\boldsymbol{x}|_\nu \asymp |A_\nu \boldsymbol{x}|_\nu = \left| \begin{pmatrix} b_{1,\nu}^{(1)} & b_{2,\nu}^{(1)} & \cdots & b_{n,\nu}^{(1)} \\ 0 & b_{2,\nu}^{(2)} & \cdots & b_{n,\nu}^{(2)} \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & b_{n,\nu}^{(n)} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \right|_\nu,$$

which implies $|y_n|_\nu = |(A_\nu \boldsymbol{x})_n|_\nu |b_{n,\nu}^{(n)}|_\nu^{-1}$, where $(A_\nu \boldsymbol{x})_n$ denotes the nth entry of $A_\nu \boldsymbol{x}$ and hence

$$|y_n|_\nu \ll |\boldsymbol{x}|_\nu / \lambda_n^{d_\nu}$$

by (iii). Similarly, we get

$$b_{n-1,\nu}^{(n-1)} y_{n-1} = (A_\nu \boldsymbol{x})_{n-1} - y_n b_{n,\nu}^{(n-1)}$$

and thus

$$|y_{n-1}|_\nu \ll (|\boldsymbol{x}|_\nu + |y_n b_{n,\nu}^{(n-1)}|_\nu) / |b_{n-1,\nu}^{(n-1)}|_\nu^{-1} \ll |\boldsymbol{x}|_\nu \lambda_{n-1}^{-d_\nu}$$

again by (iii) and using that $|b_{n,\nu}^{(n-1)}|_\nu \ll \lambda_n^{d_\nu}$. Continuing in this fashion completes the verification of (iv). □

Using this, by the same ideas and techniques as in chapter 12 of [61] one obtains very good bounds on the number of lattice points within a bounded box.

Lemma 7.3.3. *Let Λ be an $\mathcal{O}_K$-lattice of rank n and for $\boldsymbol{R} = (R_\nu)_{\nu|\infty} \in \mathbb{R}_{>0}^{s_K}$, define*

$$N(\Lambda, \boldsymbol{R}) := \#\left\{\boldsymbol{x} \in \Lambda \colon |\boldsymbol{x}|_\nu < R_\nu \text{ for all } \nu \mid \infty\right\}.$$

Then if $\lambda_1, \ldots, \lambda_n$ are the successive minima of Λ, we have

$$N(\Lambda, \boldsymbol{R}) \asymp_{K,n} \prod_{i=1}^{n} \max\left\{1, \frac{|\boldsymbol{R}|}{\lambda_i^{d_K}}\right\}.$$

Proof. For the lower bound, let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \Lambda$ be linearly independent over K such that $|\boldsymbol{a}_i|_\nu \leq \lambda_i^{d_\nu}$ for all $1 \leq i \leq n$ and $\nu \mid \infty$. Then any $\boldsymbol{x} = \sum_{i=1}^n \mu_i \boldsymbol{a}_i$ with $\mu_i \in \mathcal{O}_K$ and $|\mu_i|_\nu \ll R_\nu/\lambda_i^{d_\nu}$ will be counted by $N(\Lambda, \boldsymbol{R})$ provided the implied constant is sufficiently small with respect to n. By Lemma 7.2.6 we have

$$\#\{(\mu_1, \ldots, \mu_n) \in \mathcal{O}_K^n \colon |\mu_i|_\nu \ll R_\nu/\lambda_i^{d_\nu} \text{ for all } 1 \leq i \leq n \text{ and } \nu \mid \infty\} \asymp \prod_{i=1}^{n} \max\left\{1, \frac{|\boldsymbol{R}|}{\lambda_i^{d_K}}\right\},$$

from which the lower bound follows.

Turning to the upper bound, let Λ' be the lattice from Lemma 7.3.2. Then as $\Lambda \subset \Lambda'$ it clearly suffices to prove the claimed upper bound for $N(\Lambda', \boldsymbol{R})$ instead of $N(\Lambda, \boldsymbol{R})$. Let $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ be the basis of Λ' and write $\boldsymbol{x} = \sum_{i=1}^n y_i \boldsymbol{b}_i$ with $y_i \in \mathcal{O}_K$. Then if $|\boldsymbol{x}|_\nu \leq R_\nu$, it follows from (iv) of Lemma 7.3.2 that

$$|y_i|_\nu \ll |\boldsymbol{x}|_\nu \lambda_i^{-d_\nu}.$$

In particular, if $|\boldsymbol{x}|_\nu \leq R_\nu$, then $|y_i|_\nu \ll R_\nu \lambda_i^{-d_\nu}$ and by Lemma 7.2.6 the number of such $y_i \in \mathcal{O}_K$ is $O(\max\{1, |\boldsymbol{R}|\lambda_i^{-d_K}\})$ as desired. $\square$

7.3.2 Rational points on conics

Suppose we are given $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n \in \mathbb{R}_{>0}^{s_K}$ with $\boldsymbol{R}_i = (R_{i,\nu})_{\nu|\infty}$ and set

$$L(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n) = \{\boldsymbol{x} \in \mathcal{O}_K^n \colon |x_i|_\nu \leq R_{i,\nu} \text{ for all } \nu \mid \infty\}.$$

Let $F \in \mathcal{O}_K[x_1, x_2, x_3]$ be a quadratic form. In this section we are concerned with upper bounds for the number of elements in $\boldsymbol{x} \in L(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3)$ such that $F(\boldsymbol{x}) = 0$ which are uniform with respect to F. Given such F, we let $M \in \mathsf{Mat}_{3\times 3}(K)$ denote the underlying matrix and define $\Delta(F) \subset \mathcal{O}_K$ and $\Delta_0(F) \subset \mathcal{O}_K$ to be the ideal generated by the determinant and 2×2 minors of M respectively. The following results were proved by Browning and Swarbrick-Jones for number fields [43] and go back to work of Heath-Brown [101] and Broberg [27], who proved a slightly weaker version in the context of the rational numbers and number fields respectively.

Proposition 7.3.4. *Suppose we are given a non-singular quadratic form $F \in \mathcal{O}_K[x_0, x_1, x_2]$ and $\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3 \in \mathbb{R}_{>0}^{s_K}$. Then*

$$N_F(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3) := \#\{\boldsymbol{x} \in \mathbb{P}^2(K) \cap L(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_2) \colon F(\boldsymbol{x}) = 0\} \ll_K 1 + (|\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3|)^{1/3}.$$

Proof. If K is a number field, this is Theorem 4.7 of [43], so we may assume that K is a function field. Moreover, the proof is almost identical, so we shall be brief. Note that if $|\boldsymbol{R}_i| < 1$ for some $1 \le i \le 3$, then this forces $x_i = 0$. In particular, every point counted by $N_F(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3)$ lies on the intersection of a conic with a line. By Bézout's theorem there are at most 2 such points over K. Therefore, from now on we may and shall assume that $|\boldsymbol{R}_i| \ge 1$ for $1 \le i \le 3$. Define $R = |\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3|$ and choose prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_r \subset \mathcal{O}_K$ such that

$$cR^{1/3} \le \mathrm{N}(\mathfrak{p}_1) \le \cdots \le \mathrm{N}(\mathfrak{p}_r) \ll R^{1/3},$$

where $c > 0$ and r are constants to be determined in due course. First suppose that $\mathrm{N}(\mathfrak{p}_i) \mid \Delta(F)$ for $i = 1, \ldots, r$. Then we have $\mathrm{N}(\Delta(F)) \gg R^{r/3}$ and hence $\|F\| \gg \mathrm{N}(\Delta(F))^{1/3s_K} \gg R^{r/9s_K}$. Letting $B = \prod_{\nu|\infty} \max\{R_{1,\nu}, R_{2,\nu}, R_{3,\nu}\}$, then it is clear that $B \le R$ and that any $\boldsymbol{x} \in Z_2$ with $|x_i|_\nu \le R_{i,\nu}$ satisfies $\|\boldsymbol{x}\|_\infty \le B$. In particular, if we choose $r > 108$, it follows from Lemma 7.4.2 that $N_F(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3) \ll 1$, which is sufficient. We may therefore assume from now on that there is a prime ideal $\mathfrak{p}$ with $\mathrm{N}(\mathfrak{p}) \asymp R^{1/3}$ such that $\mathfrak{p} \nmid \Delta(F)$. Let $\mathbb{F}_\mathfrak{p} = \mathcal{O}_\mathfrak{p}/\mathfrak{p}\mathcal{O}_\mathfrak{p}$. As the reduction of F modulo $\mathfrak{p}$ is non-singular, the projective conic defined by $F = 0$ over $\mathbb{F}_\mathfrak{p}$ has precisely $\mathrm{N}(\mathfrak{p}) + 1$ points and our goal is to show that any such solution gives rise to at most 2 projective points. Given that $\mathrm{N}(\mathfrak{p}) \ll R^{1/3}$, this will be sufficient to complete the proof.

Suppose that $\boldsymbol{x}_0 \in \mathbb{F}_\mathfrak{p}^3 \setminus \{\boldsymbol{0}\}$ satisfies $F(\boldsymbol{x}_0) \equiv 0 \ (\mathfrak{p})$. Then by a Hensel lifting argument we can always find $\boldsymbol{x}_1 \in \mathcal{O}_\mathfrak{p}$ such that $\boldsymbol{x}_1 \equiv \boldsymbol{x}_0 \ (\mathfrak{p})$ and $F(\boldsymbol{x}_1) \equiv 0 \ (\mathfrak{p}^2)$. Suppose now that $\boldsymbol{x} \in Z_2$ is such that $\boldsymbol{x} \equiv \lambda \boldsymbol{x}_1 \ (\mathfrak{p})$ for some $\lambda \in \mathcal{O}_K$ and $F(\boldsymbol{x}) = 0$. It follows that there exists $\boldsymbol{z} \in \mathcal{O}_K^3$ such that $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + \pi \boldsymbol{z}$, where π is a uniformizer of $\mathcal{O}_\mathfrak{p}$. We then have

$$\pi \lambda \boldsymbol{z} \cdot \nabla F(\boldsymbol{x}_1) \equiv 0 \ (\mathfrak{p}^2).$$

Moreover, we can assume that R is sufficiently large so that $\mathfrak{p}$ is not equal to one of the fixed representatives of the class group of K. As $\boldsymbol{x} \in Z_2$, we must then have $\lambda \notin \mathfrak{p}$ and hence $\boldsymbol{z} \cdot \nabla F(\boldsymbol{x}_1) \equiv 0 \ (\mathfrak{p})$. Therefore, $\boldsymbol{x} \cdot \nabla F(\boldsymbol{x}_1) \equiv 0 \ (\mathfrak{p}^2)$ and $\boldsymbol{x}$ must belong to the set

$$\{\boldsymbol{x} \in \mathcal{O}_\mathfrak{p}^3 \colon \boldsymbol{x} \equiv \lambda \boldsymbol{x}_0 \ (\mathfrak{p}) \text{ and } \boldsymbol{x} \cdot \nabla F(\boldsymbol{x}_1) \equiv 0 \ (\mathfrak{p}^2)\},$$

which defines an $\mathcal{O}_\mathfrak{p}$-lattice $L_\mathfrak{p}$ of determinant $\mathrm{N}(\mathfrak{p})^3$ and rank 3. Let now Λ be the unique $\mathcal{O}_K$-lattice such that $\Lambda_\mathfrak{p} = L_\mathfrak{p}$ and $\Lambda_\mathfrak{q} = \mathcal{O}_\mathfrak{q}$ for all $\mathfrak{q} \neq \mathfrak{p}$. Moreover, choose elements $\gamma_1, \gamma_2, \gamma_3 \in K$ such that $|\gamma_i|_\nu \asymp \prod_{j=1,2,3, j\neq i} R_{i,\nu}$ and define

$$L' = \{(\gamma_1 x_1, \gamma_2 x_2, \gamma_3 x_3) \in K^3 \colon (x_1, x_2, x_3) \in \Lambda\},$$

which is an $\mathcal{O}_K$-lattice of determinant $\det(L') = |R|^2 \mathrm{N}(\mathfrak{p})^3$. Let Λ' be the lattice from Lemma 7.3.2 containing L' with basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3 \in K^3$ and successive minima $\lambda_1 \le \lambda_2 \le \lambda_3$. If we write $\boldsymbol{x} = y_1 \boldsymbol{b}_1 + y_2 \boldsymbol{b}_2 + y_3 \boldsymbol{b}_3$, then it follows from (iv) of Lemma 7.3.2 that $|y_3| \ll |\boldsymbol{x}|_\nu \lambda_3^{-d_\nu} \ll R_{1,\nu} R_{2,\nu} R_{3,\nu} \lambda_3^{-d_\nu}$. Moreover, it follows from (7.3.2) that $\lambda_3^{3d_K} \gg \det(\Lambda') \asymp R^2 \mathrm{N}(\mathfrak{p})^3$ and hence after taking the product over all places, we deduce

$$\mathrm{N}(y_3) \ll \frac{R^3}{R^{2/3} \mathrm{N}(\mathfrak{p})^{1/3}} \ll 1.$$

In particular, if we choose c from the beginning of the lemma sufficiently large, then we must have $y_3 = 0$. It follows that $\boldsymbol{x}$ lies on the intersection of the quadric with a line, and hence contains at most 2 projective points by Bézout's theorem. □

Lemma 7.3.5. *Let $\boldsymbol{x} \in \mathcal{O}_K$ be such that $F(\boldsymbol{x}) = 0$. If $\operatorname{char}(K) \neq 2$, then $\boldsymbol{x}$ must lie in at least one of $O(\tau(\Delta(F)))$ many lattices Γ such that $\det(\Gamma) \gg_K \frac{\mathrm{N}(\Delta(F))}{\mathrm{N}(\Delta_0(F))^{3/2}}$, where τ is the divisor function on ideals.*

Proof. If K is a number field, this is Corollary 4.6 of [43], so we may assume that $\operatorname{char}(K) > 2$ from now on. Let $\mathfrak{p} \subset \mathcal{O}_K$ be a prime such that $\mathfrak{p} \mid \Delta(F)$ and π a uniformizer of $\mathcal{O}_\mathfrak{p}$. As in the proof of [27, Lemma 4(b)], we may diagonalize F and assume that it is given by

$$F(\boldsymbol{x}) = \varepsilon_1 \pi^{\alpha_1} x_1^2 + \varepsilon_2 \pi^{\alpha_2} x_2^2 + \varepsilon_3 \pi^{\alpha_3} x_3^2,$$

where $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq 0$ and ε_i are units. In particular, if we set $a_\mathfrak{p} = \nu_\mathfrak{p}(\Delta(F))$ and $b_\mathfrak{p} = \nu_\mathfrak{p}(\Delta_0(F))$, then $a_\mathfrak{p} = \alpha_1 + \alpha_2 + \alpha_3$ and $b_\mathfrak{p} = \alpha_1 + \alpha_2$. We now claim that if $\boldsymbol{x} \in \mathcal{O}_\mathfrak{p}^3$ satisfies $F(\boldsymbol{x}) = 0$, then there are $\mathcal{O}_\mathfrak{p}$-lattices $L_1, \ldots, L_M$ with $M \leq \alpha_3 + 1$ and $\det(L_i) \geq \mathrm{N}(\mathfrak{p})^{(2\alpha_3 - \alpha_2 - \alpha_1)/2} = \mathrm{N}(\mathfrak{p})^{(2a_\mathfrak{p} - 3b_\mathfrak{p})/2}$. The statement will then follow upon letting Γ to be one of the lattices such that $\Gamma_\mathfrak{p} = L_i$ for some $1 \leq i \leq M$ and $\mathfrak{p} \mid \Delta(F)$.

Suppose that $x_i = \pi^{\xi_i} u_i$, where $u_i \in \mathcal{O}_\mathfrak{p}^\times$ for $i = 1, 2$. Then if $F(\boldsymbol{x}) = 0$, we must have

$$\varepsilon_1 u_1^2 \pi^{\alpha_1 + 2\xi_1} + \varepsilon_2 u_2^2 \pi^{\alpha_2 + 2\xi_2} \equiv 0 \ (\pi^{\alpha_3}). \tag{7.3.5}$$

We now consider two cases. First, let us assume that $\alpha_3 \leq \min_{i=1,2}\{\alpha_i + 2\xi_i\}$. Then if we set

$$L_1 = \{\boldsymbol{x} \in \mathcal{O}_\mathfrak{p}^3 \colon x_i \in (\mathfrak{p}\mathcal{O}_\mathfrak{p})^{\max\{0, \lceil \frac{\alpha_3 - \alpha_i}{2} \rceil\}}, i = 1, 2\},$$

it is clear that $L_\mathfrak{p}$ is an $\mathcal{O}_\mathfrak{p}$-lattice of rank 3. We must have $\boldsymbol{x} \in L_1$ and in addition

$$\begin{aligned} \det(L_1) &\geq \mathrm{N}(\mathfrak{p})^{\max\{0, \lceil \frac{\alpha_3 - \alpha_1}{2} \rceil\} + \max\{0, \lceil \frac{\alpha_3 - \alpha_2}{2} \rceil\}} \\ &\geq \mathrm{N}(\mathfrak{p})^{(2\alpha_3 - \alpha_1 - \alpha_2)/2}, \end{aligned}$$

which is satisfactory.

Let us now assume that $\alpha_3 > \min_{i=1,2}\{\alpha_i + 2\xi_i\}$. In this case (7.3.5) implies that $\alpha_1 + 2\xi_1 = \alpha_2 + 2\xi_2 = \eta$, say. Moreover, (7.3.5) also gives $(u_1/u_2)^2 \equiv -\varepsilon_2/\varepsilon_1 \ (\pi^{\alpha_3 - \eta})$. As $\operatorname{char}(K) \neq 2$, Hensel's lemma implies that if this equation is solvable, that there are $r_1, r_2 \in (\mathcal{O}_\mathfrak{p}/\mathfrak{p}^{\alpha_3 - \eta}\mathcal{O}_\mathfrak{p})^\times$ such that $u_1 \equiv r_i u_2 \ (\mathfrak{p}^{\alpha_3 - \eta})$. In particular, $\boldsymbol{x}$ must satisfy

$$x_i \equiv 0 \ (\mathfrak{p}^{\xi_i}) \quad \text{and} \quad x_1 \pi^{-\xi_1} \equiv r_i x_2 \pi^{-\xi_2} \ (\mathfrak{p}^{\alpha_3 - \eta}).$$

These conditions define an $\mathcal{O}_\mathfrak{p}$-lattice of determinant $\mathrm{N}(\mathfrak{p})^{\alpha_3 + \xi_1 + \xi_2} \geq \mathrm{N}(\mathfrak{p})^{(2\alpha_3 - \alpha_1 - \alpha_2)/2}$ which is sufficient. Moreover, taking into account the possible values that ξ_1 and ξ_2 can take, we see that are at most $\alpha_3 + 1$ possible lattices in total. ☐

Corollary 7.3.6. *Suppose that $\operatorname{char}(K) \neq 2$. Let $\Delta(F)$ and $\Delta_0(F)$ be the fractional ideals in K generated by the discriminant and the 2×2 minors of F. Then*

$$N_F(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3) \ll_\varepsilon \left(1 + \frac{|\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3| \mathrm{N}(\Delta_0(F))^{3/2}}{\mathrm{N}(\Delta(F))}\right)^{1/3} \tau(\Delta(F)).$$

Proof. Let Γ be one of the lattices from Lemma 7.3.5 with $\det(\Gamma) \gg \mathrm{N}(\Delta(F))/\mathrm{N}(\Delta_0(F))^{3/2}$ and choose $\gamma_1, \gamma_2, \gamma_3 \in K$ such that $|\gamma_i|_\nu \asymp \prod_{j=1,2,3, j\neq i} R_{j,\nu}$. Put $L = \{(\gamma_1 x_1, \gamma_2 x_3, \gamma_3 x_3) \in$

$K^3\colon (x_1,x_2,x_3)\in\Gamma\}$, so that L is an $\mathcal{O}_K$-lattice of determinant $R^2\det(\Gamma)$. Moreover, note that if $\boldsymbol{x}\in\Gamma$ satisfies $|x_i|_\nu\le R_{i,\nu}$, then $|\gamma_i x_i|_\nu \ll R_{1,\nu}R_{2,\nu}R_{3,\nu}$.

Let Γ' be the lattice from Lemma 7.3.2 such that $L\subset\Gamma'$ with basis $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3$. Suppose that $\boldsymbol{x}\in\Gamma$ satisfies $F(\boldsymbol{x})=0$ and write $\boldsymbol{x}=y_1\boldsymbol{b}_1+y_2\boldsymbol{b}_2+y_3\boldsymbol{b}_3$. Then according to (iv) of Lemma 7.3.2 we must have $|y_i|_\nu\ll|\boldsymbol{x}|_\nu\lambda_i^{-d_\nu}$. In particular, it follows from Lemma 7.3.1 and Proposition 7.3.4 that

$$\begin{aligned}\#\left\{\boldsymbol{y}\in\mathbb{P}^2(K)\colon \begin{array}{l} F(y_1\boldsymbol{b}_1+y_2\boldsymbol{b}_2+y_3\boldsymbol{b}_3)=0,\\ |y_i|_\nu\ll R_{1,\nu}R_{2,\nu}R_{3,\nu}\lambda_i^{d_\nu}\end{array}\right\} &\ll 1+\left(\frac{|\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3|^3}{(\lambda_1\lambda_2\lambda_3)^{d_K}}\right)^{1/3}\\ &\ll 1+\left(\frac{|\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3|}{\det(\Gamma)}\right)^{1/3}\\ &\ll 1+\left(\frac{|\boldsymbol{R}_1||\boldsymbol{R}_2||\boldsymbol{R}_3|\mathrm{N}(\Delta_0(F))^{3/2}}{\mathrm{N}(\Delta(F))}\right)^{1/3}.\end{aligned}$$

As every $\boldsymbol{x}$ with $F(\boldsymbol{x})=0$ is contained in at most $O(\tau(\Delta(F)))$ different lattices Γ as above, the result follows. □

In addition we will at some point require the following result, which was proved for number fields by Broberg [27, Proposition 7].

Lemma 7.3.7. *Let $q\in\mathcal{O}_K[x_1,x_2,x_3]$ be a non-singular quadratic form such that $q(0,x_2,x_3)$ is also non-singular. Let $\boldsymbol{R}\in\mathbb{R}_{\ge1}^{s_K}$ and $R\ge1$ be given. Then*

$$\#\{\boldsymbol{x}\in(L(\boldsymbol{R})\times\mathcal{O}_K^2)\cap Z_2\colon \|x_i\|\le R, i=2,3, q(\boldsymbol{x})=0\}\ll_K 1+|\boldsymbol{R}|(|\boldsymbol{R}|\|q\|R)^\varepsilon.$$

Proof. As $q(0,x_2,x_3)$ is non-singular, we can find a matrix $P\in\mathrm{GL}_2(K)$ with entries $p_{ij}\in\mathcal{O}_K$ satisfying $\|p_{ij}\|\ll\|q\|^B$ for some absolute constant $B>0$ and such that $q(x_1,P(x_2,x_2))$ takes the shape

$$\alpha L_1(x_1,x_2,x_3)^2+\beta L_2(x_1,x_2,x_3)^2-\gamma x_1^2,$$

with non-zero coefficients α,β,γ and L_i are linear forms such that $L_1(0,x_2,x_3)$ and $L_2(0,x_2,x_3)$ are not proportional. Upon multiplying by a suitable constant, we may assume that α,β,γ lie in $\mathcal{O}_K$ and satisfy $\|\alpha\|,\|\beta\|,\|\gamma\|\ll\|q\|^B$, with possibly a new value of B. If we set $L=K(\sqrt{\beta})$, then the equation $q(x_1,P(x_2,x_3))=0$ becomes

$$L_3(x_1,x_2,x_3)L_4(x_1,x_2,x_3)=\gamma x_1^2,$$

where L_3,L_4 are linear forms with coefficients in $\mathcal{O}_L$. Note that if $x_1=0$, then there are $O(1)$ possibilities for $(0,x_2,x_3)\in Z_3$ such that $q(0,x_2,x_3)=0$, so we may suppose $x_1\ne0$ from now on. Set $z_2=L_3(x_1,x_2,x_3)$ and $z_3=L_4(x_1,x_2,x_3)$ and observe that the assignment is injective, as $q(0,x_2,x_3)$ is non-singular. We then have $|z_i|_\nu\ll\|q\|^B\max\{R_\nu,R\}\ll\|q\|^B|\boldsymbol{R}|R$, so that Lemma 7.2.10 implies that for x_1 fixed, there are $O((\|q\|^B|\boldsymbol{R}|R\mathrm{N}(\gamma x_1^2))^\varepsilon)=O((\|q\||\boldsymbol{R}|R)^\varepsilon)$ possible (z_2,z_3) such that $z_2z_3=\gamma x_1^2$. As there are $O(|\boldsymbol{R}|)$ available $x_1\in L(\boldsymbol{R})$ by Lemma 7.2.6 this is sufficient. □

7.4 Rational points on smooth genus 1 curves

Let $E\subset\mathbb{P}^n$ be a smooth genus 1 curve of degree d and for $B\ge1$ define

$$N_E(B)\coloneqq\#\{x\in E(K)\colon H(x)\le B\},$$

where $H\colon \mathbb{P}^n(K) \to \mathbb{R}_{>0}$ denotes the usual height function. Our goal of this section is to produce a *uniform* upper bound for $N_E(B)$ and prove Proposition 7.1.4. Let $\mathscr{D}_E$ and $\mathscr{C}_E$ denote the minimal discriminant and the conductor of E, which are either ideals or divisors depending on whether K is a number field or a function field. The main input is the *rank growth hypothesis (RGH)* stated in Conjecture 7.1.1 for elliptic curves E:

$$r_E = o(\log \mathrm{N}(\mathscr{C}_E)) \quad \text{as } \mathrm{N}(\mathscr{C}_E) \to \infty, \tag{7.4.1}$$

where r_E denotes the rank of E. In his remarkable work, Brumer [46, Proposition 6.9] proved Conjecture(7.1.1) when $\mathrm{char}(K) > 3$.

Proposition 7.4.1. *Suppose* $\mathrm{char}(K) > 3$*. Then* (7.4.1) *holds with the estimate*

$$r_E \ll_K \frac{\log \mathrm{N}(\mathscr{C}_E)}{\log\log \mathrm{N}(\mathscr{C}_E)}.$$

We will prove Proposition 7.1.4 in several steps. Let us first show how (7.4.1) implies a uniform upper bound for $N_E(B)$ when E is an elliptic curve given in Weierstrass form

$$E\colon zy^2 = x^3 + axz^2 + bz^3 \tag{7.4.2}$$

with $a, b \in \mathcal{O}_K$. By abuse of notation we shall also write

$$E(x,y,z) = zy^2 - (x^3 + axz^2 + bz^3).$$

Let $h\colon E(K) \to \mathbb{R}_{\geq 0}$ be the canonical height of E and define the height of E to be

$$H_E := \prod_{\nu \in \Omega_K} \max\{1, |a|_\nu^{1/2}, |b|_\nu^{1/3}\}.$$

Provided $\mathrm{char}(K) \neq 2, 3$, it follows from work of Zimmer [214, p. 40] that

$$h(P) = \log H(P) + O(\log H_E) \tag{7.4.3}$$

for $P \in E(K)$. The Mordell–Weil group $E(K)$ is a finitely generated abelian group of rank r_E and so any point $P \in E(K)$ may be uniquely written as

$$P = T + \sum_{i=1}^{r_E} n_i T_i,$$

where $T \in E(K)_{\mathsf{tors}}$ is a torsion point and $T_1, \ldots, T_{r_E}$ are generators of $E(K)$. In this case we have

$$h(P) = h\left(\sum n_i T_i\right) = Q_E(n_1, \ldots, n_{r_E})$$

where $Q_E \in \mathbb{Z}[x_1, \ldots, x_{r_E}]$ is a positive-definite quadratic form. By work of Mazur [141] for $K = \mathbb{Q}$, Merel [143] for number fields and Levin [134] for non-isotrivial elliptic curves over function fields, we know $E(K)_{\mathsf{tors}} \ll 1$, where the implied constant only depends on the degree of the number field in characteristic 0 or the genus of the function field in positive characteristic. For isotrivial curves over function fields, it is explained in [199] that we always obtain the bound $E(K)_{\mathsf{tors}} \ll q^2$, where q is the cardinality of the field of constants K. It thus follows from (7.4.3) that for a suitable absolute constant $C > 0$ we have

$$N_E(B) \ll_K \#\left\{(n_1, \ldots, n_{r_E}) \in \mathbb{Z}^{r_E} \colon Q_E(n_1, \ldots, n_{r_E}) \leq C \log(H_E B)\right\}.$$

Lemma 4 of Heath-Brown [102] implies that

$$N_E(B) \ll 1 + (9C\log(H_E B)/B_{0,E})^{r_E/2}, \tag{7.4.4}$$

where

$$B_{0,E} = \min\{h(P)\colon h(P) \neq 0\}.$$

Hindry–Silverman [108, Corollary 4.2] show that

$$B_{0,E} \geq \log(\mathrm{N}(\mathscr{D}_E))\exp(-A\log(\mathrm{N}(\mathscr{D}_E))/\log(\mathrm{N}(\mathscr{C}_E))),$$

where $\mathscr{D}_E$ is the minimal discriminant of E and $A > 0$ is an absolute constant only depending on K. Note that strictly speaking Corollary 4.2 of Hindry–Silverman is only stated for number fields. However, it is an immediate consequence of Theorem 4.1 in the same work, which is valid for any global field. For now assume that $\mathrm{N}(\mathscr{C}_E) \geq C(\varepsilon)$ where $C(\varepsilon)$ is a constant chosen sufficiently large so that $r_E \leq \varepsilon \log \mathrm{N}(\mathscr{C}_E)/A$ holds. Following the analysis of Heath-Brown [102] on page 23 word by word, we arrive at the estimate

$$N_E(B) \ll_\varepsilon (BH_E)^\varepsilon. \tag{7.4.5}$$

For the case when $\mathrm{N}(\mathscr{C}_E) < C(\varepsilon)$ our treatment differs according to whether K is a function field or a number field. If K is a number field then there are only finitely many elliptic curves of a given conductor (cf. [185, p. IX.6]). Thus we find

$$\max_{E\colon \mathrm{N}(\mathscr{C}_E)\leq C(\varepsilon)} r_E \ll_{\varepsilon,K} 1,$$

and

$$\min_{E\colon \mathrm{N}(\mathscr{C}_E)\leq C(\varepsilon)} B_{0,E} \gg_{\varepsilon,K} 1,$$

where the lower bound follows from the fact that the quadratic forms that determine the height of a point are positive definite. Hence (7.4.4) delivers (7.4.5) for elliptic curves E with $\mathrm{N}(\mathscr{C}_E) \leq C(\varepsilon)$.

If K is now a function field, then Theorem 6.1 in [5] shows that unless $H_E = 1$ we have

$$B_{0,E} \gg_K p^{-2e}\log \mathrm{N}(\mathscr{D}_E),$$

where p^e denotes the inseparability degree of the field extension $K/\mathbb{F}_q(j(E))$, which by convention is 1 if $j(E) \in \mathbb{F}_q$. If $j(E)$ is non-constant, then e is the maximal non-negative integer such that $j(E) \in K^{p^e}$. In particular, we have $p^e \ll \log \mathrm{N}(j(E)) \ll \log \mathrm{N}(\mathscr{D}_E)$. Moreover, since $\log \mathrm{N}(\mathscr{C}_E) \ll_\varepsilon 1$ again implies $r_E \ll_\varepsilon 1$, and noting $\log \mathrm{N}(\mathscr{D}_E) \ll \log H_E$ we obtain (7.4.5) by substituting the aforementioned estimates into (7.4.4). Finally, if $\mathrm{N}(\mathscr{D}_E) = 1$ then E has everywhere good reduction. This implies that E is isotrivial, meaning that its j-invariant $j(E)$ is an element of the field of constants $\mathbb{F}_q$. In particular it follows from the theory of elliptic curves over function fields that E is isomorphic to a twist of a constant elliptic curve over K. Since twisting by any non-constant element of K increases the norm $\mathrm{N}(\mathscr{D}_E)$, which is for example proved in Proposition 5.7.1 in [56] for quadratic twists (the other twists follow analogously). We deduce that the number of elliptic curves E with everywhere good reductions is bounded by $O_{\mathbb{F}_q}(1)$. In particular we have $B_{0,E} \gg_{\mathbb{F}_q} 1$ for all such curves and thus obtain (7.4.5).

Therefore it only remains to remove the dependence on H_E in (7.4.5) to complete the proof of Proposition 7.1.4 for elliptic curves in Weierstrass form. To do so, we reprove Theorem 4 of Heath-Brown [104] over arbitrary global fields.

Lemma 7.4.2. *Let $F \in \mathcal{O}_K[x_1, x_2, x_3]$ be an irreducible form of degree d whose coefficient vector lies in Z_{N-1}, where $N = (d+1)(d+2)/2$. Then*

$$\#\{\boldsymbol{x} \in \mathbb{P}^2(K) \colon F(\boldsymbol{x}) = 0,\ H(\boldsymbol{x}) \leq B\} \leq d^2 \quad \text{or} \quad \|F\| \ll B^{dN/s_K}.$$

Proof. The proof is identical to that of Heath-Brown, and so we shall be brief. Write $N = (d+1)(d+2)/2$ and suppose $N_F(B) \geq d^2 + 1$. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in Z_2$ be such that $H(\boldsymbol{x}_i) \leq B$ and $F(\boldsymbol{x}_i) = 0$. Consider the $(d^2+1) \times N$ matrix M whose ith row consists of all possible monomials of degree d formed by the coordinates of $\boldsymbol{x}_i$. If $\boldsymbol{f} \in Z_{N-1}$ is the coefficent-vector associated to F, then we have $M\boldsymbol{f} = 0$. Moreover, as $\boldsymbol{f} \neq 0$, M must have rank at most $N-1$ and so there exists a non-zero solution $\boldsymbol{g} \in Z_{N-1}$ constructed out of the minors of M, so that $\|\boldsymbol{g}\| \ll B^{dN/s_K}$. Let G be the form of degree d associated to $\boldsymbol{g}$. Since the hypersurfaces defined by F and G intersect in d^2+1 points, Bézout's theorem implies that G is a multiple of F, so that $\|F\| \ll \|G\| \ll B^{dN/s_K}$ since both $\boldsymbol{f}$ and $\boldsymbol{g}$ belong to Z_{N-1}. □

It is clear that E given as in (7.4.2) is defined by a form whose coefficients define the unit ideal and upon multiplying it by a unit $u \in \mathcal{O}_K$, we may assume that its coefficient vector lies in Z_9. We then have

$$H_E \leq \prod_{\nu \mid \infty} \max\{1, |a|_\nu, |b|_\nu|\} = \prod_{\nu \mid \infty} \max\{|u|_\nu, |ua|_\nu, |ub|_\nu\} \ll \|E\|^{s_K}.$$

In particular, Lemma 7.4.2 implies that either $N_E(B) \leq 9$ or $H_E \ll B^{10}$. The former case is clearly sufficient and if the latter holds, then (7.4.5) hands us the estimate

$$N_E(B) \ll B^{\varepsilon}, \tag{7.4.6}$$

for any elliptic curve in Weierstrass form where the implied constant only depends on the ground field K.

We will now use (7.4.6) to deduce Proposition 7.1.4. First let us suppose that $E \subset \mathbb{P}^2$ is a smooth genus 1 curve given as the vanishing locus of a non-singular cubic form $G \in \mathcal{O}_K[x, y, z]$. If $N_E(B) = 0$, then we clearly have the desired upper bound. If not, we can find $t_0 \in E(K)$ with $H(t_0) \leq B$. and use t_0 to construct a birational map $\theta \colon \mathbb{P}^2 \to \mathbb{P}^2$ defined over K transforming E into an elliptic curve E' defined in Weierstrass form. It is clear that θ has coefficents that are rational functions in the coefficients of G and the coordinates of t_0. It thus follows for any $x \in E(K)$ with $H(x) \leq B$ that $H(\theta(x)) \leq CH(t_0)^A \|G\|^A$ for some absolute constant $C > 0$ and hence

$$N_E(B) \ll N_{E'}(CB^A \|G\|^A).$$

By Lemma 7.4.2 we have $N_E(B) \leq 9$ or $\|G\| \ll B^A$, so that in the latter case (7.4.6) yields

$$N_E(B) \ll (C\|G\|^A B^A)^{\varepsilon} \ll B^{\varepsilon},$$

which is sufficient for the proof of Proposition 7.1.4 for plane elliptic curves.

Next we suppose $E \subset \mathbb{P}^3$. If E is contained in any plane, say $H \subset \mathbb{P}^3$, we can find a point $p \in \mathbb{P}^3(K) \setminus H(K)$ with $H(p) \ll 1$. Projecting away from p gives a map $\varphi \colon \mathbb{P}^3 \setminus \{p\} \to \mathbb{P}^2(K)$ that restricts to an isomorphism $E \to \varphi(E) \subset \mathbb{P}^2(K)$. Since $H(p) \ll 1$, it follows that there exists some absolute constant $C > 0$ such that $H(\varphi(x)) \leq CH(x)$ for any $x \in \mathbb{P}^3(K) \setminus \{p\}$, so that

$$N_E(B) \leq N_{\varphi(E)}(CB) \ll B^{\varepsilon},$$

by what we have shown for planar curves. We may therefore assume that E is not contained in any hyperplane. Suppose there exists $t_0 \in E(K)$ with $H(t_0) \leq B$. Then there exists a change of variables $L\colon \mathbb{P}^3 \to \mathbb{P}^3$ sending t_0 to the point $t_1 = [1, 0, 0, 0]$, so that $H(L(x)) \ll H(t_0)^A H(x) \ll B^A H(x)$. Let E' be the image of E under this linear transformation. Any elliptic curve in $\mathbb{P}^3$ that is not contained in a hyperplane may be defined as the complete intersection of two quadrics. In particular, we may assume that E' is given by

$$x_0 L_1(x_1, x_2, x_3) = q_1(x_1, x_2, x_3) \quad \text{and} \quad x_0 L_2(x_1, x_2, x_3) = q_2(x_1, x_2, x_3) \tag{7.4.7}$$

for some linear forms $L_1, L_2 \in \mathcal{O}_K[x_2, x_3, x_4]$ and quadratic forms $q_1, q_2 \in \mathcal{O}_K[x_2, x_3, x_4]$. Since E' is non-singular, the Jacobian criterion implies that L_1 and L_2 are not proportional. In particular, we can eliminate x_0 from (7.4.7) to produce an equation

$$L_2(x_1, x_2, x_3) q_1(x_1, x_2, x_3) = L_1(x_1, x_2, x_3) q_2(x_1, x_2, x_3)$$

which gives a plane elliptic curve E'' containing the point corresponding to $L_1(x_1, x_2, x_3) = L_2(x_1, x_2, x_3) = 0$. In particular, the map $E' \setminus \{t_1\} \to E''$ defines a birational map $E' \dashrightarrow E''$ which is one-to-one except when $L_1(x_1, x_2, x_3) = L_2(x_1, x_2, x_3) = 0$. There are are most $O(1)$ such points and thus

$$N_E(B) \ll N_{E''}(CB^A) \ll B^{\varepsilon},$$

again by our estimate for planar curves.

The only remaining cases are when $n > 3$, which we shall reduce to the case $n = 3$ shortly. So suppose $E \subset \mathbb{P}^n$ is a non-singular genus 1 curve of degree d. If E is contained in a hyperplane $H \subset \mathbb{P}^n$, we can find a point $p \in \mathbb{P}^n(K) \setminus H(K)$ such that $H(p) \ll 1$. Projection away from p defines an isomorphism $\varphi\colon E \to E' \subset \mathbb{P}^{n-1}$. Moreover, it is easy to see that $H(p) \ll 1$ implies $H(\varphi(x)) \ll H(x)$. We may therefore assume that E is not contained in any hyperplane from now on. Let $S(E) \subset \mathbb{P}^n$ denote the secant variety of E, which by definition is the closure of all lines that meet E in two points. It is well known that $\dim(S(E)) \leq 3$, $S(E)$ is irreducible and that projection away from a point $p \in \mathbb{P}^n$ restricts to a closed immersion of E if and only if $p \notin S(E)$. Theorem 4.3 of [58] gives a formula for the degree of $S(E)$ that only depends on d and n, and since $n > \dim(S(E))$ we may use (7.2.3) to deduce the existence of a point $p \in \mathbb{P}^n(K) \setminus S(E)(K)$ such that $H(p) \ll_{d,n,K} 1$ and $\varphi\colon \mathbb{P}^n \setminus \{p\} \to \mathbb{P}^{n-1}$ restricts to an isomorphism $E \to E' \subset \mathbb{P}^{n-1}$, where $E' = \varphi(E)$ is a non-singular genus one curve of degree at most d. Since $H(p) \ll 1$, we again have $H(\varphi(x)) \ll_{d,n,K} H(x)$. Regardless of whether E is contained in a hyperplane or not, we have shown the existence of an elliptic curve $E' \subset \mathbb{P}^{n-1}$ of degree at most d such that

$$N_E(B) \leq N_{E'}(CB)$$

for some constant C only depending on d, n and K. As $n > 3$, it is clear that we can continue this process until $E' \subset \mathbb{P}^3$, a case that we already dealt with. Therefore, we have completed our proof of Proposition 7.1.4.

Remark. It is well known that any elliptic curve over a global field K with $\mathrm{char}(K) \neq 2, 3$ can be put into short Weierstrass form [185, Chapter III, Proposition 3.1]. This is usually proved via the Riemann-Roch theorem. However, for our applications it is important that we keep track of how the height of points is affected, which we could achieve by working with projections instead.

7.5 Conic bundles

The aim of this section is to prove Theorem 7.1.3. That is, we shall prove that if $\mathrm{char}(K) \neq 2$ and X is a del Pezzo surface of degree $d = 4$ or $d = 5$ admitting a conic bundle structure, then we have

$$N_U(B) \ll B^{1+\varepsilon}.$$

From now on until the end of this work, all implied constants are allowed to depend on the del Pezzo surface under consideration.

7.5.1 del Pezzo surfaces of degree 5 with a conic bundle structure

According to Section 5 of [77] a del Pezzo surface of degree 5 with a conic bundle structure may be realised as a nonsingular hypersurface defined by

$$sQ_1(x_0, x_1, x_2) + tQ_2(x_0, x_1, x_2) = 0 \tag{7.5.1}$$

inside $\mathbb{P}^1 \times \mathbb{P}^2$. At the beginning of Section 5 in [77] the authors make the assumption that the field K is perfect. This is not necessary in order to show that the X can be realised as described above, in particular their proof for (7.5.1) goes through as long as $\mathrm{char}(K) \neq 2$. Since X is nonsingular the determinant

$$\Delta(s,t) = \det(sM_1 + tM_2)$$

defines a separable cubic form in (s,t), where M_i are the 3×3 matrices defining Q_i. The associated height function is given by

$$H([s,t],[x_0,x_1,x_2]) = H([s,t])H([x_0,x_1,x_2]),$$

where the factors on the right hand side denote the usual height in $\mathbb{P}^1$ and $\mathbb{P}^2$, respectively. Furthermore, the exceptional curves on X are given by the points such that $Q_1(\boldsymbol{x}) = Q_2(\boldsymbol{x}) = 0$ or $\Delta(s,t) = 0$. We proceed in two cases.

Case 1: $H([s,t]) \leq B^{1/2}$. Since we wish to count points such that $H(\boldsymbol{x})H([s,t]) \leq B$ we can cover the contribution to the counting function $N_U(B)$ by

$$\sum_{\substack{\|(s,t)\|_\infty \leq B^{1/2} \\ \Delta(s,t) \neq 0}} N_{(s,t)}\left(\frac{B}{\|(s,t)\|_\infty}\right),$$

where the sum runs over $(s,t) \in Z_1$ and where

$$N_{(s,t)}(R) := \#\left\{\boldsymbol{x} \in Z_2 \colon \|\boldsymbol{x}\| \ll R^{1/s_K},\ sQ_1(\boldsymbol{x}) + tQ_2(\boldsymbol{x}) = 0\right\}.$$

Note that we are permitted to exclude the zeroes of the discriminant from this summation, since they lie on the exceptional locus. We can apply Proposition 7.3.4 to see that the contribution is bounded above by

$$\sum_{\|(s,t)\|_\infty \leq B^{1/2}} B^\varepsilon \left(1 + \frac{B\mathrm{N}(\Delta_0(s,t))^{1/2}}{\|(s,t)\|_\infty \mathrm{N}(\Delta(s,t))^{1/3}}\right).$$

The proof of Lemma 7 in [27], which is a basic application of elimination theory, carries over identically to our setting and shows that we have $\mathrm{N}(\Delta_0(s,t))^{1/2} \ll 1$. For the remaining range we will divide the values of $\|(s,t)\|$ and $\mathrm{N}(\Delta(s,t))$ into dyadic intervals. If $\|(s,t)\| \sim R^{1/s_K}$ and $\mathrm{N}(\Delta(s,t)) \sim S$ then Lemma 7.2.17 shows that there are at most $R^{1+\varepsilon}S^{1/3}$ many points (s,t) that lie within such a dyadic interval. Hence the contribution from this dyadic interval is bounded by

$$R^{1+\varepsilon}S^{1/3}\left(1+\frac{B}{RS^{1/3}}\right) \ll B^{1+\varepsilon},$$

since $R \leq B^{1/2}$ and $S \ll R^3 \ll B^{3/2}$. The number of dyadic intervals that we have to consider is bounded by $\log^2(B)$ and since each individual contribution is bounded by $B^{1+\varepsilon}$ this suffices to show that the contribution to $N_U(B)$ is indeed bounded by $B^{1+\varepsilon}$.

Case 2: $H([s,t]) \geq B^{1/2}$. In particular we must have $H([x_0,x_1,x_2]) \leq B^{1/2}$. Let $([s,t],[x_0,x_1,x_2])$ be a point of the del Pezzo surface in question. Hence by (7.5.1) a representative of $[s,t]$ takes the shape

$$(s,t) = (Q_1(\boldsymbol{x}), -Q_2(\boldsymbol{x})).$$

Writing $\mathfrak{d}$ for the ideal generated by $Q_1(\boldsymbol{x})$ and $Q_2(\boldsymbol{x})$, we then have

$$H([s,t]) = \frac{\|Q_1(\boldsymbol{x}), Q_2(\boldsymbol{x})\|_\infty}{\mathrm{N}(\mathfrak{d})}.$$

In what follows, given an ideal $\mathfrak{d} \subset \mathcal{O}_K$ such that $\mathrm{N}(\mathfrak{d}) \ll B$ we will count the number of $\boldsymbol{x} \in Z_2$ with $Q_i(\boldsymbol{x}) \equiv 0 \bmod \mathfrak{d}$. In particular, given $\alpha \ll B^{1/2}$ we may divide the height of $\boldsymbol{x}$ into dyadic intervals. We are therefore led to consider sums of the shape

$$\sum_{\substack{\boldsymbol{x}\in Z_2\colon \|\boldsymbol{x}\|_\infty \sim\alpha \\ (Q_1(\boldsymbol{x}),Q_2(\boldsymbol{x}))=\mathfrak{d} \\ \|Q_i(\boldsymbol{x})\|_\infty \leq \frac{B\mathrm{N}(\mathfrak{d})}{\alpha}}} 1. \tag{7.5.2}$$

In view of this we define the set

$$\widetilde{V}_{\mathfrak{d}} = \left\{\boldsymbol{x} \in (\mathcal{O}_K/\mathfrak{d})^3 : (\boldsymbol{x},\mathfrak{d}) \in Z_3,\ Q_1(\boldsymbol{x}) \equiv Q_2(\boldsymbol{x}) \equiv 0 \bmod \mathfrak{d}\right\},$$

where the condition $(\boldsymbol{x},\mathfrak{d}) \in Z_3$ is meant to indicate that the ideal generated by x_0, x_1, x_2 and $\mathfrak{d}$ is equal to one of the fixed representatives $\mathfrak{a}_i$ of the ideal class group of K. Further we define

$$V_{\mathfrak{d}} = \widetilde{V}_{\mathfrak{d}}/(\mathcal{O}_K/\mathfrak{d})^\times.$$

Given $\boldsymbol{y} \in V_{\mathfrak{d}}$ we further define

$$\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \coloneqq \{\boldsymbol{z} \in \mathcal{O}_K^3\colon \boldsymbol{z} \equiv \lambda\boldsymbol{y} \bmod \mathfrak{d} \text{ for some } \lambda \in \mathcal{O}_K\}.$$

Changing the order of summation in (7.5.2), using the notation introduced above it is easy to see that it suffices to estimate

$$S_\alpha \coloneqq \sum_{\mathrm{N}(\mathfrak{d})\ll\alpha^2} \sum_{\boldsymbol{y}\in V_{\mathfrak{d}}} \sum_{\substack{\boldsymbol{x}\in Z_2\colon \|\boldsymbol{x}\|_\infty\sim\alpha \\ \|Q_i(\boldsymbol{x})\|_\infty \leq \frac{B\mathrm{N}(\mathfrak{d})}{\alpha} \\ \boldsymbol{x}\in\Lambda_{\mathfrak{d}}(\boldsymbol{y})}} 1. \tag{7.5.3}$$

To this end, the following two lemmas help us deal with $V_{\mathfrak{d}}$ and the lattice $\Lambda_{\mathfrak{d}}(\boldsymbol{y})$. Since the del Pezzo surface in question is non-singular it follows from the definition (7.5.1) that $Q_1(\boldsymbol{x}) = Q_2(\boldsymbol{x}) = 0$ defines a non-singular intersection in $\mathbb{P}^2$.

Lemma 7.5.1. *Given an ideal $I \subset \mathcal{O}_K$ we have*

$$\#V_I \ll \mathrm{N}(I)^{\varepsilon}.$$

Proof. The proof is the same as Lemma 2 in [30] adjusted to our more general setting and therefore we will be brief. Note first that via multiplicativity and homogeneity it suffices to show

$$\#\widetilde{V}_{\mathfrak{p}^s} \ll \mathrm{N}(\mathfrak{p}^s),$$

for any prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ and any positive integer s. Note also that unless $\mathfrak{p}$ is one of the $O(1)$ many fixed representatives for the class group of K the condition $(\boldsymbol{x}, \mathfrak{p}^s) \in Z_3$ is equivalent to saying that $(x_1, x_2, x_3, \mathfrak{p}^s)$ generates $\mathcal{O}_K$. Also consider the isomorphism

$$\mathcal{O}_K/\mathfrak{p}^s \cong \mathcal{O}_\nu/(\pi)^s,$$

where $\mathcal{O}_\nu \subset K_\nu$, for ν the place corresponding to $\mathfrak{p}$, and where π is a uniformizer for K_ν. Then we find that it suffices to consider

$$\rho(\pi^s) = \#\left\{\boldsymbol{x} \in (\mathcal{O}_\nu/(\pi^s))^3 \colon \langle x_1, x_2, x_3, \pi\rangle = \mathcal{O}_\nu,\ Q_i(\boldsymbol{x}) \equiv 0 \bmod (\pi^s),\ \text{for } i = 1, 2\right\}.$$

For ease of notation write ψ for the additive character on K_ν. Using the orthogonality relation from Lemma 7.2.5 we find

$$\rho(\pi^s) = \frac{1}{\mathrm{N}(\pi)^{2s}} \sum_{\boldsymbol{b} \bmod \pi^s} \sideset{}{^*}\sum_{\boldsymbol{x} \bmod \pi^s} \psi\left(\frac{b_1 Q_1(\boldsymbol{x}) + b_2 Q_2(\boldsymbol{x})}{\pi^s}\right), \tag{7.5.4}$$

where $\sum^*$ indicates that we only sum over tuples $\boldsymbol{x}$ such that $\langle \boldsymbol{x}, \pi^s\rangle = \mathcal{O}_\nu$, or equivalently $\pi \nmid x_i$. Extracting common factors between π^s and $\boldsymbol{b}$ in the display above we obtain

$$\rho(\pi^s) = \frac{1}{\mathrm{N}(\pi)^{2s}} \sum_{0 \le i < s} \mathrm{N}(\pi)^{3i} S(s-i) + \mathrm{N}(\pi)^s \left(1 - \frac{1}{\mathrm{N}(\pi)^3}\right),$$

where

$$S(k) = \sideset{}{^*}\sum_{\boldsymbol{b} \bmod \pi^k} \sideset{}{^*}\sum_{\boldsymbol{x} \bmod \pi^k} \psi\left(\frac{F(\boldsymbol{b}, \boldsymbol{x})}{\pi^k}\right),$$

and where $F(\boldsymbol{b}, \boldsymbol{x}) = b_1 Q_1(\boldsymbol{x}) + b_2 Q_2(\boldsymbol{x})$. It suffices to show $S(k) = O(1)$ for $k \ge 2$ and $S(1) = O(\mathrm{N}(\pi)^3)$. Regarding $S(1)$ note that by Bézout's theorem we find that $\rho(\pi) \ll \mathrm{N}(\pi)$ since $\mathcal{O}_\nu/(\pi) \cong \mathcal{O}_K/\mathfrak{p} \cong \mathbb{F}_{\mathrm{N}(\mathfrak{p})}$. Note that there is a finite number of primes $\mathfrak{p}$ the two quadrics might share a common component and so Bézout's theorem does not apply for these cases. However, by choosing the implied constant large enough since there are only finitely many primes involved we still obtain $\rho(\pi) \ll \mathrm{N}(\pi)$ for all primes. Substituting this into (7.5.4) for $s = 1$ we indeed find $S(1) = O(\mathrm{N}(\pi)^3)$. If $k \ge 2$ then after introducing a dummy sum over $a \in (\mathcal{O}_\nu/\pi^k)^\times$, and making a change of variables $\boldsymbol{b} \mapsto a\boldsymbol{b}$ we may evaluate the arising Ramanujan sums to obtain

$$S(k) = \left(1 - \frac{1}{\mathrm{N}(\pi)}\right)^{-1} \Big(\mathcal{N}(k) - \mathrm{N}(\pi)^4 \mathcal{N}(k-1)\Big),$$

where $\mathcal{N}(k)$ denotes the number of solutions to $F(\boldsymbol{b}, \boldsymbol{x}) \equiv 0 \bmod \pi^k$ such that $\pi \nmid \boldsymbol{b}$ and $\pi \nmid \boldsymbol{x}$. Writing the discriminant $D \in \mathcal{O}_K$ of the pair of quadrics Q_i as the resultant of the 5 quadratic forms appearing in $\nabla F(\boldsymbol{b}, \boldsymbol{x})$ (see [82, Chapter 13]) via elimination theory we obtain polynomials $G_{ij}(\boldsymbol{y})$ with coefficients in $\mathcal{O}_K$ where $\boldsymbol{y} = (\boldsymbol{b}, \boldsymbol{x})$ and a positive integer R such that

$$Dy_i^R = \sum_{1 \le j \le 5} G_{ij}(\boldsymbol{y}) \frac{\partial F}{\partial y_i}, \quad \text{for } 1 \le i \le 5.$$

Since Q_i define a non-singular intersection over K we have that $D \neq 0$. Writing $\delta_\pi = \nu_{\mathfrak{p}}(D)$ we thus obtain that if $\pi^m \mid \nabla F(\boldsymbol{b}, \boldsymbol{x})$ and $\pi \nmid \boldsymbol{x}$ and $\pi \nmid \boldsymbol{b}$ then $m \le \delta_\pi$. Since $D \in \mathcal{O}_K$ is nonzero we first note that $\delta_\pi = 0$ for all but finitely many $\mathfrak{p}$. In particular, if $\delta_\pi > 0$ and $2 \le k \le 2\delta_\pi + 1$ then we may choose our implied constant sufficiently big, only depending on Q_i to get $S(k) = O(1)$. If $k \ge 2\delta + 2$ then we may apply a standard Hensel lifting argument to show that

$$C_m(k+1) = \mathrm{N}(\pi)^4 C_m(k),$$

where $C_m(k)$ is the number of $\boldsymbol{y} \bmod \pi^k$ with $\pi \nmid \boldsymbol{b}$ and $\pi \nmid \boldsymbol{x}$ such that $\pi^k \mid F(\boldsymbol{y})$ and $\pi^m \parallel \nabla F(\boldsymbol{y})$. We note that since $\mathcal{O}_\nu$ is a principal ideal domain the lifting argument goes through completely analogously. Finally, noting that

$$\mathcal{N}(k) = \sum_{0 \le m \le \delta} C_m(k)$$

yields $S(k) = 0$ if $k \ge 2\delta + 2$, and thus $S(k) = O(1)$ for all $k \ge 2$, thereby completing the proof. □

Lemma 7.5.2. *Given $\boldsymbol{y} \in \mathcal{O}_K^3$ such that $(y_1, y_2, y_3, \mathfrak{d})$ generates one of the fixed representatives $\mathfrak{a}_i$ of the ideal class group, the lattice $\Lambda_{\mathfrak{d}}(\boldsymbol{y})$ has rank 3 and*

$$\det \Lambda_{\mathfrak{d}}(\boldsymbol{y}) \asymp_K \mathrm{N}(\mathfrak{d})^2.$$

Proof. First note that $\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \supset (\mathfrak{d}\mathcal{O}_K)^3$ and therefore

$$\det \Lambda_{\mathfrak{d}}(\boldsymbol{y}) = [\mathcal{O}_K^3 \colon \Lambda_{\mathfrak{d}}(\boldsymbol{y})] = \frac{[\mathcal{O}_K^3 \colon (\mathfrak{d}\mathcal{O}_K)^3]}{[\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \colon (\mathfrak{d}\mathcal{O}_K)^3]} = \frac{\mathrm{N}(\mathfrak{d})^3}{[\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \colon (\mathfrak{d}\mathcal{O}_K)^3]}.$$

Hence it suffices to establish $[\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \colon (\mathfrak{d}\mathcal{O}_K)^3] \asymp \mathrm{N}(\mathfrak{d})$. First note that it is clear that $\lambda\boldsymbol{y} + (\mathfrak{d}\mathcal{O}_K)^3$ where λ runs through all possible elements of $\mathcal{O}_K/\mathfrak{d}$ exhausts all possible cosets in $\Lambda_{\mathfrak{d}}(\boldsymbol{y})/(\mathfrak{d}\mathcal{O}_K)^3$ and hence $[\Lambda_{\mathfrak{d}}(\boldsymbol{y}) \colon (\mathfrak{d}\mathcal{O}_K)^3] \le \mathrm{N}(\mathfrak{d})$.

For a lower bound regarding the index, given $c > 0$ we consider elements $r \in \mathcal{O}_K$ such that $\|r\| < c\mathrm{N}(\mathfrak{d})^{1/s_K}$. We claim that if we choose c small enough, only depending on K, then the elements $r\boldsymbol{y} + (\mathfrak{d}\mathcal{O}_K)^3$ where r runs through elements as above are all distinct cosets. Since the number of elements r with the above property is $\gg \mathrm{N}(\mathfrak{d})$ this suffices in order to prove the lemma. If not of all the above cosets are distinct then upon taking their difference we see that there exists some $r_0 \in \mathcal{O}_K \setminus \{0\}$ with $\|r_0\| \ll c\mathrm{N}(\mathfrak{d})^{1/s_K}$ and $r_0 y_i \in \mathfrak{d}$ for $i = 1, 2, 3$. In particular we have

$$\langle r_0 y_1, r_0 y_2, r_0 y_3, r_0 \mathfrak{d} \rangle \subset \mathfrak{d}$$

Since the ideal generated by y_i and $\mathfrak{d}$ is one of the fixed representatives of the class group of K there exists some nonzero $z \in \langle y_1, y_2, y_3, \mathfrak{d} \rangle$ such that $\|z\| = O_K(1)$. Hence $r_0 z \in \mathfrak{d}$ and so after multiplying r_0 with a unit if necessary we find

$$\mathrm{N}(\mathfrak{d})^{1/s_K} \le \|r_0 z\|_\infty^{1/s_K} \ll \|r_0 z\|.$$

Therefore we have

$$\mathrm{N}(\mathfrak{d})^{1/s_K} \ll \|r_0 z\| \ll \|r_0\| \|z\| \ll \|r_0\| \ll c\mathrm{N}(\mathfrak{d})^{1/s_K},$$

and so upon choosing $c > 0$ small enough we obtain a contradiction. Thus the claim and consequently the Lemma transpire. $\square$

We now return to estimating S_α as it was defined in (7.5.3). Since there are only B^ε many different values of α we need to consider, it suffices to show $S_\alpha \ll B^{1+\varepsilon}$.

We denote the successive minima of $\Lambda_{\mathfrak{d}}(\boldsymbol{y})$ by $\lambda_1 \leq \lambda_2 \leq \lambda_3$, as they were defined in Section 7.3.1. Note first that for every $\boldsymbol{y} \in V_{\mathfrak{d}}$ using Lemma 7.3.1 and Lemma 7.2.7 we may take a representative $\boldsymbol{y} \in \mathcal{O}_K^3$ such that $|\boldsymbol{y}|_\nu \ll \mathrm{N}(\mathfrak{d})^{d_\nu/d_K}$. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{O}_K^3$ with $|\boldsymbol{x}_i|_\nu \ll 1$ for all $\nu \mid \infty$ and $i = 1, 2$ such that $\boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2$ are linearly independent. Moreover, choose $\delta \in \mathfrak{d}$ such that $|\delta|_\nu \ll \mathrm{N}(\mathfrak{d})^{d_\nu/d_K}$, which is possible by Lemma 7.3.1. Then the set $\{\boldsymbol{y}, \boldsymbol{y}+\delta\boldsymbol{x}_1, \boldsymbol{y}+\delta\boldsymbol{x}_2\}$ constitutes a linearly independent set inside $\Lambda_{\mathfrak{d}}(\boldsymbol{y})$. The ν-adic absolute value of each of these vectors is bounded by $O(\mathrm{N}(\mathfrak{d})^{d_\nu/d_K})$ for all $\nu \mid \infty$. We deduce

$$\lambda_3^{d_K} \ll \prod_{\nu\mid\infty} \mathrm{N}(\mathfrak{d})^{d_\nu/d_K} = \mathrm{N}(\mathfrak{d}),$$

and so Lemma 7.3.1 yields

$$\frac{1}{(\lambda_1\lambda_2)^{d_K}} \ll \frac{1}{\mathrm{N}(\mathfrak{d})},$$

since $\det \Lambda_{\mathfrak{d}}(\boldsymbol{y}) \asymp \mathrm{N}(\mathfrak{d})^2$ by Lemma 7.5.2. We will use this fact as well as Lemma 7.3.3 in order to estimate S_α. If $\alpha < \lambda_2^{d_K}$ then the number of lattice points with $\|\boldsymbol{x}\|_\infty \sim \alpha$ is bounded by $O(1)$ since we only consider the contribution from primitive points. Thus S_α is bounded by

$$\sum_{\mathrm{N}(\mathfrak{d})\ll\alpha^2} \sum_{\boldsymbol{y}\in V_{\mathfrak{d}}} 1 \ll \alpha^{2+\varepsilon} \ll B^{1+\varepsilon}.$$

If $\lambda_2^{d_K} \leq \alpha < \lambda_3^{d_K}$, then via Lemma 7.3.3 the number of lattice points that we count is bounded by $O(\alpha^2/(\lambda_1\lambda_2)^{d_K})$. Thus we obtain

$$S_\alpha \ll \sum_{\mathrm{N}(\mathfrak{d})\ll\alpha^2} \sum_{\boldsymbol{y}\in V_{\mathfrak{d}}} \frac{\alpha^2}{\mathrm{N}(\mathfrak{d})} \ll \alpha^{2+\varepsilon} \ll B^{1+\varepsilon}.$$

Finally, if $\lambda_3^{d_K} \leq \alpha$ then we divide the contribution into the range $\alpha^3/B \ll \mathrm{N}(\mathfrak{d}) \ll \alpha^2$ and the range $\mathrm{N}(\mathfrak{d}) \ll \alpha^3/B$. For the first range we may employ Lemma 7.3.3 to find that the contribution is bounded by

$$\sum_{\alpha^3/B\ll\mathrm{N}(\mathfrak{d})\ll\alpha^2} \frac{\alpha^3}{\mathrm{N}(\mathfrak{d})^2} \ll \alpha^3(\alpha^{-2} + B/\alpha^3)\alpha^\varepsilon \ll B^{1+\varepsilon}.$$

It remains to handle

$$\sum_{\mathrm{N}(\mathfrak{d})\ll\alpha^3/B} \sum_{\boldsymbol{y}\in V_{\mathfrak{d}}} \sum_{\substack{\boldsymbol{x}\in Z_2\,:\, \|\boldsymbol{x}\|_\infty\sim\alpha \\ \|Q_i(\boldsymbol{x})\|_\infty \leq \frac{B\mathrm{N}(\mathfrak{d})}{\alpha} \\ \boldsymbol{x}\in\Lambda_{\mathfrak{d}}(\boldsymbol{y})}} 1. \tag{7.5.5}$$

At this point it does not suffice to just count the points contained in a ball inside the lattice, and we need to take advantage of the additional restriction given by the quadratic forms. We deal with this in the following Lemma.

Lemma 7.5.3. *Let $\Lambda \subset K^3$ be a lattice with successive minima $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and let $Q \in \mathcal{O}_K[x_1, x_2, x_3]$ be a quadratic form of rank at least 2. Let $\alpha, R \geq 1$ be real numbers such that $R^{1/2} \ll \alpha \ll R$. Consider*

$$N(\alpha, R) \coloneqq \#\left\{\boldsymbol{x} \in Z_2 \cap \Lambda \colon \|\boldsymbol{x}\|_\infty < \alpha,\ \|Q(\boldsymbol{x})\|_\infty < R\right\}.$$

We have the bound

$$N(\alpha, R) \ll \frac{\alpha^{4+\varepsilon}}{R^2} + \frac{\alpha^{3+\varepsilon}}{R\lambda_1^{d_K}} + \frac{\alpha^{2+\varepsilon}}{(\lambda_1\lambda_2)^{d_K}} + \frac{\alpha^{1+\varepsilon}R}{\det\Lambda} + \frac{\alpha}{(\det\Lambda)^{1/3}}.$$

Deferring the proof for now, applying Lemma 7.5.3 to the inner sum in (7.5.5) with $R = B\mathrm{N}(\mathfrak{d})/\alpha$ we find that

$$S_\alpha \ll B^\varepsilon \sum_{\mathrm{N}(\mathfrak{d}) \ll \alpha^3/B} \left(\frac{\alpha^6}{B^2\mathrm{N}(\mathfrak{d})^2} + \frac{\alpha^4}{B\mathrm{N}(\mathfrak{d})\lambda_1^{d_K}} + \frac{\alpha^2}{(\lambda_1\lambda_2)^{d_K}} + \frac{B}{\mathrm{N}(\mathfrak{d})} + \frac{\alpha}{\mathrm{N}(\mathfrak{d})^{2/3}}\right).$$

Using the fact that $\lambda_1 \gg 1$ as well as $1/(\lambda_1\lambda_2)^{d_K} \ll \mathrm{N}(\mathfrak{d})^{-1}$ one may easily check that the above expression is bounded by $B^{1+\varepsilon}$. Once we prove Lemma 7.5.3 this concludes the proof of Theorem 7.1.3 for $d = 5$.

Proof of Lemma 7.5.3. First note that the contribution to $N(\alpha, R)$ from the vectors $\boldsymbol{x} \in Z_2 \cap \Lambda$ such that $Q(\boldsymbol{x}) = 0$ is bounded above by $O(1 + \alpha/\det(\Lambda)^{1/3})$ by using the same argument as in the proof of Corollary 7.3.6.

For the remaining contribution we begin by decomposing the possible local absolute values that the quadratic forms may take into dyadic intervals. Note first that if $Q(\boldsymbol{x}) \neq 0$ then $\prod_\nu |Q(\boldsymbol{x})|_\nu \geq 1$. Further, if $\boldsymbol{x}$ is counted by $N(\alpha, R)$ then $|Q(\boldsymbol{x})|_\nu \ll \alpha^{2/s_K}$ holds for all ν. We deduce that the vectors counted by $N(\alpha, R)$ satisfy

$$\alpha^{-2} \leq |Q(\boldsymbol{x})|_\nu \ll \alpha^{2/s_K},$$

for all $\nu \mid \infty$. Fix now a place $\omega \mid \infty$ and let $\boldsymbol{r} = (r_\nu)_{\nu \neq \omega}$. We define

$$N(\alpha, \boldsymbol{r}) = \#\left\{\boldsymbol{x} \in \Lambda \colon |\boldsymbol{x}|_\nu \ll \alpha^{1/s_K},\ |Q(\boldsymbol{x})|_\nu \sim r_\nu \text{ and } \|Q(\boldsymbol{x})\|_\infty \leq R\right\}.$$

It suffices to obtain establish the desired upper bound for $N(\alpha, \boldsymbol{r})$ whenever $\alpha^{-2} \leq r_\nu \leq R\alpha^{2s_K}$ is satisfied, since the number of dyadic decompositions we require in order to cover the set of points counted by $N(\alpha, R)$ is bounded by α^ε. Writing $r_\omega = 2^{1-s_K}R/\prod_{\nu \neq \omega} r_\nu$ we clearly see that

$$N(\alpha, \boldsymbol{r}) \leq \#\left\{\boldsymbol{x} \in \Lambda \colon |\boldsymbol{x}|_\nu \ll \alpha^{1/s_K}, |Q(\boldsymbol{x})|_\nu \leq r_\nu\right\}$$

Consider the set

$$S(\alpha, (r_\nu)_\nu) = \prod_\nu \left\{\boldsymbol{x} \in K_\nu^3 \colon |\boldsymbol{x}|_\nu \ll \alpha^{1/s_K}, |Q(\boldsymbol{x})|_\nu \leq r_\nu\right\}.$$

Writing $\Delta \colon K^3 \hookrightarrow \prod_\nu K_\nu^3$ for the diagonal embedding, we see that

$$N(\alpha, \boldsymbol{r}) \ll \#(\Delta^{-1}(S(\alpha, (r_\nu)_\nu)) \cap \Lambda).$$

Given $\nu \mid \infty$ such that K_ν is not isomorphic to the complex numbers, consider the box

$$\mathcal{B}_\nu = \left\{\boldsymbol{x} \in K_\nu^3 \colon |\boldsymbol{x}|_\nu \leq r_\nu/\alpha^{1/s_K}\right\}.$$

If $K_\nu \cong \mathbb{C}$ then write

$$\mathcal{B}_\nu = \left\{ \boldsymbol{x} \in K_\nu^3 \colon \max_i |(\mathrm{Re}(x_i), \mathrm{Im}(x_i)| \leq \sqrt{r_\nu/\alpha^{1/s_K}} \right\},$$

where the absolute value in the definition of $\mathcal{B}_\nu$ is taken to be the usual absolute value on $\mathbb{R}$. Note that for complex ν we have

$$\{\boldsymbol{x} \in K_\nu^3 \colon |\boldsymbol{x}|_\nu \leq r_\nu/\alpha^{1/s_K}\} \subset \mathcal{B}_\nu \subset \{\boldsymbol{x} \in K_\nu^3 \colon |\boldsymbol{x}|_\nu \leq 2r_\nu/\alpha^{1/s_K}\}.$$

Define

$$\mathcal{B} = \prod_\nu \mathcal{B}_\nu.$$

Since $S(\alpha, (r_\nu)_\nu)$ defines a bounded set we may cover it with M, say, translates of $\mathcal{B}$, denoted by $\mathcal{B}_i$. We may choose the $\mathcal{B}_i$ such that the pairwise intersection of these boxes has trivial measure, and also such that $\mathcal{B}_i \cap S(\alpha, (r_\nu)_\nu) \neq \emptyset$.

By translating $\Delta^{-1}(\mathcal{B}_i)$ by a point contained in $\Lambda \cap \Delta^{-1}(\mathcal{B}_i)$ if necessary, then via translation invariance of the lattice we find

$$\#(\Lambda \cap \Delta^{-1}(\mathcal{B}_i)) \ll \# \left\{ \boldsymbol{x} \in \Lambda \colon |\boldsymbol{x}|_\nu \ll \frac{r_\nu}{\alpha^{1/s_K}} \text{ for all } \nu \mid \infty \right\},$$

for all $i = 1, \ldots, M$. Lemma 7.3.3 therefore delivers

$$\#(\Lambda \cap \Delta^{-1}(\mathcal{B}_i)) \ll 1 + \frac{\prod_\nu r_\nu}{\alpha \lambda_1^{d_K}} + \frac{(\prod_\nu r_\nu)^2}{\alpha^2 (\lambda_1 \lambda_2)^{d_K}} + \frac{(\prod_\nu r_\nu)^3}{\alpha^3 \det(\Lambda)},$$

and thus

$$N(\alpha, \boldsymbol{r}) \ll M \left(1 + \frac{\prod_\nu r_\nu}{\alpha \lambda_1^{d_K}} + \frac{(\prod_\nu r_\nu)^2}{\alpha^2 (\lambda_1 \lambda_2)^{d_K}} + \frac{(\prod_\nu r_\nu)^3}{\alpha^3 \det(\Lambda)} \right). \tag{7.5.6}$$

To obtain an upper bound on the number of boxes M that we need in order to cover this region, consider $(\boldsymbol{x}_\nu)_\nu \in \mathcal{B}_i \cap S(\alpha, (r_\nu)_\nu)$. Then all points inside $\mathcal{B}_i$ are of the form $(\boldsymbol{x}_\nu)_\nu + (\boldsymbol{y}_\nu)_\nu$ with $|\boldsymbol{y}_\nu|_\nu \ll r_\nu/\alpha^{1/s_K}$. Therefore we have that

$$|\boldsymbol{x}_\nu + \boldsymbol{y}_\nu|_\nu \ll \alpha^{1/s_K} + \frac{r_\nu}{\alpha^{1/s_K}} \ll \alpha^{1/s_K},$$

since we only consider $(r_\nu)_\nu$ such that $r_\nu \ll \alpha^{2/s_K}$ holds. Further we find

$$\begin{aligned} |Q(\boldsymbol{x}_\nu + \boldsymbol{y}_\nu)|_\nu &= |Q(\boldsymbol{x}_\nu) + \boldsymbol{x}_\nu^T \nabla Q(\boldsymbol{y}_\nu) + Q(\boldsymbol{y}_{\boldsymbol{\nu}})|_\nu \\ &\ll r_\nu + \alpha^{1/s_K} \frac{r_\nu}{\alpha^{1/s_K}} + \left(\frac{r_\nu}{\alpha^{1/s_K}} \right)^2 \\ &\ll r_\nu. \end{aligned}$$

Since i was arbitrary, we deduce that there exists a constant $C > 0$, only depending on the quadratic form Q and K such that

$$\bigcup_{i=1}^M \mathcal{B}_i \subset S(C\alpha, C(r_\nu)_\nu).$$

Hence

$$M \ll \left(\frac{\alpha}{\prod_\nu r_\nu} \right)^3 \operatorname{vol} S(\alpha, (r_\nu)_\nu). \tag{7.5.7}$$

We will now compute the volume of $S(Q,\alpha,(r_\nu)_\nu)$. Note first that

$$\operatorname{vol} S(\alpha,(r_\nu)_\nu) = \prod_{\nu|\infty} \operatorname{vol} S_\nu(\alpha, r_\nu),$$

where

$$S_\nu(\alpha, r_\nu) = \left\{ \boldsymbol{x} \in K_\nu^3 \colon |\boldsymbol{x}|_\nu \ll \alpha^{1/s_K}, |Q(\boldsymbol{x})|_\nu \ll r_\nu \right\}.$$

We consider two different cases. Firstly, if Q is anisotropic over K_ν then there exists some constant $D>0$ depending only on Q such that $|Q(\boldsymbol{x})|_\nu \geq D$ holds for all $\boldsymbol{x} \in K_\nu^3$ such that $|\boldsymbol{x}|_\nu = 1$. Therefore $|Q(\boldsymbol{x})|_\nu \geq D|\boldsymbol{x}|_\nu^2$ holds, and so $\boldsymbol{x} \in S_\nu(\alpha, R)$ implies that we must have $|\boldsymbol{x}|_\nu \ll r_\nu^{1/2}$. As a result we easily deduce

$$\operatorname{vol} S_\nu(\alpha, R) \ll r_\nu^{3/2} \ll \alpha^{1/s_K} r_\nu$$

via recalling that we only consider r_ν such that $r_\nu \ll \alpha^{2/s_K}$. If Q is isotropic on the other hand, then by the classical theory of quadratic forms after a linear transformation of the coordinates in K_ν we may assume that it is of the shape

$$Q(x,y,z) = yz - Dx^2,$$

for some constant $D \in K_\nu$ depending on Q that can be 0 if the rank of Q is 2. Note first that away from the nullset $z=0$ we clearly have

$$\operatorname{vol}\left\{ y \in K_\nu \colon |yz - Dx^2|_\nu \ll r_\nu, |y|_\nu \ll \alpha^{1/s_K} \right\} \ll \min\left\{ \alpha^{1/s_K}, \frac{r_\nu}{|z|_\nu} \right\},$$

for any $x \in K_\nu$. Hence we find

$$\operatorname{vol} S_\nu(\alpha, r_\nu) \ll \int_{|x|_\nu \ll \alpha^{1/s_K}} \int_{|z|_\nu \ll \alpha^{1/s_K}} \min\left\{ \alpha^{1/s_K}, \frac{r_\nu}{|z|_\nu} \right\} dz dx \ll r_\nu \alpha^{1/s_K + \varepsilon}.$$

We conclude that

$$\operatorname{vol} S(\alpha, R) \ll \alpha^{1+\varepsilon} \prod_{\nu|\infty} r_\nu.$$

From (7.5.7) it thus follows that

$$M \ll \frac{\alpha^{4+\varepsilon}}{(\prod_\nu r_\nu)^2}.$$

Finally, recall $\prod_\nu r_\nu \asymp R$. Using all of this along with (7.5.6) and recalling the contribution from $Q(\boldsymbol{x}) = 0$ we have

$$N(\alpha, R) \ll \frac{\alpha^{4+\varepsilon}}{R^2} + \frac{\alpha^{3+\varepsilon}}{R\lambda_1^{d_K}} + \frac{\alpha^{2+\varepsilon}}{(\lambda_1\lambda_2)^{d_K}} + \frac{\alpha^{1+\varepsilon} R}{\det \Lambda} + \frac{\alpha}{(\det \Lambda)^{1/3}}$$

which completes the proof of this Lemma. □

7.5.2 del Pezzo surfaces of degree 4

Our proof is very similar to the one given by Browning and Swarbrick-Jones [43] except for the fact that we avoid the Thue–Siegel–Roth theorem.

A del Pezzo surface of degree 4 may be written as a complete intersection of two quadrics inside $\mathbb{P}^4$. After a change of variables if necessary, if X contains a conic we may write the system of quadrics as

$$x_0x_1 - x_2x_3 = Q(x_0, x_1, x_2, x_3) + x_4^2 = 0$$

inside $\mathbb{P}^4$, where Q is a quadratic form defined over $\mathcal{O}_K$. Taking $U \subset X$ to be the Zariski open set obtained after removing the exceptional locus (in this case it consists of 16 lines) we find two conic fibrations $\pi_i \colon U \to \mathbb{P}^1$ explicitly given by

$$\pi_1(\boldsymbol{x}) = \begin{cases} [x_0, x_2] & \text{if } (x_0, x_2) \neq \mathbf{0}, \\ [x_3, x_1] & \text{if } (x_3, x_1) \neq \mathbf{0}, \end{cases}$$

and

$$\pi_2(\boldsymbol{x}) = \begin{cases} [x_0, x_3] & \text{if } (x_0, x_3) \neq \mathbf{0}, \\ [x_2, x_1] & \text{if } (x_2, x_1) \neq \mathbf{0}. \end{cases}$$

In particular, this gives rise to a well-defined morphism $\pi \colon X \to \mathbb{P}^1 \times \mathbb{P}^1$. Denoting the Segre embedding by $\psi \colon \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3$ one can easily verify that we have

$$H(\pi_1(\boldsymbol{x}))H(\pi_2(\boldsymbol{x})) = H(\psi(\pi(\boldsymbol{x})))$$

for all $\boldsymbol{x} \in X(K)$. One may further easily check that

$$\psi \circ \pi([x_0, x_1, x_2, x_3, x_4]) = [x_0, x_3, x_2, x_1]$$

holds, whenever $[x_0, x_1, x_2, x_3, x_4] \in U(K)$, and so the degree of $\psi \circ \pi \colon U \to \mathbb{P}^3$ is 1. The functoriality of heights [182, Section 2.3] therefore shows that we have

$$H(\pi_1(\boldsymbol{x}))H(\pi_2(\boldsymbol{x})) \ll H(\boldsymbol{x}),$$

for any $\boldsymbol{x} \in U(K)$. In particular it follows that given $\boldsymbol{x} \in U(K)$ we must have $H(\pi_i(\boldsymbol{x})) \ll H(\boldsymbol{x})^{1/2}$ for $i = 1$ or $i = 2$. We deduce that

$$N_U(B) \leq n_1(B) + n_2(B),$$

where

$$n_i(B) = \#\left\{\boldsymbol{x} \in U(K) \colon H(\boldsymbol{x}) \leq B,\ H(\pi_i(\boldsymbol{x})) \ll B^{1/2}\right\}.$$

It suffices to show $n_1(B) \ll B^{1+\varepsilon}$ since dealing with $n_2(B)$ is identical. Given $(s,t) \in Z_1$ denote by $n_1(B, s, t)$ the cardinality of the points in the fibre $\pi_1^{-1}([s,t]) \cap U(k)$ of height at most $O(B^{1/2})$. In order to show the desired bound for $n_1(B)$ via a dyadic decomposition argument it suffices to show the same bound for

$$n_1(R, B) \coloneqq \sum_{\substack{(s,t) \in Z_1 \\ \|(s,t)\|_\infty \sim R}} n_1(B, s, t).$$

where $R \ll B^{1/2}$. Given $(s,t) \in Z_1$ we have that a representative $(ux, yv, xv, yu, z) \in Z_4$ lies in the fibre $\pi_1^{-1}([s,t]) \cap U(k)$ precisely when

$$Q(sx, yt, xt, ys) + az^2 = 0. \tag{7.5.8}$$

For fixed (s,t) we may regard the above as a ternary quadratic form in (x,y,z). Then the discriminant $\Delta(s,t)$ defines a separable, homogeneous polynomial with $\deg \Delta(s,t) = 4$. Note also that there are no points in $\pi_1([s,t])^{-1} \cap U(k)$ such that $\Delta(s,t) = 0$. We may bound $n_1(B,s,t)$ by the number of $(x,y,z) \in Z_2$ such that (7.5.8) and

$$|(sx, yt, xt, ys, z)|_\nu \ll B^{1/s_K}$$

is satisfied for all $\nu \mid \infty$. Thus we may employ Corollary 7.3.6 in order to find

$$n_1(R,B) \ll \sum_{\substack{(s,t)\in Z_1 \\ \|(s,t)\|_\infty \sim R \\ \Delta(s,t)\neq 0}} B^\varepsilon \left(1 + \frac{B\mathrm{N}(\Delta_0(s,t))^{1/2}}{R^{2/3}\mathrm{N}(\Delta(s,t))^{1/3}}\right).$$

First note that the number of $(s,t) \in Z_1$ such that $\|(s,t)\|_\infty \sim R$ is bounded by $R^2 \ll B$. Further, an argument of Broberg [27, Lemma 7] shows that we have $\mathrm{N}(\Delta_0(s,t)) \ll 1$ for all $(s,t) \in Z_1$ such that $\Delta(s,t) \neq 0$. We are thus left with estimating

$$\sum_{\substack{(s,t)\in Z_1 \\ \|(s,t)\|_\infty \sim R \\ \Delta(s,t)\neq 0}} \frac{B^{1+\varepsilon}}{R^{2/3}\mathrm{N}(\Delta(s,t))^{1/3}}. \tag{7.5.9}$$

We will proceed by further dividing the value of $\mathrm{N}(\Delta(s,t))$ into dyadic intervals. Say $\mathrm{N}(\Delta(s,t)) \sim S$, then we clearly have $1 \leq S \ll R^4$ and thus the number of dyadic intervals we have to consider is bounded by $O(B^\varepsilon)$. We can apply Lemma 7.2.18 to find that there are at most $O(R^{1+\varepsilon}(1+S/R^3))$ many points $(s,t) \in Z_1$ such that $\|(s,t)\|_\infty \sim R$ and $\mathrm{N}(\Delta(s,t)) \sim S$. Hence if $R \leq S$ then we may bound (7.5.9) by

$$\frac{B^{1+\varepsilon}R}{R^{2/3}S^{1/3}}\left(1+\frac{S}{R^3}\right) \ll \frac{B^{1+\varepsilon}R^{1/3}}{S^{1/3}} + \frac{B^{1+\varepsilon}S^{2/3}}{R^{8/3}} \ll B^{1+\varepsilon,}$$

since $S \ll R^4$, which is satisfactory. It therefore remains to bound

$$\sum_{\substack{(s,t)\in Z_1 \\ \|(s,t)\|_\infty \sim R \\ \mathrm{N}(\Delta(s,t))\ll R}} n_1(B,s,t).$$

We would like to apply Lemma 7.3.7 in order to bound $n_1(B,s,t)$ so we need to make sure the conditions are satisfied. Let $q(x,y,z) = Q(sx,ty,tx,sy) + az^2$. Then we have

$$q(0,y,z) = y^2 Q(0,t,0,s) + az^2,$$

and

$$q(x,0,z) = x^2 Q(s,0,t,0) + az^2.$$

Given some $(s,t) \in K^2$ both of the above binary forms are singular only if $Q(s,0,t,0) = Q(0,t,0,s) = 0$. This can happen for at most $O(1)$ primitive points $(s,t) \in Z_1$ unless $Q(s,0,t,0)$ and $Q(0,t,0,s)$ are both identically zero. If this were the case, however, then it is straightforward to check that X has a singular point. We may bound the contribution to $\sum_{(s,t)\in Z_1} n_1(B,s,t)$ such that $Q(s,0,t,0) = Q(0,t,0,s) = 0$ clearly by $O(B^{1+\varepsilon})$. Otherwise, at least one of $q(0,y,z)$ or $q(x,0,z)$ is non-singular, whence it follows from Lemma 7.3.7 that for such (s,t) we have

$$n_1(B,s,t) \ll \frac{B^{1+\varepsilon}}{R}.$$

Finally, via Lemma 7.2.18 the number of $(s,t) \in Z_1$ such that $\|(s,t)\|_\infty \sim R$ and $\mathrm{N}(\Delta(s,t)) \ll R$ holds is bounded above by $R^{1+\varepsilon}$. We conclude that

$$\sum_{\substack{(s,t)\in Z_1 \\ \|(s,t)\|_\infty \sim R \\ \mathrm{N}(\Delta(s,t))\ll R}} n_1(B,s,t) \ll R^{1+\varepsilon}\frac{B^{1+\varepsilon}}{R} \ll B^{1+\varepsilon},$$

as desired.

7.6 del Pezzo surfaces of degree 3–5

In this section we will prove Theorem 7.1.2 for the cases $d = 3,4,5$. Let X be a del Pezzo surface of degree $3 \le d \le 5$ over K and U the complement of the exceptional curves. The anti-canonical divisor $-K_X$ induces an embedding $X \subset \mathbb{P}^d$ realising X as a smooth non-degenerate surface of degree d. Let $\boldsymbol{c} \in \mathbb{P}^d(K)$ and denote by $H_{\boldsymbol{c}} \subset \mathbb{P}^d(K)$ the hyperplane defined by $\boldsymbol{x} \cdot \boldsymbol{c} = 0$. Moreover, we define $X_{\boldsymbol{c}} = H_{\boldsymbol{c}} \cap X$. By the adjunction formula, we have

$$2p_a(X_{\boldsymbol{c}}) - 2 = 0,$$

where p_a denotes the arithmetic genus. It follows that if $X_{\boldsymbol{c}}$ is smooth and geometrically irreducible, then either $X(K) = \emptyset$ or $X_{\boldsymbol{c}}$ defines an elliptic curve over K.

From Lemma 7.2.12 we know that every $\boldsymbol{x} \in \mathbb{P}^d(K)$ with $H(\boldsymbol{x}) \le B$ lies in $H_{\boldsymbol{c}}$ for some $\boldsymbol{c} \in \mathbb{P}^d(K)$ with $H(\boldsymbol{c}) \ll B^{1/d}$. We thus define

$$N_{\boldsymbol{c}}(B) := \{\boldsymbol{x} \in (X_{\boldsymbol{c}} \cap U)(K) \colon H(\boldsymbol{x}) \le B\}$$

and infer that

$$N(B) \le \sum_{\substack{\boldsymbol{c}\in\mathbb{P}^d(K) \\ H(\boldsymbol{c})\ll B^{1/d}}} N_{\boldsymbol{c}}(B). \tag{7.6.1}$$

Let $1 \le e \le d$ denote the maximum of the degrees of the irreducible components of $X_{\boldsymbol{c}}$ defined over K. It will be convenient to consider separately the contribution from each e. Note that if $e \le 2$, then $X_{\boldsymbol{c}}$ is either a union of lines, in which case we do not count its rational points, or it contains a conic defined over K. In particular, in this case X admits a conic bundle structure over K and we have the superior upper bound $N_U(B) \ll B^{1+\varepsilon}$ from Theorem 7.1.3 for $d = 4,5$, so that it suffices to consider the contribution from $3 \le e \le d$ when $d = 4$ or $d = 5$.

$\boldsymbol{e = d}$

Let us first consider the contribution from those $\boldsymbol{c}$ such that $X_{\boldsymbol{c}}$ is non-singular. In this case Proposition 7.1.4 gives $N_{\boldsymbol{c}}(B) \ll B^\varepsilon$. Moreover, the number of available $\boldsymbol{c}$ is $O(B^{(d+1)/d})$ by Lemma 7.2.6 so that we get an overall contribution of $O(B^{(d+1)/d+\varepsilon})$ to (7.6.1), which is sufficient.

Next we assume that $X_{\boldsymbol{c}}$ is irreducible but singular. In this case $N_{\boldsymbol{c}}(B) \ll B^{2/d}$ by Proposition 7.2.13. In addition, $X_{\boldsymbol{c}}$ being singular implies that $\boldsymbol{c} \in X^*(K)$. We know that $\deg(X^*) = 12$ by Proposition 7.2.3, so that by Proposition 7.2.14 the number of such $\boldsymbol{c}$ with $H(\boldsymbol{c}) \ll B^{1/d}$ is $O(B^{(d-2)/d})$. Therefore, we get a contribution of $O(B^{(d+1)/d})$, which is again satisfactory.

$e = d - 1$

This case turns out to be the most difficult one and we have to treat it individually for each value of d. Note that this case occurs precisely when $X_{\boldsymbol{c}} = L \cup D$, where $L \subset X$ is a line and $D \subset X_{\boldsymbol{c}}$ is an irreducible curve over K of degree $d-1$.

$d = 5$

We now treat the contribution from those hyperplanes for which $X_{\boldsymbol{c}} = D \cup L$, where D is an irreducible quartic curve and L a line contained in X. There are at most 10 lines in X defined over D, and so we may restrict our attention to a specific one. After a suitable change of variables we can assume that L is given by $x_2 = \cdots = x_5 = 0$. Furthermore, any hyperplane containing L takes the shape $H_{\boldsymbol{c}} = \{c_2x_2 + \cdots + c_5x_5 = 0\}$ for some $[0, 0, c_2, \ldots, c_5] \in \mathbb{P}^5(K)$. We may take a representative $\boldsymbol{c} \in Z_5$ and assume without loss of generality that $\|c_5\| = \|\boldsymbol{c}\| \asymp H(\boldsymbol{c})^{1/s_K}$.

Let $p_1 = [1, 0, \ldots, 0]$ and consider the projection $X \setminus \{p_1\} \to \mathbb{P}^4$. Since p is a non-singular point of X, the closure of the image $Y \subset \mathbb{P}^4$ is a surface of degree 4. In fact, the morphism corresponds to blowing up the point p_1 and then contracting the strict transforms of the resulting (-2)-curve and so Y is a singular del Pezzo surface of degree 4. More explicitly, since p_1 lies on a line, the point $p_0 = [1, 0, 0, 0, 0]$ will be a singularity. Any (possibly singular) del Pezzo surface of degree 4 inside $\mathbb{P}^4$ can be written as the intersection of two quadrics. Since $p_0 \in Y$, we can write $Y = V(Q_1, Q_2)$ with

$$\begin{aligned} Q_1(x) &= x_1L_1(x_2, x_3, x_4, x_5) - q_1(x_2, x_3, x_4, x_5), \\ Q_2(x) &= x_1L_2(x_2, x_3, x_4, x_5) - q_2(x_2, x_3, x_4, x_5) \end{aligned} \tag{7.6.2}$$

and where $L_i, q_i \in \mathcal{O}_K[x_2, \ldots, x_5]$ are linear and quadratic forms respectively. The Jacobian criterion applied to the singular point p_1 implies that L_1 and L_2 are proportional, so that there exist $J \in \mathcal{O}_K[x_2, \ldots, x_5]$ and $a \in \mathcal{O}_K$ such that $L_1 = J$ and $L_2 = aJ$.

Let now $Z \subset \mathbb{P}^3$ be the quadric surface defined by $aq_1(x_2, x_3, x_4, x_5) = q_2(x_2, x_3, x_4, x_5)$. The projection $[x_1, \ldots, x_5] \mapsto [x_2, \ldots, x_5]$ then defines a morphism $Y \setminus \{p_0\} \to Z$. Moreover, it is easily seen that

$$\begin{aligned} \psi[(x_2, \ldots, x_5]) &= [q_1(x_2, \ldots, x_5), x_2J(x_2, \ldots, x_5), \ldots, x_5J(x_2, \ldots, x_5)] \\ &= [aq_2(x_2, \ldots, x_5), x_2J(x_2, \ldots, x_5), \ldots, x_5J(x_2, \ldots, x_5)] \end{aligned}$$

is a well-defined inverse to the projection map from $Z \to Y \setminus \{p_0\}$ away from the locus $S \subset \mathbb{P}^3$ defined by $J = q_1 = q_2 = 0$.

As there are $O(1)$ rational points $[x_0, \ldots, x_5]$ in $D \setminus L$ above a point $[x_1, \ldots, x_5]$, it suffices to count $\boldsymbol{x} \in \mathbb{P}^4(K)$ that lie on on the image of D under the projection map with $H(\boldsymbol{x}) \leq B$. Since $D \subset H_{\boldsymbol{c}}$, any such rational point takes the shape

$$[t_1, c_5t_2, c_5t_3, c_5t_4, -(c_2t_2 + c_3t_3 + c_4t_4)]$$

with $\boldsymbol{t} = [t_1, \ldots, t_4] \in \mathbb{P}^3(K)$. For $g \in K[x_2, x_3, x_4, x_5]$, define

$$g_{\boldsymbol{c}}(t_2, t_3, t_4) = g(c_5t_2, c_5t_3, c_5t_4, -(c_2t_2 + c_3t_3 + c_4t_4)).$$

Similarly, for $\boldsymbol{t} \in \mathbb{P}^3(K)$, we define

$$H^{(\boldsymbol{c})}(\boldsymbol{t}) = H(t_1, c_5t_2, c_5t_3, c_5t_4, -(c_2t_2 + c_3t_3 + c_4t_4)).$$

Upon setting

$$N'_{\boldsymbol{c}}(B) = \#\{\boldsymbol{t} \in \mathbb{P}^3(K)\colon H^{(\boldsymbol{c})}(\boldsymbol{t}) \leq B, aq_{1,\boldsymbol{c}}(t_2,t_3,t_4) = q_{2,\boldsymbol{c}}(t_2,t_3,t_4) = t_1 J(t_2,t_3,t_4)\},$$

it transpires from (7.6.2) that $N_{\boldsymbol{c}}(B) \ll N'_{\boldsymbol{c}}(B)$. Our goal is to show the following estimate, where by abuse of notation we denote by $H_{\boldsymbol{c}}$ also the hyperplane in $\mathbb{P}^3$ defined by $c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 = 0$.

Proposition 7.6.1. *Assume that*

(i) $H_{\boldsymbol{c}} \cap S = \emptyset$, *where* $S \subset \mathbb{P}^3$ *is the variety defined by* $J = q_1 = q_2 = 0$,

(ii) $H_{\boldsymbol{c}}$ *and* Z *intersect transversally,*

(iii) $H_{\boldsymbol{c}}$ *and* $Z \cap \{J = 0\}$ *intersect transversally.*

Then

$$N'_{\boldsymbol{c}}(B) \ll B^{1/2+\varepsilon}\left(\|c_5\|_\infty^{-1/2} + \|c_5\|_\infty^{1/2}\mathrm{N}(\Delta(\boldsymbol{c}))^{-1/3}\right),$$

where $\Delta(\boldsymbol{c})$ *is the discriminant of the quadratic form* $aq_{1,\boldsymbol{c}} - q_{2,\boldsymbol{c}}$.

Suppose for a moment that Proposition 7.6.1 is proven already. After splitting $|c_5|_\nu$ for $\nu \mid \infty$ into dyadic intervals, one readily verifies that

$$\sum_{\substack{\boldsymbol{c}\in Z_3\\ \|\boldsymbol{c}\|_\infty \ll B^{1/5}}} \|c_5\|_\infty^{-1/2} \ll B^{7/10}.$$

Moreover, we can also split $\mathrm{N}(\Delta(\boldsymbol{c}))$ into dyadic intervals, say $\mathrm{N}(\Delta(\boldsymbol{c})) \sim \beta$ and $\|\boldsymbol{c}\|_\infty \sim \alpha$ with $\alpha \ll B^{1/5}$ and $\beta \ll \alpha^6$. As $\Delta(\boldsymbol{c})$ is homogeneous of degree 6, we can use Lemma 7.2.17 to deduce that

$$\alpha^{1/2}\beta^{-1/3} \sum_{\substack{\boldsymbol{c}\in Z_3\\ \|\boldsymbol{c}\|_\infty \sim \alpha\\ \mathrm{N}(\Delta(\boldsymbol{c}))\sim\beta}} 1 \ll \alpha^{7/2}\beta^{-1/6} \ll \alpha^{7/2}.$$

Summing over α gives $\sum_{\alpha \ll B^{1/5}} \alpha^{7/2} \ll B^{7/10}$. As there are $O(\alpha^{6\varepsilon}) = O(B^\varepsilon)$ possibilities for β, in total we obtain a contribution of $B^{6/5+\varepsilon}$ under the assumption that the conditions of Proposition 7.6.1 are satisfied. To deal with the remaining cases, we need the following lemma.

Lemma 7.6.2. *There is a Zariski closed subset* $W \subset \mathbb{P}^3$ *whose irreducible components have dimension at most 2 such that if* $\boldsymbol{c} \in \mathbb{P}^3$ *fails (i), (ii) or (iii) in Proposition 7.6.1, then* $\boldsymbol{c}$ *lies in* W.

Proof. We begin with a preliminary observation. We claim that the variety $S \subset \mathbb{P}^3$ is 0-dimensional. Suppose for a contradiction that there exists a positive dimensional irreducible component $E \subset S$. Then by looking at the defining equations of Y, it becomes clear that the closure E' of $\mathbb{A}^1 \times E$ is contained in Y. However, E' has dimension 2 and Y is an irreducible surface, so they must coincide. It is clear that E' is contained in the hyperplane defined by L inside $\mathbb{P}^4$, but del Pezzo surfaces are not contained in any hyperplane under the anti-canonical embedding.

If $\boldsymbol{c} \in \mathbb{P}^3$ fails (i), then $H_{\boldsymbol{c}} \cap S \neq \emptyset$. From what we have just shown, it follows that S is a union of $O(1)$ points in $\mathbb{P}^3$. In particular, $H_{\boldsymbol{c}} \cap S \neq \emptyset$ implies that $H_{\boldsymbol{c}}$ contains one of these points, which forces $\boldsymbol{c}$ to lie in one of $O(1)$ irreducible subvarieties of codimension at least 1 in $\mathbb{P}^3$.

Next, assume that $\boldsymbol{c}$ does not satisfy (ii). This implies that $\boldsymbol{c}$ lies on the dual variety of Z or intersects the singular locus of Z. As Z is is birational to X, it is again a (possibly singular) del Pezzo surface and hence has at most isolated singularities. This again implies that $\boldsymbol{c}$ lies on $O(1)$ irreducible subvarieties of codimension at least 1 in $\mathbb{P}^2$.

Finally, suppose that $\boldsymbol{c}$ fails (iii). Note that $Z \cap \{J = 0\}$ cannot be a double line, as it corresponds birationally to a hyperplane section containing L of X. As X is smooth, this implies that hyperplane sections are reduced. It follows that $Z \cap \{J = 0\}$ is isomorphic to a plane conic. If it is smooth, then the failure of (iii) implies that $\boldsymbol{c}$ lies on the dual variety of $Z \cap \{J = 0\}$ inside $\mathbb{P}^3$, which has codimension at least 1. When it is not smooth, then it is a union of two lines that intersect in a unique point $P \in \mathbb{P}^3$ and $H_{\boldsymbol{c}}$ has tangential intersection if and only if $P \in H_{\boldsymbol{c}}$, which again forces $\boldsymbol{c}$ to lie on a linear subspace of $\mathbb{P}^3$ of dimension 2. □

Suppose now that $\boldsymbol{c}$ does not meet the requirements of Proposition 7.6.1. By Proposition 7.2.13 we have $N_{\boldsymbol{c}}(B) \ll B^{1/2}$. Moreover, it follows from Lemma 7.6.2 that $\boldsymbol{c}$ lies on $O(1)$ irreducible subvarieties of dimension at most 2 inside $\mathbb{P}^3$. By (7.2.3) the number of such $\boldsymbol{c} \in \mathbb{P}^3(K)$ with $H(\boldsymbol{c}) \ll B^{1/5}$ is $O(B^{3/5})$. Hence we get an overall contribution of $O(B^{11/10})$ to (7.6.1) from this case, which is sufficient. To complete the case $e = 4$ and $d = 5$, we are therefore left with proving Proposition 7.6.1.

Proof of Proposition 7.6.1. Let $\boldsymbol{t} = (t_2, t_3, t_4) \in Z_2$ be a representative for $\boldsymbol{t} \in \mathbb{P}^2(K)$ such $aq_{1,\boldsymbol{c}}(\boldsymbol{t}) = q_{2,\boldsymbol{c}}(\boldsymbol{t})$ and write

$$\mathfrak{a} = \langle q_{1,\boldsymbol{c}}(\boldsymbol{t}), c_5 t_2 J_{\boldsymbol{c}}(\boldsymbol{t}), c_5 t_3 J_{\boldsymbol{c}}(\boldsymbol{t}), c_5 t_4 J_{\boldsymbol{c}}(\boldsymbol{t}), -J_{\boldsymbol{c}}(\boldsymbol{t})(c_2 t_2 + c_3 t_3 + c_4 t_4)\rangle.$$

Let $\mathfrak{C}$ be the ideal generated by the resultant of the homogeneous forms $q_{1,\boldsymbol{c}}, q_{2,\boldsymbol{c}}, J_{\boldsymbol{c}}$. Note that if $\mathfrak{C} = 0$, then $J_{\boldsymbol{c}}, q_{1,\boldsymbol{c}}$ and $q_{2,\boldsymbol{c}}$ share a common root in $\mathbb{P}^1$, which implies that $H_{\boldsymbol{c}}$ has non-empty intersection with S. As we assume that (i) holds, this is impossible. Since $\|\boldsymbol{c}\|_\infty \ll B^{1/5}$, we immediately get $\mathrm{N}(\mathfrak{C}) \ll B^A$. In addition, because $\boldsymbol{t} \in Z_2$, we clearly have $\mathfrak{a} \mid \mathfrak{A}\mathfrak{C}\langle c_5\rangle$, where $\mathfrak{A} = \mathfrak{a}_1 \cdots \mathfrak{a}_h$ and $\mathfrak{a}_1, \dots, \mathfrak{a}_h$ are the fixed representatives of the class group of K.

We then have $H^{(\boldsymbol{c})}(\psi(\boldsymbol{t})) \leq B$ if and only if

$$\|q_{1,\boldsymbol{c}}(\boldsymbol{t}), c_5 t_2 J_{\boldsymbol{c}}(\boldsymbol{t}), c_5 t_3 J_{\boldsymbol{c}}(\boldsymbol{t}), c_5 t_4 J_{\boldsymbol{c}}(\boldsymbol{t}), -J_{\boldsymbol{c}}(\boldsymbol{t})(c_2 t_2 + c_3 t_3 + c_4 t_4)\|_\infty \leq B\mathrm{N}(\mathfrak{a}). \quad (7.6.3)$$

Note that in the above display we may also replace $q_{1,\boldsymbol{c}}$ by $aq_{2,\boldsymbol{c}}$. As we assume that $q_{1,\boldsymbol{c}}, q_{2,\boldsymbol{c}}$ and $J_{\boldsymbol{c}}$ do not have a common root in $\overline{K}$, the inequality (7.6.3) together with Lemma 7.2.11 implies that

$$\|\boldsymbol{t}\|_\infty \ll (B\mathrm{N}(\mathfrak{a}))^A.$$

Let us now write $\mathfrak{a} = \mathfrak{b}_1\mathfrak{b}_2$, where $\mathfrak{b}_1 = \langle \mathfrak{a}, J_{\boldsymbol{c}}(\boldsymbol{t})\rangle$, so that in particular $\mathfrak{b}_1 \mid \mathfrak{C}$ and $\mathfrak{b}_2 \mid \mathfrak{A}\langle c_5\rangle$. From (7.6.3) we get that

$$\|J_{\boldsymbol{c}}(\boldsymbol{t})\|_\infty \|c_5\|_\infty \|(t_2, t_3, t_4)\|_\infty \leq H^{(\boldsymbol{c})}(\psi(\boldsymbol{t}))\mathrm{N}(\mathfrak{a}) \leq B\mathrm{N}(\mathfrak{a}),$$

and so we must have that $J_{\boldsymbol{c}}(s,t) \leq B^{1/2}\mathrm{N}(\mathfrak{b}_1)\|c_5\|_\infty^{-1/2}$ or $\|\boldsymbol{t}\|_\infty \leq B^{1/2}\mathrm{N}(\mathfrak{b}_2)\|c_5\|_\infty^{-1/2} \ll B^{1/2}\|c_5\|_\infty^{1/2}$, where we used that $\mathfrak{b}_2 \mid \mathfrak{A}\langle c_5\rangle$. In particular, it follows from our discussion so far that

$$N'_{\boldsymbol{c}}(B) \leq \sum_{\mathfrak{b}_1|\mathfrak{C}}\sum_{\mathfrak{b}_2|\mathfrak{A}\langle c_5\rangle}\left(n_1((B\mathrm{N}(\mathfrak{a}))^A, B^{1/2}\mathrm{N}(\mathfrak{b}_1)\|c_5\|_\infty^{-1/2}) + n_2(B^{1/2}\|c_5\|_\infty^{1/2})\right),$$

where for positive reals R_1, R_2 we have set

$$n_1(R_1,R_2) = \#\{\boldsymbol{t}\in Z_2\colon \|\boldsymbol{t}\|_\infty \ll R_1, \|J_{\boldsymbol{c}}(\boldsymbol{t})\|_\infty \ll R_2, \mathfrak{b}_1 \mid J_{\boldsymbol{c}}(\boldsymbol{t}), aq_{1,\boldsymbol{c}}(\boldsymbol{t}) = q_{2,\boldsymbol{c}}(\boldsymbol{t})\}$$

and

$$n_2(R_1) = \#\{\boldsymbol{t}\in Z_2\colon \|\boldsymbol{t}\|_\infty \ll R_1, aq_{1,\boldsymbol{c}}(\boldsymbol{t}) = q_{2,\boldsymbol{c}}(\boldsymbol{t})\}.$$

Let us first focus on estimating $n_1(R_1,R_2)$. Let $C_{\boldsymbol{c}}$ be the conic defined by $aq_{1,\boldsymbol{c}}(\boldsymbol{t}) = q_{2,\boldsymbol{c}}(\boldsymbol{t})$ inside $\mathbb{P}^2(K)$ and note that since we assume that (ii) holds, $C_{\boldsymbol{c}}$ is geometrically irreducible. There are two possibilities: Either there is no point $\boldsymbol{t}\in C_{\boldsymbol{c}}(K)$ with $H(\boldsymbol{t}) \ll R_1$, in which case $n_1(R_1,R_2) = 0$, or such a point exists and we can use it to obtain a parameterisation $\mathbb{P}^1(K) \to C_{\boldsymbol{c}}(K)$. Explicitly, this is done by sending a line through the point to its unique residual intersection point with the conic. In this way we obtain quadratic forms $g_1, g_2, g_2 \in \mathcal{O}_K[u,v]$ without a common factor such that the map $\psi\colon \mathbb{P}^1 \to C_{\boldsymbol{c}}$ given by $[s,t] \mapsto [g_1(s,t), g_2(s,t), g_3(s,t)]$ gives a bijection of $\mathbb{P}^1(K)$ with $C_{\boldsymbol{c}}(K)$. As we assume that the height of the initial point is bounded by R_1, it is clear that we can take the quadratic forms in such a way that $\|g_i\| \ll R_1^A$.

Let us define $Q(s,t) = J_{\boldsymbol{c}}(g_1(s,t), g_2(s,t), g_3(s,t))$. We claim that $Q(s,t)$ is square-free. To see this, note that if it has a double root $[s_0,t_0]$, then $P_0 = [g_1(s_0,t_0), g_2(s_0,t_0), g_3(s_0,t_0)]$ will satisfy $J_{\boldsymbol{c}}(P_0) = 0$ and $P_0 \in C_{\boldsymbol{c}}$. In particular, it will be a double point of $C_{\boldsymbol{c}} \cap \{J_{\boldsymbol{c}} = 0\} = Z \cap \{J = 0\} \cap H_{\boldsymbol{c}}$. As we assume that (iii) holds, this is impossible, which verifies the claim.

It is clear that $\|Q\| \ll \|J_{\boldsymbol{c}}\| \max_{i=1,2,3}\|g_i\| \ll B^A R_1^A$, so that if $\|Q(s,t)\|_\infty \ll R_2$ holds, then we must also have $\|(s,t)\|_\infty \ll (BR_1R_2)^A$ for any $(s,t)\in Z_1$ by Lemma 7.2.11. In particular, we have

$$n_1(R_1,R_2) \leq \#\{(s,t)\in Z_1\colon \|(s,t)\|_\infty \ll (BR_1R_2)^A, \mathfrak{b}_1 \mid Q(s,t), \|Q(s,t)\|_\infty \ll R_2\}.$$

As Q is square-free, we can invoke Corollary 7.2.16 to bound this last quantity and deduce that

$$n_1((B\mathrm{N}(\mathfrak{a}))^A, B^{1/2}\mathrm{N}(\mathfrak{b}_1)\|c_5\|_\infty^{-1/2}) \ll (B^A\mathrm{N}(\mathfrak{a})^A\|Q\|)^\varepsilon \frac{B^{1/2}}{\|c_5\|_\infty^{1/2}} \ll B^{1/2+\varepsilon}\|c_5\|_\infty^{-1/2},$$

where we used that $\mathrm{N}(\mathfrak{a}), \mathrm{N}(\mathfrak{b}_1), \|Q\| \ll B^A$.

Next we turn to $n_2(R_1)$. Let $\Delta(\boldsymbol{c})$ be the discriminant of the quadratic form $aq_{1,\boldsymbol{c}} - q_{2,\boldsymbol{c}}$ and let $\Delta_0(\boldsymbol{c})$ be the ideal generated by the 2×2 minors of the matrix underlying $aq_{1,\boldsymbol{c}} - q_{2,\boldsymbol{c}}$. Note that as we assume that (ii) holds, we must have $\Delta(\boldsymbol{c}) \neq 0$. Applying Corollary 7.3.6 directly to $n_2(R_1)$ gives

$$n_2(B^{1/2}\|c_5\|_\infty^{1/2}) \ll \left(1 + \frac{B^{1/2}\|c_5\|_\infty^{1/2}\mathrm{N}(\Delta_0(\boldsymbol{c}))^{1/2}}{\mathrm{N}(\Delta(\boldsymbol{c}))^{1/3}}\right)\mathrm{N}(\Delta(\boldsymbol{c}))^\varepsilon.$$

We clearly have $\mathrm{N}(\Delta(\boldsymbol{c}))^\varepsilon \ll B^\varepsilon$. Moreover, we claim that $\mathrm{N}(\Delta_0(\boldsymbol{c}))$ is in fact bounded. To see this, first assume that the 2×2 minors $M_{ij}(\boldsymbol{c})$ have a common zero in $\overline{K}$. This would imply that a hyperplane section of the quadric surface Z is a double line. However, hyperplane sections of Z correspond birationally to hyperplane sections containing L of our original quintic del Pezzo surface X and hence are reduced. By Hilbert's Nullstellensatz we can find forms $f_{ijk} \in \mathcal{O}_K[x_1,\dots,x_4]$ and $a_k \in \mathcal{O}_K \setminus \{0\}$ such that

$$a_k x_k^d = \sum f_{ijk} M_{ij}(x_1,\dots,x_4)$$

as an identity in $\mathcal{O}_K[x_1,\dots,x_4]$ for $k=1,\dots,4$. It follows that

$$\mathrm{N}(\Delta_0(\boldsymbol{c})) \ll \mathrm{N}(a_1c_1^d,\dots,a_4c_4^d) \ll 1$$

for $\boldsymbol{c}\in Z_3$, as claimed. Using the usual divisor bound for ideals, it follows from our discussion so far that

$$\begin{aligned} N_{\boldsymbol{c}}(B) &\ll \sum_{\mathfrak{b}_1|\mathfrak{C}}\sum_{\mathfrak{b}_2|\mathfrak{c}_5}\left(B^{1/2+\varepsilon}\|c_5\|_\infty^{-1/2} + \frac{B^{1/2+\varepsilon}\|c_5\|_\infty^{1/2}}{\mathrm{N}(\Delta(\boldsymbol{c}))^{1/3}}\right)\\ &\ll B^{1/2+\varepsilon}\left(\|c_5\|_\infty^{-1/2} + \|c_5\|_\infty^{1/2}\mathrm{N}(\Delta(\boldsymbol{c}))^{-1/3}\right), \end{aligned}$$

which is what we wanted to show. □

$d=4$

Next we assume $X_{\boldsymbol{c}} = D\cup L$, where $L\subset X$ is a line and $D\subset X$ is an irreducible cubic curve, both defined over K. There are at most 16 lines in X, and so we may restrict our attention to a specific one. After a suitable change of variables L is given by $x_2=x_3=x_4=0$ if we work with coordinates $x_0,\dots,x_4$ on $\mathbb{P}^4$. It follows that X is defined by

$$x_0L_1 + x_1K_1 + Q_1 = x_0L_2 + x_1K_2 + Q_2 = 0, \tag{7.6.4}$$

where $L_1, L_2, K_1, K_2 \in \mathcal{O}_K[x_2,x_3,x_4]$ are linear and $Q_1,Q_2\in\mathcal{O}_K[x_2,x_3,x_3]$ are quadratic forms respectively. If L_1 and L_2 are proportional, then $[1,0,\dots,0]$ is a singular point of X, which is impossible. We can therefore eliminate x_0 from (7.6.4) to obtain the equation

$$C(x_2,x_3,x_4) + x_1Q(x_2,x_3,x_4) = 0 \tag{7.6.5}$$

for some cubic form $C\in\mathcal{O}_K[x_2,x_3,x_4]$. Any hyperplane containing L is defined by $\boldsymbol{c}\in\mathbb{P}^4(K)$ of the shape $\boldsymbol{c}=[0,0,c_2,c_3,c_4]$. Without loss of generality, we may assume that $(0,0,c_2,c_3,c_4)\in Z_4$ is a representative of $\boldsymbol{c}$ such that $\|c_4\|=\|\boldsymbol{c}\|$, so that $\|c_4\|\asymp H(\boldsymbol{c})^{1/s_K}$. Any rational point $\boldsymbol{t}\in H_{\boldsymbol{c}}$ takes the shape $\boldsymbol{t}=[t_0,t_1,c_4t_2,c_4t_3,-(c_2t_2+c_3t_3)]$ with $[t_0,\dots,t_3]\in\mathbb{P}^3(K)$. In particular, the equation (7.6.5) transforms into

$$C_{\boldsymbol{c}}(t_2,t_3) = t_1Q_{\boldsymbol{c}}(t_2,t_3), \tag{7.6.6}$$

where we write $G_{\boldsymbol{c}}(t_2,t_3) = G(c_4t_2,c_4t_3,-c_2t_2-c_3t_3)$ for any $G\in\mathcal{O}_K[x,y,z]$. Moreover, it suffices to count $\boldsymbol{t}'=[t_1,t_2,t_3]\in\mathbb{P}^2(K)$ such that (7.6.6) holds, since there are $O(1)$ available $\boldsymbol{t}$ lying above a given $\boldsymbol{t}'$, and we redefine $N_{\boldsymbol{c}}(B)$ to be this quantity. Our goal is to establish the following estimate.

Proposition 7.6.3. *If $Q_{\boldsymbol{c}}$ is square-free, then*

$$N_{\boldsymbol{c}}(B) \ll \frac{B^{2/3+\varepsilon}}{\|c_4\|_\infty^{2/3}}.$$

Suppose for a moment that Proposition 7.6.3 holds. The contribution from such $\boldsymbol{c}$ to (7.6.1) is then

$$B^{2/3+\varepsilon} \sum_{\substack{\boldsymbol{c}\in Z^2 \\ \|\boldsymbol{c}\|_\infty \ll B^{1/4}}} \|c_4\|_\infty^{-2/3} \ll B^{5/4+\varepsilon},$$

which is satisfactory. It remains to deal with the case when $Q_{\boldsymbol{c}}$ has a repeated root. Note that the cubic surface Y defined by (7.6.5) is birational to X and in particular again a (possibly singular) del Pezzo surface. If the curve Z defined $Q = 0$ defines a double line in $\mathbb{P}^2$, then one can check via the Jacobian criterion that the singular locus of Y has dimension at least 1, which is impossible. Thus Z has only isolated singularities. Now if $Q_{\boldsymbol{c}}$ has a double root, then $H_{\boldsymbol{c}}$ intersects the singular locus of Z, or $\boldsymbol{c}$ lies on the dual variety of Z inside $\mathbb{P}^2$. Either case defines a variety inside $\mathbb{P}^2$ of dimension at most 1. In particular, there are most $O(B^{1/2})$ available $\boldsymbol{c}$ with $H(\boldsymbol{c}) \ll B^{1/4}$ by (7.2.3). Moreover, Proposition 7.2.13 implies that $N_{\boldsymbol{c}}(B) \ll B^{2/3}$, so that we have a contribution of $O(B^{7/6})$, which is sufficient.

Proof of Proposition 7.6.3. If D is irreducible, but not geometrically irreducible, then it contains $O(1)$ points by Bézout's theorem, which is sufficient. So we assume from now on that D is geometrically irreducible. In particular, $Q_{\boldsymbol{c}}$ and $C_{\boldsymbol{c}}$ do not share a common factor. The equation (7.6.6) therefore defines a singular cubic plane curve $C' \subset \mathbb{P}^2$ with a singularity at $t_0 = [1,0,0]$ and hence is rational over K. We obtain an explicit bijective morphism $\mathbb{P}^1 \to C'$ by sending a line to the unique residual intersection point with C' and t_0. Explicitly, this is given by $[s,t] \mapsto [C_{\boldsymbol{c}}(s,t), sQ_{\boldsymbol{c}}(s,t), tQ_{\boldsymbol{c}}(s,t)]$. It follows that $N_{\boldsymbol{c}}(B)$ is bounded from above by

$$\#\{[s,t] \in \mathbb{P}^1(K) \colon H([C_{\boldsymbol{c}}(s,t), c_4sQ_{\boldsymbol{c}}(s,t), c_4tQ_{\boldsymbol{c}}(s,t), -Q_{\boldsymbol{c}}(s,t)(c_2s + c_3t)]) \ll B\}$$

and we define the last quantity to be $N'_{\boldsymbol{c}}(B)$. As $\|\boldsymbol{c}\| \ll B^{1/4s_K}$, it is clear that $\|Q_{\boldsymbol{c}}\|, \|C_{\boldsymbol{c}}\| \ll B^A$. Thus if $(s,t) \in Z_1$ is counted by $N'_{\boldsymbol{c}}(B)$, then Lemma 7.2.11 implies that $\|(s,t)\|_\infty \ll B^A$. Let $\mathfrak{C}'$ be the ideal generated by the resultant of $Q_{\boldsymbol{c}}(s,t)$ and $C_{\boldsymbol{c}}(s,t)$ and define $\mathfrak{C} = \mathfrak{C}'\mathfrak{a}_1 \cdots \mathfrak{a}_h$. By construction, as $(s,t) \in Z_1$, we must then have $\langle C_{\boldsymbol{c}}(s,t), sQ_{\boldsymbol{c}}(s,t), tQ_{\boldsymbol{c}}(s,t)\rangle \mid \mathfrak{C}$ for any $(s,t) \in Z_1$ and $\mathrm{N}(\mathfrak{C}) \ll \max\{\|Q_{\boldsymbol{c}}\|, \|C_{\boldsymbol{c}}\|\}^A \ll B^A$. Suppose now that

$$\mathfrak{a} = \langle C_{\boldsymbol{c}}(s,t), c_4sQ_{\boldsymbol{c}}(s,t), c_4tQ_{\boldsymbol{c}}(s,t), Q_{\boldsymbol{c}}(s,t)(c_2s + c_3t)\rangle.$$

If we write $\mathfrak{b}_1 = \langle \mathfrak{a}, Q_{\boldsymbol{c}}(s,t)\rangle$ and $\mathfrak{a} = \mathfrak{b}_1\mathfrak{b}_2$, then we must have $\mathfrak{b}_1 \mid \mathfrak{C}$ and $\mathfrak{b}_2 \mid \langle c_4, c_2s + c_3t\rangle$. Moreover, for any $(s,t) \in Z_1$ we have

$$\begin{aligned} H([C_{\boldsymbol{c}}(s,t),&c_4sQ_{\boldsymbol{c}}(s,t), c_4tQ_{\boldsymbol{c}}(s,t), -Q_{\boldsymbol{c}}(s,t)(c_2s + c_3t)]) \\ &= \mathrm{N}(\mathfrak{a})^{-1}\|C_{\boldsymbol{c}}(s,t), c_4sQ_{\boldsymbol{c}}(s,t), c_4tQ_{\boldsymbol{c}}(s,t), -Q_{\boldsymbol{c}}(s,t)(c_2s + c_3t)\|_\infty \\ &\geq \mathrm{N}(\mathfrak{a})^{-1}\|c_4\|_\infty\|Q_{\boldsymbol{c}}(s,t)\|_\infty\|(s,t)\|_\infty \end{aligned}$$

In particular, if (s,t) is counted by $N'_{\boldsymbol{c}}(B)$, then we must have

$$\|Q_{\boldsymbol{c}}(s,t)\|_\infty \leq \mathrm{N}(\mathfrak{b}_1)B^{2/3}\|c_4\|_\infty^{-2/3} \quad \text{or} \quad \|(s,t)\|_\infty \leq \mathrm{N}(\mathfrak{b}_2)B^{1/3}\|c_4\|_\infty^{-1/3}. \tag{7.6.7}$$

Let us first bound the contribution from those (s,t) for which the first alternative holds. As we have $\mathfrak{b}_1 \mid Q_{\boldsymbol{c}}(s,t)$, by Corollary 7.2.16 we have that the contribution is up to a constant at most

$$B^\varepsilon\|Q_{\boldsymbol{c}}\|^\varepsilon \frac{B^{2/3}}{\|c_4\|_\infty^{2/3}} \ll \frac{B^{2/3+\varepsilon}}{\|c_4\|_\infty^{2/3}}.$$

Let us now deal with the contribution form those (s,t) for which the second alternative in (7.6.7) holds. For any such (s,t), we must have $\mathfrak{b}_2 \mid c_2 s + c_3 t$. As $\mathfrak{b}_2 \mid c_4$ and $\mathrm{N}(\langle c_2, c_3, c_4\rangle) \ll 1$, it follows that the cardinality of the image of Z_1 in

$$\#\{(s,t) \in (\mathcal{O}_K/\mathfrak{b}_2)^2 \colon c_2 s + c_3 t \equiv 0 \mod \mathfrak{b}_2\}$$

is $O(1)$. Denoting the the image of Z_1 in the set above by $\mathcal{T}$, any (s,t) must be congruent to some element in $\mathcal{T}$ modulo $\mathfrak{b}_2$, and hence the number of such (s,t) with $\|(s,t)\|_\infty \ll B^{1/3}\mathrm{N}(\mathfrak{b}_2)\|c_4\|_\infty^{-1/3}$ is $O(B^{2/3}\|c_4\|_\infty^{-2/3})$. As $\mathfrak{b}_1 \mid \mathfrak{b}$, $\mathfrak{b}_2 \mid c_4$ and $\mathrm{N}(\mathfrak{C}) \ll B^A$, $\mathrm{N}(c_4) \ll B^{1/4}$, the familiar divisor bound for ideals shows that number of available $\mathfrak{b}_1$ and $\mathfrak{b}_2$ is $O(B^\varepsilon)$. In summary, we have

$$N'_{\boldsymbol{c}}(B) \ll \frac{B^{2/3+\varepsilon}}{\|c_4\|_\infty^{2/3}},$$

which is what we wanted to show. □

$d = 3$

Suppose that $X_{\boldsymbol{c}}$ contains a line $L \subset X$. Since there are at most 27 lines in X defined over K, it suffices to bound $N_{\boldsymbol{c}}(B)$ when $X_{\boldsymbol{c}}$ contains a fixed line $L \subset X$. After a suitable change of variables L is given by $x_1 = x_2 = 0$, in which case we may assume that F is of the shape

$$F(\boldsymbol{x}) = x_1 Q_1(\boldsymbol{x}) + x_2 Q_2(\boldsymbol{x})$$

for some quadratic forms $Q_1, Q_2 \in \mathcal{O}_K[x_1, \dots, x_4]$. Moreover, the fact that $L \subset V_{\boldsymbol{c}}$ also implies that $\boldsymbol{c} = [s,t,0,0]$. Any $\boldsymbol{x} \in \mathbb{P}^3(K)$ that lies on L is necessarily of the form $\boldsymbol{x} = [tx_1, -sx_1, x_2, x_3]$. In particular, if $\boldsymbol{x}$ satisfies $F(\boldsymbol{x}) = 0$, then either $x_1 = 0$, in which case $\boldsymbol{x} \in L$, or $\boldsymbol{x}$ is a root of

$$Q_{s,t}(\boldsymbol{x}) \coloneqq tQ_1(tx_1, -sx_1, x_2, x_3) - sQ_2(tx_1, -sx_1, x_2, x_3).$$

It is clear that if $(x_1, x_2, x_3) \in Z_2$ satisfies $H(tx_1, -sx_1, x_2, x_3) \le B$, then we must have $\|x_1\|_\infty \ll B\|(s,t)\|_\infty^{-1}$. Let $\Delta(s,t)$ be the discriminant of $Q_{s,t}$ considered as a ternary quadratic form in (x_1, x_2, x_3), so that $\Delta(s,t)$ is homogeneous of degree 5. In addition, define $\Delta_0(s,t)$ to be the ideal generated by the 2×2 minors of the matrix underlying $Q_{s,t}$. Then Lemma 7 of [27] shows that $\mathrm{N}(\Delta_0(s,t)) \ll 1$, and hence Proposition 7.3.4 implies

$$N_{s,t}(B) \ll B^\varepsilon + \frac{B^{1+\varepsilon}}{\|(s,t)\|_\infty^{1/3}\mathrm{N}(\Delta(s,t))^{1/3}},$$

where we used the divisor estimate $\tau(\mathrm{N}(\Delta(s,t))) \ll B^\varepsilon$, since $\mathrm{N}(\Delta(s,t)) \ll B^{5/3}$. There are $O(B^{2/3})$ available $(s,t) \in Z_1$ with $H(s,t) \ll B^{1/3}$ and hence if B^ε dominates in the estimate above, we get a contribution of $O(B^{2/3+\varepsilon})$, which is sufficient. Let us now put $\|(s,t)\|_\infty$ and $\mathrm{N}(\Delta(s,t))$ into dyadic intervals, say $\|(s,t)\|_\infty \asymp \alpha$ and $\mathrm{N}(\Delta(s,t)) \asymp \beta$, where $\alpha \ll B^{1/3}$ and $\beta \ll B^{5/3}$. By Lemma 7.2.17 the number of available (s,t) is $O(\alpha\beta^{1/5})$, and hence if the second term dominates we get a contribution of

$$B^{1+\varepsilon}\alpha^{2/3}\beta^{-2/15}.$$

Summing over dyadic intervals shows that the overall contribution is $O(B^{11/9+\varepsilon})$, which is sufficient and therefore completes our treatment of cubic surfaces.

$e = d - 2$

Note that this case is only relevant when $e = 5$ and hence $d = 3$. Suppose now that $X_{\boldsymbol{c}} = B \cup C$, where C is irreducible of degree 3. If B is a conic, we are in the context of Theorem 7.1.3, so that we may assume that B is the union of two skew lines L_1, L_2. There are $O(1)$ pairs of skew lines in X, and so we may restrict our attention to a fixed pair. Since L_1 and L_2 are skew, any hyperplane containing them must contain the three-dimensional linear space they span. This forces $\boldsymbol{c}$ to lie on a line in $\mathbb{P}^5$, and hence the number of such hyperplanes of height $O(B^{1/5})$ is $O(B^{2/5})$. Moreover, by Proposition 7.2.13 the curve C contains $O(B^{2/3})$ rational points of height B, so that the contribution is $O(B^{2/5+2/3}) = O(B^{16/15})$, which is again satisfactory and thus completes our proof of Theorem 7.1.2 for $d = 3, 4, 5$.

7.7 del Pezzo surfaces of degree 2

The anti-canonical model of a smooth del Pezzo surface of degree 2 is given by a hypersurface in weighted projective space $\mathbb{P}(2,1,1,1)$ in the variables (y, u, v, w) of the form

$$X : y^2 = g(u, v, w),$$

where $g \in \mathcal{O}_K[u, v, w]$ is a non-singular quartic form. For $[y, u, v, w] \in \mathbb{P}(2,1,1,1)(K)$, define

$$H(y, u, v, w) = \prod_{\nu \in \Omega_K} \max\{|y|_\nu^{1/2}, |u|_\nu, |v|_\nu, |w|_\nu\}.$$

Let $U \subset X$ be the complement of the exceptional curves of X. The counting function with respect to the anti-canonical height function of X is then given by

$$N(B) = \#\left\{[y, u, v, w] \in U(K) \colon H([y, u, v, w]) \leq B\right\}.$$

Let $\pi \colon X \to \mathbb{P}^2$ be the map induced by the anti-canonical divisor, that is $\pi([y, u, v, w]) = [u, v, w]$. Given $\boldsymbol{c} \in \mathbb{P}^2(K)$ we let $H_{\boldsymbol{c}} \subset \mathbb{P}^2$ be the line defined by $\boldsymbol{x} \cdot \boldsymbol{c} = 0$. We may consider the pullback of $H_{\boldsymbol{c}}$ under π. Writing $X_{\boldsymbol{c}} = \pi^{-1}(H_{\boldsymbol{c}})$, the defining curve for $X_{\boldsymbol{c}}$ can be written as

$$y^2 = g_{\boldsymbol{c}}(x, t)$$

for a suitable binary quartic form $g_{\boldsymbol{c}}$. If we write

$$N_{\boldsymbol{c}}(B) = \#\left\{\boldsymbol{t} \in (X_{\boldsymbol{c}} \cap U)(K) \colon H(\boldsymbol{t}) \leq B\right\},$$

then using Lemma 7.2.12 we find

$$N(B) \leq \sum_{\substack{\boldsymbol{c} \in \mathbb{P}^2(K) \\ H(\boldsymbol{c}) \ll B^{1/2}}} N_{\boldsymbol{c}}(B).$$

We will split the study of $N_{\boldsymbol{c}}(B)$ into three cases.

First consider the case where $g_{\boldsymbol{c}}(x, t)$ has no multiple factors. If $X_{\boldsymbol{c}}$ has no rational points of height at most B then trivially $N_{\boldsymbol{c}}(B) = 0$ for such $\boldsymbol{c}$. Otherwise, let $\boldsymbol{x}_0$ be a rational point of $X_{\boldsymbol{c}}$ with $H(\boldsymbol{x}_0) \leq B$. Then $X_{\boldsymbol{c}}$ is non-singular and defines an elliptic curve, as is shown in [53, Chapter 8]. In particular, it is shown there that there exists an isomorphism (which is independent of $\boldsymbol{c}$) that birationally maps $X_{\boldsymbol{c}}$ to an elliptic curve in Weierstrass normal form

$$E_{\boldsymbol{c}} \colon ZY^2 = X^3 + A_{\boldsymbol{c}} XZ^2 + B_{\boldsymbol{c}} Z^3.$$

Note that any rational point $\boldsymbol{x}$ of $X_{\boldsymbol{c}}$ is mapped to a rational solution $P(\boldsymbol{x})$, say. Since the birational transformation only depends polynomially on the coefficients of $g_{\boldsymbol{c}}$ there exist absolute constants $A, C > 0$ such that

$$\|P(\boldsymbol{x})\| \leq C\|g_{\boldsymbol{c}}\|^A \|\boldsymbol{x}_0\|^A \|\boldsymbol{x}\|^A.$$

Finally, since different rational points are mapped to different rational points under this birational transformation we find

$$N_{\boldsymbol{c}}(B) \ll N(E_{\boldsymbol{c}}, CB^A),$$

where $N(E, R)$ denotes the number of rational points of an elliptic curve $E \subset \mathbb{P}^2$ up to height R with respect to the usual height on $\mathbb{P}^2$. Thus, by Proposition 7.1.4 we find that in this case

$$N_{\boldsymbol{c}}(B) \ll B^{\varepsilon}.$$

There are $O(B^{3/2})$ rational points $\boldsymbol{c} \in \mathbb{P}^2(K)$ with $H(\boldsymbol{c}) \ll B^{1/2}$, so that the total contribution in this case is bounded by $O_{\varepsilon}(B^{3/2+\varepsilon})$.

Next, we consider the case when $g_{\boldsymbol{c}}(x, t)$ has a multiple factor but $X_{\boldsymbol{c}}$ is geometrically irreducible. This implies that $g_{\boldsymbol{c}} = L^2 Q$, where $L, Q \in \overline{K}[x, t]$. Note that in fact L and hence also Q must have coefficients in $\mathcal{O}_K$, because we assume $\mathrm{char}(K) \neq 2$. Indeed, if not then Q must also be square, which is impossible as then $y^2 - g_{\boldsymbol{c}}$ is reducible. Hence we may assume that L and Q are both defined over K. There are most $O(1)$ possible choices $(x, t) \in \mathbb{P}^1(K)$ such that $L(x, t)Q(x, t) = 0$, which also forces $y = 0$. It thus suffices to bound the contribution from those (y, x, t) for which $L(x, t)Q(x, t) \neq 0$. If we write $z = y/L(x, t)$, then the equation $y^2 = L^2(x, t)Q(x, t)$ implies that $z^2 = Q(x, t)$. Moreover, if $H(y, x, t) \leq B$, then z will have a representative in $\mathcal{O}_K$ with $\|z\|_\infty \ll B^2$ and (x, t) a representative in $\mathcal{O}_K^2$ with $\|(x, t)\|_\infty \ll B$. It follows that

$$N_{\boldsymbol{c}}(B) \leq \#\{(z, x, t) \in \mathbb{P}^2(K) \colon \|(x, t)\|_\infty \ll B, \|z\|_\infty \ll B^2, z^2 = Q(x, t)\}.$$

The binary form $Q(x, t)$ must be square-free, as else $y^2 - L^2 Q$ would be reducible, a case that we excluded by assumption. In particular, $z^2 = Q(x, t)$ defines an irreducible conic in $\mathbb{P}^2$ and we can use Proposition 7.3.4 to deduce that

$$N_{\boldsymbol{c}}(B) \ll B^{4/3}.$$

We will now bound the number of $\boldsymbol{c} \in \mathbb{P}^2(K)$ such that $g_{\boldsymbol{c}}$ has a multiple factor. This happens precisely if the hyperplane $H_{\boldsymbol{c}}$ has singular intersection with the quartic curve $V(g) \subset \mathbb{P}^2$ and hence $\boldsymbol{c}$ must lie on the dual curve $V(g)^* \subset \mathbb{P}^2$ of $V(g)$. As $V(g)$ is a smooth quartic curve and $\mathrm{char}(K) \neq 2, 3$, the dual curve is an irreducible curve of degree 12. By Proposition 7.2.13 it follows that the number of $\boldsymbol{c} \in V(g)^*$ with $H(\boldsymbol{c}) \ll B^{1/2}$ is $O(B^{1/12})$. Therefore in total we get a contribution of $O(B^{17/12})$ from this case.

Finally and lastly we assume that $y^2 - g_{\boldsymbol{c}}(x, t)$ is geometrically reducible. This happens precisely if $g_{\boldsymbol{c}}$ is a perfect square as a polynomial in $\overline{K}[x, t]$. In particular, any such $H_{\boldsymbol{c}}$ is a bitangent vector of $V(g)$. The 56 exceptional curves in X correspond precisely to the preimages of the 28 bitangents of $V(g)$ under π, and so any such point on $X_{\boldsymbol{c}}$ is excluded from our count.

We conclude that in total we obtain $N(B) \ll_{\varepsilon} B^{3/2+\varepsilon}$, which suffices for Theorem 7.1.2 when $d = 2$.

7.8 del Pezzo surfaces of degree 1

Via the anticanonial embedding, a non-singular del Pezzo surface X of degree 1 may be realised as a hypersurface in weighted projective space $\mathbb{P}(3,2,1,1)$ in the variables (y,x,u,v) given by an equation of the form

$$X : y^2 = x^3 + g(u,v)x + h(u,v),$$

where g and h are binary forms of degrees 4 and 6, respectively. An anti-canonical height function is given by

$$H(\boldsymbol{t}) = \prod_{\nu\in\Omega_K} \max\{|y|_\nu^{1/3}, |x|_\nu^{1/2}, |u|_\nu, |v|_\nu\}$$

for $\boldsymbol{t} = [y,x,u,v] \in \mathbb{P}(3,2,1,1)(K)$. The counting function of interest is then given by

$$N(B) = \#\{\boldsymbol{t} \in X(K) \colon H(\boldsymbol{t}) \leq B\}.$$

In this case we are aiming for an upper bound of the form $O(B^{2+\varepsilon})$, and so we do not have to remove any of the exceptional curves or singular members of $|-K_X|$. Moreover, the point $[1,1,0,0]$ is the unique base point of the anti-canonical linear system, which induces the rational map $[y,x,u,v] \mapsto [u,v] \in \mathbb{P}^1$ and so we may restrict our attention to counting those points for which $(u,v) \neq \mathbf{0}$.

Note that upon replacing $[y,x,u,v] \in \mathbb{P}(3,2,1,1)(K)$ with $[\mu^3 y, \mu^2 x, \mu u, \mu v]$ for a suitable unit $\mu \in \mathcal{O}_K^\times$, we see from Lemma 7.2.7 that every element of $\mathbb{P}(3,2,1,1)(K)$ with height at most B has a representative $(y,x,u,v) \in \mathcal{O}_K^4$ with $\|y\|^{1/3}, \|x\|^{1/2}, \|(u,v)\| \ll B^{1/s_K}$. It thus follows that

$$N(B) \leq \sum_{\substack{(u,v)\in\mathcal{O}_K^2 \\ 0<\|(u,v)\|\ll B^{1/s_K}}} N_{u,v}(B),$$

where

$$N_{u,v}(B) \coloneqq \#\{(y,x) \in \mathcal{O}_K^2 \colon \|y\| \ll B^{3/s_K}, \|x\| \ll B^{2/s_K}, y^2 = x^3 + g(u,v)x + h(u,v)\}.$$

Define $\Delta(u,v) = 4g(u,v)^2 + 27h(u,v)^3$. Note that this is a homogeneous polynomial of degree 12. If $(u,v) \in \mathcal{O}_K^2$ is such that $\Delta(u,v) \neq 0$ then $ZY^2 = X^3 + g(u,v)XZ^2 + h(u,v)Z^3$ defines an elliptic curve in $\mathbb{P}^2$ and any integral point $(y,x) \in \mathcal{O}_K^2$ gives rise to a unique rational point on the elliptic curve. By Proposition 7.1.4 we thus find $N_{(u,v)}(B) \ll B^\varepsilon$ in this case. Since the number of $(u,v) \in \mathcal{O}_K^2$ with $\|u,v\| \ll B^{1/s_K}$ is $O(B^2)$ by Lemma 7.2.6, we get a total contribution of $O_\varepsilon(B^{2+\varepsilon})$, which is sufficient.

It remains to estimate $\sum_{(u,v)} N_{u,v}(B)$ where the sum runs over $(u,v) \in \mathcal{O}_K^2$ that satisfy $\Delta(u,v) = 0$. Note first that there are at most twelve solutions $[u_i, v_i]$, $i = 1,\dots,12$ to $\Delta(u,v) = 0$ in $\mathbb{P}^1(K)$. Further, any $(u,v) \in \mathcal{O}_K^2$ such that $[u,v] = [u_i,v_i]$ in $\mathbb{P}^1(K)$ must be an $\mathcal{O}_K$-multiple of (u_i,v_i). By Lemma 7.2.6 we know that there are most $O(B\|(u_i,v_i)\|_\infty^{-1}) = O(B)$ pairs $(u,v) = \lambda(u_i,v_i)$ with $\|\lambda\| \ll B^{1/s_K}$ and $\mathrm{N}(\lambda) \ll B\|(u,v)\|_\infty^{-1}$. We thus find that the number of summands that we currently consider is bounded by $O(B)$, where the implied constant is independent of X. Moreover, it is clear that

$$N_{u,v}(B) \leq \#\{(y,x) \in \mathcal{O}_K^2 \colon \|(y,x)\| \ll B^{3/s_K}, y^2 = x^3 + g(u,v)x + h(u,v)\}$$

and that the affine curve defined by $y^2 = x^3 + g(u,v)x + h(u,v) \subset \mathbb{A}^2$ is irreducible. The affine version of the Bombieri-Pila bound generalised to global fields by Paredes and Sasyk [155, Theorem 1.9], implies that the last quantity is bounded by $O(B)$, where the implied constant is independent of X. Hence we obtain a total contribution of $O(B^2)$, which completes the proof of Theorem 7.1.2 for $d = 1$.

CHAPTER 8

Canonical singularities on moduli spaces of rational curves via the circle method

This chapter is based on [86].

8.1 Introduction

The interplay between geometry and number theory has a long and rich history. Examples such as Deligne's resolution of the Weil conjectures [66, 67], which serves as a powerful tool for estimating exponential sums and makes crucial use of the heavy machinery of algebraic geometry, shows that the impact of geometry on number theory cannot be underestimated. Rather surprisingly, sometimes this flow of information can be reversed and tools from analytic number theory can be used to give information about objects of geometric interest. In particular, building on ideas of Ellenberg and Venkatesh, Browning and Vishe [42] have demonstrated how one can use the circle method over function fields to deduce crude geometric properties of the moduli space $\mathcal{M}_{0,0}(X, e)$ of rational curves of degree e on a smooth hypersurface $X \subset \mathbb{P}^{n-1}$ of degree d when n is large compared to d. The aim of this chapter is to enrich this flow of information and develop a suitable form of the circle method to show that $\mathcal{M}_{0,0}(X, e)$ has only canonical singularities under suitable assumptions on n, d and e.

Let

$$\overline{\mu} = n(e + 1) - de - 5. \tag{8.1.1}$$

A naive heuristic based on Riemann–Roch leads one to expect that $\dim \mathcal{M}_{0,0}(X, e) = \overline{\mu}$. Assuming that $n \geq d + 3$ and X is *general*, Riedl and Yang [163] have shown that this is indeed the case and also established the irreducibility of $\mathcal{M}_{0,0}(X, e)$. Assuming $d \geq 3$, the methods based on analytic number theory from Browning and Vishe [44] allow one to deal with *any* smooth hypersurface at the cost of requiring the more stringent assumption $n > (5d - 4)2^{d-1}$. This was later refined to $n > (2d - 1)2^{d-1}$ by Browning and Sawin [42]. Moreover, by developing a motivic version of the circle method, Bilu and Browning [13] established a stabilisation result for $\mathrm{Mor}_e(\mathbb{P}^1, X)$ — the space of degree e morphisms from $\mathbb{P}^1$ to X — in the Grothendieck ring of varieties, which implies that $\mathcal{M}_{0,0}(X, e)$ is irreducible and of the expected dimension for smooth X as soon as $n > (d - 1)2^{d-1}$.

Apart from this crude geometric information about $\mathcal{M}_{0,0}(X, e)$, very little is known about their singularities. When $d \geq 3$ and $n \geq 2d + 1$, Harris, Roth and Starr [94] have shown that

$\mathcal{M}_{0,0}(X,e)$ is generically smooth. Moreover, Browning and Sawin [42] give upper bounds for the dimension of the singular locus of $\mathcal{M}_{0,0}(X,e)$. In this work we are concerned with the qualitative nature of the singularities that can occur on $\mathcal{M}_{0,0}(X,e)$. Even for smooth Fano hypersurfaces one cannot hope in general that $\mathcal{M}_{0,0}(X,e)$ is smooth, as shown by Example 2.2 of [42], and so a natural question is how bad the singularities can be.

When $d=1$ or $d=2$, then work of Kontsevich [126] and Pandharipande [121] shows that the spaces $\mathcal{M}_{0,0}(X,e)$ are smooth and irreducible for all $e \geq 1$. Our main result is as follows.

Theorem 8.1.1. *Suppose that $X \subset \mathbb{P}^{n-1}$ is a smooth hypersurface of degree $d \geq 3$ over a field of characteristic 0. If*

$$n > \begin{cases} (d^2+d-4)2^{d-1} & \text{when } e=1, \\ (de+1)(d-1)2^{d-1} & \text{when } e \geq 2, \end{cases}$$

then $\mathcal{M}_{0,0}(X,e)$ has only canonical singularities.

Turning to the particular case $e=1$, the space $F_1(X) = \mathcal{M}_{0,0}(X,1)$ is the *Fano variety variety of lines* of X. It is a classical result due to Altman and Kleiman [1] that the Fano variety of lines of *any* smooth cubic hypersurface $X \subset \mathbb{P}^{n-1}$ is smooth for $n \geq 5$. If $X \subset \mathbb{P}^{n-1}$ is a hypersurface of degree $d \geq 4$, then $F_1(X)$ is still known to be smooth, if one assumes that X is general and $d \leq 2n-5$ [124, Theorem V.4.3]. In particular, Theorem 8.1.1 gives new information about $F_1(X)$ for smooth X as soon as $d \geq 4$ and $n > (d^2+d-4)2^{d-1}$.

Theorem 8.1.1 gives a partial answer to a question by Starr [190], where he asked which type of singularities can occur on the spaces $\mathcal{M}_{0,0}(X,e)$. In the same work he proved that for $e \geq 2$ and $d \geq 3$ the spaces $\mathcal{M}_{0,0}(X,e)$ have only canonical singularities if $d+e+1 \leq n$ and X is *general*. While Theorem 8.1.1 puts stronger constraints on the number of variables than Starr's result, it has the advantage that it applies to *any* smooth hypersurface and not just generic hypersurfaces.

In forthcoming work with Hase-Liu [87], we generalise Theorem 8.1.1 to moduli spaces of higher genus curves and also prove that these moduli spaces only have terminal singularities. Along the way we are able to remove the dependence of n on e and show that the results hold provided $n \gg d^3 2^d$ and e is sufficiently large with respect to g.

Outline

We will now give a brief overview of the main steps in the proof of Theorem 8.1.1. Instead of working with $\mathcal{M}_{0,0}(X,e)$ directly, it suffices to work with the naive parameter space $\mathrm{Mor}_e(\mathbb{P}^1, X)$ of degree e morphisms from $\mathbb{P}^1$ to X. Our proof relies on studying the jet schemes $J_m(\mathrm{Mor}_e(\mathbb{P}^1,X))$ for any integer $m \geq 0$. The link to canonical singularities is provided by a result due to Mustaţă [150], which states that a variety Y has canonical singularities if and only if $J_m(Y)$ is irreducible for all $m \geq 0$. More precisely, we will deduce Theorem 8.1.1 from the following result.

Theorem 8.1.2. *Assume that $X \subset \mathbb{P}^{n-1}$ is a smooth hypersurface of degree $d \geq 3$ over a field K with $\mathrm{char}(K) > d$ if it is positive. Then providing*

$$n > \begin{cases} (d^2+d-4)2^{d-1} & \text{when } e=1, \\ (de+1)(d-1)2^{d-1} & \text{when } e \geq 2, \end{cases}$$

the mth jet scheme $J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$ is irreducible and of dimension $(m+1)(\overline{\mu}+3)$ for all $m \geq 0$.

For any variety Y, the fiber of the morphism $J_m(Y) \to Y$ above a smooth point is isomorphic to $\mathbb{A}^{m \dim Y}$. In addition, points on $\mathcal{M}_{0,0}(X, e)$ correspond to PGL_2 orbits on $\mathrm{Mor}_e(\mathbb{P}^1, X)$, so that the expected dimension of $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is $\overline{\mu} + \dim \mathrm{PGL}_2 = \overline{\mu} + 3$. In particular, Theorem 8.1.2 confirms the naive expectation that the dimension of $J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$ is $(m+1)(\overline{\mu}+3)$.

Our strategy to prove that $J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$ is irreducible and of the expected dimension follows the path paved by Browning and Vishe [44]. More precisely, after performing a spreading out argument, it suffices to prove the corresponding result over a finite field $\mathbb{F}_q$. By appealing to the Lang–Weil estimate this is equivalent to understanding the number of $\mathbb{F}_q$-points on $J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$ as q goes to to infinity, which in turn correspond to $\mathbb{F}_q[t][s]/(s^{m+1})$-points on X whose degree in t is bounded by e, which we solve by developing a suitable version of the circle method. This counting problem can be interpreted as counting the number of $\mathbb{F}_q[t]$-points of bounded height on a system of $m+1$ equations in $n(m+1)$ variables. Classical applications of the circle method require the number of variables to grow roughly quadratically in the number of equations. In particular, since for n, d and e we want to count points for *all* $m \geq 0$, this approach seems hopeless and Browning and Sawin [42] wrote in their work that it does not seem possible that their method will prove that $\mathcal{M}_{0,0}(X, e)$ has canonical singularities. Nevertheless, we succeed using a modified version of the circle method, as we shall now explain.

The key point is that we do not treat the equations for $J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$ as a system of equations, but keep working with it as a single equation over $\mathbb{F}_q[t][s]/(s^{m+1})$. Our situation may thus be compared to that over number fields: if one wants to count solutions to a single equation over a number field, but uses Weil restriction to consider it as a system of equations over $\mathbb{Q}$, then a naive application of the circle method would require the number of variables to grow quadratically in the degree of the number field. However, as demonstrated by Skinner [186], if things are set up correctly, one can essentially still treat it as one equation and obtain completely analogous results compared to $\mathbb{Q}$.

In our approach we perform harmonic analysis over $\mathbb{F}_q[s]/(s^{m+1})((t^{-1}))$ and then continue along the standard lines of attack of the circle method. This includes a division into minor and major arcs followed by a suitable form of Weyl differencing to bound the exponential sums that arise in this process. The extraneous factor of $de+1$ in Theorem 8.1.1 arises because the bounds we obtain for the exponential sums involved only depend on rational approximation of the parameter modulo s. In particular, our arcs are in some sense too coarse. The reason for this is that Diophantine approximation is significantly more difficult modulo higher powers of s due to the presence of zero divisors and non-constant units.

The results mentioned thus far have been generalised to the setting of smooth complete intersections of the same degree by Browning, Vishe and Yamagishi [45]. Given $p_1, \dots, p_b \in \mathbb{P}^1$ and $y_1, \dots, y_b \in X$, they further generalised their method to deal with the parameter space $\mathrm{Mor}_e(\mathbb{P}^1, X, p_1, \dots, p_b; y_1, \dots, y_b)$ of degree e morphisms $f \colon \mathbb{P}^1 \to X$ satisfying $f(b_i) = p_i$ for $i = 1, \dots, b$. Their approach is based on a function field version of Rydin Myerson's work [166, 167] that allows one to handle systems of forms with the circle method for which the number of variables only grows linearly in the number of equations. It seems plausible that our work extends to the setting of [45] and it would be interesting to see whether Rydin Myerson's approach could be useful in the study of canonical singularities. In another direction, Hase-Liu [95] has extended the circle method approach to shed light on the geometry of $\mathrm{Mor}_e(C, X)$ when C is a curve of genus $g \geq 1$. We believe that a combination of Hase-Liu's approach with the one developed in this paper should be capable of establishing that $\mathrm{Mor}_e(C, X)$ only has canonical singularities when $g \geq 1$ for suitable ranges of e and n.

8.2 Preliminaries

We will typically work with the space $\mathrm{Mor}_e(\mathbb{P}^1, X)$ of degree e morphisms from $\mathbb{P}^1$ to X instead of $\mathcal{M}_{0,0}(X, e)$. We begin by recalling some basic facts about parameter spaces of morphisms and jet schemes.

8.2.1 Parameter spaces of morphisms

Let $X \subset \mathbb{P}^{n-1}$ be a smooth hypersurface over a field K defined as the vanishing of a homogeneous polynomial $F \in K[x_1, \ldots, x_n]$ of degree d. Recall that a morphism $f \colon \mathbb{P}^1 \to X$ of degree e over K is given by a tuple $(f_1(s,t), \ldots, f_n(s,t))$ of binary forms $f_1, \ldots, f_n \in K[s,t]$ of degree e such that $f_1(s,t), \ldots, f_n(s,t)$ do not all share a common non-constant factor and $F(f_1(s,t), \ldots, f_n(s,t))$ vanishes identically in s and t. By identifying the space of n-tuples of binary forms of degree e modulo the action of $K^\times$ with $\mathbb{P}^{n(e+1)-1}$, we can thus realise $\mathrm{Mor}_e(\mathbb{P}^1, X)$ as an open subscheme of the closed subset of $\mathbb{P}^{n(e+1)-1}$ defined by $F(f_1(s,t), \ldots, f_n(s,t)) = 0$.

8.2.2 Jet schemes

We now recall some basic facts about jet schemes, all of which can be found in [71, Section 2]. Let Y be a scheme of finite type over K. For every integer $m \geq 0$, there exists a scheme $J_m(Y)$ of finite type over K such that for any k-algebra A the set of A-valued points $J_m(Y)(A)$ is given by $Y(A[s]/(s^{m+1}))$. The scheme $J_m(Y)$ is unique up to a canonical isomorphism and is called the *mth jet scheme* of Y.

Jet schemes are functorial in the sense that a morphism $f \colon Y \to Z$ of K-schemes induces a morphism $f_m \colon J_m(Y) \to J_m(Z)$ for every $m \geq 0$. Moreover, if f is an open or a closed immersion, then so is f_m. In particular, if $Y \subset \mathbb{P}^{n-1}$ is an open subset of a closed subvariety, then we can think of $J_m(Y)$ as being contained $J_m(\mathbb{P}^{n-1})$.

8.2.3 Spreading Out

We will now begin to relate the proof of Theorem 8.1.1 to a counting problem. We use the notation from Subsection 8.2.1 and recall that $\overline{\mu}$ defined in (8.1.1) is the expected dimension of $\mathcal{M}_{0,0}(X, e)$. Since two morphisms in $\mathrm{Mor}_e(\mathbb{P}^1, X)$ give rise to the same rational curve if and only if they differ by an element in PGL_2, the expected dimension of $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is

$$\mu := \overline{\mu} + 3 = n(e+1) - de - 2. \tag{8.2.1}$$

Our main ingredient is the following result due to Mustață [150].

Theorem 8.2.1. *Let Y be a local complete intersection scheme over a field of characteristic 0. Then Y has only canonical singularities if and only if the mth jet scheme Y is irreducible for all $m \geq 0$.*

We want to apply Theorem 8.2.1 to $\mathrm{Mor}_e(\mathbb{P}^1, X)$. Note that under the hypotheses of Theorem 8.1.1, it follows from [42, Theorem 1.1] that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ is a local complete intersection. Hence Mustaţă's result is applicable and Theorem 8.1.2 implies that $\mathrm{Mor}_e(\mathbb{P}^1, X)$ only has canonical singularities. Now to deduce Theorem 8.1.1 from Theorem 8.1.2, recall that $\mathcal{M}_{0,0}(X, e)$ is a Deligne–Mumford stack that can be realised as the stack quotient $[\mathrm{Mor}_e(\mathbb{P}^1, X)/\operatorname{PGL}_2]$. As PGL_2 is a smooth group scheme, the natural projection map $\mathrm{Mor}_e(\mathbb{P}^1, X) \to \mathcal{M}_{0,0}(X, e)$ is a smooth atlas. Having canonical singularities is a smooth local property and so $\mathcal{M}_{0,0}(X, e)$ only has canonical singularities if $\mathrm{Mor}_e(\mathbb{P}^1, X)$ does.

Next, we briefly recall the spreading out process as described in [44, Section 2]. To do so, define $X_e = \mathrm{Mor}_e(\mathbb{P}^1, X)$ and $X_{e,m} = J_m(\mathrm{Mor}_e(\mathbb{P}^1, X))$, so that $X_{e,0} = X_e$. For all $m \geq 0$, the schemes $X_{e,m}$ are defined over a finitely generated $\mathbb{Z}$-algebra Λ, that can be explicitly realised by adjoining the coefficients of F to $\mathbb{Z}$. For any maximal ideal $\mathfrak{m} \subset \Lambda$ the quotient $\Lambda/\mathfrak{m}$ is a finite field and $X_{e,m}$ will be irreducible once we can show that $X_{e,m} \times \mathrm{Spec}(\Lambda/\mathfrak{m})$ is irreducible for any maximal ideal $\mathfrak{m} \subset \Lambda$. By inverting $d!$ and enlarging Λ if necessary, we may moreover assume that the reduction of X modulo $\mathfrak{m}$ is smooth and $\mathrm{char}(\Lambda/\mathfrak{m}) > d$.

It follows from [151, Corollary 2.7] that any irreducible component of $X_{e,m}$ has dimension at least $(m+1)\dim X_e$. In addition, by Chevalley's upper semicontinuity there exists an open subset $U \subset \mathrm{Spec}(\Lambda)$ such that $\dim X_{e,m} \leq \dim(X_{e,m} \times \mathrm{Spec}(\Lambda/\mathfrak{m}))$ for all maximal ideals $\mathfrak{m} \subset U$.

In view of the Lang–Weil estimates, to show that $X_{e,m}$ is irreducible and of the expected dimension, if $\dim X_e = \mu$, it therefore suffices to show that

$$\lim_{q\to\infty} \frac{\# X_{e,m}(\mathbb{F}_q)}{q^{(m+1)\mu}} \leq 1. \tag{8.2.2}$$

We are now ready to introduce our main counting function. Given a non-singular homogeneous form $F \in \mathbb{F}_q[x_1, \dots, x_n]$ of degree d, we define

$$N_m(e) := \#\{\boldsymbol{x} \in (\mathbb{F}_q[s]/(s^{m+1})[t])^n \colon F(\boldsymbol{x}) \equiv 0 \bmod s^{m+1}, \deg_t \boldsymbol{x} \leq e\}. \tag{8.2.3}$$

Lemma 8.2.2. *Let $e \geq 1$. Suppose that*

$$N_m(e) = q^{(m+1)(\mu+1)}(1 + o(1))$$

as $q \to \infty$ for all $m \geq 0$. Then

$$\lim_{q\to\infty} q^{-(m+1)\mu} \# X_{e,m}(\mathbb{F}_q) \leq 1$$

for all $m \geq 0$.

Proof. Identifying again the space of n-tuples of binary forms of degree e up to multiplication by scalars with $\mathbb{P}^{n(e+1)-1}$, we clearly have

$$X_e \subset \{\boldsymbol{x} \in \mathbb{P}^{n(e+1)-1} \colon F(\boldsymbol{x}) = 0\},$$

since we merely dropped the coprimality condition in the definition of X_e. By functoriality of jet schemes, we can identify $X_{e,m}(\mathbb{F}_q)$ as a subset of the set of $\mathbb{F}_q$-points of the mth jet scheme of $\mathbb{P}^{n(e+1)-1}$, which is given by $\mathbb{P}^{n(e+1)-1}(\mathbb{F}_q[s]/(s^{m+1}))$. It follows that

$$X_{e,m}(\mathbb{F}_q) \subset \{\boldsymbol{x} \in \mathbb{P}^{n(e+1)-1}(\mathbb{F}_q[s]/(s^{m+1})) \colon F(\boldsymbol{x}) \equiv 0 \bmod s^{m+1}\}.$$

As $\mathbb{F}_q[s]/(s^{m+1})$ is a local ring, the collection of the $\mathbb{F}_q[s]/(s^{m+1})$-points of $\mathbb{P}^{n(e+1)-1}$ is

$$(\mathbb{F}_q[s]/(s^{m+1}))^{n(e+1)} \setminus \{\mathbf{0}\})/(\mathbb{F}_q[s]/(s^{m+1}))^{\times}.$$

Since an element of $\mathbb{F}_q[s]/(s^{m+1})$ is a unit if and only if its reduction modulo s lies in $\mathbb{F}_q^{\times}$, we have $\#(\mathbb{F}_q[s]/(s^{m+1}))^{\times} = (q-1)q^m$ and thus

$$\begin{aligned}\#X_{e,m}(\mathbb{F}_q) &\leq \#\{\boldsymbol{x} \in \mathbb{P}^{n(e+1)-1}(\mathbb{F}_q[s]/(s^{m+1}))\colon F(\boldsymbol{x}) \equiv 0 \bmod s^{m+1}\} \\ &= \frac{\#\{\boldsymbol{x} \in ((\mathbb{F}_q[s]/(s^{m+1}))[t])^n\colon F(\boldsymbol{x}) \equiv 0 \bmod s^{m+1}, \boldsymbol{x} \not\equiv \mathbf{0} \bmod s, \deg_t(\boldsymbol{x}) \leq e\}}{(q-1)q^m},\end{aligned}$$

where have identified binary forms of degree e with polynomials in t of degree at most e. Dropping the condition $\boldsymbol{x} \not\equiv \mathbf{0} \bmod s$, we therefore have

$$\#X_{e,m}(\mathbb{F}_q) \leq \frac{N_m(e)}{(q-1)q^m}.$$

In particular, if

$$N_m(e) = q^{(m+1)(\mu+1)}(1+o(1))$$

for all $m \geq 0$ as $q \to \infty$, then

$$\#X_{e,m} \leq q^{(m+1)\mu}(1+o(1))$$

holds as $q \to \infty$. □

In particular, Theorem 8.1.2 and hence also Theorem 8.1.1 will be a consequence of the following result, whose proof will be carried out in the final section.

Proposition 8.2.3. *Let $d \geq 3$ and assume that $e \geq 1$. Suppose that*

$$n > \begin{cases} (d^2+d-4)2^{d-1} & \text{when } e = 1, \\ (de+1)(d-1)2^{d-1} & \text{when } e \geq 2, \end{cases}$$

and that $F \in \mathbb{F}_q[x_1, \ldots, x_n]$ is non-singular of degree d. If $d < \mathrm{char}(\mathbb{F}_q)$, then for any $m \geq 0$, we have $N_m(e) = q^{(m+1)(\mu+1)}(1+o(1))$ uniformly in q.

8.3 Harmonic analysis

For $m \geq 0$, let $K_\infty = \mathbb{F}_q((t^{-1}))$ and $K_\infty^{(m)} = K_\infty[s]/(s^{m+1})$. We define $\mathcal{O} = \mathbb{F}_q[t]$ and $\mathcal{O}_m = (\mathbb{F}_q[s]/(s^{m+1}))[t]$, which consists of elements of the form $x = x_0 + sx_1 + \cdots s^m x_m$ with $x_i \in \mathcal{O}$ for $i = 0, \ldots, m$. Let $\alpha = \sum_{i \leq M} a_i t^i \in K_\infty$ be such that $a_m \neq 0$. Then we denote by $|\alpha|_0$ the norm on K_∞ given by $|\alpha|_0 = q^M$ and we define $\|\alpha\|_0 = |\sum_{i \leq -1} a_i t^i|_0$ to be the distance to the nearest integer. We extend both to all of K_∞ by $|0|_0 = \|0\|_0 = 0$. Any $\alpha \in K_\infty^{(m)}$ can be written as

$$\alpha = \alpha_0 + s\alpha_1 + \cdots + s^m \alpha_m \tag{8.3.1}$$

with $\alpha_i \in K_\infty$. We then define a norm on $K_\infty^{(m)}$ via $|\alpha|_m = \max |\alpha_i|_0$ and similarly set $\|\alpha\|_m = \max\{\|\alpha_i\|_0\}$. Moreover, we let

$$\mathbb{T}^{(m)} = \{\alpha \in K_\infty^{(m)}\colon |\alpha|_m < 1\}$$

be the analogue of the unit interval. There are reduction maps $\pi_l\colon K_\infty^{(m)} \to K_\infty^{(l)}$ for any $m > l$ and by abuse of notation we shall write $|\alpha|_l = |\pi_l(\alpha)|_l$ for $\alpha \in K_\infty^{(m)}$ and similarly $\|\alpha\|_l = \|\pi_l(\alpha)\|_l$.

Let $e_q\colon \mathbb{F}_q \to \mathbb{C}^\times$ be the additive character defined by $e_q(x) = \exp\Big(2\pi i \operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)/p\Big)$ and let $\psi\colon K_\infty \to \mathbb{C}^\times$ be the standard additive character defined by

$$\sum a_i t^i \mapsto e_q(a_{-1}),$$

where $\sum a_i t^i \in K_\infty$ and $a_i \in \mathbb{F}_q$. We define a character on $K_\infty^{(m)}$ by

$$\psi_m(\alpha) = \prod_{i=0}^{m} \psi(\alpha_i),$$

if α is given by (8.3.1). Let $\mathrm{d}\alpha$ be the usual Haar measure on K_∞ normalised in such a way that $\int_{\mathbb{T}^{(0)}} \mathrm{d}\alpha = 1$. It is clear that $K_\infty^{(m)}$ is an $(m+1)$-dimensional K_∞-vector space and thus we can extend the Haar measure from K_∞ to $K_\infty^{(m)}$ by $\mathrm{d}\alpha = \mathrm{d}\alpha_0 \cdots \mathrm{d}\alpha_m$ if α is given by (8.3.1). We begin by recording some useful orthogonality relations.

Lemma 8.3.1. *Let $x \in \mathcal{O}_m$ and $N \geq 0$. Then*

$$\int_{\{\alpha \in K_\infty^{(m)}\colon |\alpha|_m < q^{-N}\}} \psi_m(\alpha x) \mathrm{d}\alpha = \begin{cases} q^{-(m+1)N} & \text{if } |x|_m < q^N, \\ 0 & \text{else.} \end{cases}$$

Proof. It is clear that $B_N \coloneqq \{\alpha \in K_\infty^{(m)}\colon |\alpha|_m < q^{-N}\}$ is a compact subgroup of $K_\infty^{(m)}$ under addition and that the restriction of ψ_m to B_N defines a continuous character on it. It therefore follows from standard harmonic analysis arguments that

$$\int_{B_N} \psi_m(\alpha x) \mathrm{d}\alpha = \begin{cases} \mathrm{vol}(B_N) & \text{if } \psi_m(\alpha x) = 1 \text{ for all } \alpha \in B_N, \\ 0 & \text{else.} \end{cases}$$

A straightforward computation shows that $\mathrm{vol}(B_N) = q^{-(m+1)N}$, so that it suffices to show that $\alpha \mapsto \psi_m(\alpha x)$ restricts to the trivial character on B_N if and only if $|x|_m < q^N$. If $|x|_m < q^N$, then certainly $\psi_m(\alpha x) = 1$ identically in α. So let us suppose that $x = x_0 + \cdots + x_l t^l$ with $l \geq N$, where $x_i \in \mathbb{F}_q[s]/(s^{m+1})$ for $0 \leq i \leq l$ and $x_l \neq 0$. Let us write $x_l = s^k(y_l + s y_l')$ with $0 \leq k \leq m$, where $y_l \in \mathbb{F}_q^\times$ and $y_l' \in \mathbb{F}_q[s]/(s^{m-k})$. If we define $\alpha = s^{m-k} t^{-l-1}$, then by assumption $|\alpha|_m < q^{-N}$ and the coefficient of t^{-1} in αx is by construction $s^m y_l$, whence $\psi_m(\alpha x) = e_q(y_l) \neq 1$. □

Lemma 8.3.2. *Let $\alpha \in K_\infty^{(m)}$ and $N \geq 1$. Then*

$$\sum_{\substack{x \in \mathcal{O}_m \\ |x|_m < q^N}} \psi_m(\alpha x) = \begin{cases} q^{(m+1)N} & \text{if } \|\alpha\|_m < q^{-N} \\ 0 & \text{else.} \end{cases}$$

Proof. The collection of $x \in \mathcal{O}_m$ with $|x|_m < q^N$ forms a discrete subgroup $\mathcal{O}_m^{<N}$ of $\mathcal{O}_m$ of order $q^{(m+1)N}$ and $x \mapsto \psi_m(\alpha x)$ an additive character on it. So as in the proof of the previous lemma, it suffices to show that $\psi_m(\alpha x)$ is the trivial character on $\mathcal{O}_m^{<N}$ if and only if $\|\alpha\|_m < q^{-N}$. The if part holds trivially. In addition, since translation by an element in $\mathcal{O}_m$ does not affect the value of ψ_m, we may assume without loss of generality that $\alpha \in \mathbb{T}^{(m)}$ is such that $|\alpha|_m \geq q^{-N}$. So suppose $\alpha = \sum_{i \leq -M} \alpha_i t^i$ with $M \leq N$, where $\alpha_i \in \mathbb{F}_q[s]/(s^{m+1})$ and $\alpha_{-M} \neq 0$. We can then write $\alpha_{-M} = s^k(a_0 + s a_1)$ for some $a_0 \in \mathbb{F}_q^\times$, $a_1 \in \mathbb{F}_q[s]/(s^{m-k})$ and $0 \leq k \leq m$. If we define $x = s^{m-k} t^{M-1}$, then by assumption $|x|_m < q^N$ and by construction we have $\psi_m(\alpha x) = e_q(a_0) \neq 1$ as desired. □

8.4 Circle method

Recall that for $F \in \mathbb{F}_q[x_1, \dots, x_n]$ homogeneous of degree d, we defined the counting function

$$N_m(e) \coloneqq \#\{\boldsymbol{x} \in \mathcal{O}_m^n \colon |\boldsymbol{x}|_m \le q^e, F(\boldsymbol{x}) \equiv 0 \bmod s^{m+1}\}.$$

If we set

$$S(\alpha) = \sum_{\substack{\boldsymbol{x} \in \mathcal{O}_m^n \\ |\boldsymbol{x}|_m \le q^e}} \psi_m(\alpha F(\boldsymbol{x}))$$

for $\alpha \in \mathbb{T}^{(m)}$, then it follows from Lemma 8.3.1 with $N = 0$ that

$$N_m(e) = \int_{\mathbb{T}^{(m)}} S(\alpha) \mathrm{d}\alpha.$$

Our approach to studying $N_m(e)$ is to identify a set of major arcs, which will give the main contribution to $N_m(e)$, and a set of minor arcs on which $|S(\alpha)|$ is sufficiently small. More precisely, let $\alpha = \alpha_0 + \dots + s^m \alpha_m$ with $\alpha_i \in \mathbb{T}^{(0)}$ and $M = \lceil \frac{de+1}{2} \rceil$. For $-1 \le J \le M$, we then define

$$\mathfrak{M}(J) = \bigcup_{\substack{r \in \mathcal{O} \text{ monic} \\ |r| \le q^J}} \{\alpha \in \mathbb{T}^{(m)} \colon \|\alpha_0 r\|_0 < q^{J-de-1}\}, \tag{8.4.1}$$

where we note that $\mathfrak{M}(-1) = \emptyset$. It follows from Dirichlet's approximation theorem over K_∞ as recorded in [35, Lemma 5.7] that every $\alpha \in \mathbb{T}^{(m)}$ lies in $\mathfrak{M}(J+1) \setminus \mathfrak{M}(J)$ for some $-1 \le J \le M-1$. In addition, it is easy to see that

$$\mathrm{vol}(\mathfrak{M}(J)) \le q^{2J-de-1}. \tag{8.4.2}$$

8.4.1 Weyl differencing

Suppose that

$$F(\boldsymbol{x}) = \sum_{i_1, \dots, i_d = 1}^{n} c_{i_1, \dots, i_d} x_{i_1} \cdots x_{i_d}$$

is a non-singular form of degree d with symmetric coefficients $c_{i_1, \dots, i_d} \in \mathbb{F}_q$. Associated to F are the multilinear forms

$$\psi_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) = d! \sum_{i_1, \dots, i_{d-1} = 1}^{n} c_{i_1, \dots, i_{d-1}, i} x_{i_1}^{(1)} \cdots x_{i_{d-1}}^{(d-1)}$$

for $1 \le i \le n$. Let $V \subset \mathbb{A}^{(d-1)n}$ be the variety defined by $\psi_1 = \dots = \psi_n = 0$ and denote by V_m the mth jet scheme of V.

Lemma 8.4.1. *Let F be a non-singular form of degree d and suppose that $\mathrm{char}(K) > d$. If we define $m_0 = \lceil \frac{m+1}{d-1} \rceil$, then any irreducible component of V_m has dimension at most $(m+1)n(d-1) - nm_0$.*

Proof. We identify $\boldsymbol{x}^{(i)} \in \mathbb{A}^{(m+1)n}$ for $1 \le i \le d-1$ with $\boldsymbol{x}_0^{(i)} + s\boldsymbol{x}_1^{(i)} + \dots + s^m \boldsymbol{x}_m^{(i)}$, where $\boldsymbol{x}_j^{(i)} \in \mathbb{A}^n$. Since $V \subset \mathbb{A}^{(d-1)n}$, we have $V_m \subset J_m(\mathbb{A}^{(d-1)n}) = \mathbb{A}^{(m+1)n(d-1)}$. Under this description, V_m is given by

$$\{(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) \in \mathbb{A}^{(m+1)n(d-1)} \colon \psi_i(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(d-1)}) \equiv 0 \bmod s^{m+1} \text{ for } 1 \le i \le n\}.$$

Let now

$$\Delta = \{(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) \in \mathbb{A}^{(m+1)n(d-1)} \colon \boldsymbol{x}^{(1)} = \cdots = \boldsymbol{x}^{(d-1)}\}$$

be the diagonal, which has dimension $(m+1)n$. Suppose that $(\boldsymbol{x}, \ldots, \boldsymbol{x}) \in \Delta \cap V_m$. Then as $\mathrm{char}(K) > d$ we must have $\nabla F(\boldsymbol{x}) \equiv \mathbf{0} \bmod s^{m+1}$. If $\boldsymbol{x} \neq 0$, we can write $\boldsymbol{x} = s^l \boldsymbol{x}'$ with $\boldsymbol{x}' \neq \mathbf{0} \bmod s$. Then as F is non-singular, this implies $l(d-1) \geq m+1$, as otherwise the reduction of $\boldsymbol{x}'$ modulo s would produce a singular point of F over K. This is equivalent to $\boldsymbol{x}_0 = \cdots = \boldsymbol{x}_{m_0-1} = 0$ and hence

$$\dim \Delta \cap V_m \leq n(m+1-m_0).$$

Therefore,

$$\dim V_m \leq \dim(\Delta \cap V_m) + (m+1)n(d-1) - \dim \Delta = n(m+1-m_0) + (m+1)n(d-2),$$

from which the result follows. □

For $\alpha \in \mathbb{T}^{(m)}$ and $r_1, r_2 \geq 0$, let

$$M_m(\alpha, r_1, r_2) = \#\left\{ \underline{\boldsymbol{x}} \in \mathcal{O}_m^{n(d-1)} \colon \begin{array}{l} |\boldsymbol{x}^{(1)}|_m, \ldots, |\boldsymbol{x}^{(d-1)}|_m < q^{r_1 - r_2}, \\ \|\alpha \psi_i(\underline{\boldsymbol{x}})\|_m < q^{-r_1-(d-1)r_2} \text{ for } 1 \leq i \leq n \end{array} \right\}.$$

We will now employ Weyl differencing to obtain upper bounds for $S(\alpha)$. Replacing $\mathbb{F}_q[t]$ with $\mathcal{O}_m$, the argument leading to [131, Corollary 4.3.2] goes through verbatim and gives together with Lemma 8.3.2 the following result.

Lemma 8.4.2. *Let $\alpha \in \mathbb{T}^{(m)}$. Then*

$$|S(\alpha)|^{2^{d-1}} \leq q^{(e+1)(m+1)(2^{d-1}-d+1)n} M_m(\alpha, e+1, 0).$$

We require an analogue of Davenport's shrinking lemma, which follows in our situation directly from the analogous statement over $\mathbb{F}_q[t]$.

Lemma 8.4.3. *Let $L_1, \ldots, L_n \in K_\infty^{(m)}[x_1, \ldots, x_n]$ be linear forms and set*

$$K_m(a,b) := \#\{\boldsymbol{x} \in \mathcal{O}_m^n \colon |\boldsymbol{x}|_m < q^a, \|L_i(\boldsymbol{x})\|_m < q^{-b} \text{ for } 1 \leq i \leq n\}$$

for $a, b \geq 1$. Then for $0 \leq r < a \leq b$, we have

$$K_m(a,b) \leq q^{(m+1)nr} K(a-r, b+r).$$

Proof. We can write

$$L_i(\boldsymbol{x}) = \sum_{j=0}^{m} s^j L_{i,j}(\boldsymbol{x})$$

for $1 \leq i \leq n$ and some $L_{i,j} \in K_\infty[x_1, \ldots, x_n]$. If we write $\boldsymbol{x} = \boldsymbol{x}_0 + s\boldsymbol{x}_1 + \cdots + s^m \boldsymbol{x}_m$ with $\boldsymbol{x}_i \in \mathbb{F}_q[t]^n$ for $0 \leq i \leq m$, then it follows that

$$L_i(\boldsymbol{x}) = \sum_{k=0}^{m} s^k \widetilde{L}_{i,k}(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m),$$

as an identity in $K_\infty^{(m)}[x_1,\dots,x_n]$, where $\widetilde{L}_{i,k}$ is now the linear form over K_∞ in $(m+1)n$ variables given by

$$\widetilde{L}_{i,k}(\boldsymbol{x}_0,\dots,\boldsymbol{x}_m)=\sum_{j+l=k}L_{i,j}(\boldsymbol{x}_l).$$

With this notation, for any $1\le i\le n$, we have that

$$\|L_i(\boldsymbol{x})\|_m<q^{-b}\quad\text{if and only if}\quad \left\|\widetilde{L}_{i,k}(\boldsymbol{x}_0,\dots,\boldsymbol{x}_m)\right\|_0<q^{-b}\text{ for }0\le k\le m.$$

In particular, we can rewrite $K_m(a,b)$ as a problem over $\mathbb{F}_q[t]$ via

$$K_m(a,b)=\#\left\{(\boldsymbol{x}_0,\dots,\boldsymbol{x}_m)\in\mathbb{F}_q[t]^{n(m+1)}:\begin{array}{c}|\boldsymbol{x}_0|_0,\dots,|\boldsymbol{x}_m|_0<q^a,\\ \left\|\widetilde{L}_{i,k}(\boldsymbol{x}_0,\dots,\boldsymbol{x}_m)\right\|_0<q^{-b}\\ \text{for }1\le i\le n,0\le k\le m\end{array}\right\}.$$

We can now apply Lemma 5.3 of [42] to deduce that $K_m(a,b)\le q^{(m+1)nr}K_m(a-r,b+r)$ as desired. ☐

We now have everything at hand to reveal our main estimate for the exponential sums involved.

Lemma 8.4.4. *Let $\alpha\notin\mathfrak{M}(J)$ and suppose $l\in\mathbb{Z}_{\ge0}$ is such that*

$$l\le 1+\frac{J}{d-1}\quad\text{and}\quad l\le e+1.$$

Let $m_0=\lceil\frac{m+1}{d-1}\rceil$. Then there exists a constant $c>0$ depending only on d,n and m such that

$$|S(\alpha)|\le cq^{(e+1)n(m+1)-lnm_0/2^{d-1}}.$$

Proof. Recall from Lemma 8.4.2 that

$$|S(\alpha)|^{2^{d-1}}\le q^{(e+1)(m+1)(2^{d-1}-d+1)n}M_m(\alpha,e+1,0).\tag{8.4.3}$$

We will now give an upper bound for $M_m(\alpha,e+1,0)$ under the assumption that $\alpha\notin\mathfrak{M}(J)$. Set $r=e+1-l$. We can apply Lemma 8.4.3 $d-1$ times with this choice of r to $M_m(\alpha,e+1,0)$, by fixing all but one of $\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)}$. With the notation from the lemma, we always have $a=e+1\ge r=e+1-l$ and in the ith step $b=e+1+l\ge a$, so that the hypotheses of Lemma 8.4.3 are satisfied. This yields

$$M_m(\alpha,e+1,0)\le q^{(m+1)n(d-1)r}M_m(\alpha,e+1,r).$$

Suppose that $(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)})\in\mathcal{O}_m^{n(d-1)}$ is counted by $M_m(\alpha,e+1,r)$ and define $\rho=\psi_i(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(d-1)})$. Then upon writing $\rho=\rho_0+\dots+\rho_ms^m$ with $\rho_j\in\mathbb{F}_q[t]$ for $j=0,\dots,m$, we have $|\rho_j|_0\le q^{(d-1)(l-1)}$. Moreover, if $\alpha=\alpha_0+\dots+s^m\alpha_m$ with $\alpha_j\in K_\infty$, then

$$\|\alpha\rho\|_m=\|\alpha_0\rho_0+s(\alpha_0\rho_1+\alpha_1\rho_0)+\dots+s^m(\alpha_0\rho_m+\cdots\alpha_m\rho_0)\|_m$$

and hence $\|\alpha\rho\|_m<q^{-(e+1)-(d-1)r}=q^{(l-1)(d-1)-ed-1}$ holds if and only if

$$\|\alpha_0\rho_0\|_0,\dots,\|\alpha_0\rho_m+\dots+\alpha_m\rho_0\|_0<q^{(l-1)(d-1)-de-1}.\tag{8.4.4}$$

Note that (8.4.4) does not change if we multiply ρ by an element in $\mathbb{F}_q^\times$. In particular, we may assume that ρ_0 is monic. Therefore, as $\alpha\notin\mathfrak{M}(J)$ and $l\le 1+J/(d-1)$, the inequality

$\|\alpha_0\rho_0\|_0 < q^{(l-1)(d-1)-de-1}$ can only hold if $\rho_0 = 0$ by the definition of $\mathfrak{M}(J)$ in (8.4.1). Thus (8.4.4) implies $\|\alpha_0\rho_1\|_0 < q^{(l-1)(d-1)-de-1}$ and hence again $\rho_1 = 0$. Continuing in this fashion we get $\rho_0 = \cdots = \rho_m = 0$ and thus $\psi_i(\boldsymbol{x}^{(0)}, \ldots, \boldsymbol{x}^{(d-1)}) = 0$. It follows that

$$M_m(\alpha, e{+}1, r) \leq \# \left\{ (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) \in \mathcal{O}_m^{n(d-1)} \colon \begin{array}{l} |\boldsymbol{x}^{(1)}|_m, \ldots, |\boldsymbol{x}^{(d-1)}|_m \leq q^{e-r}, \\ \psi_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) = 0 \text{ for } 1 \leq i \leq n \end{array} \right\}.$$

If $\boldsymbol{x}^{(i)} = \boldsymbol{x}_0^{(i)} + \cdots + s^m \boldsymbol{x}_m^{(i)}$ with $\boldsymbol{x}_j^{(i)} \in \mathcal{O}$, then by definition of jet schemes we have that $\psi_i(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(d-1)}) = 0$ for $i = 1, \ldots, n$ if and only if $(\boldsymbol{x}_j^{(i)})_{1 \leq i \leq d-1, 0 \leq j \leq m} \in \mathcal{O}_0^{(d-1)n(m+1)}$ lies on the mth jet scheme V_m of V. In particular, the upper bound for the dimension from Lemma 8.4.1 combined with uniform estimates for the number of $\mathbb{F}_q[t]$-points in terms of the dimension and the degree [34, Lemma 2.8] imply that

$$\begin{aligned} M_m(\alpha, e+1, 0) &\leq q^{(m+1)n(d-1)r} \#\{\underline{\boldsymbol{x}} \in \mathcal{O}_0^{(m+1)(d-1)n} \colon \underline{\boldsymbol{x}} \in V_m, |\underline{\boldsymbol{x}}|_0 < q^l\} \\ &\leq c_{d,n,m} q^{(m+1)n(d-1)(e+1-l)+l((m+1)n(d-1)-nm_0)}, \end{aligned}$$

for some constant $c_{d,n,m} > 0$ that is independent of q. Once combined with (8.4.3), a straightforward computation shows

$$|S(\alpha)|^{2^{d-1}} \leq c_{d,n,m} q^{(e+1)n(m+1)2^{d-1} - lnm_0},$$

from which the statement of the lemma follows. □

8.4.2 Deduction of Proposition 8.2.3

We will now assume that $n > (de+1)(d-1)2^{d-1}$ if $e \geq 1$ and $n > (d^2+d-4)2^{d-1}$ if $e = 1$. Note that since we also assume that $d \geq 3$, we have $d^2 + d - 4 \geq (d+1)(d-1)$, so that $n > (de+1)(d-1)2^{d-1}$ also holds for $e = 1$. Our goal is to prove

$$N_m(e) = q^{(m+1)(n(e+1)-de-1)}(1 + o(1)) \tag{8.4.5}$$

for all $m \geq 0$ as $q \to \infty$. We will proceed by an induction on m. The case $m = 0$ is already handled by Browning and Sawin [42, Section 5.3] under the weaker assumption $n > (2d-1)2^{d-1}$. We may therefore assume $m > 0$ from now on.

Major arcs

Let us now evaluate the contribution from $J = 0$ in (8.4.1). Note that if $\alpha = \alpha_0 + \cdots + s^m \alpha_m$ with $\alpha_i \in \mathbb{T}^{(0)}$, then $\alpha \in \mathfrak{M}(0)$ implies $|\alpha_0|_0 < q^{-de-1}$. In particular, if $\boldsymbol{x} \in \mathcal{O}_m^n$ is such that $|\boldsymbol{x}|_m \leq q^e$, then we have $|\alpha_0 F(\boldsymbol{x})|_m \leq |\alpha_0|_m q^{de} < q^{-1}$. It follows that if we write $\alpha' = \alpha_1 + \cdots + s^{m-1}\alpha_m$, then we must have $\psi_m(\alpha F(\boldsymbol{x})) = \psi_m(s\alpha' F(\boldsymbol{x})) = \psi_{m-1}(\alpha' F(\boldsymbol{x}))$, which only depends on $\boldsymbol{x}$ modulo s^m. Therefore,

$$\begin{aligned} \int_{\mathfrak{M}(0)} S(\alpha) \mathrm{d}\alpha &= q^{-de-1} \int_{\mathbb{T}^{(m-1)}} S(\alpha') \mathrm{d}\alpha' \\ &= q^{n(e+1)-de-1} N_{m-1}(e) \\ &= q^{(m+1)(n(e+1)-de-1)}(1 + o(1)) \end{aligned}$$

as $q \to \infty$ by the induction hypothesis.

Minor arcs

Recall the definition of the major arcs $\mathfrak{M}(J)$ in (8.4.1). As noted earlier, every $\alpha \in \mathbb{T}^{(m)} \setminus \mathfrak{M}(0)$ lies in $\mathfrak{M}(J+1) \setminus \mathfrak{M}(J)$ for some $0 \leq J \leq M-1$, where $M = \lceil \frac{de+1}{2} \rceil$. Let $m_0 = \lceil \frac{m+1}{d-1} \rceil$ and $l = 1 + \lfloor \frac{J}{d-1} \rfloor$. One can check that as long as $e \geq 1$ and $d \geq 2$, the hypotheses of Lemma 8.4.4 are met. In particular, it follows from (8.4.2) and Lemma 8.4.4 that

$$\begin{aligned}\int_{\mathfrak{M}(J+1)\setminus\mathfrak{M}(J)} |S(\alpha)| \mathrm{d}\alpha &\ll q^{2J+2-de-1+n(e+1)(m+1)-lnm_0/2^{d-1}} \\ &= q^{(m+1)(n(e+1)-de-1)+m(de+1)+2J+2-lnm_0/2^{d-1}}.\end{aligned}$$

If we write $J = J_1(d-1) + J_2$ with $0 \leq J_2 \leq d-2$, then the exponent becomes

$$(m+1)(n(e+1)-de-1) + m(de+1) + 2J_2 + 2 - nm_0/2^{d-1} + J_1(2(d-1) - nm_0/2^{d-1}).$$

As $m_0 \geq 1$ and our assumptions on n imply that $2(d-1) - n/2^{d-1} < 0$, the term is maximal at $J_1 = 0$ and $J_2 = d-2$, as we shall henceforth assume. Since there are at most $M = O_{d,e}(1)$ choices for J, in order to establish (8.4.5) it suffices to show that

$$E \coloneqq m(de+1) + 2d - 2 - nm_0/2^{d-1} < 0. \tag{8.4.6}$$

Firstly, suppose that $m+1 \leq d-1$, so that $m_0 = 1$. We then have

$$E \leq (d-2)(de+1) + 2(d-1) - n/2^{d-1}.$$

If $e = 1$, we assume that $n > (d^2+d-4)2^{d-1}$, which implies

$$\begin{aligned}E &< (d-2)(d+1) + 2(d-1) - (d^2+d-4) \\ &= 0,\end{aligned}$$

while for $e \geq 2$ the assumption $n > (de+1)(d-1)2^{d-1}$ gives

$$\begin{aligned}E &< (d-2)(de+1) + 2(d-1) - (de+1)(d-1) \\ &= -(de+1) + 2(d-1) \\ &\leq -(2d+1) + 2(d-1) \\ &= -3,\end{aligned}$$

which is both satisfactory.

Secondly, suppose that $m+1 > d-1$ and hence $m_0 = \lceil \frac{m+1}{d-1} \rceil \geq \frac{m+1}{d-1}$. Therefore,

$$\begin{aligned}E &\leq m(de+1) + 2(d-1) - n(m+1)/2^{d-1}(d-1) \\ &= m(de+1-n/2^{d-1}(d-1)) + 2(d-1) - n/2^{d-1}(d-1).\end{aligned}$$

Under our assumptions on n and d, the coefficient of m is negative, so that $m \geq d-1$ and $n > (de+1)(d-1)2^{d-1}$ implies

$$\begin{aligned}E &\leq (d-1)(de+1-n/2^{d-1}(d-1)) + 2(d-1) - n/2^{d-1}(d-1) \\ &< d-1-n/2^{d-1}(d-1),\end{aligned}$$

which is negative since we assume that $n > (d-1)(de+1)2^{d-1} \geq (d-1)^2 2^{d-1}$ and therefore completes our treatment of (8.4.5).

Bibliography

[1] A. B. Altman and S. L. Kleiman. Foundations of the theory of Fano schemes. *Compositio Math.* **34** (1977), pp. 3–47.

[2] N. Arala. The analytic Hasse Principle for certain singular intersections of quadrics in $\mathbb{P}^9$. *Preprint* (2023), arXiv: `2310.15969`.

[3] N. Arala, J. R. Getz, J. Hou, C.-H. Hsu, H. Li, and V. Y. Wang. A nonabelian circle method. *Preprint* (2024), arXiv: `2407.11804`.

[4] E. Artin. *Quadratische Körper im Gebiete der höheren Kongruenzen*. PhD thesis. Universiät Leipzig, 1921.

[5] P. Autissier, M. Hindry, and F. Pazuki. Regulators of elliptic curves. *Int. Math. Res. Not. IMRN* (2021), pp. 4976–4993.

[6] V. V. Batyrev and Y. Tschinkel. Manin's conjecture for toric varieties. *J. Algebraic Geom.* **7** (1998), pp. 15–53.

[7] V. V. Batyrev and Y. Tschinkel. Rational points on some Fano cubic bundles. *C. R. Acad. Sci. Paris Sér. I Math.* **323** (1996), pp. 41–46.

[8] R. Beheshti, B. Lehmann, E. Riedl, and S. Tanimoto. Rational curves on del Pezzo surfaces in positive characteristic. *Trans. Amer. Math. Soc. Ser. B* **10** (2023), pp. 407–451.

[9] J. Bergström, A. Diaconu, D. Petersen, and C. Westerland. Hyperelliptic curves, the scanning map, and moments of families of quadratic L-functions. *Preprint* (2023), arXiv: `2302.07664`.

[10] C. Bernert. Small solutions to homogeneous and inhomogeneous cubic equations. *Preprint* (2023), arXiv: `2310.02039`.

[11] C. Bernert and L. Hochfilzer. Cubic forms over imaginary quadratic number fields and pairs of rational cubic forms. *Preprint* (2023), arXiv: `2307.10294`.

[12] M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y. Zhao. Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves. *J. Amer. Math. Soc.* **33** (2020), pp. 1087–1099.

[13] M. Bilu and T. Browning. A motivic circle method. *Preprint* (2023), arXiv: `2304.09645`.

[14] B. J. Birch, H. Davenport, and D. J. Lewis. The addition of norm forms. *Mathematika* **9** (1962), pp. 75–82.

[15] B. J. Birch and H. P. F. Swinnerton-Dyer. The Hasse problem for rational surfaces. *J. Reine Angew. Math.* **274/275** (1975), pp. 164–174.

[16] B. J. Birch. Forms in many variables. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **265** (1962), pp. 245–263.

[17] E. Bombieri and J. Vaaler. On Siegel's lemma. *Invent. Math.* **73** (1983), pp. 11–32.

[18] D. Bonolis and T. Browning. Uniform bounds for rational points on hyperelliptic fibrations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **24** (2023), pp. 173–204.

[19] D. Bonolis, T. Browning, and Z. Huang. Density of rational points on some quadric bundle threefolds. *Preprint* (2022), arXiv: 2204.09322.

[20] J. Bourgain, C. Demeter, and L. Guth. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math. (2)* **184** (2016), pp. 633–682.

[21] D. Bourqui. Fonction zêta des hauteurs des variétés toriques déployées dans le cas fonctionnel. *J. Reine Angew. Math.* **562** (2003), pp. 171–199.

[22] D. Bourqui. Fonction zêta des hauteurs des variétés toriques non déployées. *Mem. Amer. Math. Soc.* **211** (2011), pp. viii+151.

[23] J. Brandes. Forms representing forms and linear spaces on hypersurfaces. *Proc. Lond. Math. Soc. (3)* **108** (2014), pp. 809–835.

[24] R. de la Bretèche and T. Browning. Manin's conjecture for quartic del Pezzo surfaces with a conic fibration. *Duke Math. J.* **160** (2011), pp. 1–69.

[25] R. de la Bretèche. Nombre de points de hauteur bornée sur les surfaces de del Pezzo de degré 5. *Duke Math. J.* **113** (2002), pp. 421–464.

[26] R. de la Bretèche and É. Fouvry. L'éclaté du plan projectif en quatre points dont deux conjugués. *J. Reine Angew. Math.* **576** (2004), pp. 63–122.

[27] N. Broberg. Rational points on cubic surfaces. In: *Rational points on algebraic varieties*. Springer, 2001, pp. 13–35.

[28] T. Browning and D. R. Heath-Brown. Counting rational points on hypersurfaces. *J. Reine Angew. Math.* **584** (2005), pp. 83–115.

[29] T. Browning and D. R. Heath-Brown. Rational points on quartic hypersurfaces. *J. Reine Angew. Math.* **629** (2009), pp. 37–88.

[30] T. Browning and R. Munshi. Rational points on singular intersections of quadrics. *Compos. Math.* **149** (2013), pp. 1457–1494.

[31] T. Browning and S. M. Prendiville. Improvements in Birch's theorem on forms in many variables. *J. Reine Angew. Math.* **731** (2017), pp. 203–234.

[32] T. Browning and E. Sofos. Counting rational points on quartic del Pezzo surfaces with a rational conic. *Math. Ann.* **373** (2019), pp. 977–1016.

[33] T. Browning and P. Vishe. Cubic hypersurfaces and a version of the circle method for number fields. *Duke Math. J.* **163** (2014), pp. 1825–1883.

[34] T. Browning and P. Vishe. Rational points on cubic hypersurfaces over $\mathbb{F}_q(t)$. *GAFA* **25** (2015), pp. 671–732.

[35] T. Browning. *Cubic forms and the circle method*. Vol. 343. Progress in Mathematics. Birkhäuser/Springer, Cham, 2021, pp. xiv+166.

[36] T. Browning. Revisiting the Manin-Peyre conjecture for the split del Pezzo surface of degree 5. *New York J. Math.* **28** (2022), pp. 1193–1229.

[37] T. Browning, R. Dietmann, and D. R. Heath-Brown. Rational points on intersections of cubic and quadric hypersurfaces. *Journal of the Institute of Mathematics of Jussieu* **14** (2015), pp. 703–749.

[38] T. Browning, J. Glas, and V. Y. Wang. Optimal sums of three cubes in $\mathbb{F}_q[t]$. *Preprint* (2024), arXiv: `2408.03668`.

[39] T. Browning, J. Glas, and V. Y. Wang. Sums of three cubes over a function field. *Preprint* (2024), arXiv: `2402.07146`.

[40] T. Browning and D. R. Heath-Brown. Forms in many variables and differing degrees. *Journal of the European Mathematical Society* **19** (2017), pp. 357–394.

[41] T. Browning and W. Sawin. A geometric version of the circle method. *Annals of Mathematics* **191** (2020), pp. 893–948.

[42] T. Browning and W. Sawin. Free rational curves on low degree hypersurfaces and the circle method. *Algebra Number Theory* **17** (2023), pp. 719–748.

[43] T. Browning and M. Swarbrick Jones. Counting rational points on del Pezzo surfaces with a conic bundle structure. *Acta Arith.* **163** (2014), pp. 271–298.

[44] T. Browning and P. Vishe. Rational curves on smooth hypersurfaces of low degree. *Algebra Number Theory* **11** (2017), pp. 1657–1675.

[45] T. Browning, P. Vishe, and S. Yamagishi. Rational curves on complete intersections and the circle method. *Preprint* (2024), arXiv: `2404.11123`.

[46] A. Brumer. The average rank of elliptic curves. I. *Invent. Math.* **109** (1992), pp. 445–472.

[47] A. Cafure and G. Matera. Improved explicit estimates on the number of solutions of equations over a finite field. *Finite Fields Appl.* **12** (2006), pp. 155–185.

[48] F. Campana. Connexité rationnelle des variétés de Fano. *Ann. Sci. École Norm. Sup. (4)* **25** (1992), pp. 539–545.

[49] L. Caporaso, J. Harris, and B. Mazur. Uniformity of rational points. *J. Amer. Math. Soc.* **10** (1997), pp. 1–35.

[50] M. Car. Sommes d'exponentielles dans $\mathbb{F}_{2^h}((X^{-1}))$. *Acta Arithmetica* **62** (1992), pp. 303–328.

[51] M. Car and J. Cherly. Sommes de cubes dans l'anneau $\mathbb{F}_{2^h}[X]$. *Acta Arithmetica* **65** (1993), pp. 227–241.

[52] M. Car and L. Gallardo. Sums of cubes of polynomials. *Acta Arithmetica* **112** (2004), pp. 41–50.

[53] J. W. S. Cassels. *Lectures on elliptic curves*. Vol. 24. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1991, pp. vi+137.

[54] J.-L. Colliot-Thélène. Points Rationnels sur les Fibrations. In: *Higher Dimensional Varieties and Rational Points*. Ed. by K. Böröczky, J. Kollár, and T. Szamuely. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 171–221.

[55] R. Conceição, D. Ulmer, and J. F. Voloch. Unboundedness of the number of rational points on curves over function fields. *New York J. Math.* **18** (2012), pp. 291–293.

[56] I. Connell. Elliptic curve handbook. *Preprint* McGill University, available at "`http://www.math.mcgill.ca/connell/public/ECH1/`". (1999),

[57] I. Coskun and J. Starr. Rational curves on smooth cubic hypersurfaces. *Int. Math. Res. Not. IMRN* (2009), pp. 4626–4641.

[58] M. Dale. Terracini's lemma and the secant variety of a curve. *Proc. London Math. Soc. (3)* **49** (1984), pp. 329–339.

[59] H. Davenport. Cubic forms in 29 variables. *Proc. Roy. Soc. London Ser. A* **266** (1962), pp. 287–298.

[60] H. Davenport. Cubic forms in sixteen variables. *Proc. Roy. Soc. London Ser. A* **272** (1963), pp. 285–303.

[61] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Second. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2005, pp. xx+140.

[62] H. Davenport. Cubic forms in thirty-two variables. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **251** (1959), pp. 193–232.

[63] H. Davenport. *Private letter to Keith Matthews*. 1964. URL: http://www.numbertheory.org/thesis/remarks_by_davenport.pdf.

[64] O. Debarre. *Higher-dimensional algebraic geometry*. Universitext. Springer-Verlag, New York, 2001, pp. xiv+233.

[65] R. Dedekind and H. Weber. Theorie der algebraischen Functionen einer Veränderlichen. *J. Reine Angew. Math.* **92** (1882), pp. 181–290.

[66] P. Deligne. La conjecture de Weil: I. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* **43** (1974), pp. 273–307.

[67] P. Deligne. La conjecture de Weil: II. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* **52** (1980), pp. 137–252.

[68] R. Dietmann. Weyl's inequality and systems of forms. *Q. J. Math.* **66** (2015), pp. 97–110.

[69] W. Duke, J. Friedlander, and H. Iwaniec. Bounds for automorphic L-functions. *Inventiones Mathematicae* **112** (1993), pp. 1–8.

[70] B. Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. J. Math.* **82** (1960), pp. 631–648.

[71] L. Ein and M. Mustaţă. Jet schemes and singularities. In: *Algebraic geometry—Seattle 2005. Part 2*. **80**. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 2009, pp. 505–546.

[72] J. S. Ellenberg, A. Venkatesh, and C. Westerland. Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields. *Ann. of Math. (2)* **183** (2016), pp. 729–786.

[73] A. Entin. Monodromy of hyperplane sections of curves and decomposition statistics over finite fields. *Int. Math. Res. Not. IMRN* (2021), pp. 10409–10441.

[74] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.* **73** (1983), pp. 349–366.

[75] D. Flores. A quantitative Hasse principle for weighted quartic forms. *Mathematika* **70** (2024), Paper No. e12236, 24.

[76] J. Franke, Y. I. Manin, and Y. Tschinkel. Rational points of bounded height on Fano varieties. *Inventiones Mathematicae* **95** (1989), pp. 421–435.

[77] C. Frei, D. Loughran, and E. Sofos. Rational points of bounded height on general conic bundle surfaces. *Proc. Lond. Math. Soc. (3)* **117** (2018), pp. 407–440.

[78] C. Frei and M. Madritsch. Forms of differing degrees over number fields. *Mathematika* **63** (2017), pp. 92–123.

[79] W. Fulton. *Intersection theory*. Second. Vol. 2. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Springer-Verlag, Berlin, 1998, pp. xiv+470.

[80] W. Fulton and R. Lazarsfeld. Connectivity and its applications in algebraic geometry. *Lecture Notes in Math* **862** (1981), pp. 26–92.

[81] L. Gallardo. On the restricted Waring problem over $\mathbb{F}_{2^n}[t]$. *Acta Arithmetica* **92** (2000), pp. 109–113.

[82] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1994, pp. x+523.

[83] J. R. Getz. Secondary terms in asymptotics for the number of zeros of quadratic forms over number fields. *J. Lond. Math. Soc. (2)* **98** (2018), pp. 275–305.

[84] S. R. Ghorpade and G. Lachaud. Number of solutions of equations over finite fields and a conjecture of Lang and Weil. In: *Number theory and discrete mathematics (Chandigarh, 2000)*. Trends Math. Birkhäuser, Basel, 2002, pp. 269–291.

[85] J. Glas. Complete intersections of cubic and quadric hypersurfaces over $\mathbb{F}_q(t)$. *Preprint* (2023), arXiv: `2306.02718`.

[86] J. Glas. Canonical singularities on moduli spaces of rational curves via the circle method. *Preprint* (2024), arXiv: `2405.16648`.

[87] J. Glas and M. Hase-Liu. Terminal singularities on moduli spaces of curves. *To appear* (2024).

[88] J. Glas and L. Hochfilzer. On a question of Davenport and diagonal cubic forms over $\mathbb{F}_q(t)$. *Preprint* (2022), arXiv: `2208.05422`.

[89] J. Glas and L. Hochfilzer. Rational points on del Pezzo surfaces of low degree. *Preprint* (2024), arXiv: `2401.04759`.

[90] U. Görtz and T. Wedhorn. *Algebraic Geometry I: Schemes*. Springer, 2020.

[91] M. Greenberg and D. Polaski. *Lectures on Forms in Many Variables*. Math Lecture Notes Series. W. A. Benjamin, 1969.

[92] M. A. Hanselmann. *Rational points on quartic hypersurfaces*. PhD thesis. Ludwig-Maximilians-Universität München, 2012.

[93] G. H. Hardy and S. Ramanujan. Asymptotic formulaæ in combinatory analysis. *Proceedings of the London Mathematical Society* **2** (1918), pp. 75–115.

[94] J. Harris, M. Roth, and J. Starr. Rational curves on hypersurfaces of low degree. *J. Reine Angew. Math.* **571** (2004), pp. 73–106.

[95] M. Hase-Liu. A higher genus circle method and an application to geometric Manin's conjecture. *Preprint* (2024), arXiv: `2402.10498`.

[96] H. Hasse. Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen. *Journal für die Reine und Angewandte Mathematik* **152** (1923), pp. 129–148.

[97] H. Hasse. Über die Kongruenzzetafunktionen. *Unter Benutzung von Mitteilungen von Prof. Dr. F. K. Schmidt und Prof. Dr. E. Artin. Sitzungsber. Preuss. Akad. Wissenschaften* (1934).

[98] D. R. Heath-Brown. Cubic forms in ten variables. *Proceedings of the London Mathematical Society* **3** (1983), pp. 225–257.

[99] D. R. Heath-Brown. A multiple exponential sum to modulus p^2. *Canad. Math. Bull.* **28** (1985), pp. 394–396.

[100] D. R. Heath-Brown. A new form of the circle method, and its application to quadratic forms. *Journal für die reine und angewandte Mathematik* **1996** (1996), pp. 149–206.

[101] D. R. Heath-Brown. The density of rational points on cubic surfaces. *Acta Arith.* **79** (1997), pp. 17–30.

[102] D. R. Heath-Brown. Counting rational points on cubic surfaces. *Astérisque* **251** (1998), pp. 13–30.

[103] D. R. Heath-Brown. The circle method and diagonal cubic forms. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **356** (1998), pp. 673–699.

[104] D. R. Heath-Brown. The density of rational points on curves and surfaces. *Ann. of Math. (2)* **155** (2002), pp. 553–595.

[105] D. R. Heath-Brown. Cubic forms in 14 variables. *Invent. Math.* **170** (2007), pp. 199–230.

[106] D. Heath-Brown. The circle method and diagonal cubic forms. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **356** (1998), pp. 673–699.

[107] H. A. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.* **19** (2006), pp. 527–550.

[108] M. Hindry and J. H. Silverman. The canonical height and integral points on elliptic curves. *Invent. Math.* **93** (1988), pp. 419–450.

[109] L. Hochfilzer. Systems of bihomogeneous forms of small bidegree. *Acta Arith.* **213** (2024), pp. 325–368.

[110] A. Holme. The geometric and numerical properties of duality in projective algebraic geometry. *Manuscripta Math.* **61** (1988), pp. 145–162.

[111] C. Hooley. On Waring's problem. *Acta Mathematica* **157** (1986), pp. 49–97.

[112] C. Hooley. On nonary cubic forms. *J. Reine Angew. Math.* **386** (1988), pp. 32–98.

[113] C. Hooley. On nonary cubic forms. II. *J. Reine Angew. Math.* **415** (1991), pp. 95–165.

[114] C. Hooley. On nonary cubic forms. III. *J. Reine Angew. Math.* **456** (1994), pp. 53–63.

[115] C. Hooley. On nonary cubic forms: IV. *J. Reine Angew. Math.* **680** (2013), pp. 23–39.

[116] C. Hooley. On octonary cubic forms. *Proceedings of the London Mathematical Society* **109** (2014), pp. 241–281.

[117] L. Hua. *Additive theory of prime numbers*. eng. Vol. 13. Translations of mathematical monographs 13. XIII, 190 S. Providence, Rhode Island: American Mathematical Society, 1965.

[118] J. Igusa. *An introduction to the theory of local zeta functions*. Vol. 14. American Mathematical Soc., 2007.

[119] N. M. Katz. Estimates for "singular" exponential sums. *International Mathematics Research Notices* **1999** (1999), pp. 875–899.

[120] N. M. Katz and P. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*. Vol. 45. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 1999, pp. xii+419.

[121] B. Kim and R. Pandharipande. The connectedness of the moduli space of maps to homogeneous spaces. In: *Symplectic geometry and mirror symmetry (Seoul, 2000)*. World Sci. Publ., River Edge, NJ, 2001, pp. 187–201.

[122] S. L. Kleiman. Tangency and duality. In: *Proceedings of the 1984 Vancouver conference in algebraic geometry*. **6**. CMS Conf. Proc. Amer. Math. Soc., Providence, RI, 1986, pp. 163–225.

[123] H. D. Kloosterman. On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$. *Acta Math.* **49** (1927), pp. 407–464.

[124] J. Kollár. *Rational curves on algebraic varieties*. Vol. 32. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics. Springer-Verlag, Berlin, 1996, pp. viii+320.

[125] J. Kollár, Y. Miyaoka, and S. Mori. Rational connectedness and boundedness of Fano manifolds. *J. Differential Geom.* **36** (1992), pp. 765–779.

[126] M. Kontsevich. Enumeration of rational curves via torus actions. In: *The moduli space of curves (Texel Island, 1994)*. **129**. Progr. Math. Birkhäuser Boston, Boston, MA, 1995, pp. 335–368.

[127] R. M. Kubota. Waring's problem for $\mathbf{F}_q[x]$. *Dissertationes Math. (Rozprawy Mat.)* **117** (1974), p. 60.

[128] S. Lang. *Algebraic number theory*. Second. Vol. 110. Graduate Texts in Mathematics. Springer-Verlag, New York, 1994, pp. xiv+357.

[129] S. Lang and A. Weil. Number of points of varieties in finite fields. *Amer. J. Math.* **76** (1954), pp. 819–827.

[130] C. Le Rudulier. Points algébriques de hauteur bornée sur une surface. *Bull. Soc. Math. France* **147** (2019), pp. 705–748.

[131] S. A. Lee. *On the applications of the circle method to function fields, and related topics*. PhD thesis. University of Bristol, 2013.

[132] B. Lehmann, A. K. Sengupta, and S. Tanimoto. Geometric consistency of Manin's conjecture. *Compos. Math.* **158** (2022), pp. 1375–1427.

[133] B. Lehmann and S. Tanimoto. On exceptional sets in Manin's conjecture. *Res. Math. Sci.* **6** (2019), Paper No. 12, 41.

[134] M. Levin. On the Group of Rational Points on Elliptic Curves Over Function Fields. *American Journal of Mathematics* **90** (1968), pp. 456–462.

[135] R. Lidl and H. Niederreiter. *Finite Fields*. 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996.

[136] Y.-R. Liu and T. D. Wooley. Waring's problem in function fields. *Journal für die reine und angewandte Mathematik* **2010** (2010), pp. 1–67.

[137] Y. I. Manin. *Cubic forms*. Second. Vol. 4. North-Holland Mathematical Library. Algebra, geometry, arithmetic, Translated from the Russian by M. Hazewinkel. North-Holland Publishing Co., Amsterdam, 1986, pp. x+326.

[138] A. Mânzăţeanu. Counting points on $\mathrm{Hilb}^m\mathbb{P}^2$ over function fields. *Preprint* (2019), arXiv: 1905.04772.

[139] O. Marmon and P. Vishe. On the Hasse principle for quartic hypersurfaces. *Duke Math. J.* **168** (2019), pp. 2727–2799.

[140] J. V. Matijasevič. Diophantine sets. *Uspehi Mat. Nauk* **27** (1972), pp. 185–222.

[141] B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.* (1977). With an appendix by Mazur and M. Rapoport, pp. 33–186.

[142] R. B. McFeat. Geometry of numbers in adele spaces. *Dissertationes Math. (Rozprawy Mat.)* **88** (1971), p. 49.

[143] L. Merel. Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Invent. Math.* **124** (1996), pp. 437–449.

[144] J.-F. Mestre. Courbes elliptiques et formules explicites. In: *Seminar on number theory, Paris 1981–82 (Paris, 1981/1982)*. **38**. Progr. Math. Birkhäuser Boston, Boston, MA, 1983, pp. 179–187.

[145] J. Miller, P. Patzt, D. Petersen, and O. Randal-Williams. Uniform twisted homological stability. *Preprint* (2024), arXiv: 2402.00354.

[146] H. Minkowski. Mémoire sur la théorie des formes quadratiques à coefficients entiers.. *Preprint* French. Mém. Sav. Étr. (2) 29, No. 2. 180 p. (1887). (1887),

[147] Y. Miyaoka. On the Kodaira dimension of minimal threefolds. *Math. Ann.* **281** (1988), pp. 325–332.

[148] L. J. Mordell. On the rational solutions of the indeterminate equations of the third and fourth degrees. *Proc. Cambridge Philos. Soc.* **21** (1922/23), pp. 179–192.

[149] R. Munshi. Pairs of quadrics in 11 variables. *Compos. Math.* **151** (2015), pp. 1189–1214.

[150] M. Mustaţă. Jet schemes of locally complete intersection canonical singularities. *Invent. Math.* **145** (2001). With an appendix by David Eisenbud and Edward Frenkel, pp. 397–424.

[151] M. Mustaţă. The dimension of jet schemes of singular varieties. In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. II*. Kyung Moon Sa, Seoul, 2014, pp. 673–693.

[152] K. H. Nguyen. On a uniform bound for exponential sums modulo p^m for Deligne polynomials. *Preprint* (2021), arXiv: 2111.11898.

[153] M. Northey and P. Vishe. On the Hasse principle for complete intersections. *Preprint* (2024),

[154] C. F. Osgood. Errata: "Effective bounds on the "Diophantine approximation" of algebraic functions over fields of arbitrary characteristic and applications to differential equations" (Nederl. Akad. Wetensch. Proc. Ser. A **7**8 (1975), no. 2, 105–119). *Indag. Math.* **37** (1975). Nederl. Akad. Wetensch. Proc. Ser. A **7**8, p. 401.

[155] M. Paredes and R. Sasyk. Uniform bounds for the number of rational points on varieties over global fields. *Algebra Number Theory* **16** (2022), pp. 1941–2000.

[156] E. Peyre. Hauteurs et mesures de Tamagawa sur les variétés de Fano. *Duke Mathematical Journal* **79** (1995), pp. 101–218.

[157] E. Peyre. Points de hauteur bornée sur les variétés de drapeaux en caractéristique finie. *Acta Arith.* **152** (2012), pp. 185–216.

[158] E. Peyre. Liberté et accumulation. *Doc. Math.* **22** (2017), pp. 1615–1659.

[159] E. Peyre. Chapter V: Beyond heights: slopes and distribution of rational points. In: *Arakelov geometry and Diophantine applications*. **2276**. Lecture Notes in Math. Springer, Cham, 2021, pp. 215–279.

[160] B. Poonen and J. F. Voloch. Random Diophantine equations. In: *Arithmetic of higher-dimensional algebraic varieties (Palo Alto, CA, 2002)*. **226**. Progr. Math. With appendices by Jean-Louis Colliot-Thélène and Nicholas M. Katz. Birkhäuser Boston, Boston, MA, 2004, pp. 175–184.

[161] D. Ramakrishnan and R. J. Valenza. *Fourier analysis on number fields*. Vol. 186. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999, pp. xxii+350.

[162] I. Reiner. *Maximal orders*. Vol. 28. London Mathematical Society Monographs. New Series. Corrected reprint of the 1975 original, With a foreword by M. J. Taylor. The Clarendon Press, Oxford University Press, Oxford, 2003, pp. xiv+395.

[163] E. Riedl and D. Yang. Kontsevich spaces of rational curves on Fano hypersurfaces. *J. Reine Angew. Math.* **748** (2019), pp. 207–225.

[164] M. Rosen. *Number theory in function fields*. Vol. 210. Graduate Texts in Mathematics. Springer-Verlag, New York, 2002, pp. xii+358.

[165] S. L. Rydin Myerson. Systems of forms in many variables. *Preprint* (2017), arXiv: `1709.08917`.

[166] S. L. Rydin Myerson. Quadratic forms and systems of forms in many variables. *Inventiones Mathematicae* **213** (2018), pp. 205–235.

[167] S. L. Rydin Myerson. Systems of cubic forms in many variables. *Journal für die reine und angewandte Mathematik* **2019** (2019), pp. 309–328.

[168] P. Salberger. Uniform bounds for rational points on cubic hypersurfaces. In: *Arithmetic and geometry*. **420**. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 2015, pp. 401–421.

[169] P. Salberger. Counting rational points on projective varieties. *Proc. Lond. Math. Soc. (3)* **126** (2023), pp. 1092–1133.

[170] P. Samuel. *Lectures on old and new results on algebraic curves*. Vol. No. 36. Tata Institute of Fundamental Research Lectures on Mathematics. Notes by S. Anantharaman. Tata Institute of Fundamental Research, Bombay, 1966, pp. ii+127+iii.

[171] E. Sato. A criterion for uniruledness in positive characteristic. *Tohoku Math. J. (2)* **45** (1993), pp. 447–460.

[172] W. Sawin. Freeness alone is insufficient for Manin-Peyre. *Preprint* (2020), arXiv: `2001.06078`.

[173] W. Sawin and M. Shusterman. Möbius cancellation on polynomial sequences and the quadratic Bateman-Horn conjecture over function fields. *Invent. Math.* **229** (2022), pp. 751–927.

[174] W. Sawin and M. Shusterman. On the Chowla and twin primes conjectures over $\mathbb{F}_q[T]$. *Ann. of Math. (2)* **196** (2022), pp. 457–506.

[175] S. H. Schanuel. Heights in number fields. *Bull. Soc. Math. France* **107** (1979), pp. 433–449.

[176] D. Schindler. Bihomogeneous forms in many variables. *J. Théor. Nombres Bordeaux* **26** (2014), pp. 483–506.

[177] D. Schindler. A variant of Weyl's inequality for systems of forms and applications. In: *Advances in the theory of numbers*. **77**. Fields Inst. Commun. Fields Inst. Res. Math. Sci., Toronto, ON, 2015, pp. 207–218.

[178] D. Schindler and A. Skorobogatov. Norms as products of linear polynomials. *J. Lond. Math. Soc. (2)* **89** (2014), pp. 559–580.

[179] F. K. Schmidt. Analytische Zahlentheorie in Körpern der Charakteristik p. *Math. Z.* **33** (1931), pp. 1–32.

[180] W. M. Schmidt. The density of integer points on homogeneous varieties. *Acta Math.* **154** (1985), pp. 243–296.

[181] E. S. Selmer. The Diophantine equation $ax^3 + by^3 + cz^3 = 0$. *Acta Math.* **85** (1951), 203–362 (1 plate).

[182] J.-P. Serre. *Lectures on the Mordell-Weil theorem*. Ed. by M. Brown and M. Waldschmidt. Third. Aspects of Mathematics. With a foreword by Brown and Serre. Friedr. Vieweg & Sohn, Braunschweig, 1997, pp. x+218.

[183] J.-P. Serre. *Topics in Galois theory*. Second. Vol. 1. Research Notes in Mathematics. With notes by Henri Darmon. A K Peters, Ltd., Wellesley, MA, 2008, pp. xvi+120.

[184] T. Shioda. An example of unirational surfaces in characteristic p. *Math. Ann.* **211** (1974), pp. 233–236.

[185] J. H. Silverman. *The arithmetic of elliptic curves*. Vol. 106. Graduate Texts in Mathematics. Springer-Verlag, New York, 1986, pp. xii+400.

[186] C. M. Skinner. Forms over number fields and weak approximation. *Compositio Math.* **106** (1997), pp. 11–29.

[187] A. N. Skorobogatov. Beyond the Manin obstruction. *Invent. Math.* **135** (1999), pp. 399–424.

[188] A. N. Skorobogatov. Diagonal quartic surfaces. In: *Oberwolfach Reports*. **33**. 2009, pp. 76–79.

[189] L. M. Spreafico. Axiomatic theory for transversality and Bertini type theorems. *Archiv der Mathematik* **70** (1998), pp. 407–424.

[190] J. Starr. The Kodaira dimension of spaces of rational curves on low degree hypersurfaces. *Preprint* (2003), arXiv: `math/0305432`.

[191] M. Swarbrick Jones. A note on a theorem of Heath-Brown and Skorobogatov. *Q. J. Math.* **64** (2013), pp. 1239–1251.

[192] J. T. Tate Jr. *Fourier analysis in number fields and Hecke's zeta functions*. Thesis (Ph.D.)–Princeton University. ProQuest LLC, Ann Arbor, MI, 1950, p. 60.

[193] R. Taylor and A. Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)* **141** (1995), pp. 553–572.

[194] J. L. Thunder. Siegel's lemma for function fields. *Michigan Mathematical Journal* **42** (1995), pp. 147–162.

[195] Z. Tian. Hasse principle for three classes of varieties over global function fields. *Duke Mathematical Journal* **166** (2017), pp. 3349–3424.

[196] Z. Tian and L. Zhang. Weak approximation for cubic hypersurfaces and degree 4 del Pezzo surfaces. *International Mathematics Research Notices* **2018** (2018), pp. 762–784.

[197] T. H. Tran. Secondary terms in asymptotics for the number of zeros of quadratic forms. *Preprint* (2019), arXiv: `1910.14530`.

[198] Y. Tschinkel. *Arithmetic of Algebraic Surfaces*. PhD thesis. M.I.T., 1992.

[199] D. Ulmer. Elliptic curves over function fields. In: *Arithmetic of L-functions*. **18**. IAS/Park City Math. Ser. Amer. Math. Soc., Providence, RI, 2011, pp. 211–280.

[200] R. C. Vaughan. *The Hardy-Littlewood method*. Second. Vol. 125. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1997, pp. xiv+232.

[201] R. Vaughan. On Waring's problem for cubes. *Journal für die reine und angewandte Mathematik* **1986** (1986), pp. 122–170.

[202] P. Vishe. Rational points on complete intersections over $\mathbb{F}_q(t)$. *Proceedings of the London Mathematical Society* **126** (2023), pp. 556–619.

[203] V. Y. Wang. Diagonal cubic forms and the large sieve. *Preprint* (2021), arXiv: `2108.03395`.

[204] V. Y. Wang. Sums of cubes and the Ratios Conjectures. *Preprint* (2023), arXiv: `2108.03398`.

[205] V. Y. Wang. Special cubic zeros and the dual variety. *Preprint* (2024), arXiv: `2108.03396`.

[206] A. Weil. On the Riemann hypothesis in function fields. *Proc. Nat. Acad. Sci. U.S.A.* **27** (1941), pp. 345–347.

[207] A. Weil. *Basic number theory*. Classics in Mathematics. Reprint of the second (1973) edition. Springer-Verlag, Berlin, 1995, pp. xviii+315.

[208] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)* **141** (1995), pp. 443–551.

[209] T. D. Wooley. On simultaneous additive equations. II. *J. Reine Angew. Math.* **419** (1991), pp. 141–198.

[210] T. D. Wooley. On simultaneous additive equations, IV. *Mathematika* **45** (1998), pp. 319–335.

[211] T. D. Wooley. The cubic case of the main conjecture in Vinogradov's mean value theorem. *Adv. Math.* **294** (2016), pp. 532–561.

[212] T. D. Wooley. Nested efficient congruencing and relatives of Vinogradov's mean value theorem. *Proc. Lond. Math. Soc. (3)* **118** (2019), pp. 942–1016.

[213] S. Yamagishi. Birch's theorem on forms in many variables with a Hessian condition. *Preprint* (2023), arXiv: `2304.02620`.

[214] H. G. Zimmer. On the difference of the Weil height and the Néron-Tate height. *Math. Z.* **147** (1976), pp. 35–51.