LOCAL-GLOBAL PRINCIPLES FOR LINEAR SPACES ON HYPERSURFACES

Julia Brandes

School of Mathematics

May 2014

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science

Abstract

In this thesis we study various aspects of the problem of finding rational linear spaces on hypersurfaces. This problem can be approached by the Hardy–Littlewood circle method, establishing a Local-Global Principle provided that the hypersurface is 'sufficiently non-singular' and the number of variables is large enough. However, the special structure of the linear spaces allows us to obtain some improvement over previous approaches. A generalised version is also addressed, which allows us to count linear spaces under somewhat more flexible conditions.

We then investigate the local solubility. In particular, by adopting a new approach to the analysis of the density of p-adic solutions arising in applications of the circle method, we show that under modest conditions the existence of non-trivial p-adic solutions suffices to establish positivity of the singular series. This improves on earlier approaches due to Davenport, Schmidt and others, which require the existence of non-singular p-adic solutions.

Finally, we exhibit the strength of our methods by deriving unconditional results concerning the existence of linear spaces on systems of cubic and quintic equations.

Acknowledgements

I would like to thank my supervisor Trevor Wooley, to whom I am deeply indebted for his constant commitment and unwavering support, for his valuable advice not only in mathematical questions, and not least for providing guidance through some difficult periods, often spending a great amount of time and effort. With his patient and liberal sharing of ideas, but also his cheerfulness and easy-going attitude, he not only made this thesis possible, but also played an important part in making the process of compiling it thoroughly enjoyable.

I have had useful conversations about various mathematical ideas, among others with Tim Browning, Jörg Brüdern, Rainer Dietmann and Scott Parsell as well as my fellow PhD students at Bristol. Thanks in particular to Siu-lun Alan Lee for often lending a first ear to my ideas or questions.

I would also like to thank everybody in Howard House and in particular all my fellow postgraduates for creating a friendly and inspiring working environment and making the past three years an unforgettable experience and one of the happiest periods of my life. Without the active social life at the maths department, including the two weekly pub trips with the number theory group and with the graduate students, doing a PhD would not have been the same.

My special thanks go to my former flatmates Elena Couce and Wenwei Zhang for having become my second family in Bristol and for being just as crazy and fun as they are.

Finally, I would like to thank my family for supplying all the means tangible and intangible that are necessary for finding my way at university and in life and for their unfaltering and ever-caring support and encouragement.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Julia Brandes

Date: May 2014

Contents

Notation

Although we believe that most of our notation is self-explanatory, we would like to recall a few items that will be used throughout the thesis.

We will follow the usage in analytic number theory in writing $e(x) = e^{2\pi i x}$ for the exponential function. In a similar vein, whenever the letter ϵ occurs, the respective statement is true for all $\epsilon > 0$. We will therefore not trace the particular 'value' of each ϵ, which can consequently change from statement to statement. Furthermore, whenever we write $\sum_{n=a}^{b} f(n)$ where b is not an integer, the sum is to be understood to mean $\sum_{a \leq n \leq b} f(n)$.

We will abuse vector notation extensively, so any statement involving vectors should be read entrywise. In this vein we will write $|\mathbf{x}| \leq P$ to mean $|x_i| \leq P$ for all entries i. Similarly, the greatest common divisor $(a, \mathbf{x})$ is to be read as the greatest common divisor $(a, x_1, \ldots, x_n)$ of a and all entries of $\mathbf{x}$.

Finally, the Vinogradov and Landau symbols will be used throughout. Thus when for two functions $f(x)$ and $g(x)$ there exists a constant C such that $|f(x)| < C|g(x)|$ for all values x, we will write $f(x) = O(g(x))$ or, more concisely, $f(x) \ll g(x)$. If additionally $g(x) \ll f(x)$, this will be written as $f(x) \asymp g(x)$. Furthermore, we write $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to \infty$ and $f(x) \sim g(x)$ if $f(x)/g(x) \to 1$ as $x \to \infty$.

Chapter 1

Introduction

1.1 Linear spaces on hypersurfaces

In this thesis we consider a wide range of questions connected to the problem of finding linear spaces on hypersurfaces. Most generally, for a given set of forms $\psi^{(1)}, \ldots, \psi^{(R)} \in \mathbb{Z}[t_1, \ldots, t_m]$ of degree d, we are interested in representations of the shape

$$F^{(\rho)}(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m) = \psi^{(\rho)}(t_1, \ldots, t_m), \qquad 1 \leq \rho \leq R,$$

by general forms $F^{(1)}, \ldots, F^{(R)} \in \mathbb{Z}[x_1, \ldots, x_s]$ of the same degree. This is a generalisation of the classical problem of solving polynomial equations in the integers, which has been studied since antiquity but still remains of interest today, in particular since Matijasevič [45], using ideas of Robinson [55], ruled out any possibility of algorithmically determining whether a given equation has integer solutions. Nonetheless, research has continued to determine solubility for as large a class of polynomials as possible. One method that has proved particularly successful over the years is the Hardy–Littlewood circle method, which is a very versatile instrument relying on Fourier analysis. First developed in order to thoroughly understand a specific class of equations, it has been generalised to a wide range of polynomials and has not yet exhausted its potential.

In this thesis, we present an adaptation of the circle method to the linear spaces problem. It turns out that it is possible to capture the structure of the linear space and exploit it, thus gaining much better control over the problem than in earlier approaches. The main theorem can be modified so as to extend to spaces whose generators are subject to varying size restrictions, thus adding in flexibility and opening the door for a number of potential applications.

Two of the main obstacles in the implementation of the method are addressed separately. Firstly, the method relies on the fact that, if an equation is expected to be soluble over $\mathbb{Q}$, it should certainly have solutions in all fields containing $\mathbb{Q}$, but it has proved frustratingly difficult to show that this necessary condition is satisfied. We obtain some improvement in the treatment of this problem by using a geometric argument that allows us to exclude some difficult cases. Secondly, the circle method fails when the system of the $F^{(\rho)}$ is too singular, and even though methods have been developed to address this problem, they are either very costly or very technical and specific to certain special cases. However, with our improved bounds in the application of the circle method, we are able to make some headway in this direction.

1.2 History: Waring's Problem

One of the most lasting topics in mathematics, and simultaneously one of its most influential driving forces, is the theory of solving equations. The first attempts to methodically do so are lost in the mists of time, but there is evidence that all ancient civilisations have been developing algorithms to this purpose, and indeed the modern name of *diophantine equations* designating a polynomial equation over the integers goes back to the Greek mathematician Diophantus of Alexandria, author of the oldest surviving textbook[1] on the topic and known in popular science through his supposed epitaph encoding his

[1] The book owes its lasting importance not least to its most famous reader Pierre de Fermat, who postulated his famous Last Theorem in a marginal note in his copy of the *Arithmetica*.

age in a mathematical riddle[1].

Early modern age saw a new surge in the study of diophantine equations. In the late 18th century, Edward Waring [71, p. 349] formulated a series of mathematical statements, including the following.

Omnis integer numerus vel est cubus, vel e duobus, tribus, 4, 5, 6, 7, 8, vel novem cubis compositus, est etiam quadrato-quadratus vel e duobus, tribus, &c. usque ad novemdecim compositus, & sic deinceps.

Every natural number is either a cube or composed of two, three, 4, 5, 6, 7, 8, or nine cubes; it is also a biquadrate or composed of two, three, &c. up to nineteen biquadrates, & so forth.[2]

The original list does not contain any proofs, but whereas proofs for all the other claims have been subsequently supplied, this one has proved elusive. In fact, it was not until 1909 that Hilbert [33] succeeded in establishing that the number of k-th powers needed in order to write any natural number n as a sum of such powers is finite. The argument he used leads to an existence proof, but fails to supply satisfactory upper bounds on the number of k-th powers required, so the focus turned towards the challenge of obtaining a quantitative understanding of the problem.

Definition 1.1. *The number $g(k)$ is defined to be the smallest integer g with the property that for every natural number n and for every $s \geq g$ the equation*

$$n = x_1^k + x_2^k + \ldots + x_s^k \tag{1.2.1}$$

has a solution $(x_1, \ldots, x_s)$ in the non-negative integers.

The values provided by Waring have subsequently been confirmed. In 1908, Wieferich [72] proved[3] that $g(3) = 9$, and in 1986 Balasubramanian, Deshouillers and Dress [4, 5] confirmed that $g(4) = 19$. For the sake of complete-

[1] The original version is recorded in the Anthologia Graeca [53, 14.126].

[2] Translation by the author.

[3] A gap in his proof was filled by Kempner [38].

ness we mention that $g(2) = 4$ is Lagrange's Theorem and has been known since 1770. More in general, a fairly straightforward argument shows that

$$g(k) \geq \left[\left(\frac{3}{2} \right)^k \right] + 2^k - 2, \tag{1.2.2}$$

and one probably has equality (see [68, Chapter 1] for the argument and references).

However, it became also apparent that these values are attained only a finite number of times. In the case of $k = 3$ this was pointed out by Landau [39] as an immediate reaction to Wieferich's work, and eventually Linnik [43] succeeded to show that all but finitely many integers can be written as a sum of at most seven cubes. Similar results have been established for $k = 4$. This phenomenon seems to indicate that the large bounds on $g(k)$ given in (1.2.2) arise from size restrictions rather than the actual number theoretic structure of the problem and reflect the fact that, if n is very small, the x_i can take only a very limited number of values, thus imposing a large measure of rigidity on the problem. It is therefore more natural to ignore those anomalies and instead consider the number of k-th powers needed to represent every large enough integer.

Definition 1.2. *The number $G(k)$ is defined to be the smallest integer g with the property that there exists an integer N such that for every natural number $n \geq N$ and for every $s \geq g$ the equation* (1.2.1) *has a solution $(x_1, \ldots, x_s)$ in the non-negative integers.*

Obviously we have $G(k) \leq g(k)$, while elementary methods yield a lower bound of $G(k) \geq k + 1$. It may be noteworthy that, whereas the quantity $g(k)$ is by now almost completely understood (see [68, Chapter 1]), the only k for which exact values of $G(k)$ have been established are $G(2) = 4$ by Lagrange's Theorem and $G(4) = 16$ due to Davenport [12]; even for $k = 3$ all that is currently known is $4 \leq G(3) \leq 7$.

Regarding higher powers, the most efficient tool is the circle method, which gives estimates on the number $R_{s,k}(n)$ of solutions of (1.2.1). This was in its original shape first conceived in 1916 by Hardy and Ramanujan [30] and applied to Waring's problem a few years later by Hardy and Littlewood [28]; in its modern shape it goes back to a more tractable reformulation due to Vinogradov [70] from 1928. Notice that, since all numbers involved are positive, the x_i cannot exceed $P = [n^{1/k}]$ if we want (1.2.1) to hold. On the other hand, if all the variables $x_1, \ldots, x_s$ are constrained within some box $[1, P]^s$, one has P^s possible inputs, while the values that can be taken lie all inside the interval $[s, s \cdot P^k]$, so every integer n inside this interval should on average be hit roughly P^{s-k} times. Thus one would expect an asymptotic formula of the shape

$$R_{s,k}(n) \asymp P^{s-k} \asymp n^{s/k-1}. \tag{1.2.3}$$

We let $\widetilde{G}(k)$ be the least number of variables required so that (1.2.3) holds. Asymptotic formulae of the shape (1.2.3) have first been established by Hardy and Littlewood in 1920 [28], and they supplied the bound $\widetilde{G}(k) \leq (k-2)2^{k-1}+5$ in subsequent work [29]. This result has been subject to numerous improvements over the years (for the history we refer to [69]), of which we mention the classical bound of $\widetilde{G}(k) \leq 2^k + 1$ due to Hua [37], which features in most textbooks on the circle method, and the very latest result by Wooley [79, Cor. 1.7], which establishes $\widetilde{G}(k) \leq 2k^2 - 2k - 12$.

For the sake of completeness we remark that better results for $G(k)$ can be obtained by counting only those solutions that are particularly easy to count, thus bounding $R_{s,k}(n)$ away from zero. This idea has been pursued in various guises, and the derived bounds are generally of a magnitude of $G(k) \ll k \log k$ with gradual improvements on the implied constant and the error terms. Again we do not give details of the history and refer instead to sources such as Vaughan's and Wooley's survey article [69] or the extensive bibliography in Vaughan's book [68], but the best bound currently known is given by

$$G(k) \leq k \left(\log k + \log \log k + 2 + O\left(\frac{\log \log k}{\log k} \right) \right).$$

and is due to Wooley [73, Theorem 1.4].

1.3 Generalisations of Waring's problem (1): Forms in many variables

It is a natural question to ask whether results similar to those described above generalise to more general polynomials, or indeed whether there are statements one can make that apply to the widest possible range of polynomials and systems of polynomial equations. For homogeneous problems, the question is therefore under what conditions one can show that the number of zeros obeys an asymptotic formula similar to the one given in (1.2.3). We remark at this point that we will follow the convention and denote the degree by the letter d in the setting of general forms, whereas in the situation of diagonal forms it is traditionally denoted by k.

A major difference between Waring's problem and the treatment of the general situation is that, whereas the Waring equation is non-singular, a general form may be highly singular, and this has an effect on the possible number of solutions, as is easily seen by considering a system of forms that are linearly dependent, or also degenerate forms such as $x_1(x_2^{d-1}+\ldots+x_s^{d-1})$. This indicates that it may be useful to exclude cases that are too singular.

Definition 1.3. *The* relative number of variables s^* *of a system* $F^{(1)},\ldots,F^{(R)}$ *of forms is defined as the difference between the absolute number* s *of variables and the dimension of the* singular locus[1] *associated to* $\mathbf{F}$ *which is given by*

$$\mathfrak{V}=\left\{\mathbf{x}\in\mathbb{C}^s\colon \operatorname{rank}\left(\frac{\partial F^{(\rho)}(\mathbf{x})}{\partial x_i}\right)_{\substack{1\le i\le s\\ 1\le\rho\le R}}\le R-1\right\}.$$

The following is a theorem by Birch [8], following in the wake of work on cubic forms by Davenport [13–15]. Birch's Theorem has later been generalised

[1] *Our definition of a singular locus is that of Birch [8], contrary to the usage in algebraic geometry, where the singular locus is given by* $\operatorname{Sing}_{\text{alg-geom}}(\mathbf{F})=\mathfrak{V}\cap\{\mathbf{F}(\mathbf{x})=\mathbf{0}\}$. *It is therefore better to think of* $\mathfrak{V}$ *as the set of* potentially *singular points, so that the singular locus as defined in algebraic geometry contains all those potentially singular points that actually lie on the variety. Note that the two notions coincide when* $R=1$, *and in general one has* $\dim\mathfrak{V}\le\dim\operatorname{Sing}_{\text{alg-geom}}(\mathbf{F})+R-1$.

by Schmidt [62]. Consider an R-tuple of forms $F^{(1)}, \dots, F^{(R)} \in \mathbb{Z}[x_1, \dots, x_s]$ of degree d and denote by $N_{s,R}^{(d)}(P)$ the number of $|\mathbf{x}| \leq P$ that solve the simultaneous equations $F^{(1)}(\mathbf{x}) = \ldots = F^{(R)}(\mathbf{x}) = 0$.

Birch's Theorem. *Let $F^{(1)}, \dots, F^{(R)}$ be a set of forms of equal degree d in s variables, and suppose that*

$$s^* > 2^{d-1}(d-1)R(R+1).$$

Then the number $N_{s,R}^{(d)}(P)$ of simultaneous zeros of $F^{(1)}, \dots, F^{(R)}$ contained in $[-P, P]^s$ is given by

$$N_{s,R}^{(d)}(P) = P^{s-Rd}\chi_\infty \prod_{p \text{ prime}} \chi_p + O(P^{s-Rd-\delta})$$

for some $\delta > 0$, where the constants χ_∞ and χ_p are non-negative and encode the real and p-adic solubility, respectively.

The number of variables required in this theorem is roughly on par with the classical approach to Waring's problem by the circle method (see Hua [37]), and indeed all of the later work on diagonal equations such as Waring's problem depends crucially on the very simple shape of the problem and will therefore not be applicable in more general situations. Here, on the contrary, we are confronted with a plethora of possible polynomials, and indeed the strength of Birch's theorem is the wide range of its applicability, even allowing for small singularities. Note that for highly singular polynomials the statement of the theorem is not true.

One should notice that there are some improvements to Birch's treatment. In 1983, Heath-Brown [31] was able to prove that every non-singular cubic form in at least ten variables has a non-trivial zero. Ten variables is also the barrier required to ensure that a cubic form has solutions in all p-adic fields, as has been shown by Lewis [40]. Nonetheless, Heath-Brown's result has been further improved by Hooley [34], who could show that any non-singular form in at least 9 variables, now subject to local solubility, possesses a non-trivial zero, and who very recently [35] was able to lower the barrier even to eight, conditionally on a suitable form of the Riemann hypothesis for Hasse–Weil L-functions. This is quite a remarkable achievement in the light of the fact that even in the diagonal case we still need at least seven variables to represent zero (Baker [3, Theorem 1]).

Since Birch's Theorem is very central to this thesis, we try to describe some key elements of its proof. The method used is a version of the Hardy–Littlewood circle method and has been adapted to more general polynomials by Davenport [13–15] and Birch [8]. The method hinges on the fundamental Fourier-theoretic identity that for all intergers x one has

$$\int_0^1 e^{2\pi i\alpha x}\mathrm{d}\alpha = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{else.} \end{cases}$$

Replacing x by a polynomial $F(\mathbf{x})$ yields an indicator function on the solutions $\mathbf{x}$ of the equation $F(x_1, \dots, x_s) = 0$, and by summing over all possible inputs $\mathbf{x}$ in a given box $[-P, P]^s$ one obtains an expression for the number of solutions given by

$$N_{s,1}^{(d)}(P) = \sum_{-P\leq\mathbf{x}\leq P} \int_0^1 e\left(\alpha F(x_1, \dots, x_s)\right) \mathrm{d}\alpha. \tag{1.3.1}$$

By writing

$$T(\alpha) = \sum_{-P\leq\mathbf{x}\leq P} e\left(\alpha F(x_1, \dots, x_s)\right),$$

Equation (1.3.1) takes the shape

$$N_{s,1}^{(d)}(P) = \int_0^1 T(\alpha)\mathrm{d}\alpha.$$

The strategy is now to understand the size of $T(\alpha)$ for various values of α. Notice that $e(\alpha F(\mathbf{x})) = 1$ whenever $\alpha F(\mathbf{x})$ is an integer, so if α is a rational number with small denominator q, say, the value of $F(\mathbf{x})$ will be divisible by q for fairly many (roughly P^s/q) values $x_1, \dots, x_s$, so the size of $|T(\alpha)|$ will be comparatively large. On the other hand, as the denominator of α increases, the values of $e(\alpha F(\mathbf{x}))$ get more and more equidistributed on the unit circle, and one expects a lot of cancellation to happen as one averages over all $\mathbf{x}$ in the box $[-P,P]^s$. By consequence, the contribution from these α should be negligible. This dichotomy can be exploited by dividing the unit interval into two sets that respect the different behaviour of $T(\alpha)$ with respect to α, namely the *major arcs* $\mathfrak{M}$ of all α close to a rational number with a small denominator, and the *minor arcs* $\mathfrak{m}$ where $|T(\alpha)|$ is small. This allows us to write

$$N(P) = \int_{\mathfrak{M}} T(\alpha)\mathrm{d}\alpha + \int_{\mathfrak{m}} T(\alpha)\mathrm{d}\alpha, \tag{1.3.2}$$

where

$$\left| \int_{\mathfrak{m}} T(\alpha)\mathrm{d}\alpha \right| \ll \int_{\mathfrak{m}} |T(\alpha)|\,\mathrm{d}\alpha \ll \sup_{\alpha\in\mathfrak{m}} |T(\alpha)|$$

is expected to be small, while on the major arcs one should have an approximation $\alpha = a/q + \beta$ which is well-controlled.

Making this idea explicit is not easy. By Cauchy–Schwarz, one has

$$|T(\alpha)|^2 \leq \sum_{-P\leq \mathbf{x}\leq P} \left| \sum_{\mathbf{h}} e\big(\alpha\left(F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x})\right)\big) \right|,$$

where the term in the exponential sum is now only of degree $d-1$ in $\mathbf{x}$. Iterating this procedure and applying the Cauchy–Schwarz inequality gradually reduces the degree, until eventually the remaining exponential sum is linear in $\mathbf{x}$. Linear exponential sums, however, are bounded non-trivially only if the linear forms do not vanish and the coefficient α is not close to a rational number with a small denominator, and in this case, the quality of the upper bound is related directly to the quality of the approximations to α. Hence, if we suppose that $|T(\alpha)|$ is large, this forces α to satisfy some diophantine approximation condition. If, on the other hand, the linear form vanishes identically, that implies that the original form F has been highly singular from the beginning. Altogether one obtains the following case disctinction.

Lemma 1.1 (Tripartite Weyl's inequality, first version)**.** *For any choice of parameters k and $\theta \in (0,1]$ there are three possibilities.*

(A) The exponential sum is bounded above by $|T(\alpha)| \ll P^{s-k\theta}$*, or*

(B) there are integers q *and* a *satisfying*

$$q \ll P^{(d-1)\theta} \quad \text{and} \quad \left|\alpha - \frac{a}{q}\right| \ll \frac{P^{-d+(d-1)\theta}}{q}, \qquad \text{or}$$

(C) the singular locus associated to F *has dimension at least* $s - 2^{d-1}k$*.*

The first case here corresponds to a minor arcs estimate, the second one describes a diophantine approximation defining a set of major arcs $\mathfrak{M}(\theta)$, while the last one occurs only if the system is highly singular and has to be excluded by demanding that $s^* > 2^{d-1}k$. One can see that there is a tradeoff here in the optimal choice of k and θ. It is necessary to choose θ small enough to allow a good control of the error term on the major arcs, while k should be as small as possible so as to not impose too harsh a condition on the number of variables. Meanwhile, one needs $k\theta > d$ so that the contribution from the minor arcs is smaller than the expected main term. It turns out, however, that if k is at least as large as $2(d-1)$, one can to some large extent avoid this problem by successively pruning down the major arcs, which enables us to choose θ as small as we like. As a consequence one has the following reformulation of the tripartite Weyl inequality.

Lemma 1.2 (Tripartite Weyl's inequality, second version)**.** *Let* k *be a parameter satisfying* $k > 2(d-1)$*, and suppose* $s^* > 2^{d-1}k$*. Then for any* $0 < \theta < d/(2(d-1))$ *we may take the major arcs to be the set of all* α *with a rational approximation* a/q *satisfying*

$$q \ll P^{(d-1)\theta} \quad \text{and} \quad \left|\alpha - \frac{a}{q}\right| \ll \frac{P^{-d+(d-1)\theta}}{q},$$

and the minor arcs contribution is given by

$$\int_{\mathfrak{m}} |T(\alpha)| \mathrm{d}\alpha \ll P^{s-d-\delta}$$

for some $\delta > 0$ *depending on* k *and* θ*.*

As a consequence of (1.3.2) and Lemma 1.2 we have

$$N_{s,1}^{(d)}(P) = \int_{\mathfrak{M}} T(\alpha)\mathrm{d}\alpha + o(P^{s-d-\delta}).$$

It therefore remains to analyse the major arcs. By writing $\alpha = a/q + \beta$, one can replace

$$\begin{aligned} T\left(\frac{a}{q} + \beta\right) &\sim q^{-1} \sum_{\mathbf{x}=1}^{q} e\left(\frac{aF(\mathbf{x})}{q}\right) \times \int_{[-P,P]^s} e(\beta F(\boldsymbol{\zeta}))\mathrm{d}\boldsymbol{\zeta} \\ &= q^{-1} S_q(a) \times v(P, \beta) \end{aligned}$$

with an error depending on θ that is acceptable if θ has been chosen sufficiently small. Thus the discrete and continuous contributions can be separated and one has

$$\int_{\mathfrak{M}} T(\alpha)\mathrm{d}\alpha \sim \sum_{q=1}^{P^{(d-1)\theta}} q^{-1} \sum_{\substack{a=1 \\ (a,p)=1}}^{q} S_q(a) \int_{-q^{-1}P^{-d+(d-1)\theta}}^{q^{-1}P^{-d+(d-1)\theta}} v(P, \beta)\mathrm{d}\beta.$$

One can show that for $k > 2(d-1)$, the integral can be extended to the entire real line, thus yielding the *singular integral* $\mathfrak{J}(P)$. Exploiting the homogeneity of F, integration by parts yields

$$\int_{\mathbb{R}} v(P, \beta)\mathrm{d}\beta = P^{s-d} \int_{\mathbb{R}} v(1, \beta)\mathrm{d}\beta,$$

and one observes that

$$\int_{\mathbb{R}} v(1, \beta)\mathrm{d}\beta = \int_{\mathbb{R}} \int_{[-1,1]^s} e(\beta F(\boldsymbol{\zeta}))\mathrm{d}\boldsymbol{\zeta}\mathrm{d}\beta$$

is a (rescaled) continuous version of (1.3.1) which can be interpreted as the density of solutions to $F(\mathbf{x}) = 0$ in the real unit box.

Similarly, the sum in the first term converges for $k > 2(d-1)$, and its completion is called the *singular series* and denoted by $\mathfrak{S}$. Furthermore, it is not hard to show that S_q is multiplicative in q, which allows us to define constants χ_p by writing

$$\mathfrak{S} = \sum_{q=1}^{\infty} q^{-1} S_q = \prod_{p \text{ prime}} \sum_{l=0}^{\infty} p^{-l} S_{p^l} = \prod_{p \text{ prime}} \chi_p.$$

Since

$$\sum_{l=1}^{L} p^{-l} \sum_{\substack{a=1 \\ (a,p)=1}}^{q} S_{p^l}(a) = \sum_{l=1}^{L} \sum_{\mathbf{x}=1}^{p^l} p^{-l} \sum_{\substack{a=1 \\ (a,p)=1}}^{p^l} e\left(\frac{aF(\mathbf{x})}{p^l}\right) = \sum_{\mathbf{x}=1}^{p^L} p^{-L} \sum_{a=1}^{p^L} e\left(\frac{aF(\mathbf{x})}{p^L}\right)$$

is the discrete analogue of the counting function in (1.3.1) modulo p^L, one sees that for any p the quantity χ_p is going to be greater than zero only if the equation $F(x_1, \ldots, x_s) = 0$ has a positive density of solutions in the p-adic numbers. Hence altogether one obtains

$$N_{s,1}^{(d)}(P) P^{s-d} \chi_\infty \prod_{p \text{ prime}} \chi_p + O(P^{s-d-\delta}),$$

provided that $s^* > 2^d(d-1)$. This proves Birch's Theorem in the case $R = 1$.

It should be remarked that the method is nigh identical if the system to be solved is of the shape $F(\mathbf{x}) = n$ for some integers n, as the solution n will occur only in the final shape of the local densities $\chi_\infty(n)$ and $\chi_p(n)$.

1.4 Generalisations of Waring's problem (2): The multidimensional problem

Another possible generalisation of Waring's problem is obtained by replacing the variables x_i by linear forms

$$L_i(\mathbf{t}) = x_{i,1}t_1 + x_{i,2}t_2 + \ldots + x_{i,m}t_m$$

and the integer n by a homogeneous polynomial $\psi(\mathbf{t})$ of degree k. In this case one seeks solutions $L_1, \ldots, L_s$ to the equation

$$(L_1(\mathbf{t}))^k + \ldots + (L_s(\mathbf{t}))^k = \psi(\mathbf{t}) \tag{1.4.1}$$

in $(\mathbb{Z}[t_1, \ldots, t_m])^s$. Notice that for $m = 1$ this setup reduces to the traditional version of Waring's problem.

In a more general setting over $\mathbb{C}$, this problem can be addressed with methods from algebraic geometry and has almost completely been solved by Alexander and Hirschowitz [1]. In particular, they proved that with a small number of known and well-understood exceptions, every homogeneous polynomial $\psi(t_1, \dots, t_m)$ that can be related to a set of points in general position (see [46] for details) has the expected number of representations as a sum of s powers of linear polynomials, provided that

$$s \geq \frac{1}{m}\binom{m-1+k}{k},$$

and this bound is sharp. Less is known for those forms that do not fulfil the stated generality condition, and finding a lower bound for s that applies to both general and exceptional polynomials is still an unsolved problem even in the complex setting (see the discussion in the introduction of [54], for instance). For our purposes, these results are of interest inasmuch they purvey inherited upper bounds for the number of representations that exist over $\mathbb{Q}$ as a subfield of $\mathbb{C}$, but since the particular structure of $\mathbb{Q}$ is forfeited by the embedding into the complex numbers, the Alexander–Hirschowitz Theorem is unfit to deliver any real number-theoretic information.

By expanding Equation (1.4.1) and equating coefficients of $t_1, \dots, t_m$, one sees that every linear form solving the equation translates bijectively to a point solution $\mathbf{x}_1, \dots, \mathbf{x}_m \in [-P, P]^{ms}$ of a system of r equations, where

$$r = \binom{m+k-1}{k} \tag{1.4.2}$$

is the number of monomials in ψ. This puts us into a situation where the methods for systems of forms outlined in the previous section are largely applicable, and since the problem has a total of ms variables and r equations of degree k, one expects the number of solutions of (1.4.1) to be $\sim cP^{ms-rk}$ with some constant c accounting for the local solution densities. This problem has been studied extensively by Parsell [47–51], who was able to derive a formula of the expected shape for the number of solutions of (1.4.1) provided the number of variables is large enough and derived lower bounds in other cases; the strongest result currently available establishes an asymptotic formula under the condition

$$s \geq 2\binom{k+m}{m}(k+1) - 2k - 1$$

and is due to Parsell, Prendiville, and Wooley [52, Theorem 1.4].

As is the case with Waring's Problem, the methods employed in [52] rely strongly on the comparatively simple geometry of the problem, so it would be of interest to investigate the behaviour of the multidimensional version of Birch's Theorem where these more efficient methods break down. This generalised problem has so far been studied only in the quadratic case, where matrix algebra, modular forms and dynamical systems provide a different set of methods, and the only attempt to tackle this question by the circle method is a very recent paper by Dietmann and Harvey [23]. Our first result gives a formula for the number of representations of a homogeneous polynomial by another homogeneous polynomial, and is the special case of a more general theorem concerning systems of multiple equations.

Theorem 1.1. *Let $F \in \mathbb{Z}[x_1, \ldots, x_s]$ and $\psi \in \mathbb{Z}[t_1, \ldots, t_m]$ be homogeneous polynomials of degree $d \geq 2$ and suppose that $m \geq 2$. Further, let*

$$s^* > 3 \cdot 2^{d-1}(d-1)(r+1).$$

Then there exist nonnegative constants $\chi_\infty(\psi)$ and $\chi_p(\psi)$ for every prime p such that the number $N_{s,\psi}(P)$ of solutions $\mathbf{x}_1, \ldots, \mathbf{x}_m \in [-P, P]^{ms}$ of

$$F(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) = \psi(t_1, \ldots, t_m) \tag{1.4.3}$$

identically in $t_1, \ldots, t_m$ is given by

$$N_{s,\psi}(P) = P^{ms-rd}\chi_\infty(\psi) \prod_{p \text{ prime}} \chi_p(\psi) + o\left(P^{ms-rd}\right).$$

This theorem is the main result of a paper [10] and will be explained in detail in Chapter 2. One remarks that the number of variables required bears some resemblance to what was obtained in Birch's Theorem. This is to be expected, as we follow his methods to arrive at the conclusion of Theorem 1.1. On closer inspection, however, a strict analogy would lead to a condition more of the shape

$$(ms)^* > 2^{d-1}(d-1)r(r+1),$$

where $(ms)^*$ is the difference between ms and the dimension of the singular locus of the expanded system of r equations, and indeed, this is roughly what Dietmann and Harvey [23] obtain in the quadratic case. In our Theorem 1.1, however, one sees that one factor of r on the right hand side has been replaced by the absolute constant 3, so Theorem 1.1 does somewhat better than one would originally expect. In particular, in the case $d = 2$ we obtain a bound of the magnitude $s \gg m^2$, which is significantly stronger than the recent result of Dietmann and Harvey [23] which has $s \gg m^4$, based on the older methods. The reason for this lies in the fact that when we expand Equation (1.4.3) in order to reduce the problem of finding lines on a hypersurface to that of counting point solutions of a system of equations, each pointwise equation of the expanded system carries some of the structural information of the original polynomial equation it is derived from, and consequently all r pointwise equations derived from the same polynomial equation are structurally very similar. This will cause the expanded system of r pointwise equations to collapse at some point during the proof, and it is here where we are able to save the factor $r \sim m^d$. This simplification, however, in turn creates some technical complications which give rise to the factor 3; this factor has no deeper justification and it would be desirable in the future to find a way to avoid this.

The most important special case of Theorem 1.1 is that of counting linear spaces on hypersurfaces and corresponds to the case of representing the zero polynomial. In the light of possible applications that may be of interest in the context of algebraic geometry, one can ask whether statements of this kind can be generalised so as to allow different size constraints for the generators $\mathbf{x}_1, \dots, \mathbf{x}_m$ of the linear space. We are therefore interested in counting the number $N(P_1, \dots, P_m)$ of integer vectors $\mathbf{x}_1, \dots, \mathbf{x}_m$ that solve the equation

$$F(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) = 0 \tag{1.4.4}$$

identically in t, where the variables are subject to size constraints of the shape $|\mathbf{x}_i| \leq P_i$ for all i. It turns out that the same methods that are used to prove Theorem 1.1 continue to be applicable, albeit with some minor modifications.

Theorem 1.2. *Let $d \geq 2$ and $m \geq 2$ be positive integers and $P_1 \leq \ldots \leq P_m$ large, and let*

$$s^* > 2^{d-1} \max \left\{ 3(d-1)(r+1), rd \left(\frac{\log P_m}{\log P_1} \right) \right\}.$$

Then the number $N(P_1, \dots, P_m)$ of solutions $|\mathbf{x}_i| \leq P_i$ of (1.4.4) *identically in $t_1, \dots, t_m$ is given by*

$$N(P_1, \dots, P_m) = \left(\prod_{i=1}^m P_i \right)^{s-rd/m} \chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(\left(\prod_{i=1}^m P_i \right)^{s-rd/m} \right),$$

where the constants χ_∞ and χ_p are those corresponding to the case $\psi = 0$ in Theorem 1.1.

A similar result for the case $d = 2$ is implicit in Dietmann and Harvey's work on representing quadratic forms by quadratic forms [23]. Theorem 1.2 will be proved in Chapter 3, where we will also address a modification of the result which counts solutions to (1.4.4) with the $\mathbf{x}_i$ contained in certain sublattices of $\mathbb{Z}^s$.

1.5 The local solubility problem

Birch's Theorem and its multidimensional version in Theorem 1.1 are instances of the Hasse principle, as they establish asymptotic behaviour for a counting problem under the condition that the equations in question be locally soluble. In fact, since $\mathbb{Q}$ is a subset of both $\mathbb{R}$ and $\mathbb{Q}_p$ for all primes p, rational zeros can exist only if the equation is soluble over the real and the p-adic numbers. The opposite implication is, however, not true, and there are known counter-examples (e.g. $3x^3 + 4y^3 + 5z^3 = 0$ due to Selmer [64]; see also the extensive research on the Brauer–Manin obstruction), so proving the validity of the Hasse principle is a necessary first step in proving that a form represents zero. On the other hand, this means that in order to obtain actual information about whether a counting problem is indeed solvable in the integers, it is indispensable to understand the local solubilities.

Definition 1.4. *We write $\gamma_d^p(R, m)$ for the least integer γ such that every set of R homogeneous polynomials with integer coefficients of degree d in $s \geq \gamma$ variables contains an m-dimensional linear space in the p-adic numbers.*

For instance, we have $\gamma_3^p(R, 1) \leq 5300R(3R+1)^2$ due to Schmidt [59], and

$$\gamma_d^p(R, m) \leq (R^2d^2 + mR)^{2^{d-2}} d^{2^{d-1}} \tag{1.5.1}$$

due to Wooley [77, Theorem 2.4].

In this notation we have information about the local solubility constants χ_∞ and χ_p.

Lemma 1.3. *One has*

(i) $\chi_\infty > 0$ *whenever the degree d is odd and* $s^* \geq 3 \cdot 2^{d-1}(d-1)R(R+1)$, *and*

(ii) $\chi_p > 0$ *provided that* $s^* > 2^{d-1}(d-1)R\gamma_d^p(R, m)$.

The case $m = 1$ is essentially due to Schmidt (see the corollary of Proposition I in [62]), and even though the generalisations to higher-dimensional spaces are fairly straightforward, we will give a complete account of the argument in the attempt to keep this thesis as self-contained as possible. Together with (1.5.1), this gives us asymptotic estimates on the number of rational linear spaces on hypersurfaces.

Theorem 1.3. *Let $F^{(1)}, \ldots, F^{(R)} \in \mathbb{Z}[x_1, \ldots, x_s]$ be forms of equal odd degree $d \geq 3$, and $m \geq 2$ an integer, and let r be as in* (1.4.2)*. Furthermore, suppose that*

$$s^* > 3 \cdot 2^{d-1}(d-1)d^{2^{d-1}} R \max\left\{Rr + 1, \left(R^2 d^2 + Rm\right)^{2^{d-2}}\right\}.$$

Then the number $N^{(d)}_{s,R,m}(P)$ of solutions $\mathbf{x}_1, \ldots, \mathbf{x}_m \in [-P, P]^{ms}$ of

$$F^{(\rho)}(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) = 0 \qquad (1 \leq \rho \leq R)$$

identically in $t_1, \ldots, t_m$ is given by

$$N^{(d)}_{s,R,m}(P) = P^{ms-Rrd} \chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(P^{ms-Rrd}\right),$$

and the product of the local densities $\chi_\infty \prod_p \chi_p$ is positive.

In order to put this result into context, the most relevant seems to be the work of Dietmann [21], who was the first to establish polynomial growth in m for the number of variables necessary in order to guarantee the existence of an affine m-space on a single form F, provided that F is non-singular. Apart from imposing a looser nonsingularity condition, our Theorem 1.3 supersedes Dietmann's bound of

$$s \geq 2^{5+2^{d-1}d} d! d^{2^d+1} m^{d(1+2^{d-1})}$$

by a power of $2d$ in m, due mainly to our more careful perusal of Schmidt's methods [62].

It is also obvious from these results that, whereas the real solubility condition can easily be bypassed by restricting ourselves to equations of odd degree, the p-adic situation is much harder to control, so that the number of variables required to guarantee positivity of the χ_p is, as things stand, bound to dominate the overall number of variables. It is therefore of great interest to obtain an improved treatment of the p-adic solubility of the problem. It turns out that one essential weakness of all treatments so far is that they fail to exploit our knowledge of p-adic solubility in the optimal way. This is due to the fact that the most efficient treatment of the singular series makes use of Hensel's Lemma in order to obtain the expected density of p-adic solutions from a non-singular solution modulo some power of p, which then implies that the singular series will be positive as soon as there exists at least one non-singular p-adic solution. Unfortunately, our knowledge concerning p-adic solubility deals only with the existence of non-trivial solutions without addressing the question of whether or not these solutions are non-singular. In Chapter 4 we develop a geometric approach for making the transition from non-trivial to non-singular p-adic points at the cost of inflating the number of variables by no more than what is required by the geometry of the problem. One of the main theorems that will be proved in the course of this thesis is the following.

Theorem 1.4. *The constants χ_p occurring in Birch's theorem are positive, provided that $s^* \geq \gamma_d^p(R, 1)$ for all p.*

This bound clearly supersedes the previous treatment of the case $m = 1$ due to Schmidt, which we recorded in Lemma 1.3. In contrast, the bound in Theorem 1.4 is much simpler, such as one would naïvely expect, and indeed the effect of our new result is that, from a philosophical point of view and as far as p-adic solubility is concerned, the non-diagonal problem is now on an equal footing with Waring-type situations.

On combining Theorem 1.4 with Birch's Theorem and (1.5.1) one obtains an asymptotic formula for the number of points on hypersurfaces, where the main term is positive.

Corollary 1. *Let $F^{(1)}, \ldots, F^{(R)}$ and $N_{s,R}^{(d)}(P)$ be as in Birch's Theorem. Then $N_{s,R}^{(d)}(P) \asymp P^{s-Rd}$, provided that d is odd and $s^* \geq (Rd^2)^{2^{d-1}}$.*

This follows directly from the fact that

$$(Rd^2)^{2^{d-1}} > 2^{d-1}(d-1)R(R+1)$$

for all admissible parameters d and R. For comparison, recall that Schmidt's approach [62, Prop. I] via Lemma 1.3 yields

$$s^* > 2^{d-1}(d-1)R(Rd^2)^{2^{d-1}};$$

as in the case of Theorem 1.4 we see an improvement of order R.

1.6 Unconditional estimates

So far all of our interest has been revolving around polynomial equations that are not too singular, so it is not unreasonable to ask what happens if this non-singularity condition is violated. This is indeed an important question, as the presence of singularities can have a profound impact on the behaviour of the number of solutions. In fact, assuming singularities of a certain kind can impose powerful extra conditions and thus make singular forms more amenable to available methods than non-singular ones. This phenomenon is being exploited with good success for certain classes of forms of small degree in relatively few variables (an overview and bibliography of recent progress can be found in Dan Loughran's PhD Thesis [44, p. 62]). In very general situations with many variables, however, the presence of singularities seems to present a nuisance rather than an advantage, and the methods so far available struggle to accommodate highly singular cases.

Definition 1.5. *We define $\gamma_d(R,m)$ to be the least integer γ such that every set of R homogeneous polynomials with integer coefficients of degree d in $s \geq \gamma$ variables contains a rational m-dimensional linear space.*

The question is therefore whether there is an upper bound to $\gamma_d(R,m)$ for any given values of d, R and m. This is manifestly not the case whenever the degree d is even, since any form that is definite cannot have any zero different from the trivial one even in $\mathbb{R}$. Avoiding this complication, however, Birch [6] proved in 1957 that $\gamma_d(R,m)$ is finite whenever the degree d is odd, following up on previous work by Brauer [11] who established that every homogeneous polynomial in sufficiently many variables over the p-adic numbers $\mathbb{Q}_p$ represents zero non-trivially. Unfortunately, while Birch's method is in principle constructive, the bounds it yields – "not even astronomical" in Birch's own words[1] – are far from useful. Wooley [76] has made Birch's methods explicit, and even though he managed to improve on them, the bounds he obtains are towers of exponentials of exponentially growing height.

Faced with these difficulties, interest focussed on cubic forms as the simplest interesting special case. It turns out that in this case singular forms can be related fairly efficiently to systems of linear equations, and indeed virtually simultaneously[2] with Birch's work Davenport [13] published his proof that $\gamma_3(1,1) \leq 32$. This number he subsequently managed to bring down to 29 in [14] and finally 16 in [15], where the problem rested until eventually in 2007 Heath-Brown [32] established $\gamma_3(1,1) \leq 14$, which is the best known bound up to this day.

For systems of R forms one has to employ an iterative procedure that generates an additional factor R but nonetheless yields results that are, in the light of the general scarcity of knowledge in the area, quite competitive. The major stumbling block is, again, our insufficient understanding of the p-adic problem. In fact, in 1982 in an important series of papers [57–60] Schmidt found an estimate on $\gamma_3^*(R,1)$ of size R^3, thus establishing an asymptotic number of solutions by Birch's theorem under the condition that $s^* \gg R^4$, and used the iterative approach mentioned above to derive $\gamma_3(R,1) \leq (10R)^5$. This result can be extended fairly easily to linear spaces via the correspondence between linear spaces as solutions of one polynomial equation and points on an

[1]see [7, p. 458] for the quote.

[2]In fact, in the cubic case Birch was beaten narrowly by not only Davenport but also Lewis [41], whose proof that $\gamma_3(R,m) < \infty$ appeared in the same issue of Mathematika as Birch's paper [6] and with a note by the editor (Davenport?) that these two and Davenport's own work had been submitted within only a few months. See also the comment in [16, p. 1188] for the chronology.

expanded system of equations, yielding $\gamma_3(R,m) \ll R^5 m^{14}$, which is recorded in [42, p.293].

There is, however, an alternative approach directly via the multi-dimensional version of Birch's Theorem. In a previous implementation by Dietmann [20, Theorem 2], this approach led to $\gamma_3(R,m) \ll R^6 m^4 + R^5 m^6$, but our work in Chapter 2 allows us to prove the following.

Theorem 1.5. *$\gamma_3(R,m) \ll R^6 + R^3 m^3$.*

This supersedes Dietmann's bound and improves on Schmidt's bound whenever $R < m^{14}$.

The question remains what happens for forms of degree higher than three. Partly because of the somewhat discouraging nature of the results by Birch [6] and Wooley [76, Theorem 1], this problem has not been attempted very much; there is, however, a theorem by Dietmann [21, Theorem 2] who used a version of the multidimensional Birch Theorem to prove that $\gamma_5(1,m) \ll m^{439}$. Due to the improvements obtained in Chapter 2, we are now able to reduce this to $\gamma_5(1,m) \ll m^{48}$. The result can be extended to $R > 1$, but the resulting bound of $\gamma_5(R,m)$ grows doubly exponentially in R and is therefore inferior to Wooley's result in [74, Theorem 2] as soon as $R \gg \log m$. All of these results will be discussed in greater detail in Chapter 5.

Chapter 2

Forms representing forms and linear spaces on hypersurfaces

2.1 Background and History

The problem of determining whether two given forms represent one another is a classical one and has triggered important developments in the history of modern number theory. The oldest work in this area goes back to Gauß [26, Section V, §§282–284], who in the framework of his general theory of binary and ternary quadratic forms also addressed the question of whether the former can be represented by the latter. The problem has been studied more comprehensively by Siegel, who in the 1930s wrote a series of papers [65–67] to deduce what is now known as Siegel's mass formula, which provides an averaged local-global principle for the representation of quadratic forms by quadratic forms.

Analysing forms individually, Hsia, Kitaoka and Kneser [36] proved that for every definite matrix $A \in \mathbb{Z}^{s\times s}$ and for every positive definite matrix $B \in \mathbb{Z}^{m\times m}$ whose minimum is not too small there exists an $(s \times m)$-matrix X such that $X^t AX = B$, provided that $s \geq 2m+3$ and B is locally representable by A. This result has been improved by Ellenberg and Venkatesh [25], who showed that this is possible for positive definite matrices B with square-free discriminant[1] and whose minimum is not too small, provided that $s \geq m+5$, and again they require that the problem be locally soluble.

While these latter results are concerned only with the existence of solutions, and also Siegel's original work does not generally give quantitative estimates for the number $N(A,B)$ of representations of a matrix B by A, Dietmann and Harvey [23] applied the circle method and obtained an asymptotic formula of the shape

$$N(A,B) = \chi_\infty(A,B) \prod_{p \text{ prime}} \chi_p(A,B) + O\left(\det(B)^{\frac{s-m-1}{2}-\delta}\right),$$

where the constants are given by

$$\chi_\infty(A,B) = \det(A)^{-m/2} \det(B)^{(s-m-1)/2} \sqrt{\pi}^{ms-\frac{m(m+1)}{2}} \prod_{j=m+1}^{s} \Gamma\left(\frac{j-m}{2}\right)^{-1}$$

and

$$\chi_p(A,B) = \lim_{t\to\infty} (p^t)^{ms-\frac{m(m+1)}{2}} \operatorname{Card}\{X \pmod{p^t} \colon X^t A X \equiv B \pmod{p^t}\}.$$

The condition on s they require depends quite delicately on the matrix B and is given by

$$s > 2\left(\frac{m(m+1)}{2}+\beta\right)\left(\frac{m(m+1)}{2}+1\right),$$

where $\beta \geq 0$ is a parameter characterising the eccentricity of B, which will typically be very small and vanishes altogether when B is a multiple of the unit matrix.

For higher degree the problem is less well studied, with most of the effort focussing on the case when $R=1$ and F is diagonal. A special case of multidimensional representation problems is the question of finding linear spaces contained in hypersurfaces, corresponding to representing the null-polynomial $\psi = 0$ by F, and in fact the proofs are, as far as our methods are concerned,

[1] This condition can be somewhat relaxed, see [63].

identical. One weakness of the methods should be addressed, namely that they do not, strictly speaking, count linear spaces, but rather parametrisations thereof, and since every linear space may have one or many parametrisations with a given height, one falls prey to double-counting. It stands to hope that in the future a more sophisticated approach involving a suitable inclusion-exclusion argument will provide a way of avoiding this.

In the case $m = 2$, Arkhipov and Karatsuba [2, Theorems 1 and 4] were able to obtain the expected asymptotic formula $\sim cP^{2s-(d+1)d}$ for the number of representations with variables bounded by P, where the constant c is non-negative and encodes the local solubilities, provided that $s \gg d^3 \log d$, and to establish a lower bound of the expected order of magnitude if $s \gg d^2 \log d$. This latter result has later been sharpened by Parsell [48, Theorem 4], who obtained the same lower bound under the condition that

$$s \geq \frac{14}{3} d^2 \log d + \frac{10}{3} d^2 \log \log d + O(d^2),$$

and in subsequent work [50, Theorem 4] showed that, in general, one has an asymptotic formula provided the number of variables satisfies

$$s \geq dmr \left(\frac{4}{3} \log(dr) + \log(md) + 2 \log \log d + 8 \right),$$

where

$$r = \binom{m+d-1}{m-1}. \tag{2.1.1}$$

This result has been superseded by Parsell, Prendiville, and Wooley [52, Theorem 1.4] thanks to a recent breakthrough in the methods commonly used for Waring's problem (see [78] and [79]), which yields an asymptotic formula for the number of solutions under the relaxed condition that

$$s \geq 2 \binom{d+m}{m} (d+1) - 2d - 1. \tag{2.1.2}$$

Meanwhile, Parsell [51, Theorem 1.2] obtained specific results for certain small values of d and m that give better bounds than what can be obtained by (2.1.2). In particular, one can take

$$s \geq d(d-1)2^{d-2} + 2d(d+1) + 1$$

in the case $m = 2$ and

$$s \geq \min\{2m^3 + 6m^2 - 20m + 29, (5/3)m^3 + 5m^2 + (10/3)m + 1\}$$

for $d = 3$. One should note that as a consequence of this it is sufficient to take $s \geq 29$ for lines on diagonal cubic surfaces.

While the problem of counting linear spaces on diagonal hypersurfaces has been addressed in some detail, the only similar results on general, not necessarily diagonal hypersurfaces apart from the abovementioned result by Dietmann and Harvey are due to Dietmann, who established an asymptotic formula with positive main term if $s \geq 2^{4d+1}d!rd(rd^2)^{2^{d-1}}$ and the form F is non-singular [21]. In the cubic case (see [20]) he considers systems of R equations and relaxes the non-singularity condition by demanding that no form in the rational pencil of the forms vanishes on a rational space whose codimension is $O(R^3m^6 + R^5m^5)$. However, both of these results treat the problem in a rather cursory manner without making an effort to distill and exploit the underlying structure of the problem.

This gives the motivation for our Theorem 1.1. The statement is somewhat more general than the one given in Section 1.4, as we will consider systems of R simultaneous representation problems. In order to give a rigorous enunciation of the results, it is useful to introduce some notation. Let P be a large positive integer, write $\boldsymbol{\psi}$ for the R-tuple $\left(\psi^{(1)}, \ldots, \psi^{(R)}\right)$ and denote by $N^{(d)}_{s,R,m}(P; \mathbf{F}; \boldsymbol{\psi})$ the number of integral solutions of the equations

$$F^{(\rho)}(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) = \psi^{(\rho)}(t_1, \ldots, t_m) \qquad (1 \leq \rho \leq R) \tag{2.1.3}$$

identically in $t_1, \ldots, t_m$ with $\mathbf{x}_i \in [-P, P]^s$ for $1 \leq i \leq m$.

Theorem 2.1. *Let $d \geq 2$, R, and $m \geq 2$ be positive integers, and let*

$$s^* > 3 \cdot 2^{d-1}(d-1)R(Rr+1).$$

Then there exist nonnegative constants $\chi_\infty(\boldsymbol{\psi})$ and $\chi_p(\boldsymbol{\psi})$ for every prime p such that

$$N^{(d)}_{s,R,m}(P;\mathbf{F};\boldsymbol{\psi}) = P^{ms-Rrd}\chi_\infty(\boldsymbol{\psi}) \prod_{p \text{ prime}} \chi_p(\boldsymbol{\psi}) + o\left(P^{ms-Rrd}\right).$$

Note that Theorem 2.1 does not compare directly with Dietmann's results [20, 21] quoted above, as we do not require at this point that the local factors χ_∞ and χ_p are indeed positive. These factors will be studied more closely in Chapter 4, where it will transpire that Theorem 2.1 enables us to save quite significantly over Dietmann's results. Similarly, in comparison to Dietmann's and Harvey's result [23] we save a factor r in the number of variables required, and one can indeed show that this would also be the savings over Dietmann's results in [20] and [21], were he to abandon the requirement that all local solubility conditions be met.

2.2 Notation and Setting

For $1 \leq \rho \leq R$ let $F^{(\rho)} \in \mathbb{Z}[x_1, \ldots, x_s]$ be given by

$$F^{(\rho)}(\mathbf{x}) = \sum_{\mathbf{i} \in \{1,\ldots,s\}^d} c^{(\rho)}_{\mathbf{i}} x_{i_1} \cdot \ldots \cdot x_{i_d}$$

with symmetric coefficients $c^{(\rho)}_{\mathbf{i}} \in \mathbb{Z}/d!$, and define the multilinear form $\Phi^{(\rho)}$ associated to $F^{(\rho)}$ by

$$\Phi^{(\rho)}\left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}\right) = \sum_{\mathbf{i} \in \{1,\ldots,s\}^d} c^{(\rho)}_{\mathbf{i}} x^{(1)}_{i_1} \cdot \ldots \cdot x^{(d)}_{i_d}.$$

Thus one has

$$F^{(\rho)}(\mathbf{x}) = \Phi^{(\rho)}(\mathbf{x}, \ldots, \mathbf{x}).$$

In order to count solutions to (2.1.3), one needs to understand expressions of the shape

$$F^{(\rho)}(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m). \tag{2.2.1}$$

This requires an appropriate kind of index notation. Write J for the set of multi-indices $(j_1, j_2, \ldots, j_d) \in \{1, 2, \ldots, m\}^d$, where we allow repetitions in the tuples $(j_1, j_2, \ldots, j_d)$ but disregard order. The number of these is r, which is the parameter defined in (2.1.1). By means of the Multinomial Theorem, Equation (2.2.1) can be written as

$$F^{(\rho)}\left(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m\right) = \sum_{\mathbf{j}\in J} A(\mathbf{j}) t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d} \Phi^{(\rho)}(\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_d}), \tag{2.2.2}$$

where the factors $A(\mathbf{j})$ take account of the multiplicity of each term and are defined as follows. To every $\mathbf{j} \in J$ one can associate numbers $\mu_1(\mathbf{j}), \ldots, \mu_m(\mathbf{j})$ between 0 and d such that

$$t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d} = t_1^{\mu_1(\mathbf{j})} t_2^{\mu_2(\mathbf{j})} \cdot \ldots \cdot t_m^{\mu_m(\mathbf{j})}. \tag{2.2.3}$$

In other words, the $\mu_i(\mathbf{j})$ count the multiplicity with which any given $\mathbf{x}_i$ appears in the term with index $\mathbf{j}$. In this notation, the factors $A(\mathbf{j})$ are given by the multinomial coefficients

$$A(\mathbf{j}) = \binom{d}{\mu_1(\mathbf{j}), \mu_2(\mathbf{j}), \ldots, \mu_m(\mathbf{j})}.$$

Let $\psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(R)}$ be homogeneous polynomials of degree d in m variables, defined by

$$\psi^{(\rho)}(t_1, \ldots, t_m) = \sum_{\mathbf{j}\in J} n_{\mathbf{j}}^{(\rho)} A(\mathbf{j}) t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d}$$

for $n_{\mathbf{j}}^{(\rho)} \in \mathbb{Z}/d!$, and write

$$\boldsymbol{\psi} = \left(\psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(R)}\right).$$

For the sake of brevity and compactness, we will use the shorthand notation

$$(\mathbf{x}_1, \ldots, \mathbf{x}_m) = \overline{\mathbf{x}}$$

and

$$\Phi^{(\rho)}(\mathbf{x}_{j_1},\dots,\mathbf{x}_{j_d}) = \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}).$$

We can view Equation (2.1.3) as a polynomial equation in $t_1,\dots,t_m$, so after expanding and sorting by coefficients one obtains a system of equations

$$\Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}) = n^{(\rho)}_{\mathbf{j}} \quad (\mathbf{j} \in J, 1 \le \rho \le R),$$

which is amenable to a circle method approach. In order to set up a practicable notation, we write $\boldsymbol{\alpha}^{(\rho)} = (\alpha^{(\rho)}_{\mathbf{j}})_{\mathbf{j}\in J}$ for given $1 \le \rho \le R$ and let

$$\mathfrak{F}^{(\rho)}\big(\mathbf{x}_1,\dots,\mathbf{x}_m;\boldsymbol{\alpha}^{(\rho)}\big) = \sum_{\mathbf{j}\in J} \alpha^{(\rho)}_{\mathbf{j}} \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}). \tag{2.2.4}$$

Furthermore, let

$$\big(\boldsymbol{\alpha}^{(1)},\boldsymbol{\alpha}^{(2)},\dots,\boldsymbol{\alpha}^{(R)}\big) = \underline{\boldsymbol{\alpha}},$$

then we can write the sum over the expressions in (2.2.4) as

$$\mathfrak{F}\left(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}}\right) = \sum_{\rho=1}^{R} \mathfrak{F}^{(\rho)}\big(\mathbf{x}_1,\dots,\mathbf{x}_m;\boldsymbol{\alpha}^{(\rho)}\big). \tag{2.2.5}$$

Since it will be useful at some point to sort the components of $\underline{\boldsymbol{\alpha}}$ by the $\mathbf{j}$ as opposed to the ρ, we seize the opportunity to define $\underline{\alpha}_{\mathbf{j}} = \big(\alpha^{(1)}_{\mathbf{j}},\dots,\alpha^{(R)}_{\mathbf{j}}\big)$ for all $\mathbf{j} \in J$. The same notational conventions will be observed for the coefficients $n^{(\rho)}_{\mathbf{j}}$ of the target polynomials $\boldsymbol{\psi}$.

The expression in (2.2.5) collects all the Rr terms that arise from expanding each of the R equations as a sum of r multilinear forms, and thus allows us to define the exponential sum in a very compact notation as

$$T(\underline{\boldsymbol{\alpha}}) = \sum_{\overline{\mathbf{x}}} e(\mathfrak{F}\left(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}}\right)).$$

In general, the sum will be over a box $-P \le \mathbf{x}_i \le P$ for all $1 \le i \le m$, but in special cases we will write $T(\underline{\boldsymbol{\alpha}}, X)$ or $T(\underline{\boldsymbol{\alpha}}, \mathfrak{B})$ to denote an ms-dimensional hypercube with sidelength $2X$ or a domain $\mathfrak{B} \subset \mathbb{Z}^{ms}$, respectively. Altogether, classical orthogonality relations imply that the number of simultaneous representations of $\psi^{(\rho)}$ by $F^{(\rho)}$ contained in the hypercube $[-P,P]^{ms}$ is described by the integral

$$\begin{aligned} N_{s,R,m}^{(d)}(P;\mathbf{F};\boldsymbol{\psi}) &= \int_{[0,1)^{Rr}} T(\underline{\boldsymbol{\alpha}}) e(-\underline{\boldsymbol{\alpha}} \cdot \underline{\mathbf{n}}) \mathrm{d}\underline{\boldsymbol{\alpha}} \\ &= \sum_{\substack{\mathbf{x}_1,\ldots,\mathbf{x}_m \\ |\mathbf{x}_i| \leq P}} \int_{[0,1)^{Rr}} e\left(\mathfrak{F}\left(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}}\right) - \underline{\boldsymbol{\alpha}} \cdot \underline{\mathbf{n}}\right) \mathrm{d}\underline{\boldsymbol{\alpha}}. \end{aligned} \tag{2.2.6}$$

For the sake of convenience we will in the future suppress most of the parameters and use the more concise notation $N_{s,\boldsymbol{\psi}}(P)$.

It should be noted, however, that although expressions as in (2.2.4) and (2.2.5) aim to simultaneously solve rR equations, the single equations can be reassembled and can thus be read in the way of our original problem of finding m-dimensional linear spaces on the intersection of R hypersurfaces. In fact, one can view the coefficients $\alpha_{\mathbf{j}}^{(\rho)}$ as absorbing the factors $A(\mathbf{j}) t_{j_1} \cdot \ldots \cdot t_{j_d}$ arising in (2.2.2), and write, somewhat imprecisely,

$$\begin{aligned} \mathfrak{F}^{(\rho)}\big(\overline{\mathbf{x}};\boldsymbol{\alpha}^{(\rho)}\big) - \boldsymbol{\alpha}^{(\rho)} \cdot \mathbf{n}^{(\rho)} &= \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}}^{(\rho)} A(\mathbf{j}) t_{j_1} \cdot \cdots \cdot t_{j_d} \big(\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}}) - n_{\mathbf{j}}^{(\rho)}\big) \\ &= \alpha^{(\rho)} \left(F^{(\rho)}(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m) - \psi^{(\rho)}(t_1,\ldots,t_m)\right). \end{aligned}$$

Thus (2.2.6) can be read either as finding simultaneous solutions of Rr equations or equivalently as counting m-dimensional linear spaces on R hypersurfaces, and while for the greater part of the analysis we will stick to the former interpretation, it will be convenient to switch to the latter one when analysing the singular series more carefully.

2.3 Weyl differencing

The proof of Theorem 2.1 is largely along the lines of the classical arguments of Birch [8] and Schmidt [62], with most of the analysis and notation following Birch, while imitating Schmidt's arguments in the treatment of the singular series and singular integral.

The first step is to establish an inequality of Weyl type as presented in Davenport's book [17, Chapters 12 and 13], or in a more general version, in [8]. Although this is fairly standard, we will give a rather detailed exposition because it is here that the specific shape of the forms assembled in $\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})$ comes into play.

Lemma 2.1. *Let $1 \leq k \leq d$ and j_l $(l = 1,\dots,k)$ be integers with $1 \leq j_l \leq m$. Then*

$$|T(\underline{\boldsymbol{\alpha}})|^{2^k} \ll P^{\left((2^k-1)m-k\right)s} \sum_{\mathbf{h}_1,\dots,\mathbf{h}_k\in[-P,P]^s} \sum_{\overline{\mathbf{x}}} e(\Delta_{j_k,\mathbf{h}_k}\cdots\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})),$$

where the discrete differencing operator $\Delta_{i,\mathbf{h}}$ is defined by its action on the form $\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})$ as

$$\Delta_{i,\mathbf{h}}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}}) = \mathfrak{F}(\mathbf{x}_1,\dots,\mathbf{x}_i+\mathbf{h},\dots,\mathbf{x}_m;\underline{\boldsymbol{\alpha}}) - \mathfrak{F}(\mathbf{x}_1,\dots,\mathbf{x}_i,\dots,\mathbf{x}_m;\underline{\boldsymbol{\alpha}}) \quad (2.3.1)$$

and the sum over the $\overline{\mathbf{x}}$ extends over suitable boxes of sidelength at most $2P$.

Proof. As is usual with Weyl differencing arguments, we proceed by induction. The case $k = 1$ follows from a simple application of Cauchy's inequality and one has

$$|T(\underline{\boldsymbol{\alpha}})|^2 \ll P^{(m-1)s} \sum_{\substack{\mathbf{x}_i \\ i\neq j_1}} \left(\sum_{\mathbf{h}_1} \sum_{\mathbf{x}_{j_1}} e\left(\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})\right) \right).$$

Here the summation over $\mathbf{x}_{j_1}$ extends over the intersection of the boxes given by $|\mathbf{x}_{j_1}| \leq P$ and $|\mathbf{x}_{j_1}+\mathbf{h}_1| \leq P$, which is again a box whose sides are of length at most $2P$.

Now let us assume that the lemma is true for a given k. Again by Cauchy's inequality, one finds

$$\begin{aligned}|T(\underline{\boldsymbol{\alpha}})|^{2^{k+1}} &\ll P^{2\left((2^k-1)m-k\right)s}\left|\sum_{\mathbf{h}_1,\ldots,\mathbf{h}_k}\sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_k,\mathbf{h}_k}\cdots\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})\right)\right|^2 \\ &\ll P^{\left(2^{k+1}-2\right)ms-2ks}P^{ks+(m-1)s} \\ &\qquad \times \sum_{\mathbf{h}_1,\ldots,\mathbf{h}_k}\sum_{\substack{\mathbf{x}_i \\ i\neq j_{k+1}}}\left|\sum_{\mathbf{x}_{j_{k+1}}} e\left(\Delta_{j_k,\mathbf{h}_k}\cdots\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})\right)\right|^2 \\ &\ll P^{\left((2^{k+1}-1)m-(k+1)\right)s} \\ &\qquad \times \sum_{\mathbf{h}_1,\ldots,\mathbf{h}_{k+1}}\sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_{k+1},\mathbf{h}_{k+1}}\cdots\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})\right)\end{aligned}$$

as required. □

For the sake of notational brevity in the following considerations, we will write $\hat{\mathbf{h}}$ for the $(d-1)$-tuple $(\mathbf{h}_1,\ldots,\mathbf{h}_{d-1})$. Our final estimate of the exponential sum $T(\underline{\boldsymbol{\alpha}})$ is an application of the above.

Lemma 2.2. *For any* $\mathbf{j}\in J$ *one has the estimate*

$$|T(\underline{\boldsymbol{\alpha}})|^{2^{d-1}} \ll P^{(2^{d-1}m-d)s}\sum_{\hat{\mathbf{h}}}\prod_{i=1}^{s}\min\left(P,\left\|M(\mathbf{j})\sum_{\rho=1}^{R}\alpha_{\mathbf{j}}^{(\rho)}B_i^{(\rho)}(\hat{\mathbf{h}})\right\|^{-1}\right),$$

where the functions $B_i^{(\rho)}$ *are given by*

$$\Phi^{(\rho)}(\mathbf{x},\mathbf{h}_1,\ldots,\mathbf{h}_{d-1}) = \sum_{i=1}^{s} x_i B_i^{(\rho)}(\mathbf{h}_1,\ldots,\mathbf{h}_{d-1})$$

and the coefficients $M(\mathbf{j})$ *are defined by means of* (2.2.3) *as*

$$M(\mathbf{j}) = \mu_1(\mathbf{j})!\mu_2(\mathbf{j})!\cdot\ldots\cdot\mu_m(\mathbf{j})!.$$

Proof. Inserting $k=d-1$ in the above lemma gives

$$|T(\underline{\boldsymbol{\alpha}})|^{2^{d-1}} \ll P^{\left((2^{d-1}-1)m-(d-1)\right)s}\sum_{\hat{\mathbf{h}}}\sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_{d-1},\mathbf{h}_{d-1}}\cdots\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\underline{\boldsymbol{\alpha}})\right).$$

By the definition (2.3.1) of $\Delta_{j,\mathbf{h}}$ and the polynomial structure of $\mathfrak{F}$, every differencing step reduces the degree of the resulting form by one, and therefore this last expression depends only linearly on the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$. In particular, for any given ρ all the forms assembled in $\mathfrak{F}^{(\rho)}(\overline{\mathbf{x}}; \boldsymbol{\alpha}^{(\rho)})$ are instances of the same multilinear form $\Phi^{(\rho)}$ associated to the original form $F^{(\rho)}$, and this structure is naturally preserved by the differencing procedure. Writing $R(\hat{\mathbf{h}})$ for the terms independent of $\overline{\mathbf{x}}$, one obtains

$$\begin{aligned}
&\sum_{\hat{\mathbf{h}}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_{d-1},\mathbf{h}_{d-1}} \cdots \Delta_{j_1,\mathbf{h}_1} \mathfrak{F}(\overline{\mathbf{x}}; \underline{\boldsymbol{\alpha}})\right) \\
=&\sum_{\hat{\mathbf{h}}} \sum_{\substack{\mathbf{x}_i \\ i \neq j_d}} \sum_{\mathbf{x}_{j_d}} e\left(\sum_{\rho=1}^{R} \sum_{k=1}^{m} M(\mathbf{j}) \alpha^{(\rho)}_{(j_1,\ldots,j_{d-1},k)} \Phi^{(\rho)}(\mathbf{x}_k, \mathbf{h}_1, \ldots, \mathbf{h}_{d-1}) + R(\hat{\mathbf{h}})\right) \\
\ll& P^{s(m-1)} \sum_{\hat{\mathbf{h}}} \left| \sum_{\mathbf{x}_{j_d}} e\left(M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha^{(\rho)}_{\mathbf{j}} \Phi^{(\rho)}(\mathbf{x}_{j_d}, \mathbf{h}_1, \ldots, \mathbf{h}_{d-1})\right)\right|.
\end{aligned}$$

Since the $\mathbf{x}_j$ are contained in suitable boxes contained in $[-P, P]^s$, by standard arguments one arrives at the estimate

$$|T(\underline{\boldsymbol{\alpha}})|^{2^{d-1}} \ll P^{(2^{d-1}m-d)s} \sum_{\hat{\mathbf{h}}} \prod_{i=1}^{s} \min\left(P, \left\| M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha^{(\rho)}_{\mathbf{j}} B^{(\rho)}_i(\hat{\mathbf{h}}) \right\|^{-1}\right),$$

which is indeed the required expression. □

Note that whereas the exponential sum $T(\underline{\boldsymbol{\alpha}})$ is independent of $\mathbf{j}$, this is not true for the expression on the right hand side of the equation. This is due to the fact that the Weyl differencing argument forces us to close in onto one index $\mathbf{j}$, but since it is immaterial which difference is taken in each step, the estimate holds for any index $\mathbf{j} \in J$. This pays off when we read the estimate in Lemma 2.2 from the right, as it will save us the averaging over the indices $\mathbf{j}$, and one obtains independent estimates of equal quality for all $\mathbf{j} \in J$. This is the main point in which the linear space setting behaves differently from the general situation as treated in [8], as it allows us to treat what is technically a system of r distinct equations as a single one, and this different behaviour will ultimately yield the improvements this work obtains over previous estimates.

We use Lemma 2.2 to generate a tripartite case distinction. Let us assume that we have $|T(\underline{\boldsymbol{\alpha}})| \gg P^{ms-k\theta}$ for some parameters $k, \theta > 0$, and let $\mathbf{j}$ be fixed. In this case Lemma 2.2 yields

$$\left(P^{ms-k\theta}\right)^{2^{d-1}} \ll P^{(2^{d-1}m-d)s} \sum_{\hat{\mathbf{h}}} \prod_{i=1}^{s} \min\left(P, \left\| M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\hat{\mathbf{h}}) \right\|^{-1}\right),$$

which is equivalent to

$$\sum_{\mathbf{h}_1,\ldots,\mathbf{h}_{d-1}} \prod_{i=1}^{s} \min\left(P, \left\| M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\hat{\mathbf{h}}) \right\|^{-1}\right) \gg P^{ds-2^{d-1}k\theta}.$$

Note that this expression is independent of m. This implies that the following arguments can be extracted directly from standard references. In particular, one can apply the geometry of numbers in order to count how often the minimum takes on a nontrivial value (see [17], chapters 12 and 13 for a detailed exposition). The subsequent lemma is the analogue of [17, Lemma 13.3] or [8, Lemma 2.4], respectively.

Lemma 2.3. *Suppose that $|T(\underline{\boldsymbol{\alpha}})| \gg P^{ms-k\theta}$ for some parameters $k, \theta > 0$, and let $N(X, Y)$ denote the number of $(d-1)$-tuples $\mathbf{h}_1, \ldots, \mathbf{h}_{d-1}$ in the box $|\mathbf{h}_k| \le X$ satisfying*

$$\left\| M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\hat{\mathbf{h}}) \right\| < Y \tag{2.3.2}$$

for all $i = 1, \ldots, s$ and $\mathbf{j} \in J$. Then we have the estimate

$$N(P^\theta, P^{-d+(d-1)\theta}) \gg P^{(d-1)s\theta - 2^{d-1}k\theta - \epsilon}.$$

From Lemma 2.3 one infers that either the exponential sum is small, or for all i and $\mathbf{j}$ the quantity $M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\hat{\mathbf{h}})$ is often close to an integer. The latter can be effected in two ways, as it will occur either if the forms $B_i^{(\rho)}$ tend to vanish for geometric reasons, or by genuine (i. e. non-zero) solutions to the diophantine approximation problem that is implicit in (2.3.2). This yields a threefold case distinction, which lies at the heart of all circle method arguments concerning general homogeneous polynomials.

Lemma 2.4. *Let $0 < \theta \le 1$ and k be parameters, and let $\underline{\boldsymbol{\alpha}} \in [0, 1)^{rR}$. Then there are three possibilities.*

(A) The exponential sum $T(\underline{\boldsymbol{\alpha}})$ *is bounded by*

$$|T(\underline{\boldsymbol{\alpha}})| \ll P^{ms-k\theta}.$$

(B) For every $\mathbf{j} \in J$ *one finds* $(q_{\mathbf{j}}, \underline{a}_{\mathbf{j}}) \in \mathbb{Z}^{R+1}$ *satisfying*

$$0 < q_{\mathbf{j}} \ll P^{(d-1)R\theta} \qquad \text{and} \qquad \left|\alpha_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} - a_{\mathbf{j}}^{(\rho)}\right| \ll P^{-d+(d-1)R\theta}$$

for all $1 \leq \rho \leq R$.

(C) The number of $(d-1)$*-tuples* $(\mathbf{h}_1, \ldots, \mathbf{h}_{d-1}) \leq P^{\theta}$ *that satisfy*

$$\operatorname{rank}\left(B_i^{(\rho)}(\mathbf{h}_1, \ldots, \mathbf{h}_{d-1})\right)_{i,\rho} \leq R-1 \tag{2.3.3}$$

is asymptotically greater than $\left(P^{\theta}\right)^{(d-1)s-2^{d-1}k-\epsilon}$.

Proof. This follows by the same argument as in [8, Lemma 2.5]. Suppose that the estimate in (A) does not hold, so that by Lemma 2.3 for every $\mathbf{j} \in J$ we have

$$\left\|M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\hat{\mathbf{h}}_{\mathbf{j}})\right\| < P^{-d+(d-1)\theta}$$

for at least $\gg P^{(d-1)s\theta-2^{d-1}k\theta-\epsilon}$ choices of $\hat{\mathbf{h}} \leq P^{\theta}$. Fixing some index $\mathbf{j} \in J$ and writing $B(\mathbf{j})$ for the $(R \times s)$-matrix $\left(M(\mathbf{j}) B_i^{(\rho)}(\hat{\mathbf{h}}_{\mathbf{j}})\right)_{i,\rho}$, this means we can find integer vectors $\mathbf{A}(\mathbf{j})$ and $\boldsymbol{\delta}(\mathbf{j}) \in \mathbb{Z}^s$ with the property that

$$B(\mathbf{j})\underline{\alpha}_{\mathbf{j}} - \mathbf{A}(\mathbf{j}) = \boldsymbol{\delta}(\mathbf{j}) \ll P^{-d+(d-1)\theta}. \tag{2.3.4}$$

If the matrix $B(\mathbf{j})$ is of full rank for some $(d-1)$-tuple $\hat{\mathbf{h}}_{\mathbf{j}} = (\mathbf{h}_1, \ldots, \mathbf{h}_{d-1})$, then we can find a non-vanishing $(R \times R)$-minor whose absolute value we denote by $q_{\mathbf{j}}$. We remark here for further reference the obvious fact that $q_{\mathbf{j}}$ is independent of any particular index ρ. This allows us to implement a discrete matrix inversion in order to generate approximations of the $\underline{\alpha}_{\mathbf{j}}$.

We may assume without loss of generality that the non-vanishing minor $B_0(\mathbf{j})$ of $B(\mathbf{j})$ is the first one, and write $\mathbf{A}_0(\mathbf{j}), \boldsymbol{\delta}_0(\mathbf{j}) \in \mathbb{Z}^R$ for the corresponding portions of $\mathbf{A}(\mathbf{j})$ and $\boldsymbol{\delta}(\mathbf{j})$. Then we can find integer solutions $a_{\mathbf{j}}^{(1)}, \ldots, a_{\mathbf{j}}^{(R)}$ to the system

$$B_0(\mathbf{j})\underline{a}_{\mathbf{j}} = q_{\mathbf{j}}\mathbf{A}_0(\mathbf{j}). \tag{2.3.5}$$

Combining (2.3.4) and (2.3.5), we obtain

$$B_0(\mathbf{j})\left(q_{\mathbf{j}}\underline{\alpha}_{\mathbf{j}} - \underline{a}_{\mathbf{j}}\right) = q_{\mathbf{j}}\boldsymbol{\delta}_0(\mathbf{j}).$$

By Cramer's rule this returns the required bound, and the proof is complete on noting that

$$0 < q_{\mathbf{j}} \ll \max_{i,\rho} |B_i^{(\rho)}(\hat{\mathbf{h}}_{\mathbf{j}})|^R \ll P^{R(d-1)\theta}.$$

□

The reader should take note here that, since the estimates obtained in Lemma 2.2 hold independently of the index $\mathbf{j}$, the differencing variables $\hat{\mathbf{h}}_{\mathbf{j}}$ that generate the rational approximation need not be the same for all $\mathbf{j} \in J$, and therefore the approximations will in general depend on the $\mathbf{j}$-component of the $\alpha_{\mathbf{j}}^{(\rho)}$ that is being approximated. On the other hand, it is clear that the arguments themselves, in particular the upper bound for the values of $B_i^{(\rho)}(\hat{\mathbf{h}}_{\mathbf{j}})$, are independent of the index $\mathbf{j}$ chosen in the beginning of the proof, so we can find rational approximations of the same quality for all vectors $\underline{\alpha}_{\mathbf{j}}$, and their denominators vary with $\mathbf{j} \in J$ but, as we remarked earlier, are independent of ρ.

The condition (2.3.3) of case (C) in Lemma 2.4 is tantamount to a system of simultaneous equations in $s(d-1)$ variables and thus defines a variety which we call $\mathfrak{B}$. By [8, Lemma 3.2], its dimension is

$$\dim \mathfrak{B} \geq (d-1)s - 2^{d-1}k. \tag{2.3.6}$$

Since, however, the variety $\mathfrak{B}$ is not particularly easy to handle, we follow Birch [8, Lemma 3.3] and replace it by the singular locus $\mathfrak{V}$. Let $\mathfrak{D}$ denote the diagonal

$$\mathfrak{D} = \{\mathbf{h}_1, \ldots, \mathbf{h}_{d-1} \in \mathbb{Z}^s : \mathbf{h}_1 = \ldots = \mathbf{h}_{d-1}\} \subseteq \mathbb{Z}^{(d-1)s}.$$

Obviously $\dim \mathfrak{D} = s$, and we have $\mathfrak{V} = \mathfrak{D} \cap \mathfrak{B}$. By the Affine Dimension Theorem this means that

$$\dim \mathfrak{V} \geq \dim \mathfrak{B} + \dim \mathfrak{D} - (d-1)s,$$

which together with (2.3.6) gives

$$s - \dim \mathfrak{V} \leq (d-1)s - \dim \mathfrak{B} \leq 2^{d-1}k.$$

Notice that $s - \dim \mathfrak{V} = s^*$ by Definition 1.3. This allows us to exclude the third case in Lemma 2.4 by choosing the number of variables sufficiently large.

Lemma 2.5. *Let $\underline{\boldsymbol{\alpha}} \in [0,1)^{Rr}$ and let $0 < \theta \leq 1$ and k be parameters with*

$$s^* > 2^{d-1}k. \tag{2.3.7}$$

Then the alternatives are the following.

(A) The exponential sum $T(\underline{\boldsymbol{\alpha}})$ is bounded by

$$|T(\underline{\boldsymbol{\alpha}})| \ll P^{ms-k\theta}.$$

(B) For every $\mathbf{j} \in J$ one finds $(q_{\mathbf{j}}, \underline{a}_{\mathbf{j}}) \in \mathbb{Z}^{R+1}$, satisfying

$$0 < q_{\mathbf{j}} \ll P^{(d-1)R\theta} \qquad \text{and} \qquad \left|\alpha_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} - a_{\mathbf{j}}^{(\rho)}\right| \ll P^{-d+(d-1)R\theta}$$

for all $1 \leq \rho \leq R$.

2.4 Major Arcs dissection

Lemma 2.5 suggests a major arcs dissection in terms of the parameter θ, a notion that can be made rigorous by specifying the implicit constant. Let C be sufficiently large in terms of the coefficients of the $F^{(\rho)}$. We define the major arcs $\mathfrak{M}(P,\theta)$ to be the set of all $\underline{\boldsymbol{\alpha}} \in [0,1)^{Rr}$ that have a rational approximation satisfying

$$\begin{aligned} &0 \leq a_{\mathbf{j}}^{(\rho)} < q_{\mathbf{j}} \leq CP^{(d-1)R\theta}, \\ &\left|\alpha_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} - a_{\mathbf{j}}^{(\rho)}\right| \leq CP^{-d+(d-1)R\theta} \quad (1 \leq \rho \leq R), \end{aligned} \tag{2.4.1}$$

and the minor arcs

$$\mathfrak{m}(P,\theta) = [0,1)^{Rr} \setminus \mathfrak{M}(P,\theta)$$

to be the complement thereof. In the interest of readability, we will omit the parameter P during most of the analysis, specifying it only in cases where ambiguities might be likely to arise. It is, however, worthwhile to note that this definition respects the case distinction of Lemma 2.5, that is, for every $\underline{\boldsymbol{\alpha}} \in [0,1)^{rR}$ one has either a rational approximation as in (2.4.1) or the estimate in case (A) holds true.

The major arcs are disjoint if $2R(d-1)\theta < d$. Indeed, if there were a point $\underline{\boldsymbol{\alpha}}$ with the property that at least one component $\alpha_{\mathbf{j}}^{(\rho)}$ has two approximations $a_{\mathbf{j}}^{(\rho)}/q_{\mathbf{j}}$ and $b_{\mathbf{j}}^{(\rho)}/p_{\mathbf{j}}$ both satisfying (2.4.1), then this would imply

$$1 \leq \left| a_{\mathbf{j}}^{(\rho)} p_{\mathbf{j}} - b_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} \right| \leq p_{\mathbf{j}} \left| q_{\mathbf{j}} \alpha_{\mathbf{j}}^{(\rho)} - a_{\mathbf{j}}^{(\rho)} \right| + q_{\mathbf{j}} \left| p_{\mathbf{j}} \alpha_{\mathbf{j}}^{(\rho)} - b_{\mathbf{j}}^{(\rho)} \right| \ll P^{-d+2(d-1)R\theta},$$

which is a contradiction for large P as soon as $2R(d-1)\theta < d$. It follows that the major arcs are disjoint.

Further, the volume of the major arcs is at most

$$\begin{aligned} \operatorname{vol}(\mathfrak{M}(\theta)) &\ll \prod_{\mathbf{j} \in J} \left(\sum_{q_{\mathbf{j}}=1}^{CP^{R(d-1)\theta}} \prod_{\rho=1}^{R} \left(\sum_{a_{\mathbf{j}}^{(\rho)}=0}^{q_{\mathbf{j}}-1} \frac{P^{-d+R(d-1)\theta}}{q_{\mathbf{j}}} \right) \right) \\ &\ll \prod_{\mathbf{j} \in J} \left(\sum_{q_{\mathbf{j}}=1}^{CP^{R(d-1)\theta}} \left(P^{-d+R(d-1)\theta} \right)^{R} \right) \\ &\ll P^{-drR+(d-1)rR(1+R)\theta}. \end{aligned} \tag{2.4.2}$$

As will become apparent in the following discussion, we will need to fix the parameter θ rather small so as to allow a better error control when examining the major arcs contribution more closely. Also, in order to minimise the number of variables required in (2.3.7), we should like to choose k small. On the other hand, we require $k\theta > Rrd$ in order to get a suitable estimate on the minor arcs. This discrepancy motivates the following pruning lemma.

Lemma 2.6. *Suppose the parameters k and θ satisfy*

$$0 < \theta < \theta_0 = \frac{d}{(d-1)(R+1)}$$

and

$$k > Rr(R+1)(d-1). \tag{2.4.3}$$

Then there exists a $\delta > 0$ such that the minor arcs contribution is bounded by

$$\int_{\mathfrak{m}(P,\theta)} |T(\underline{\boldsymbol{\alpha}})| \mathrm{d}\underline{\boldsymbol{\alpha}} \ll P^{ms-Rrd-\delta}.$$

Proof. This is a straightforward adaptation of [8, Lemma 4.4]. Given θ between 0 and θ_0, we can find a parameter $\delta > 0$ such that

$$(k - Rr(R+1)(d-1))\theta > 2\delta \tag{2.4.4}$$

and a sequence θ_i with the property that

$$1 \geq \theta_0 > \theta_1 > \theta_2 > \ldots > \theta_{M-1} > \theta_M = \theta > 0$$

and subject to the condition

$$(\theta_i - \theta_{i+1})k < \delta \quad \text{for all } i. \tag{2.4.5}$$

This is always possible with

$$M = O(1). \tag{2.4.6}$$

Then on writing

$$\mathfrak{m}_i = \mathfrak{m}(\theta_i) \setminus \mathfrak{m}(\theta_{i-1}) = \mathfrak{M}(\theta_{i-1}) \setminus \mathfrak{M}(\theta_i)$$

one has

$$\mathrm{vol}(\mathfrak{m}_i) \leq \mathrm{vol}(\mathfrak{M}(\theta_{i-1})) \ll P^{-drR+(d-1)rR(1+R)\theta_{i-1}}$$

by (2.4.2). Recall that for $\underline{\boldsymbol{\alpha}} \in \mathfrak{m}(\theta)$, we are in the situation of case (A) in Lemma 2.5, so the minor arcs contribution is bounded by

$$\begin{aligned}\int_{\mathfrak{m}(\theta)\setminus\mathfrak{m}(\theta_0)} |T(\underline{\boldsymbol{\alpha}})| \, \mathrm{d}\underline{\boldsymbol{\alpha}} &= \sum_{i=1}^{M} \int_{\mathfrak{m}_i} |T(\underline{\boldsymbol{\alpha}})| \, \mathrm{d}\underline{\boldsymbol{\alpha}} \\ &\ll \sum_{i=1}^{M} \mathrm{vol}(\mathfrak{M}(\theta_{i-1})) \sup_{\underline{\boldsymbol{\alpha}} \in \mathfrak{m}(\theta_i)} |T(\underline{\boldsymbol{\alpha}})| \\ &\ll \sum_{i=1}^{M} P^{-Rrd+(d-1)rR(1+R)\theta_{i-1}} P^{ms-k\theta_i}.\end{aligned}$$

By (2.4.6), the sum is of no consequence and can be replaced by a maximum over all $i \in \{1, \dots, M\}$. Hence the exponent is

$$\begin{aligned} &- Rrd + (d-1)rR(1+R)\theta_{i-1} + ms - k\theta_i \\ &\quad = ms - Rrd + k(\theta_{i-1} - \theta_i) - (k - (d-1)rR(1+R))\theta_{i-1} \\ &\quad \leq ms - Rrd - \delta, \end{aligned}$$

where the last inequality uses (2.4.4) and (2.4.5). Finally, we observe that on $\mathfrak{m}(\theta_0)$ the result follows directly from Lemma 2.5. □

2.5 Homogenising the approximations

Lemma 2.5 (B) gives us approximations of the shape

$$\alpha_{\mathbf{j}}^{(\rho)} = a_{\mathbf{j}}^{(\rho)}/q_{\mathbf{j}} + \beta_{\mathbf{j}}^{(\rho)}$$

with denominators that are in general different for each $\mathbf{j} \in J$. It will, however, greatly facilitate the future analysis if we can find a common denominator q such that approximations of the shape

$$\alpha_{\mathbf{j}}^{(\rho)} = b_{\mathbf{j}}^{(\rho)}/q + \gamma_{\mathbf{j}}^{(\rho)}$$

and of a similar quality hold. For sufficiently small θ this is indeed possible, but in order to homogenise the set of major arcs, we have to surmount some technical difficulties.

Define

$$q = \mathrm{lcm}_{\mathbf{j} \in J}\{q_{\mathbf{j}}\} \quad \text{and} \quad b_{\mathbf{j}}^{(\rho)} = a_{\mathbf{j}}^{(\rho)} q/q_{\mathbf{j}} \quad (\mathbf{j} \in J, 1 \leq \rho \leq R), \tag{2.5.1}$$

and note that $\gcd(\underline{\mathbf{b}}, q) = 1$.

Lemma 2.7. *Let q and $b_{\mathbf{j}}^{(\rho)}$ be as above. There exist integer weights $\lambda_{\mathbf{j}}^{(\rho)}$ for all $\mathbf{j} \in J$ and $1 \leq \rho \leq R$ such that $\lambda_{\mathbf{j}}^{(\rho)} \leq q_{\mathbf{j}}$ and*

$$\gcd\Big(\sum_{\mathbf{j},\rho} \lambda_{\mathbf{j}}^{(\rho)} b_{\mathbf{j}}^{(\rho)}, q\Big) = 1. \tag{2.5.2}$$

Proof. By Euclid's algorithm there exist parameters $\lambda_{\mathbf{j}}^{(\rho)}$ such that

$$\sum_{\mathbf{j},\rho} \lambda_{\mathbf{j}}^{(\rho)} b_{\mathbf{j}}^{(\rho)} = \gcd_{\mathbf{j},\rho}\{b_{\mathbf{j}}^{(\rho)}\}$$

and hence

$$\gcd\Big(\sum_{\mathbf{j},\rho} \lambda_{\mathbf{j}}^{(\rho)} b_{\mathbf{j}}^{(\rho)}, q\Big) = \gcd(\underline{\mathbf{b}}, q) = 1.$$

These $\lambda_{\mathbf{j}}^{(\rho)}$ really live modulo $q_{\mathbf{j}}$, since by writing $\lambda_{\mathbf{j}}^{(\rho)} = c_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} + \mu_{\mathbf{j}}^{(\rho)}$ and recalling (2.5.1) one has

$$b_{\mathbf{j}}^{(\rho)} \lambda_{\mathbf{j}}^{(\rho)} = \frac{a_{\mathbf{j}}^{(\rho)} q}{q_{\mathbf{j}}} \Big(c_{\mathbf{j}}^{(\rho)} q_{\mathbf{j}} + \mu_{\mathbf{j}}^{(\rho)}\Big) \equiv \frac{a_{\mathbf{j}}^{(\rho)} q}{q_{\mathbf{j}}} \mu_{\mathbf{j}}^{(\rho)} \equiv b_{\mathbf{j}}^{(\rho)} \mu_{\mathbf{j}}^{(\rho)} \pmod{q}.$$

This allows us to take $\lambda_{\mathbf{j}}^{(\rho)} \leq q_{\mathbf{j}}$. □

The following Lemma, which is essentially a higher-dimensional analogue of [9, Lemma 2.2], may be useful in other contexts, so we will state it in a rather general fashion.

Lemma 2.8. *Let $\mathcal{B}$ be the image of $[-P, P]^n$ under some integral non-singular linear transformation A, and assume that $\mathcal{B} \subset [-X, X]^n$ for some X. Furthermore, let*

$$T(\boldsymbol{\alpha}, \mathcal{B}) = \sum_{\mathbf{x} \in \mathcal{B}} e(\boldsymbol{\alpha} \cdot \mathbf{F}(\mathbf{x}))$$

be a multidimensional exponential sum over $\mathcal{B}$. Then we have

$$\big|T(\boldsymbol{\alpha}, \mathcal{B})\big| \ll (\log P)^n \sup_{\boldsymbol{\eta} \in [0,1)^n} \big|H(\boldsymbol{\alpha}, \boldsymbol{\eta}; X)\big|,$$

where $H(\boldsymbol{\alpha}, \boldsymbol{\eta}; X)$ is defined as

$$H(\boldsymbol{\alpha}, \boldsymbol{\eta}; X) = \sum_{|\mathbf{x}| \leq X} e\big(\boldsymbol{\alpha} \cdot \mathbf{F}(\mathbf{x}) - \boldsymbol{\eta} \cdot \mathbf{x}\big).$$

The important thing to notice here is that $H(\boldsymbol{\alpha}, \boldsymbol{\eta}; P)$ is nothing but the usual exponential sum $T(\boldsymbol{\alpha}, P)$ with a linear twist characterised by the parameter $\boldsymbol{\eta} \in [0, 1)^n$. Thus its behaviour will not essentially differ from that of the usual exponential sum, and the two can be regarded as roughly the same object. This means that Lemma 2.8 enables us to treat exponential sums over rather more general convex sets than standard rectangular boxes.

Proof. By the orthogonality relations, one has

$$\sum_{\mathbf{y}\in\mathcal{B}} e\left(\boldsymbol{\alpha}\cdot\mathbf{F}(\mathbf{y})\right) = \sum_{|\mathbf{x}|\leq X} e\left(\boldsymbol{\alpha}\cdot\mathbf{F}(\mathbf{x})\right)\sum_{\mathbf{y}\in\mathcal{B}}\int_{[0,1)^n} e(\boldsymbol{\eta}\cdot\mathbf{y}-\boldsymbol{\eta}\cdot\mathbf{x})\mathrm{d}\boldsymbol{\eta}$$
$$= \int_{[0,1)^n}\sum_{|\mathbf{x}|\leq X} e\left(\boldsymbol{\alpha}\cdot\mathbf{F}(\mathbf{x})-\boldsymbol{\eta}\cdot\mathbf{x}\right)\sum_{\mathbf{y}\in\mathcal{B}} e(\boldsymbol{\eta}\cdot\mathbf{y})\mathrm{d}\boldsymbol{\eta}.$$

Thus if we write

$$D(\boldsymbol{\eta},\mathcal{B}) = \sum_{\mathbf{y}\in\mathcal{B}} e(\boldsymbol{\eta}\cdot\mathbf{y}),$$

the exponential sum $T(\boldsymbol{\alpha},\mathcal{B})$ can be expressed in terms of $H(\boldsymbol{\alpha},\boldsymbol{\eta};X)$ as

$$T(\boldsymbol{\alpha},\mathcal{B}) = \int_{[0,1)^n} H(\boldsymbol{\alpha},\boldsymbol{\eta};X)D(\boldsymbol{\eta},\mathcal{B})\mathrm{d}\boldsymbol{\eta}$$
$$\ll \sup_{\boldsymbol{\eta}\in[0,1)^n}\left|H(\boldsymbol{\alpha},\boldsymbol{\eta};X)\right|\int_{[0,1)^n} D(\boldsymbol{\eta},\mathcal{B})\mathrm{d}\boldsymbol{\eta}.$$

Now the $\mathbf{y}$ are in the image of $[-P,P]^n$ under A and can therefore be written as $\mathbf{y}=A\mathbf{x}$ with $|\mathbf{x}|\leq P$. This implies that

$$\boldsymbol{\eta}\cdot\mathbf{y} = \boldsymbol{\eta}\cdot A\mathbf{x} = A^t\boldsymbol{\eta}\cdot\mathbf{x}$$

and consequently

$$D(\boldsymbol{\eta},\mathcal{B}) = D(A^t\boldsymbol{\eta},[-P,P]^n) \ll \prod_{i=1}^n \min(P,\|(A^t\boldsymbol{\eta})_i\|^{-1}).$$

It follows that

$$\int_{[0,1)^n} D(\boldsymbol{\eta},\mathcal{B})\mathrm{d}\boldsymbol{\eta} \ll \int_{[0,1)^n}\prod_{i=1}^n \min(P,\|(A^t\boldsymbol{\eta})_i\|^{-1})\mathrm{d}\boldsymbol{\eta}$$
$$\ll (\det A)^{-1}\int_{\mathcal{C}}\prod_{i=1}^n \min(P,\|\eta_i\|^{-1})\mathrm{d}\boldsymbol{\eta},$$

where $\mathcal{C}$ is the image of $[0,1)^n$ under A^t. Obviously, the integrand is positive and 1-periodic in every direction, so we can bound the integral over $\mathcal{C}$ by a number of copies of the integral over the unit cube. Since A is integral, $\mathcal{C}$ is a parallelepiped with integral vertices, and by the 1-periodicity of the integrand we may replace the domain by $\mathrm{vol}(\mathcal{C}) = \det(A)$ copies of the unit cube. Hence we obtain

$$\int_{\mathcal{C}} \prod_{i=1}^{n} \min(P, \|\eta_i\|^{-1}) \mathrm{d}\boldsymbol{\eta} \ll \det A \int_{[0,1)^n} \prod_{i=1}^{n} \min(P, \|\eta_i\|^{-1}) \mathrm{d}\boldsymbol{\eta}$$
$$\ll \det A (\log P)^n.$$

This gives the result. □

We remark that the identical analysis can be used to show

$$\sum_{\mathbf{x} \in \mathcal{B}} e(a_{\mathbf{x}}) \ll (\log P)^n \sup_{\boldsymbol{\eta} \in [0,1)^n} \left| \sum_{|\mathbf{x}| \leq X} e(a_{\mathbf{x}} - \boldsymbol{\eta} \cdot \mathbf{x}) \right|,$$

where $\mathcal{B}$, P and X are as before and the arguments $a_{\mathbf{x}} \in \mathbb{C}$ are arbitrary.

We have now collected the technical tools necessary in order to homogenise our set of major arcs, which allows us to proceed and prove a homogenised version of Lemma 2.5.

Lemma 2.9. *Assume that* $\theta R(d-1)(r+3) < d$. *Under the conditions of Lemma 2.5, we can replace alternative* (B) *by*

(B′) There exists an integer $0 < q \ll P^{2(d-1)R\theta}$ *such that one finds* $\underline{\mathbf{a}} \in \mathbb{Z}^{Rr}$, *satisfying*

$$\left| \alpha_{\mathbf{j}}^{(\rho)} q - a_{\mathbf{j}}^{(\rho)} \right| \ll P^{-d+3(d-1)R\theta} \quad (1 \leq \rho \leq R, \mathbf{j} \in J).$$

Proof. Make the substitution

$$\mathbf{x}'_k = \begin{cases} \mathbf{x}_k - \mathbf{x}_{k+1} & \text{if } 1 \leq k \leq m-1, \\ \mathbf{x}_m & \text{if } k = m, \end{cases} \tag{2.5.3}$$

so that

$$\mathbf{x}_k = \sum_{i=k}^{m} \mathbf{x}'_i \quad (1 \leq k \leq m).$$

Furthermore, observe that the index set J is equipped with a partial order relation induced by entrywise comparison, i.e. $\mathbf{j} \leq \mathbf{j}'$ for two elements $\mathbf{j}, \mathbf{j}' \in J$ if and only if $j_k \leq j'_k$ for all $1 \leq k \leq d$.

Since the proof of Lemma 2.4 produces the same denominators $q_{\mathbf{j}}$ independently of the ρ-component of the coefficients $\alpha_{\mathbf{j}}^{(\rho)}$, it suffices, without loss of generality, to consider only the case $\rho = 1$. This allows us to avoid unnecessary complexity of the notation by dropping the index.

Observe that

$$\Phi(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_d}) = \Phi\Big(\sum_{k_1 \geq j_1} \mathbf{x}'_{k_1}, \ldots, \sum_{k_d \geq j_d} \mathbf{x}'_{k_d}\Big) = \sum_{\mathbf{k} \geq \mathbf{j}} \Phi(\mathbf{x}'_{k_1}, \ldots, \mathbf{x}'_{k_d}).$$

Now consider the weighted exponential sum

$$T(\boldsymbol{\lambda\alpha}, P) = \sum_{\substack{|\mathbf{x}_i| \leq P \\ 1 \leq i \leq m}} e\left(\sum_{\mathbf{j} \in J} \lambda_{\mathbf{j}} \alpha_{\mathbf{j}} \Phi(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_d})\right),$$

where the $\lambda_{\mathbf{j}}$ will be fixed later. This can be expressed in terms of the alternative variables $\overline{\mathbf{x}}'$ and yields

$$\begin{aligned} T(\boldsymbol{\lambda\alpha}, P) &= \sum_{\substack{\mathbf{x}'_i \\ 1 \leq i \leq m}} e\left(\sum_{\mathbf{j} \in J} \lambda_{\mathbf{j}} \alpha_{\mathbf{j}} \sum_{\mathbf{k} \geq \mathbf{j}} \Phi(\mathbf{x}'_{k_1}, \ldots, \mathbf{x}'_{k_d})\right) \\ &= \sum_{\substack{\mathbf{x}'_i \\ 1 \leq i \leq m}} e\left(\sum_{\mathbf{k} \in J} \Big(\sum_{\mathbf{j} \leq \mathbf{k}} \lambda_{\mathbf{j}} \alpha_{\mathbf{j}}\Big) \Phi(\mathbf{x}'_{k_1}, \ldots, \mathbf{x}'_{k_d})\right), \end{aligned}$$

where the sum over $\mathbf{x}'_i$ is over domains $\mathcal{B}_i$ contained in $[-2P, 2P]^s$ as determined by (2.5.3). Thus if we define $\boldsymbol{\alpha}'$ by $\alpha'_{\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{j}} \lambda_{\mathbf{k}} \alpha_{\mathbf{k}}$, this yields the identity $T(\boldsymbol{\lambda\alpha}, P) = T(\boldsymbol{\alpha}', \prod_i \mathcal{B}_i)$.

Beware that the domains $\mathcal{B}_i$ are not independent of one another, so one has to apply great care in exchanging the order of summation, and this affects our possibilities of applying Weyl's inequality severely. However, since the transformation (2.5.3) is integral and non-singular and maps $[-P, P]^{ms}$ onto a subset of $[-2P, 2P]^{ms}$, Lemma 2.8 comes to our rescue and yields

$$|T(\boldsymbol{\lambda\alpha}, P)| = |T(\boldsymbol{\alpha}', \textstyle\prod_i \mathcal{B}_i)| \ll (\log P)^{ms} \sup_{\overline{\boldsymbol{\eta}} \in [0,1)^{ms}} \left| H(\boldsymbol{\alpha}', \overline{\boldsymbol{\eta}}; 2P) \right|. \tag{2.5.4}$$

The exponential sum $T(\boldsymbol{\alpha}', P)$ and its twisted cousin $H(\boldsymbol{\alpha}', \overline{\boldsymbol{\eta}}; P)$ should be thought of as being roughly of the same order of magnitude. Indeed, since

$$\Delta_{j_2, \mathbf{h}_2} \Delta_{j_1, \mathbf{h}_1} (\overline{\boldsymbol{\eta}} \cdot \overline{\mathbf{x}}) = \Delta_{j_2, \mathbf{h}_2} (\boldsymbol{\eta}_{j_1} \cdot \mathbf{h}_1) = 0,$$

any linear twist has no effect in the deduction of Weyl's inequality, and a small modification of the proof of Lemma 2.1 readily shows that Lemma 2.2 continues to hold if $T(\boldsymbol{\alpha}', P)$ is replaced by $H(\boldsymbol{\alpha}', \overline{\boldsymbol{\eta}}; P)$ as long as $d \geq 2$. This implies that all following estimates of Section 2.3 will remain unaffected by the twist. In particular, the minor arcs estimate will hold for $H(\boldsymbol{\alpha}', \overline{\boldsymbol{\eta}}; P)$ if and only if it does so for $T(\boldsymbol{\alpha}', P)$. Collecting these arguments together, one concludes that $|T(\boldsymbol{\alpha}', P)| \ll P^{ms-k\theta}$ implies that $|H(\boldsymbol{\alpha}', \overline{\boldsymbol{\eta}}; P)| \ll P^{ms-k\theta}$ for arbitrary $\overline{\boldsymbol{\eta}}$, and by (2.5.4) it follows that $|T(\boldsymbol{\lambda\alpha}, P)| \ll P^{ms-k\theta}(\log P)^{ms}$. Since the singular case is excluded, this in turn means that whenever $\boldsymbol{\lambda\alpha}$ possesses an approximation as in case (B) of Lemma 2.5, then so does $\boldsymbol{\alpha}'$.

Now suppose that $\boldsymbol{\alpha} \in \mathfrak{M}(\theta)$ for some θ. Then it has approximations $\alpha_{\mathbf{j}} = a_{\mathbf{j}}/q_{\mathbf{j}} + \beta_{\mathbf{j}}$, and according to Lemma 2.7 we can find integer weights $\lambda_{\mathbf{j}} \leq q_{\mathbf{j}} \ll P^{(d-1)R\theta}$ satisfying (2.5.2). One has

$$\lambda_{\mathbf{j}}\alpha_{\mathbf{j}} = \lambda_{\mathbf{j}}\frac{a_{\mathbf{j}}}{q_{\mathbf{j}}} + \lambda_{\mathbf{j}}\beta_{\mathbf{j}},$$

so $\boldsymbol{\lambda\alpha}$ is certainly contained in $\mathfrak{M}(2\theta)$. By the above considerations this implies that we have rational approximations for the components $\alpha'_{\mathbf{j}}$ of $\boldsymbol{\alpha}'$, given by $a'_{\mathbf{j}}$ and $q'_{\mathbf{j}} \ll P^{2(d-1)R\theta}$ such that $|\alpha'_{\mathbf{j}}q'_{\mathbf{j}} - a'_{\mathbf{j}}| \ll P^{-d+2(d-1)R\theta}$.

Consider the last term $\alpha'_{\mathbf{m}} = \alpha'_{m,\dots,m}$. We have the approximation

$$\alpha'_{\mathbf{m}} = \frac{a'_{\mathbf{m}}}{q'_{\mathbf{m}}} + O\left(\frac{P^{-d+2(d-1)R\theta}}{q'_{\mathbf{m}}}\right). \tag{2.5.5}$$

On the other hand, $\mathbf{m} \geq \mathbf{j}$ for all $\mathbf{j} \in J$ with respect to the partial order of the $\mathbf{j}$, so inserting the definition of $\alpha'_{\mathbf{m}}$ gives the alternative approximation

$$\begin{aligned}\alpha'_{\mathbf{m}} &= \sum_{\mathbf{j}\in J} \alpha_{\mathbf{j}}\lambda_{\mathbf{j}} = \sum_{\mathbf{j}\in J}\left(\frac{\lambda_{\mathbf{j}}a_{\mathbf{j}}}{q_{\mathbf{j}}} + O\left(\frac{P^{-d+2R(d-1)\theta}}{q_{\mathbf{j}}}\right)\right)\\ &= \frac{\sum_{\mathbf{j}\in J}\lambda_{\mathbf{j}}b_{\mathbf{j}}}{q} + O\left(\frac{P^{-d+(r+1)(d-1)R\theta}}{q}\right). \end{aligned} \tag{2.5.6}$$

Our goal is now to show that the two approximations (2.5.5) and (2.5.6) are actually the same. However, if they are distinct, it follows that

$$\begin{aligned}\frac{1}{q'_{\mathbf{m}}q} &\le \left|\frac{a'_{\mathbf{m}}}{q'_{\mathbf{m}}} - \frac{\sum_{\mathbf{j}\in J}\lambda_{\mathbf{j}}b_{\mathbf{j}}}{q}\right| \\ &\le \left|\alpha'_{\mathbf{m}} - \frac{a'_{\mathbf{m}}}{q'_{\mathbf{m}}}\right| + \left|\alpha'_{\mathbf{m}} - \frac{\sum_{\mathbf{j}\in J}\lambda_{\mathbf{j}}b_{\mathbf{j}}}{q}\right| \\ &\ll \frac{qP^{-d+2(d-1)R\theta} + q'_{\mathbf{m}}P^{-d+(r+1)(d-1)R\theta}}{q'_{\mathbf{m}}q} \\ &\ll \frac{P^{-d+(r+3)(d-1)R\theta}}{q'_{\mathbf{m}}q},\end{aligned}$$

which is a contradiction if θ is sufficiently small. Choosing θ in accordance with the hypothesis of the statement of the lemma, we can thus conclude that (2.5.5) and (2.5.6) coincide. Lemma 2.7 now ensures that $\gcd(\sum_{\mathbf{j}\in J}\lambda_{\mathbf{j}}b_{\mathbf{j}}, q) = 1$, so the above approximations are both reduced fractions and one has the common denominator $q = q'_{\mathbf{m}} \ll P^{2(d-1)R\theta}$.

Finally, the bound on $|q\alpha_{\mathbf{j}} - b_{\mathbf{j}}|$ follows by observing that

$$|\alpha_{\mathbf{j}}q - b_{\mathbf{j}}| = \frac{q}{q_{\mathbf{j}}}|\alpha_{\mathbf{j}}q_{\mathbf{j}} - a_{\mathbf{j}}| \le q|\alpha_{\mathbf{j}}q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{2(d-1)R\theta}P^{-d+(d-1)R\theta}.$$

This yields the statement. □

It should be noted here that the factors 2 and 3 that arise in the homogenising process probably have little right to exist at all, and that one would expect them to succumb to a more momentous argument than the one we have been presenting here. However, the important accomplishment of this section is to avoid collecting another factor r in the homogenising process, as that would undo the gains from Section 2.3 and throw us back into the situation of treating Equation (2.1.3) as a system in rR variables without regard for symmetries.

2.6 Generating functions analysis

Our goal in this section is to show that the major arcs contribution can be interpreted as a product of local densities. In order to do so, it is somewhat inconvenient that the estimate of $\underline{\boldsymbol{\beta}} = |\underline{\boldsymbol{\alpha}} - \underline{\mathbf{a}}/q|$ depends on q. We therefore extend the major arcs slightly and define our final choice of major arcs $\mathfrak{M}'(P,\theta)$ to be set of all $\underline{\boldsymbol{\alpha}} = \underline{\mathbf{a}}/q + \underline{\boldsymbol{\beta}}$ contained in $[0,1)^{Rr}$ that satisfy

$$|\underline{\boldsymbol{\beta}}| \leq C' P^{-d+3(d-1)R\theta} \qquad \text{and} \qquad 0 \leq \underline{\mathbf{a}} < q \leq C' P^{2(d-1)R\theta} \tag{2.6.1}$$

for some suitably large constant C'. As before, we will omit the parameter P in most cases. Notice that by Lemma 2.9 this definition comprises the original major arcs as defined in (2.4.1). Henceforth all parameters $\underline{\boldsymbol{\alpha}}, \underline{\mathbf{a}}, q, \underline{\boldsymbol{\beta}}$ will be implicitly understood to satisfy the major arcs inequalities as given in (2.6.1).

Letting

$$S_q(\underline{\mathbf{a}}) = \sum_{\overline{\mathbf{x}}=1}^{q} e\left(\frac{\mathfrak{F}(\overline{\mathbf{x}}; \underline{\mathbf{a}})}{q}\right) \tag{2.6.2}$$

and

$$v_P(\underline{\boldsymbol{\beta}}) = \int_{|\overline{\boldsymbol{\xi}}| \leq P} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \underline{\boldsymbol{\beta}})\right) \mathrm{d}\overline{\boldsymbol{\xi}}, \tag{2.6.3}$$

we can replace the exponential sum by an expression that reflects the rational approximation to $\underline{\boldsymbol{\alpha}}$ and will be easier to handle.

Lemma 2.10. *Assume that $\underline{\boldsymbol{\alpha}} \in \mathfrak{M}'(P,\theta)$. Then there exists an integer vector $(\underline{\mathbf{a}}, q)$ such that*

$$\begin{aligned} T(\underline{\boldsymbol{\alpha}}) - q^{-ms} S_q(\underline{\mathbf{a}}) v_P(\underline{\boldsymbol{\beta}}) &\ll q^{ms}\left(1 + (Pq^{-1})^{ms} q P^{d-1} |\underline{\boldsymbol{\beta}}|\right) \\ &\ll P^{ms-1+5(d-1)R\theta}. \end{aligned}$$

Proof. The first estimate is essentially like in [48, Lemma 8.1]. Sorting the terms in arithmetic progressions modulo q, we have

$$T(\underline{\boldsymbol{\alpha}}) = \sum_{\overline{\mathbf{z}} \bmod q} e(\mathfrak{F}(\overline{\mathbf{z}}; \underline{\mathbf{a}}/q)) \sum_{\substack{\overline{\mathbf{w}} \in \mathbb{Z}^{ms} \\ q\overline{\mathbf{w}} + \overline{\mathbf{z}} \leq P}} e(\mathfrak{F}(q\overline{\mathbf{w}} + \overline{\mathbf{z}}; \underline{\boldsymbol{\beta}})),$$

and thus the difference $T(\underline{\boldsymbol{\alpha}}) - q^{-ms} S_q(\underline{\boldsymbol{\alpha}}) v_P(\underline{\boldsymbol{\beta}})$ is given by

$$\sum_{\overline{\mathbf{z}} \bmod q} e(\mathfrak{F}(\overline{\mathbf{z}}; \mathbf{a}/q)) H(q, \boldsymbol{\beta}, \overline{\mathbf{z}}),$$

where

$$H(q,\boldsymbol{\beta},\overline{\mathbf{z}}) = \sum_{\substack{\overline{\mathbf{w}}\in\mathbb{Z}^{ms}\\ q\overline{\mathbf{w}}+\overline{\mathbf{z}}\le P}} e(\mathfrak{F}(q\overline{\mathbf{w}}+\overline{\mathbf{z}};\underline{\boldsymbol{\beta}})) - q^{-ms}\int_{\overline{\boldsymbol{\xi}}\le P} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}};\underline{\boldsymbol{\beta}})\right)\mathrm{d}\overline{\boldsymbol{\xi}}$$
$$= \sum_{\substack{\overline{\mathbf{w}}\in\mathbb{Z}^{ms}\\ q\overline{\mathbf{w}}+\overline{\mathbf{z}}\le P}} \int_{\overline{\mathbf{w}}}^{\overline{\mathbf{w}}+1} \left\{e(\mathfrak{F}(q\overline{\mathbf{w}}+\overline{\mathbf{z}};\underline{\boldsymbol{\beta}})) - e(\mathfrak{F}(q\overline{\boldsymbol{\xi}}+\overline{\mathbf{z}};\underline{\boldsymbol{\beta}})) + O(1)\right\}\mathrm{d}\overline{\boldsymbol{\xi}}.$$

By the Mean Value Theorem, the integral is $\ll q(P^{d-1}\sum_{\mathbf{j}\in J}|\underline{\boldsymbol{\beta}}|)$. Furthermore, the sum over $\overline{\mathbf{w}}$ is $\ll (Pq^{-1})^{ms}$, and the remaining term is just $S_q(\underline{\boldsymbol{\alpha}})$ and can be bounded trivially by q^{ms}. Therefore, we have

$$T(\underline{\boldsymbol{\alpha}}) - q^{-ms}S_q(\underline{\mathbf{a}})v_P(\underline{\boldsymbol{\beta}}) \ll q^{ms}\left(1 + (Pq^{-1})^{ms}qP^{d-1}\sum_{\mathbf{j}\in J}|\underline{\boldsymbol{\beta}}|\right)$$

as claimed, and the second inequality follows on inserting the bounds on q and $\underline{\boldsymbol{\beta}}$. Notice that we can assume θ to be small enough for the first term q^{ms} to be negligible. $\square$

The next step is to integrate the expression from Lemma 2.10 over $\mathfrak{M}'(P,\theta)$ in order to determine the overall error arising from this substitution. For this purpose, define the truncated singular series and singular integral as

$$\mathfrak{S}_{\psi}(P) = \sum_{q=1}^{C'P^{2(d-1)R\theta}} q^{-ms}\sum_{\substack{\underline{\mathbf{a}}=0\\ (\underline{\mathbf{a}},q)=1}}^{q-1} S_q(\underline{\mathbf{a}})e\left(-(\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\alpha}})/q\right)$$

and

$$\mathfrak{J}_{\psi}(P) = \int_{|\underline{\boldsymbol{\beta}}|\le C'P^{-d+3(d-1)R\theta}} v_P(\underline{\boldsymbol{\beta}})e(-\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\beta}})\mathrm{d}\underline{\boldsymbol{\beta}},$$

respectively.

Lemma 2.11. *The total major arcs contribution is given by*

$$\int_{\mathfrak{M}'(\theta)} T(\underline{\boldsymbol{\alpha}})e(-\underline{\boldsymbol{\alpha}}\cdot\underline{\mathbf{n}})\mathrm{d}\underline{\boldsymbol{\alpha}} = \mathfrak{S}_{\psi}(P)\mathfrak{J}_{\psi}(P) + O\left(P^{ms-Rrd+(d-1)R(5Rr+7)\theta-1}\right).$$

Proof. The volume of the extended major arcs $\mathfrak{M}'(\theta)$ is bounded by

$$\begin{aligned}\operatorname{vol}(\mathfrak{M}'(\theta)) &\ll \sum_{q=1}^{C'P^{2R(d-1)\theta}} \prod_{\mathbf{j}\in J}\prod_{\rho=1}^{R}\Big(\sum_{a_{\mathbf{j}}^{(\rho)}=0}^{q-1} P^{-d+3R(d-1)\theta}\Big)\\ &\ll \sum_{q=1}^{C'P^{2R(d-1)\theta}} \left(qP^{-d+3R(d-1)\theta}\right)^{rR}\\ &\ll P^{-Rrd+(d-1)R(5Rr+2)\theta}.\end{aligned}$$

Together with Lemma 2.10, this implies that

$$\begin{aligned}&\sum_{q=1}^{C'P^{2(d-1)R\theta}} \sum_{\substack{\underline{\mathbf{a}}=0\\(\underline{\mathbf{a}},q)=1}}^{q-1} \int_{|\underline{\boldsymbol{\beta}}|\leq P^{-d+3(d-1)R\theta}} \Big(T(\underline{\mathbf{a}}/q+\underline{\boldsymbol{\beta}}) - q^{-ms}S_q(\underline{\mathbf{a}})v_P(\underline{\boldsymbol{\beta}})\Big)\mathrm{d}\underline{\boldsymbol{\beta}}\\ &\ll \sup_{\underline{\boldsymbol{\alpha}}=\underline{\mathbf{a}}/q+\underline{\boldsymbol{\beta}}\in\mathfrak{M}'(\theta)} \left|T(\underline{\mathbf{a}}/q+\underline{\boldsymbol{\beta}}) - q^{-ms}S_q(\underline{\mathbf{a}})v_P(\underline{\boldsymbol{\beta}})\right| \times \operatorname{vol}(\mathfrak{M}'(\theta))\\ &\ll P^{ms-1+5(d-1)R\theta}P^{-Rrd+(d-1)R(5Rr+2)\theta}.\end{aligned}$$

This proves the statement. □

We can now fix θ such that $(d-1)R(5Rr+7)\theta < 1$ for the rest of our considerations and note that with this choice Lemmata 2.6, 2.9 and 2.11 are applicable, yielding

$$N_{s,\boldsymbol{\psi}}(P) = \mathfrak{S}_{\boldsymbol{\psi}}(P)\mathfrak{J}_{\boldsymbol{\psi}}(P) + o(P^{ms-Rrd})$$

if the number of variables is large enough.

The truncated singular series and integral can be extended to infinity. Recalling the definition in (2.6.3) and (2.2.4) and using the homogeneity of $\mathbf{F}$, integration by parts reveals that

$$\begin{aligned}v_P(\underline{\boldsymbol{\beta}}) &= \int_{|\overline{\boldsymbol{\xi}}|\leq P} e\left(\sum_{\rho=1}^{R}\sum_{\mathbf{j}\in J} \beta_{\mathbf{j}}^{(\rho)}\Phi_{\mathbf{j}}^{(\rho)}(\boldsymbol{\xi})\right)\mathrm{d}\overline{\boldsymbol{\xi}}\\ &= P^{ms}\int_{|\overline{\boldsymbol{\xi}}|\leq 1} e\left(\sum_{\rho=1}^{R}\sum_{\mathbf{j}\in J} P^d\beta_{\mathbf{j}}^{(\rho)}\Phi_{\mathbf{j}}^{(\rho)}(\boldsymbol{\xi})\right)\mathrm{d}\overline{\boldsymbol{\xi}}\\ &= P^{ms}v_1(P^d\underline{\boldsymbol{\beta}}).\end{aligned} \tag{2.6.4}$$

It follows that after another integration by parts one has

$$\begin{aligned}\mathfrak{J}_{\boldsymbol{\psi}}(P) &= \int_{|\underline{\boldsymbol{\beta}}|\le C'P^{-d+3(d-1)R\theta}} v_P(\underline{\boldsymbol{\beta}})e(-\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\beta}})\mathrm{d}\underline{\boldsymbol{\beta}} \\ &= P^{ms}\int_{|\underline{\boldsymbol{\beta}}|\le C'P^{-d+3(d-1)R\theta}} v_1(P^d\underline{\boldsymbol{\beta}})e(-\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\beta}})\mathrm{d}\underline{\boldsymbol{\beta}} \\ &= P^{ms-Rrd}\int_{|\underline{\boldsymbol{\beta}}|\le C'P^{3(d-1)R\theta}} v_1(\underline{\boldsymbol{\beta}})e\left(-(\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\beta}})/P^d\right)\mathrm{d}\underline{\boldsymbol{\beta}}.\end{aligned}$$

We therefore define, if existent, the complete singular series $\mathfrak{S}_{\boldsymbol{\psi}}$ and the singular integral $\mathfrak{J}_{\boldsymbol{\psi}}$ as

$$\begin{aligned}\mathfrak{S}_{\boldsymbol{\psi}} &= \sum_{q=1}^{\infty}\sum_{\substack{\underline{\mathbf{a}}=0\\(\underline{\mathbf{a}},q)=1}}^{q-1} q^{-ms}S_q(\underline{\mathbf{a}})e\left(-(\underline{\mathbf{n}}\cdot\underline{\mathbf{a}})/q\right), \\ \mathfrak{J}_{\boldsymbol{\psi}} &= \int_{\mathbb{R}^{rR}} v_1(\underline{\boldsymbol{\beta}})e\left(-(\underline{\mathbf{n}}\cdot\underline{\boldsymbol{\beta}})/P^d\right)\mathrm{d}\underline{\boldsymbol{\beta}}.\end{aligned}$$

In either case, convergence implies that the errors $|\mathfrak{S}_{\boldsymbol{\psi}}(P)-\mathfrak{S}_{\boldsymbol{\psi}}|$ and $\left|\mathfrak{J}_{\boldsymbol{\psi}}(P)-P^{ms-Rrd}\mathfrak{J}_{\boldsymbol{\psi}}\right|$ are $o(1)$, and we will be able to replace the statement in Lemma 2.11 by

$$\int_{\mathfrak{M}'(P,\theta)} T(\underline{\boldsymbol{\alpha}})e(\underline{\boldsymbol{\alpha}}\cdot\underline{\mathbf{n}})\mathrm{d}\underline{\boldsymbol{\alpha}} = P^{ms-Rrd}\mathfrak{J}_{\psi}\mathfrak{S}_{\psi} + o\left(P^{ms-Rrd}\right).$$

It remains to show that the above definitions are permissible.

Lemma 2.12. *Suppose $k > 3(d-1)RW$ for some $W > r$, and assume* (2.3.7) *is satisfied. Then*

$$v_1(\underline{\boldsymbol{\beta}}) \ll \left(1+|\underline{\boldsymbol{\beta}}|\right)^{-W}.$$

Proof. By (2.6.4), we see that the equation

$$v_1(\underline{\boldsymbol{\beta}}) = Q^{-ms}v_Q(Q^{-d}\underline{\boldsymbol{\beta}})$$

holds for arbitrary Q. For

$$0 < \theta_1 \le \frac{d}{(d-1)R(r+3)} \tag{2.6.5}$$

(to be determined later) and for a given $\underline{\boldsymbol{\beta}} \in \mathbb{R}^{rR}$ choose Q such that for a suitable constant C one has

$$CQ^{3(d-1)R\theta_1} = |\underline{\boldsymbol{\beta}}|. \tag{2.6.6}$$

Notice that the condition (2.6.5) on θ_1 allows us to apply Lemma 2.9, so we can assume the major arcs to be homogenised. Hence with this choice of Q, the argument $Q^{-d}\underline{\boldsymbol{\beta}}$ lies just on the edge of the corresponding major arcs $\mathfrak{M}'(Q,\theta_1)$; in fact, it is best approximated by $\underline{\mathbf{a}} = \underline{\mathbf{0}}$ and $q = 1$, and one has $q^{-ms}S_1(\underline{\mathbf{0}}) = 1$. An application of Lemma 2.10 yields

$$v_Q(Q^{-d}\underline{\boldsymbol{\beta}}) = T(Q^{-d}\underline{\boldsymbol{\beta}}) + O\left(Q^{ms-1+5(d-1)R\theta_1}\right).$$

On the other hand, since $Q^{-d}\underline{\boldsymbol{\beta}}$ lies on the edge of the extended major arcs $\mathfrak{M}'(Q,\theta_1)$, it is not contained in the original set $\mathfrak{M}(Q,\theta_1)$ of major arcs. We can therefore bound the exponential sum $T(Q^{-d}\underline{\boldsymbol{\beta}})$ by the minor arcs estimate and find

$$T(Q^{-d}\underline{\boldsymbol{\beta}}) \ll Q^{ms-k\theta_1} \ll Q^{ms-3(d-1)RW\theta_1}.$$

This gives

$$v_1(\underline{\boldsymbol{\beta}}) \ll Q^{-ms}\left(Q^{ms-3(d-1)RW\theta_1} + Q^{ms-1+5(d-1)R\theta_1}\right),$$

which is optimised by picking

$$\theta_1^{-1} = 5(d-1)R + 3(d-1)RW = R(d-1)(3W+5).$$

Notice that this choice satisfies (2.6.5) as we assumed $W > r$. This allows us to rewrite (2.6.6) in the shape

$$CQ^{3/(3W+5)} = |\underline{\boldsymbol{\beta}}|,$$

whence the bound on v_1 is

$$\begin{aligned} v_1(\underline{\boldsymbol{\beta}}) &\ll \left(Q^{-3(d-1)RW\theta_1} + Q^{-1+5(d-1)R\theta_1}\right) \\ &\ll Q^{-3W/(3W+5)} \ll |\underline{\boldsymbol{\beta}}|^{-W}. \end{aligned}$$

Furthermore, one has trivially $v_1(\underline{\boldsymbol{\beta}}) \ll 1$, so on taking the maximum one retrieves the statement. □

Having thus established the convergence of the singular integral, it remains to analyse the behaviour of the singular series. As a first step in that direction, we note that by standard arguments one has

$$S_{q_1}(\underline{\mathbf{a}}_1)S_{q_2}(\underline{\mathbf{a}}_2) = S_{q_1q_2}(q_2\underline{\mathbf{a}}_1 + q_1\underline{\mathbf{a}}_2)$$

for coprime q_1, q_2, so the argument $q_2\underline{\mathbf{a}}_1 + q_1\underline{\mathbf{a}}_2$ runs over a full set of residues modulo q_1q_2 and we have the multiplicativity property

$$\sum_{\substack{\underline{\mathbf{a}}_1=1 \\ (\underline{\mathbf{a}}_1,q_1)=1}}^{q_1} S_{q_1}(\underline{\mathbf{a}}_1) \sum_{\substack{\underline{\mathbf{a}}_2=1 \\ (\underline{\mathbf{a}}_2,q_2)=1}}^{q_2} S_{q_2}(\underline{\mathbf{a}}_2) = \sum_{\substack{\underline{\mathbf{b}}=1 \\ (\underline{\mathbf{b}},q_1q_2)=1}}^{q_1q_2} S_{q_1q_2}(\underline{\mathbf{b}}). \tag{2.6.7}$$

This allows us to restrict ourselves to considering prime powers.

Lemma 2.13. *Let $d \geq 2$, p prime and l a non-negative integer, and suppose further that* (2.3.7) *holds true. Then for any $W > 0$ such that $k > (d-1)RW$, the terms of the singular series are bounded by*

$$(p^l)^{-ms}S_{p^l}(\underline{\mathbf{a}}) \ll (p^l)^{-W}.$$

Proof. We imitate the argument of Schmidt [62, Lemma 7.1]. Pick a suitable $\theta_2 < ((d-1)R)^{-1}$ such that $k\theta_2 \geq W$, and assume that the argument $\underline{\mathbf{a}}/p^l$ is on the corresponding major arcs $\mathfrak{M}(p^l, \theta_2)$. Then by the definition (2.4.1) of the major arcs one can find $0 < q_{\mathbf{j}} \ll p^{l(d-1)R\theta_2} \ll p^l$ and $\underline{b}_{\mathbf{j}} < q_{\mathbf{j}}$ for each $\mathbf{j} \in J$ subject to

$$\left| q_{\mathbf{j}} \frac{a_{\mathbf{j}}^{(\rho)}}{p^l} - b_{\mathbf{j}}^{(\rho)} \right| \ll (p^l)^{-d+(d-1)R\theta_2} \qquad (1 \leq \rho \leq R),$$

or equivalently

$$|q_{\mathbf{j}}a_{\mathbf{j}}^{(\rho)} - b_{\mathbf{j}}^{(\rho)}p^l| \ll (p^l)^{(1-d)(1-R\theta)}.$$

For $d > 1$, inserting the bound for θ_2 forces $q_{\mathbf{j}}a_{\mathbf{j}}^{(\rho)}$ to be a multiple of p^l, but since $q_{\mathbf{j}} < p^l$ it follows that p divides $a_{\mathbf{j}}^{(\rho)}$ for every set of indices. This is, however, impossible as we had $\gcd(\underline{\mathbf{a}}, p) = 1$. By Lemma 2.5 we can therefore conclude that the minor arcs estimate is true and

$$(p^l)^{-ms}S_{p^l}(\underline{\mathbf{a}}) \leq (p^l)^{-k\theta_2} \ll (p^l)^{-W}$$

as claimed. □

We now define the p-adic density as

$$\chi_p(\boldsymbol{\psi}) = \sum_{l=0}^{\infty} p^{-lms} \sum_{\substack{\underline{\mathbf{a}}=1 \\ (\underline{\mathbf{a}},p)=1}}^{p^l} S_{p^l}(\underline{\mathbf{a}}).$$

Lemma 2.14. *Suppose $k > 3(d-1)R(Rr+1)$. Then each of the local densities χ_p and χ_∞ is absolutely convergent. Furthermore, one has*

$$1/2 \leq \prod_{p>p_0} \chi_p(\boldsymbol{\psi}) \leq 3/2$$

for some suitable p_0.

Proof. The first statement is now immediate from Lemmata 2.12 and 2.13 on choosing $W > Rr + 1$ and noting that $3(Rr + 1) > r(R + 1)$, so Lemma 2.6 will be applicable.

The second statement is essentially Lemma 5.2 and the subsequent corollary of [17]. By Lemma 2.13 with $W = Rr + 1 + \delta$ one has

$$|\chi_p(\boldsymbol{\psi}) - 1| = \left| \sum_{l=1}^{\infty} p^{-lms} \sum_{\substack{\underline{\mathbf{a}}=1 \\ (\underline{\mathbf{a}},p)=1}}^{p^l} S_{p^l}(\underline{\mathbf{a}}) \right| \ll \sum_{l=1}^{\infty} (p^l)^{-1-\delta} \ll p^{-1-\delta},$$

which implies immediately that

$$\prod_p \chi_p = 1 + O\left(\sum_p p^{-1-\delta} \right).$$

This completes the proof of the Lemma. □

It therefore follows by the multiplicativity property (2.6.7) that the singular series $\mathfrak{S}_{\boldsymbol{\psi}}$ is absolutely convergent and can be expanded as an Euler product

$$\mathfrak{S}_{\boldsymbol{\psi}} = \prod_p \chi_p(\boldsymbol{\psi}),$$

where the p-adic densities $\chi_p(\boldsymbol{\psi})$ are given by

$$\begin{aligned}\chi_p(\boldsymbol{\psi}) &= \sum_{i=0}^{\infty} p^{-ims} \sum_{\substack{\underline{\mathbf{a}}=1\\(\underline{\mathbf{a}},p)=1}}^{p^i} S_{p^i}(\underline{\mathbf{a}}) e\left(-(\underline{\mathbf{n}}\cdot\underline{\mathbf{a}})/p^i\right) \\ &= \sum_{i=0}^{\infty} p^{-ims} \sum_{\substack{\underline{\mathbf{a}}=1\\(\underline{\mathbf{a}},p)=1}}^{p^i} \sum_{\overline{\mathbf{x}}=1}^{p^i} e(\mathfrak{F}(\overline{\mathbf{x}}; p^{-i}\underline{\mathbf{a}})) e\left(-(\underline{\mathbf{n}}\cdot\underline{\mathbf{a}})/p^i\right).\end{aligned} \tag{2.6.8}$$

By the discussion in [62, Section 3], this can be interpreted as a p-adic integral

$$\chi_p(\boldsymbol{\psi}) = \int_{\mathbb{Q}_p^{rR}} \int_{\mathbb{Z}_p^{ms}} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \underline{\boldsymbol{\eta}}) - \underline{\boldsymbol{\eta}}\cdot\underline{\mathbf{n}}\right) \mathrm{d}\overline{\boldsymbol{\xi}} \mathrm{d}\underline{\boldsymbol{\eta}},$$

which is the exact analogue of (2.2.6) in the p-adic numbers $\mathbb{Q}_p$.

Similarly, the singular integral

$$\mathfrak{J}_{\boldsymbol{\psi}} = \int_{\mathbb{R}^{rR}} \int_{|\overline{\boldsymbol{\xi}}|\leq 1} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \underline{\boldsymbol{\beta}}) - (\underline{\boldsymbol{\beta}}\cdot\underline{\mathbf{n}})/P^d\right) \mathrm{d}\overline{\boldsymbol{\xi}} \mathrm{d}\underline{\boldsymbol{\beta}}$$

measures solutions in the real unit box and may thus be interpreted as the density of real solutions. Thus Theorem 2.1 is established.

Chapter 3

Counting lattices on hypersurfaces

3.1 Introduction and setup

The most important special case of our work in Chapter 2 is that of counting linear spaces on hypersurfaces and corresponds to the case of representing the zero polynomial. In other words, one is interested in counting the number of $\overline{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ that solve

$$F(\mathbf{x}_1 t_1 + \dots + \mathbf{x}_m t_m) = 0 \tag{3.1.1}$$

identically in $t_1, \dots, t_m$. This equation describes a linear space over $\mathbb{Z}^s$ of affine dimension m, or, in other words, an m-dimensional integer sublattice of $\mathbb{Z}^s$. The formulation of the problem in terms of lattices arises quite naturally, and one may wonder whether the methods are specific to the setting over $\mathbb{Z}^s$ or whether one may restrict to count solutions whose generators lie in sublattices of this.

Suppose that Λ is a lattice given by $\Lambda = \mathbb{Z}^{m \times s} C$ for some non-singular $(m \times m)$-matrix C, then we are interested in the number $N_C(P)$ of $\overline{\mathbf{x}} \in \mathbb{Z}^{s \times m} C$ of height at most P that solve (3.1.1) identically in $t_1, \dots, t_m$. This problem transforms into a normal linear spaces counting problem with the additional difficulty that the generators should all be contained in an integer lattice defined by the matrix C, and we have the following theorem.

Theorem 3.1. *Let $d \geq 2$ and $m \geq 2$ be positive integers and $r = \binom{d+m-1}{m-1}$ as in* (2.1.1)*, and let P be large. Further, let $C \in \mathbb{Z}^{m\times m}$ be a non-singular matrix, and write $\gamma_{\max}$ for the largest diagonal entry of its Smith normal form. Then, provided that*

$$s^* > 2^{d-1} \max\left\{3(d-1)(r+1), \frac{rd}{1-\log\gamma_{\max}/\log P}\right\},$$

the number $N_C(P)$ of points $(\mathbf{x}_1,\dots,\mathbf{x}_m) \in \mathbb{Z}^{s\times m}C$ of height at most P solving Equation (3.1.1) *identically in $t_1,\dots,t_m$ is given by*

$$N_C(P) = \left(\frac{P^m}{\det(C)}\right)^s P^{-rd}\chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(\left(\frac{P^m}{\det(C)}\right)^s P^{-rd}\right)$$

uniformly in C, and χ_p and χ_∞ characterise the local solubility densities of the variety defined by $F=0$ and are independent of C.

By the same methods a similar result for systems of equations may be derived, but since the necessary modifications are fairly straightforward, we refrain from explicitly doing so. One notes the strong dependence of the result on the shape of the matrix C. This arises naturally, since if $\gamma_{\max} \gg P$, one of the variables is essentially fixed and the system is really of dimension $m-1$, so we should expect a different main term. On the other hand, when C is the identity matrix this recovers Theorem 2.1.

Observe that for the purpose of understanding $N_C(P)$, we can assume without loss of generality that the matrix C is diagonal. In fact, suppose $U, V \in SL(m,\mathbb{Z})$ are such that UCV is diagonal. By interpreting the vectors $\mathbf{x}_1,\dots,\mathbf{x}_m$ as a matrix X, we can view our counting function $N_C(P)$ as counting matrices $X \in \mathbb{Z}^{m\times s}C$ of height at most P for which $F(X\mathbf{t}) = 0$ is true identically in $\mathbf{t}$. This implies, however, that we may replace $\mathbf{t}$ by $V\mathbf{t}$ in the statement and equivalently demand that $F(X(V\mathbf{t})) = 0$ identically in $\mathbf{t}$. On the other hand, the condition $X \in \mathbb{Z}^{m\times s}C$ can be written as $X = YC$ with

$Y \in \mathbb{Z}^{m\times s}$, so $N_C(P)$ counts the matrices Y for which YC is of height at most P and $F(YC\mathbf{t}) = 0$ identically in $\mathbf{t}$. Since U is unimodular, right multiplication with U is a bijection, so we may equivalently set $Y = ZU$ and count matrices $Z \in \mathbb{Z}^{m\times s}$ such that ZUC is of height at most P and $F(ZUC\mathbf{t}) = 0$ identically in $\mathbf{t} \in \mathbb{Z}^m$. It follows that we may assume without loss of generality that C is of the form $C = \operatorname{diag}(\gamma_1, \dots, \gamma_m)$, and the condition

$$\overline{\mathbf{x}} \in \mathbb{Z}^{s\times m} C \cap [-P, P]^{s\times m}$$

translates into the simpler

$$\mathbf{x}_i \in \gamma_i \mathbb{Z}^s \cap [-P, P]^s \qquad (1 \le i \le m).$$

Now recall that the form F is homogeneous, so after expanding the system and removing factors this reduces to counting the number of solutions to (3.1.1) with $\mathbf{x}_i \in \mathbb{Z}^s \cap [-P_i, P_i]^s$ where $P_i = P/\gamma_i$, and Theorem 3.1 may be restated in the version alluded to in the introduction.

Theorem 3.2. *Let $d \ge 2$ and $m \ge 2$ be positive integers and let $P_1 \le \ldots \le P_m$ be large. Further, assume that*

$$s^* > 2^{d-1} \max\left\{3(d-1)(r+1), rd\left(\frac{\log P_m}{\log P_1}\right)\right\}.$$

Then the number $N(P_1, \dots, P_m)$ of $|\mathbf{x}_i| \le P_i$ solving (3.1.1) *identically in $t_1, \dots, t_m$ is given by*

$$N(P_1, \dots, P_m) = \left(\prod_{i=1}^m P_i\right)^{s-rd/m} \chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(\left(\prod_{i=1}^m P_i\right)^{s-rd/m}\right), \quad (3.1.2)$$

where χ_∞ and χ_p denote the local solubility constants as usual.

This result is somewhat reminiscent of Schindler's work on bihomogeneous forms in [56], who derives an asymptotic formula for systems of R forms $F^{(\rho)}(\mathbf{x}, \mathbf{y})$ of bidegree (d_1, d_2), where the $\mathbf{x}$ and $\mathbf{y}$ are contained in boxes of sidelength P_1 and P_2, respectively, and the number of variables $s = s_1 + s_2$ that is required in her work is

$$s^* > 2^{d_1+d_2-2} \max\{R(R+1)(d_1 + d_2 - 1), R((\log P_1/\log P_2)d_1 + d_2)\}.$$

On the first glance, our setting is more complex as we study a system of forms of distinct multidegrees, but the fact that this system arises naturally in the linear spaces setting and thus comes equipped with certain symmetries can be used to great advantage in our analysis and allows us to derive a strong result also in this seemingly more difficult situation.

A result that is more closely related is implicit in recent work by Dietmann and Harvey [23] on the representation of quadratic forms by quadratic forms, for which they require a result of the flavour of Theorem 3.2 in the case $d = 2$. They establish an asymptotic formula comparable to that in (3.1.2), provided that the number of variables satisfies

$$s^* > 2(r+1)\left(r + \sum_{i=1}^{m} \frac{\log(P_i/P_1)}{\log P_1}\right).$$

One remarks that this grows quadratically in r. We show, however, that the methods developed for the proof of Theorem 1.1 continue to be applicable, so we save a factor of r over their work.

It is not hard to see how Theorems 3.1 and 3.2 are related. As in our work in Chapter 2, we have to understand expressions of the shape

$$F\left(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m\right) = \sum_{\mathbf{j}\in J} A(\mathbf{j}) t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d} \Phi(\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_d}),$$

where we use the notation introduced in §2.2 in writing Φ for the multilinear form associated to F, J for the set of multi-indices $(j_1, \ldots, j_d) \in \{1, \ldots, m\}^d$, and $A(\mathbf{j})$ for the combinatorial factors that take into account the multiplicity of each term. This allows us to focus on the system

$$\Phi_{\mathbf{j}}(\overline{\mathbf{x}}) = \Phi(\mathbf{x}_{j_1}, \ldots \mathbf{x}_{j_d}) = 0 \qquad (\mathbf{j} \in J).$$

Writing $\boldsymbol{\alpha} = (\alpha_{\mathbf{j}})_{\mathbf{j}\in J}$, we define

$$\mathfrak{F}\big(\mathbf{x}_1, \ldots, \mathbf{x}_m; \boldsymbol{\alpha}\big) = \sum_{\mathbf{j}\in J} \alpha_{\mathbf{j}} \Phi_{\mathbf{j}}(\overline{\mathbf{x}}). \tag{3.1.3}$$

The exponential sum is now obtained by summing (3.1.3) over all $\overline{\mathbf{x}}$ in the range that is being considered. In the case of Theorem 3.2 this is the cartesian product of the intervals $[-P_i, P_i]$, which we denote by $\mathcal{P}$. Notice that this can be transformed into the language of Theorem 3.1 by by setting $P_m = P$ and $\gamma_i = P_m/P_i$ for all i. In this notation we have

$$\operatorname{Card}\mathcal{P} = \prod_{i=1}^{m} P_i = \prod_{i=1}^{m} \frac{P}{\gamma_i} = \frac{P^m}{\det(C)}.$$

The exponential sum is thus given by

$$T(\boldsymbol{\alpha}, \mathcal{P}) = \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e(\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})),$$

and classical orthogonality relations imply that the number of solutions to (3.1.1) with $\mathbf{x}_i \leq P_i$ is given by

$$N(\mathcal{P}) = \int_{[0,1)^r} T(\boldsymbol{\alpha}, \mathcal{P}) \mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^r} \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e\left(\mathfrak{F}\left(\overline{\mathbf{x}}; \boldsymbol{\alpha}\right)\right) \mathrm{d}\boldsymbol{\alpha}. \tag{3.1.4}$$

In the case of Theorem 3.1, on the other hand, the $\mathbf{x}_i$ are multiples of γ_i, and by homogeneity we may write

$$\begin{aligned} \mathfrak{F}\big(\gamma_1 \mathbf{x}_1, \ldots, \gamma_m \mathbf{x}_m; \boldsymbol{\alpha}\big) &= \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \Phi(\gamma_{j_1} \mathbf{x}_{j_1}, \ldots, \gamma_{j_d} \mathbf{x}_{j_d}) \\ &= \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \hat{\boldsymbol{\gamma}}_{\mathbf{j}} \Phi(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_d}), \end{aligned}$$

where we introduced the notation $\hat{\boldsymbol{\gamma}}_{\mathbf{j}}$ for the product $\gamma_{j_1} \cdot \ldots \cdot \gamma_{j_d}$. Absorbing the factors $\hat{\boldsymbol{\gamma}}_{\mathbf{j}}$ into the coefficients $\alpha_{\mathbf{j}}$, we see that the number of solutions is given by

$$\begin{aligned} N_C(P) &= \int_{[0,1)^r} \sum_{\overline{\mathbf{x}} C \leq P} e\left(\mathfrak{F}\left(\overline{\mathbf{x}} C; \boldsymbol{\alpha}\right)\right) \mathrm{d}\boldsymbol{\alpha} \\ &= \int_{[0,1)^r} \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e\left(\sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \hat{\boldsymbol{\gamma}}_{\mathbf{j}} \big(\Phi_{\mathbf{j}}(\overline{\mathbf{x}})\big)\right) \mathrm{d}\boldsymbol{\alpha} \\ &= \Big(\prod_{\mathbf{j} \in J} \hat{\boldsymbol{\gamma}}_{\mathbf{j}}\Big)^{-1} N(\mathcal{P}). \end{aligned}$$

The product in the last line is symmetric in the γ_i and has altogether rd factors, so its value is

$$\prod_{\mathbf{j}\in J} \hat{\boldsymbol{\gamma}}_{\mathbf{j}} = (\gamma_1 \cdot \ldots \cdot \gamma_m)^{rd/m} = \det(C)^{rd/m}. \tag{3.1.5}$$

It is therefore the counting function $N(\mathcal{P})$ given by (3.1.4) towards which we will direct our attention.

3.2 Weyl differencing

The first step is to establish an inequality of Weyl type as presented in [8] and adapted to the linear space situation in Chapter 2. Although the greater picture of this is by now fairly standard, the different ranges of the $\mathbf{x}_i$ create some technical complications which need to be attended to with due care.

As in Chapter 2, we define the discrete difference operator by its action on the form $\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})$ as

$$\Delta_{i,\mathbf{h}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha}) = \mathfrak{F}(\mathbf{x}_1,\ldots,\mathbf{x}_i+\mathbf{h},\ldots,\mathbf{x}_m;\boldsymbol{\alpha}) - \mathfrak{F}(\mathbf{x}_1,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_m;\boldsymbol{\alpha}),$$

and write for brevity

$$\Delta^{(k)}_{\mathbf{h}_\mathbf{j}} = \Delta_{k,\mathbf{h}_{j_k}} \cdots \Delta_{1,\mathbf{h}_{j_1}}.$$

This allows us to formulate our first Weyl differencing lemma.

Lemma 3.1. *Let $2 \leq k \leq d-1$. We have*

$$|T(\boldsymbol{\alpha},\mathcal{P})|^{2^k} \ll \left(\frac{P^m}{\det(C)}\right)^{(2^k-1)s} \left(\prod_{i=1}^{k} P_i^{-s}\right) \sum_{\substack{|\mathbf{h}_l|\leq P_{j_l}\\ 1\leq l\leq k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta^{(k)}_{\mathbf{h}_\mathbf{j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right),$$

and the sum over $\overline{\mathbf{x}}$ is over a suitable box contained in $\mathcal{P}$.

Proof. The proof is, as usual, by induction. The case $k=2$ follows by Cauchy–

Schwarz via

$$
\begin{aligned}
|T(\boldsymbol{\alpha},\mathcal{P})|^2 &\ll \Bigg(\sum_{\substack{|\mathbf{x}_i|\le P_i\\ i\neq j_1}} 1\Bigg) \sum_{\substack{|\mathbf{x}_i|\le P_i\\ i\neq j_1}} \Bigg| \sum_{|\mathbf{x}_{j_1}|\le P_{j_1}} e(\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha}))\Bigg|^2 \\
&\ll \Big(\prod_{i\neq j_1} P_i^s\Big) \sum_{\substack{|\mathbf{x}_i|\le P_i\\ i\neq j_1}} \Bigg(\sum_{|\mathbf{h}_1|\le P_{j_1}} \sum_{\mathbf{x}_{j_1}} e\left(\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right)\Bigg) \\
&\ll \left(\frac{P^m}{\det(C)}\right)^s P_{j_1}^{-s} \sum_{|\mathbf{h}_1|\le P_{j_1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_1,\mathbf{h}_1}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right).
\end{aligned}
$$

Note that the final summation of $\mathbf{x}_{j_1}$ is over all $|\mathbf{x}_{j_1}| \le P_{j_1}$ that also satisfy $|\mathbf{x}_{j_1} + \mathbf{h}_1| \le P_{j_1}$, which is again a box contained in $[-P_{j_1}, P_{j_1}]$. The induction step is similar. By another application of Cauchy–Schwarz, one has

$$
\begin{aligned}
&\Bigg| \sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta^{(k)}_{\mathbf{h_j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right)\Bigg|^2 \\
&\ll \Bigg(\sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k}} \sum_{\substack{|\mathbf{x}_i|\ll P_i\\ i\neq j_{k+1}}} 1\Bigg) \sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k}} \sum_{\substack{|\mathbf{x}_i|\ll P_i\\ i\neq j_{k+1}}} \Bigg| \sum_{|\mathbf{x}_{j_{k+1}}|\ll P_{j_{k+1}}} e\left(\Delta^{(k)}_{\mathbf{h_j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right)\Bigg|^2 \\
&\ll \left(\frac{P^m}{\det(C)}\right)^s P_{j_{k+1}}^{-s} \Big(\prod_{l=1}^{k} P_{j_l}^s\Big) \sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k+1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta^{(k+1)}_{\mathbf{h_j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right),
\end{aligned}
$$

and hence

$$
\begin{aligned}
|T(\boldsymbol{\alpha},\mathcal{P})|^{2^{k+1}} &\ll \left(\frac{P^m}{\det(C)}\right)^{(2^{k+1}-2)s} \Big(\prod_{l=1}^{k} P_{j_l}^{-2s}\Big) \Bigg| \sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta^{(k)}_{\mathbf{h_j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right)\Bigg|^2 \\
&\ll \left(\frac{P^m}{\det(C)}\right)^{2^{k+1}-1} \Big(\prod_{l=1}^{k+1} P_{j_l}^{-1}\Big) \sum_{\substack{|\mathbf{h}_l|\ll P_{j_l}\\ 1\le l\le k+1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta^{(k+1)}_{\mathbf{h_j}}\mathfrak{F}(\overline{\mathbf{x}};\boldsymbol{\alpha})\right).
\end{aligned}
$$

This completes the proof. □

Writing

$$
\Phi(\mathbf{x},\hat{\mathbf{h}}) = \sum_{n=1}^{s} B_n(\hat{\mathbf{h}}) x_n \tag{3.2.1}
$$

and abbreviating $\hat{\mathbf{h}}$ for the $(d-1)$-tuple $(\mathbf{h}_1, \ldots, \mathbf{h}_{d-1})$, we let

$$\Upsilon(\mathbf{j}) = \sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \prod_{n=1}^{s} \min\left(P_{j_d}, \|M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}})\|^{-1}\right),$$

where the coefficients $M(\mathbf{j})$ are defined by means of (2.2.3) as

$$M(\mathbf{j}) = \mu_1(\mathbf{j})!\mu_2(\mathbf{j})! \cdot \ldots \cdot \mu_m(\mathbf{j})!.$$

Then we can bound the exponential sum above.

Lemma 3.2. *One has*

$$|T(\boldsymbol{\alpha}, \mathcal{P})|^{2^{d-1}} \ll \left(\frac{P^m}{\det(C)}\right)^{2^{d-1}s} \left(\prod_{k=1}^{d} P_{j_k}^{-s}\right) \Upsilon(\mathbf{j}).$$

Proof. In the case $k = d-1$, we see that Lemma 3.1 yields

$$|T(\boldsymbol{\alpha}, \mathcal{P})|^{2^{d-1}} \ll \left(\frac{P^m}{\det(C)}\right)^{(2^{d-1}-1)s} \left(\prod_{k=1}^{d-1} P_{j_k}^{-s}\right) \sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h}_{\mathbf{j}}}^{(d-1)} \mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right).$$

As in the previous chapter, we notice that the differencing procedure gradually reduces the degree while preserving the structure of the system. In our case, this means that after $d-1$ applications the resulting expression is linear in the $\mathbf{x}_1, \ldots, \mathbf{x}_m$, and since all forms in the system of equations are instances of the same multilinear form Φ, we may exploit this structure in our analysis. Writing $R(\hat{\mathbf{h}})$ for the terms independent of $\overline{\mathbf{x}}$, one obtains

$$\begin{aligned} &\sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h}_{\mathbf{j}}}^{(d-1)} \mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right) \\ &\ll \sum_{\substack{|\mathbf{h}_l| \le P_{j_l} \\ 1 \le l \le d-1}} \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e\left(\sum_{k=1}^{m} M(j_1, \ldots, j_{d-1}, k)\alpha_{(j_1, \ldots, j_{d-1}, k)} \Phi(\mathbf{x}_k, \mathbf{h}_1, \ldots, \mathbf{h}_{d-1}) + R(\hat{\mathbf{h}})\right) \\ &\ll \left(\frac{P^m}{\det(C)}\right)^s P_{j_d}^{-s} \sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \left| \sum_{|\mathbf{x}_{j_d}| \ll P_{j_d}} e\left(M(\mathbf{j})\alpha_{\mathbf{j}} \Phi(\mathbf{x}_{j_d}, \hat{\mathbf{h}})\right) \right|. \end{aligned}$$

Recalling (3.2.1), the innermost sum is

$$\sum_{|\mathbf{x}_{j_d}| \ll P_{j_d}} e\left(M(\mathbf{j})\alpha_{\mathbf{j}} \Phi(\mathbf{x}_{j_d}, \hat{\mathbf{h}})\right) = \prod_{n=1}^{s} \sum_{|x_n| \ll P_{j_d}} e\left(M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}}) x_n\right).$$

Thus by standard arguments we have

$$\sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \prod_{n=1}^{s} \sum_{|x_n| \ll P_{j_d}} e\left(M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}}) x_n\right)$$
$$\ll \sum_{\substack{|\mathbf{h}_l| \ll P_{j_l} \\ 1 \le l \le d-1}} \prod_{n=1}^{s} \min\left(P_{j_d}, \|M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}})\|^{-1}\right),$$

which is $\Upsilon(\mathbf{j})$. Altogether we have

$$|T(\boldsymbol{\alpha}, \mathcal{P})|^{2^{d-1}} \ll \left(\frac{P^m}{\det(C)}\right)^{(2^{d-1}-1)s} \left(\prod_{k=1}^{d-1} P_{j_k}^{-s}\right) \left(\frac{P^m}{\det(C)}\right)^{s} P_{j_d}^{-s} \Upsilon(\mathbf{j}),$$

which returns the statement of the lemma. □

3.3 Geometry of numbers

We want to show that either for some parameters k and $\theta > 0$ we can save $P^{k\theta}$ on the trivial estimate on $T(\boldsymbol{\alpha}, \mathcal{P})$, or we have good rational approximations to the vector $\boldsymbol{\alpha}$, or the form F has been singular from the beginning.

Let $N_{\mathbf{j}}(P_{j_1}, \ldots, P_{j_{d-1}}; Q)$ denote the cardinality of the set

$$\left\{ |\mathbf{h}_k| \le P_{j_k} \quad (1 \le k \le d-1) \colon \|M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}})\|^{-1} > Q \quad \forall n \right\}. \tag{3.3.1}$$

Lemma 3.3. *Suppose*

$$|T(\boldsymbol{\alpha}, \mathcal{P})| \gg \left(\frac{P^m}{\det(C)}\right)^{s} P^{-k\theta}$$

for some k and θ to be determined later. Then one has

$$N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right) \gg P^{-2^{d-1}k\theta - \epsilon} \prod_{k=1}^{d-1} P_{j_k}^{s}.$$

Proof. Combining the hypothesis of the lemma with Lemma 3.2 we have

$$\left(\left(\frac{P^m}{\det(C)}\right)^{s} P^{-k\theta}\right)^{2^{d-1}} \ll \left(\frac{P^m}{\det(C)}\right)^{2^{d-1}s} \left(\prod_{k=1}^{d} P_{j_k}^{-s}\right) \Upsilon(\mathbf{j}).$$

Rearranging the terms, one obtains

$$\Upsilon(\mathbf{j}) \gg P^{-2^{d-1}k\theta} \prod_{k=1}^{d} P_{j_k}^{s}. \tag{3.3.2}$$

Notice that, unlike in the situation considered in Chapter 2, we obtain distinct estimates for the $\mathbf{j} \in J$.

Now for a fixed tuple $\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}$ write

$$R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}) = \mathrm{Card}\{|\mathbf{h}_1| \leq P_{j_1} : \|M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}})\| < P_{j_d}^{-1} \quad (1 \leq n \leq s)\},$$

so that

$$\sum_{\substack{|\mathbf{h}_l| \leq P_{j_l} \\ 2 \leq l \leq d-1}} R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}) = N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right).$$

Then a familiar pigeonhole argument as in the proof of [17, Lemma 13.2] implies that for any integers r_n between 1 and P_{j_d} the number of $\mathbf{h}_1$ such that

$$\frac{r_n}{P_{j_d}} < \{M(\mathbf{j})\alpha_{\mathbf{j}} B_{\mathbf{j},n}(\hat{\mathbf{h}})\} < \frac{r_n + 1}{P_{j_d}} \quad (1 \leq n \leq s)$$

is at most $R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1})$, and thus

$$\begin{aligned}
\Upsilon(\mathbf{j}) &\ll \sum_{\substack{|\mathbf{h}_l| \leq P_{j_l} \\ 2 \leq l \leq d-1}} \prod_{n=1}^{s} \sum_{r_n=1}^{P_{j_d}-1} \min\left\{P_{j_d}, \frac{P_{j_d}}{r_n}, \frac{P_{j_d}}{r_n + 1}\right\} \\
&\ll (P_{j_d} \log(P_{j_d}))^s \sum_{\substack{|\mathbf{h}_l| \leq P_{j_l} \\ 2 \leq l \leq d-1}} R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}) \\
&\ll (P_{j_d} \log(P_{j_d}))^s N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right).
\end{aligned}$$

Inserting this into (3.3.2) gives the desired result. □

We will need the following standard lemma.

Lemma 3.4. *Let $L_1, \ldots, L_n$ be linear forms given by*

$$L_i = \lambda_{i,1} x_1 + \ldots + \lambda_{i,n} x_n \quad (1 \leq i \leq n)$$

with the additional symmetry that $\lambda_{i,j} = \lambda_{j,i}$. For a parameter A let $U(Z)$ denote the number of integer solutions $x_1, \ldots, x_n$ to the system

$$|x_i| < AZ \quad \textit{and} \quad \|L_i(\mathbf{x})\| < Z/A \quad (1 \leq i \leq n).$$

Then for $0 < Z' \leq Z \leq 1$ *we have*

$$\frac{U(Z)}{U(Z')} \ll \left(\frac{Z}{Z'}\right)^n.$$

Proof. This is [17, Lemma 12.6]. □

We are now in a position to use Lemma 3.4 for our main application of the geometry of numbers. It is here that the dependence on the matrix C in the final bound for s^* makes its first appearance. The goal is to apply Lemma 3.4 on each of the variables $\mathbf{h}_k$ in such a way that $AZ = P_{j_k}$ and $AZ' = P^\theta$ for some small θ, so that in the further course of the argument we can assume the variables to lie in small boxes which are then independent of C. However, this is legitimate only in the case when $Z' \leq Z$, so we need $(AZ')/(AZ) = P^\theta/P_{j_k} \leq 1$ for all j_k. This condition amounts to $P^\theta \leq P/\gamma_{\max}$ or, taking logarithms,

$$\theta \leq 1 - \frac{\log \gamma_{\max}}{\log P}.$$

For simplicity we write

$$\eta = \log \gamma_{\max} / \log P. \tag{3.3.3}$$

Notice that in the case when C is the unit matrix we have $\eta = 0$ and therefore $\theta \leq 1$ as usual. We can now formulate our application of the geometry of numbers.

Lemma 3.5. *Suppose that* $0 < \theta \leq 1 - \eta$, *where* η *is as in* (3.3.3). *Then under the conditions of Lemma 3.3 one has*

$$N_{\mathbf{j}}\left(P^\theta, \ldots, P^\theta; \frac{P^{d-(d-1)\theta}}{\hat{\gamma}_{\mathbf{j}}}\right) \gg P^{(d-1)s\theta - 2^{d-1}k\theta - \epsilon}.$$

Proof. This follows from Lemma 3.4 by what is essentially a standard argument. For fixed $\mathbf{h}_1, \ldots \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_{d-1}$ let $U_k(Z_k)$ denote the number of $\mathbf{h}_k < A_k Z_k$ such that $\|M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}})\| < Z_k/A_k$. We will take

$$A_1 = \frac{P}{\sqrt{\gamma_{j_1}\gamma_{j_d}}}, \qquad Z_1 = \sqrt{\frac{\gamma_{j_d}}{\gamma_{j_1}}}, \qquad Z_1' = \frac{\sqrt{\gamma_{j_d}\gamma_{j_1}}}{P^{1-\theta}},$$

and then recursively

$$A_k = A_{k-1}\sqrt{\frac{P^{1-\theta}}{\gamma_{j_k}}}, \qquad Z_k = \frac{\gamma_{j_{k-1}} Z_{k-1}}{\sqrt{P^{1-\theta}\gamma_{j_k}}}, \qquad Z'_k = \frac{\gamma_{j_k} Z_k}{P^{1-\theta}},$$

so that the relations

$$\begin{aligned} A_k Z_k &= P/\gamma_{j_k}, & Z_k/A_k &= P^{-k+(k-1)\theta}\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_{k-1}}\gamma_{j_d}, \\ A_k Z'_k &= P^{\theta}, & Z'_k/A_k &= P^{-(k+1)+k\theta}\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_k}\gamma_{j_d}, \\ Z_k/Z'_k &= P^{1-\theta}/\gamma_{j_k} \end{aligned}$$

are satisfied. Notice that with our hypothesis on θ the last relation implies $Z_k \geq Z'_k$. Also, one has $Z_k \leq Z_{k-1}$ and, as we may assume without loss of generality that $\gamma_{j_d} \leq \gamma_{j_1}$, it follows that $Z_k \leq 1$ for all k. Hence Lemma 3.4 is applicable and yields

$$U_k(Z_k) \ll \left(\frac{Z_k}{Z'_k}\right)^s U_k(Z'_k) \ll (P^{1-\theta}/\gamma_{j_k})^s U_k(Z'_k). \tag{3.3.4}$$

We can now start a recursive argument. For given k between 0 and $d-1$ consider the quantity

$$\nu(k) = N_{\mathbf{j}}\left(\underbrace{P^{\theta},\ldots,P^{\theta}}_{\text{first } k \text{ entries}}, \frac{P}{\gamma_{k+1}},\ldots,\frac{P}{\gamma_{d-1}}; \frac{P^{(k+1)-k\theta}}{\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_k}\gamma_{j_d}}\right). \tag{3.3.5}$$

We can express this in terms of $U_k(Z_k)$ or $U_{k+1}(Z_{k+1})$ by the relation

$$\nu(k) = \sum_{\substack{|\mathbf{h}_l|\leq P^{\theta}\\ 1\leq l\leq k-1}} \sum_{\substack{|\mathbf{h}_l|\leq P/\gamma_{j_l}\\ k+1\leq l\leq d-1}} U_k(Z'_k) = \sum_{\substack{|\mathbf{h}_l|\leq P^{\theta}\\ 1\leq l\leq k}} \sum_{\substack{|\mathbf{h}_l|\leq P/\gamma_{j_l}\\ k+2\leq l\leq d-1}} U_{k+1}(Z_{k+1}). \tag{3.3.6}$$

The recursive relation for the ν can now be derived from (3.3.4) and (3.3.6) and is given by

$$\begin{aligned} \nu(k-1) &= \sum_{\substack{|\mathbf{h}_l|\leq P^{\theta}\\ 1\leq l\leq k-1}} \sum_{\substack{|\mathbf{h}_l|\leq P/\gamma_{j_l}\\ k+1\leq l\leq d-1}} U_k(Z_k) \\ &\ll \left(\frac{Z_k}{Z'_k}\right)^s \sum_{\substack{|\mathbf{h}_l|\leq P^{\theta}\\ 1\leq l\leq k-1}} \sum_{\substack{|\mathbf{h}_l|\leq P/\gamma_{j_l}\\ k+1\leq l\leq d-1}} U_k(Z'_k) \\ &\ll (P^{1-\theta}/\gamma_{j_k})^s \nu(k). \end{aligned} \tag{3.3.7}$$

Iterating (3.3.7), one obtains

$$\nu(d-1) \gg P^{-(1-\theta)(d-1)s}\Big(\prod_{k=1}^{d-1}\gamma_{j_k}^s\Big)\nu(0).$$

Finally, recalling the definition (3.3.5) of the $\nu(k)$, we can insert Lemma 3.3 and find

$$\begin{aligned} N_{\mathbf{j}}\left(P^\theta,\dots,P^\theta;\frac{P^{d-(d-1)\theta}}{\hat{\boldsymbol{\gamma}}_{\mathbf{j}}}\right) &\gg P^{-(d-1)(1-\theta)s}\Big(\prod_{k=1}^{d-1}\gamma_{j_k}^s\Big)N_{\mathbf{j}}\left(\frac{P}{\gamma_{j_1}},\dots,\frac{P}{\gamma_{j_{d-1}}};\frac{P}{\gamma_{j_d}}\right)\\ &\gg P^{-(d-1)(1-\theta)s}\Big(\prod_{k=1}^{d-1}\gamma_{j_k}^s\Big)P^{(d-1)s-2^{d-1}k\theta-\epsilon}\prod_{k=1}^{d-1}\gamma_{j_k}^{-s}\\ &\gg P^{(d-1)s\theta-2^{d-1}k\theta-\epsilon}, \end{aligned}$$

as claimed. □

The message of Lemma 3.5 is that if the exponential sum is large, the quantities $M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\hat{\mathbf{h}})$ are simultaneously close to an integer for many choices of $\hat{\mathbf{h}}$. This is certainly the case if the forms B_n tends to vanish for geometric reasons, and in the other case one finds genuine (i. e. non-zero) solutions to the diophantine approximation problem that is implicit in (3.3.1). This yields the standard threefold case distinction.

Lemma 3.6. *Let $0<\theta\leq 1-\eta$ and k be parameters, and let $\boldsymbol{\alpha}\in[0,1)^r$. Then there are three possibilities.*

(A) The exponential sum $T(\boldsymbol{\alpha},\mathcal{P})$ is bounded by

$$|T(\boldsymbol{\alpha},\mathcal{P})| \ll \left(\frac{P^m}{\det(C)}\right)^s P^{-k\theta}.$$

(B) For every $\mathbf{j}\in J$ one finds integers $(q_{\mathbf{j}},a_{\mathbf{j}})$ satisfying

$$0<q_{\mathbf{j}}\ll P^{(d-1)\theta} \qquad \textit{and} \qquad |\alpha_{\mathbf{j}}q_{\mathbf{j}}-a_{\mathbf{j}}|\ll P^{-d+(d-1)\theta}\hat{\boldsymbol{\gamma}}_{\mathbf{j}}.$$

(C) The number of $|\mathbf{h}_k|\leq P^\theta$ for $1\leq k\leq d-1$ that satisfy

$$B_n(\mathbf{h}_1,\dots,\mathbf{h}_{d-1})=0 \qquad (1\leq n\leq s)$$

is asymptotically greater than $(P^\theta)^{(d-1)s-2^{d-1}k-\epsilon}$.

Proof. The proof is similar to that of Lemma 2.4. Assuming that the estimate in (A) does not hold, Lemma 3.5 implies that for every $\mathbf{j} \in J$ we have

$$\left\| M(\mathbf{j})\alpha_{\mathbf{j}} B_n(\hat{\mathbf{h}}_{\mathbf{j}}) \right\| < P^{-d+(d-1)\theta}\hat{\gamma}_{\mathbf{j}} \qquad (1 \le n \le s)$$

for at least $\gg P^{(d-1)s\theta - 2^{d-1}k\theta - \epsilon}$ choices of $\hat{\mathbf{h}} \le P^\theta$. If $M(\mathbf{j})B_n(\hat{\mathbf{h}}_{\mathbf{j}})$ is non-zero for some n and some $(d-1)$-tuple $\hat{\mathbf{h}}_{\mathbf{j}} = (\mathbf{h}_1, \dots, \mathbf{h}_{d-1})$, we denote its value by $q_{\mathbf{j}}$. It follows that we can find an integer $a_{\mathbf{j}}$ with the property that

$$|\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{-d+(d-1)\theta}\hat{\gamma}_{\mathbf{j}}.$$

This establishes the statement. □

As in Chapter 2 the singular case can be excluded. This is, however, identical to the treatment in Lemma 2.5.

Lemma 3.7. *Let $\boldsymbol{\alpha} \in [0,1)^r$ and let $0 < \theta \le 1 - \eta$ and k be parameters with*

$$s^* > 2^{d-1}k. \tag{3.3.8}$$

Then the alternatives are the following.

(A) The exponential sum $T(\boldsymbol{\alpha}, \mathcal{P})$ is bounded by

$$|T(\boldsymbol{\alpha}, \mathcal{P})| \ll \left(\frac{P^m}{\det(C)} \right)^s P^{-k\theta+\epsilon}.$$

(B) For every $\mathbf{j} \in J$ one finds integers $(q_{\mathbf{j}}, a_{\mathbf{j}})$ satisfying

$$0 < q_{\mathbf{j}} \ll P^{(d-1)\theta} \qquad \textit{and} \qquad |\alpha_{\mathbf{j}} q - a_{\mathbf{j}}| \ll P^{-d+3(d-1)\theta}\hat{\gamma}_{\mathbf{j}}.$$

Proof. This is essentially Lemma 2.5. Notice that the singular case in Lemma 3.6 is the same as that in Lemma 2.4, so the methods used to derive Lemma 2.5 from Lemma 2.4 are applicable, and the singular case can be excluded by choosing $s^* > 2^{d-1}k$. □

3.4 Major arcs dissection

Let κ be sufficiently large in terms of the coefficients of F and C, then we define our first set of major arcs $\mathfrak{M}(P,\theta)$ to be the set of all $\boldsymbol{\alpha} \in [0,1)^r$ that have a rational approximation satisfying

$$0 \le a_{\mathbf{j}} < q_{\mathbf{j}} \le \kappa P^{(d-1)\theta} \qquad \text{and} \qquad |\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \le \kappa P^{-d+(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \tag{3.4.1}$$

for all $\mathbf{j} \in J$, and the minor arcs

$$\mathfrak{m}(P,\theta) = [0,1)^r \setminus \mathfrak{M}(P,\theta)$$

to be the complement thereof. Notice again that this respects the case distinction of Lemma 3.7, so one has a minor arcs estimate for all $\boldsymbol{\alpha} \in \mathfrak{m}$. In order to keep notation simple, we omit the parameter P whenever there is no danger of confusion.

Lemma 3.8. *The volume of the major arcs is*

$$\mathrm{vol}(\mathfrak{M}(\theta)) \le \left(\frac{P^m}{\det(C)}\right)^{-rd/m} P^{2(d-1)r\theta}.$$

Proof. As in the usual setting, we have

$$\begin{aligned}\mathrm{vol}(\mathfrak{M}(\theta)) &= \prod_{\mathbf{j}\in J} \sum_{q_{\mathbf{j}}} \sum_{a_{\mathbf{j}} \le q_{\mathbf{j}}} q_{\mathbf{j}}^{-1} P^{-d+(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \\ &\ll \prod_{\mathbf{j}\in J} \sum_{q_{\mathbf{j}}} P^{-d+(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \\ &\ll \left(\frac{P^m}{\det(C)}\right)^{-rd/m} P^{2(d-1)r\theta},\end{aligned}$$

where the last step follows by (3.1.5). This completes the proof of the lemma. □

The pruning lemma is essentially the same as Lemma 2.6.

Lemma 3.9. *Suppose the parameters k and θ satisfy*

$$0 < \theta < \theta_0 = \min\left\{\frac{d}{2(d-1)}, 1-\eta\right\}$$

and

$$k > \max\left\{2r(d-1), \frac{rd}{1-\eta}\right\}. \tag{3.4.2}$$

Then the minor arcs contribution is bounded by

$$\int_{\mathfrak{m}(\theta)} |T(\boldsymbol{\alpha}, \mathcal{P})| \mathrm{d}\boldsymbol{\alpha} = O\left(\left(\frac{P^m}{\det(C)}\right)^{s-rd/m} P^{-\delta}\right).$$

Proof. We follow the proof of [8, Lemma 4.4]. As a first step, we note that $k\theta_0 > rd + \delta$ and therefore

$$\int_{\mathfrak{m}(\theta_0)} |T(\boldsymbol{\alpha}, \mathcal{P})| \mathrm{d}\boldsymbol{\alpha} \ll \left(\frac{P^m}{\det(C)}\right)^{s} P^{-rd-\delta} \ll \left(\frac{P^m}{\det(C)}\right)^{s-rd/m} P^{-\delta}.$$

Given $0 < \theta < \theta_0$, we can pick δ small enough such that

$$(k - 2r(d-1))\theta > 2\delta. \tag{3.4.3}$$

This is possible by (3.4.2). Furthermore, we choose a sequence θ_i with the property that

$$1 \geq \theta_0 > \theta_1 > \theta_2 > \ldots > \theta_{M-1} > \theta_M = \theta > 0$$

and subject to the condition

$$(\theta_i - \theta_{i+1})k < \delta \quad \text{for all } i. \tag{3.4.4}$$

This is always possible with

$$M = O(1). \tag{3.4.5}$$

Then on writing

$$\mathfrak{m}_i = \mathfrak{m}(\theta_i) \setminus \mathfrak{m}(\theta_{i-1}) = \mathfrak{M}(\theta_{i-1}) \setminus \mathfrak{M}(\theta_i)$$

one has

$$\mathrm{vol}(\mathfrak{m}_i) \leq \mathrm{vol}(\mathfrak{M}(\theta_{i-1})) \ll \left(\frac{P^m}{\det(C)}\right)^{-rd/m} P^{2(d-1)r\theta}$$

by Lemma 3.8. Recall that for $\boldsymbol{\alpha} \in \mathfrak{m}(\theta_i)$, we are in the situation of case (A) in Lemma 3.7, so the minor arcs contribution is bounded by

$$\begin{aligned}\int_{\mathfrak{m}(\theta)\setminus\mathfrak{m}(\theta_0)} |T(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} &= \sum_{i=1}^{M} \int_{\mathfrak{m}_i} |T(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \\ &\ll \sum_{i=1}^{M} \mathrm{vol}(\mathfrak{M}(\theta_{i-1})) \sup_{\boldsymbol{\alpha}\in\mathfrak{m}(\theta_i)} |T(\boldsymbol{\alpha})| \\ &\ll \sum_{i=1}^{M} \left(\frac{P^m}{\det(C)}\right)^{-rd/m} P^{2(d-1)r\theta_{i-1}} \left(\frac{P^m}{\det(C)}\right)^{s} P^{-k\theta_i} \\ &\ll \sum_{i=1}^{M} \left(\frac{P^m}{\det(C)}\right)^{s-rd/m} P^{-k\theta_i+2r(d-1)\theta_{i-1}}.\end{aligned}$$

By (3.4.5), the sum is of no consequence and can be replaced by a maximum over all $i \in \{1, \dots, M\}$. By (3.4.3) and (3.4.4) the exponent is

$$\begin{aligned}-k\theta_i + 2r(d-1)\theta_{i-1} &= k(\theta_{i-1} - \theta_i) - (k - 2r(d-1))\theta_i \\ &< k(\theta_{i-1} - \theta_i) - (k - 2r(d-1))\theta < -\delta\end{aligned}$$

and we recover the enunciation. □

The next step will be to homogenise the major arcs, so as to enable us to put the exponential sum on a common denominator. Note that, unlike in the treatment of Chapter 2, we give only a crude bound here. This is possible since later on in the analysis, where a more rigorous bound would be necessary, the problem reduces to the one considered in the previous chapter, so we will be able to import bounds from our earlier work. We therefore let $q = \mathrm{lcm}_{\mathbf{j}\in J}\, q_{\mathbf{j}}$ so that trivially $q \ll P^{r(d-1)\theta}$, and

$$|\alpha_{\mathbf{j}} q - b_{\mathbf{j}}| = \frac{q}{q_{\mathbf{j}}} |\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{-d+r(d-1)\theta} \hat{\gamma}_{\mathbf{j}}.$$

It remains to show that the major arcs contribution can be interpreted as a product of local densities. This is fairly standard and similar to the treatment in the previous chapter.

3.5 Major arcs analysis

The goal of this section is to express the contribution of the major arcs in terms of real and p-adic solution densities. This treatment is similar to that of Chapter 2, and indeed we will see how the local densities themselves loose all information regarding C and reduce to the ones treated in our former work. For technical reasons, it is convenient to extend the major arcs slightly and define our final choice of major arcs $\mathfrak{M}'(\theta)$ to be set of all $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$ contained in the interval $[0,1)^r$ that satisfy

$$|\beta_{\mathbf{j}}| \leq \kappa' P^{-d+r(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \qquad \text{and} \qquad 0 \leq \mathbf{a} < q \leq \kappa' P^{r(d-1)\theta} \tag{3.5.1}$$

for some suitably large constant κ'. Henceforth all parameters $\boldsymbol{\alpha}, \mathbf{a}, q, \boldsymbol{\beta}$ will be implicitly understood to satisfy the major arcs inequalities as given in (3.5.1). Let

$$S_q(\mathbf{a}) = \sum_{\overline{\mathbf{x}}=1}^{q} e\left(\frac{\mathfrak{F}(\overline{\mathbf{x}}; \mathbf{a})}{q}\right)$$

and

$$v_{\mathcal{P}}(\boldsymbol{\beta}) = \int_{\overline{\boldsymbol{\xi}} \in \mathcal{P}^s} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \boldsymbol{\beta})\right) \mathrm{d}\overline{\boldsymbol{\xi}}.$$

We can now replace the exponential sum by an expression in terms of the approximation given by $\mathbf{a}$, q and $\boldsymbol{\beta}$.

Lemma 3.10. *Assume that $\boldsymbol{\alpha} \in \mathfrak{M}'(P, \theta)$. Then there exists an integer vector $(\mathbf{a}, q)$ such that*

$$T(\boldsymbol{\alpha}, \mathcal{P}) - q^{-ms} S_q(\mathbf{a}) v_{\mathcal{P}}(\boldsymbol{\beta}) \ll q^{ms} \left(1 + \left(\frac{(Pq^{-1})^m}{\det(C)}\right)^s q \left(\sum_{\mathbf{j} \in J} |\beta_{\mathbf{j}}| \frac{P^d}{\hat{\gamma}_{\mathbf{j}}}\right) \frac{\gamma_{\max}}{P}\right)$$
$$\ll \left(\frac{P^m}{\det(C)}\right)^s P^{-1+2r(d-1)\theta} \gamma_{\max}.$$

Proof. Sorting the terms in arithmetic progressions modulo q, we have

$$T(\boldsymbol{\alpha}, \mathcal{P}) = \sum_{\overline{\mathbf{z}} \bmod q} e(\mathfrak{F}(\overline{\mathbf{z}}; \mathbf{a}/q)) \sum_{\substack{\overline{\mathbf{w}} \in \mathbb{Z}^{ms} \\ q\overline{\mathbf{w}} + \overline{\mathbf{z}} \in \mathcal{P}^s}} e(\mathfrak{F}(q\overline{\mathbf{w}} + \overline{\mathbf{z}}; \boldsymbol{\beta})),$$

and thus the difference $T(\boldsymbol{\alpha},\mathcal{P}) - q^{-ms}S_q(\mathbf{a})v_{\mathcal{P}}(\boldsymbol{\beta})$ is given by

$$\sum_{\overline{\mathbf{z}} \bmod q} e(\mathfrak{F}(\overline{\mathbf{z}};\mathbf{a}/q))H(q,\boldsymbol{\beta},\overline{\mathbf{z}}),$$

where

$$\begin{aligned} H(q,\boldsymbol{\beta},\overline{\mathbf{z}}) &= \sum_{\substack{\overline{\mathbf{w}}\in\mathbb{Z}^{ms}\\ q\overline{\mathbf{w}}+\overline{\mathbf{z}}\in\mathcal{P}^s}} e(\mathfrak{F}(q\overline{\mathbf{w}}+\overline{\mathbf{z}};\boldsymbol{\beta})) - q^{-ms}\int_{\mathcal{P}} e(\mathfrak{F}(\overline{\boldsymbol{\zeta}};\boldsymbol{\beta}))\mathrm{d}\overline{\boldsymbol{\zeta}} \\ &= \sum_{\substack{\overline{\mathbf{w}}\in\mathbb{Z}^{ms}\\ q\overline{\mathbf{w}}+\overline{\mathbf{z}}\in\mathcal{P}^s}} \int_{\overline{\mathbf{w}}}^{\overline{\mathbf{w}}+1} \left\{ e(\mathfrak{F}(q\overline{\mathbf{w}}+\overline{\mathbf{z}};\boldsymbol{\beta})) - e(\mathfrak{F}(q\overline{\boldsymbol{\zeta}}+\overline{\mathbf{z}};\boldsymbol{\beta})) + O(1)\right\} \mathrm{d}\overline{\boldsymbol{\zeta}} \\ &\ll \left(\frac{(Pq^{-1})^m}{\det(C)}\right)^s \left\{ q\left(\sum_{\mathbf{j}\in J}|\beta_{\mathbf{j}}|\frac{P^d}{\hat{\gamma}_{\mathbf{j}}}\right)\frac{\gamma_{\max}}{P}\right\} \end{aligned}$$

by the Mean Value Theorem. The remaining term is just $S_q(\mathbf{a})$ and can be bounded trivially by q^{ms}, so altogether we have

$$T(\boldsymbol{\alpha},\mathcal{P}) - q^{-ms}S_q(\mathbf{a})v_{\mathcal{P}}(\boldsymbol{\beta}) \ll q^{ms}\left(\frac{(Pq^{-1})^m}{\det(C)}\right)^s q\left(\sum_{\mathbf{j}\in J}|\beta_{\mathbf{j}}|\frac{P^d}{\hat{\gamma}_{\mathbf{j}}}\right)\frac{\gamma_{\max}}{P}$$

as claimed, and the second inequality follows on inserting the major arcs bounds on q and $\boldsymbol{\beta}$. □

We define the truncated singular series and singular integral as

$$\mathfrak{S}(P) = \sum_{q=1}^{P^{r(d-1)\theta}} q^{-ms}\sum_{\substack{\mathbf{a}=0\\ (\mathbf{a},q)=1}}^{q-1} S_q(\mathbf{a})$$

and

$$\mathfrak{J}(\mathcal{P}) = \int_{|\beta_{\mathbf{j}}|\leq P^{-d+r(d-1)\theta}\hat{\gamma}_{\mathbf{j}}} v_{\mathcal{P}}(\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta},$$

respectively. This notation allows us to determine the overall error arising from this substitution by integrating the expression from Lemma 3.10 over $\mathfrak{M}'(P,\theta)$.

Lemma 3.11. *The total major arcs contribution is given by*

$$\int_{\mathfrak{M}'(P,\theta)} T(\boldsymbol{\alpha},\mathcal{P})\mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}(P)\mathfrak{J}(\mathcal{P}) + O\left(\left(\frac{P^m}{\det(C)}\right)^{s-rd/m} P^{(d-1)r(r+3)\theta-1}\gamma_{\max}\right).$$

Proof. In view of the information of Lemma 3.10 it suffices to compute the volume of the extended major arcs $\mathfrak{M}'(\theta)$. We have

$$\begin{aligned}\operatorname{vol}(\mathfrak{M}'(\theta)) &\ll \sum_{q=1}^{\kappa' P^{r(d-1)\theta}} \prod_{\mathbf{j}\in J} \left(\sum_{a_{\mathbf{j}}=0}^{q-1} P^{-d+r(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \right) \\ &\ll \sum_{q=1}^{\kappa' P^{r(d-1)\theta}} \left(q P^{-d+r(d-1)\theta} \right)^r \prod_{\mathbf{j}\in J} \hat{\gamma}_{\mathbf{j}} \\ &\ll \left(\frac{P^m}{\det(C)} \right)^{-rd/m} P^{(d-1)r(r+1)\theta},\end{aligned}$$

where the last step follows by (3.1.5). This, together with Lemma 3.10, implies the statement. □

We may now fix θ such that

$$0 < \theta < \frac{1-\eta}{(d-1)r(r+3)}$$

holds. Under this condition, Lemmata 3.9 and 3.11 can be combined to establish

$$N_C(P) = \mathfrak{S}(P)\mathfrak{J}(\mathcal{P}) + O\left(\left(\frac{P^m}{\det(C)} \right)^{s-rd/m} P^{(d-1)r(r+3)\theta-1+\eta} \right),$$

where our choice of θ ensures that the error term is $o(P^m/\det(C))^{s-rd/m})$ as desired.

As usual, the expected growth rate is encoded in the geometry of the problem and derives from normalising the singular integral.

Lemma 3.12.

$$\mathfrak{J}(\mathcal{P}) = \left(\frac{P^m}{\det(C)} \right)^{s-rd/m} \int_{|\boldsymbol{\beta}| \le P^{r(d-1)\theta}} v_1(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}.$$

Proof. Recall that

$$\mathfrak{F}(\overline{\mathbf{y}}; \boldsymbol{\beta}) = \sum_{\mathbf{j}\in J} \beta_{\mathbf{j}} \Phi(\mathbf{y}_{j_1}, \ldots, \mathbf{y}_{j_d}).$$

It follows that one has

$$\begin{aligned} v_{\mathcal{P}}(\boldsymbol{\beta}) &= \int_{|\mathbf{y}_i|\le P_i} e\left(\sum_{\mathbf{j}\in J}\beta_{\mathbf{j}}\Phi(\mathbf{y}_{j_1},\dots,\mathbf{y}_{j_d})\right)\mathrm{d}\overline{\mathbf{y}} \\ &= \left(\frac{P^m}{\det(C)}\right)^s \int_{|\overline{\mathbf{y}}|\le 1} e\left(\sum_{\mathbf{j}\in J}(P^d/\hat{\boldsymbol{\gamma}}_{\mathbf{j}})\beta_{\mathbf{j}}\Phi(\mathbf{y}_{j_1},\dots,\mathbf{y}_{j_d})\right)\mathrm{d}\overline{\mathbf{y}} \end{aligned}$$

and

$$\begin{aligned} &\int_{|\beta_{\mathbf{j}}|\le P^{-d+r(d-1)\theta}\hat{\boldsymbol{\gamma}}_{\mathbf{j}}}\int_{|\overline{\mathbf{y}}|\le 1} e\left(\sum_{\mathbf{j}\in J}(P^d/\hat{\boldsymbol{\gamma}}_{\mathbf{j}})\beta_{\mathbf{j}}\Phi(\mathbf{y}_{j_1},\dots,\mathbf{y}_{j_d})\right)\mathrm{d}\overline{\mathbf{y}}\mathrm{d}\boldsymbol{\beta} \\ &= \left(\frac{P^m}{\det(C)}\right)^{-rd/m}\int_{|\boldsymbol{\beta}|\le P^{r(d-1)\theta}} v_1(\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta}. \end{aligned}$$

The statement follows. □

As in Chapter 2 the truncated singular series and integral can be extended to infinity. We therefore define the complete singular series $\mathfrak{S}$ and the singular integral $\mathfrak{J}$ as

$$\mathfrak{S} = \sum_{q=1}^{\infty}\sum_{\substack{\mathbf{a}=0 \\ (\mathbf{a},q)=1}}^{q-1} q^{-ms}S_q(\mathbf{a})$$

and

$$\mathfrak{J} = \int_{\mathbb{R}^r} v_1(\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta}.$$

Notice that these are identical to those considered in Chapter 2 and converge absolutely by Lemmata 2.12, 2.13 and 2.14, provided

$$k > 3(d-1)(r+1).$$

Together with (3.3.8) and (3.4.2), we thus obtain the final bound for our number of variables

$$s^* > 2^{d-1}\max\left\{3(d-1)(r+1), \frac{rd}{1-\eta}\right\},$$

and on inserting (3.3.3) one recovers the enunciation of the theorem.

It remains to remark that again the singular series $\mathfrak{S}$ can be expanded as an Euler product

$$\mathfrak{S} = \prod_p \chi(p),$$

and the p-adic densities $\chi(p)$ have the interpretation as a p-adic integral

$$\chi(p) = \int_{\mathbb{Q}_p^r} \int_{\mathbb{Z}_p^{ms}} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \boldsymbol{\eta})\right) \mathrm{d}\overline{\boldsymbol{\xi}} \mathrm{d}\boldsymbol{\eta}.$$

Similarly, the singular integral

$$\mathfrak{J} = \int_{\mathbb{R}^r} \int_{|\overline{\boldsymbol{\xi}}| \le 1} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \boldsymbol{\beta})\right) \mathrm{d}\overline{\boldsymbol{\xi}} \mathrm{d}\boldsymbol{\beta}$$

measures solutions in the real unit box and may thus be interpreted as the density of real solutions.

Chapter 4

The local solution densities for forms in many variables

4.1 Motivation and background

Thanks to the groundbreaking work of Davenport [13–15] and Birch [8] and later Schmidt [62], the treatment of forms in many variables is now one of the most classical applications of the circle method. However, the generic outcome of the circle method being only a local-global principle, the transition to asymptotic estimates of the number of solutions depends on the additional condition that there exist solutions over the local fields that are not only non-trivial but non-singular, and for this problem no fully satisfactory methods have been available so far. Indeed, while non-singular real solubility is by now well understood at least if the degree is odd, establishing comparable statements over $\mathbb{Q}_p$ is a far more delicate problem. In addition to giving an account of the methods commonly employed to obtain information concerning the p-adic solubility and extending them to the multidimensional situation, this chapter will be devoted to showing how the gap between non-trivial and non-singular p-adic solubility can be crossed. These new methods require a somewhat more complicated non-singularity condition, which coincides with the traditional one in the case $m = 1$, so it is in this situation that the new methods will be most efficient.

Let $H_d(R, m)$ be the least integer H such that for every system $\mathbf{F} = (F^{(1)}, \dots, F^{(R)})$ of R forms of degree d in at least $s^* \geq H$ variables there exist nonnegative constants χ_∞ and χ_p for every prime p such that the number $N^{(d)}_{s,R,m}(P)$ of linear spaces of affine dimension m and height at most P contained in the complete intersection $\mathbf{F} = \mathbf{0}$ obeys the asymptotic formula

$$N^{(d)}_{s,R,m}(P) = P^{ms-Rrd}\chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(P^{ms-Rrd}\right),$$

where the parameter r is given by

$$r = \binom{d-1+m}{d}. \tag{4.1.1}$$

Similarly, let $\gamma_d^p(R, m)$ denote the least integer γ such that any system of R forms of degree d in $s \geq \gamma$ variables contains a p-adic linear space of affine dimension m. These are bounded above, so we can safely define

$$\gamma_d^*(R, m) = \sup_{p \text{ prime}} \gamma_d^p(R, m)$$

to be the least number of variables for which there are no p-adic obstructions to the solubility of the linear spaces problem. We stress here that $\gamma_d^*(R, m)$ is defined in terms of s itself, and not in terms of s^*. This circumstance will become important later on in the analysis.

We denote by $\mathfrak{V}_m$ the singular locus related to the linear spaces problem; this notion will be elaborated on in Section 4.2. For the moment we remark that $\mathfrak{V}_1 = \mathfrak{V}$. Similar to the one-dimensional case, we write $(ms)^* = ms - \dim \mathfrak{V}_m$. In this notation, the main theorem of this chapter is the following.

Theorem 4.1. *Let $\mathbf{F} = (F^{(1)}, \dots, F^{(R)})$ be forms of odd degree d. Furthermore, suppose that $s^* > H_d(R, m)$ and $(ms)^* > \gamma_d^*(R, m)$. Then one has*

$$N^{(d)}_{s,R,m}(P) = P^{ms-Rrd}\chi_\infty \prod_{p \text{ prime}} \chi_p + o(P^{ms-Rrd}) \tag{4.1.2}$$

and the product of local densities $\chi_\infty \prod_p \chi_p$ is positive.

Notice that the bound $s^* > H_d(R,m)$ follows just from the definition of $H_d(R,m)$, and the factor χ_∞ can be shown to be positive by an argument due to Schmidt (see the treatment in [60, §2] and [61, Lemma 2 and §11]), which generalises easily to higher dimensions but which we will reproduce at the end of the chapter nonetheless. The significance of Theorem 4.1 lies therefore in the bound $(ms)^* > \gamma_d^*(R,m)$ for non-singular p-adic solubility. This bound clearly supersedes the previous treatment of the case $m = 1$ due to Schmidt (see the Supplement to Proposition I in [62]), which requires

$$s^* > 2^{d-1}(d-1)R\gamma_d^*(R,1) \tag{4.1.3}$$

in order to ensure that the p-adic densities be positive.

As to numerical values, recall that Birch's Theorem provides the upper bound

$$H_d(R,1) \leq 2^{d-1}(d-1)R(R+1), \tag{4.1.4}$$

whereas [77, Cor. 1.1] implies that

$$\gamma_d^*(R,1) \leq (Rd^2)^{2^{d-1}}.$$

Thus Theorem 4.1 implies that in the case of Birch's Theorem we have asymptotic behaviour as soon as d is odd and

$$s^* > \max\left\{2^{d-1}(d-1)R(R+1), (Rd^2)^{2^{d-1}}\right\}.$$

Notice that the second term dominates for all positive values of R and d.

Unfortunately, the case $m > 1$ is more complicated due to the more involved nature of the non-singularity condition in Theorem 4.1. In fact, we expect $(ms)^* = s^*$ to be true, so the bound on the number of variables in Theorem 4.1 would become $s^* > \max\{H_d(R,m), \gamma_d^*(R,m)\}$, but the proof of this presents some difficulties. Nonetheless it is desirable to have a formulation avoiding the more complicated non-singularity condition $\mathfrak{V}_m$. This can be achieved by an adaptation of Schmidt's bound (4.1.3) to linear spaces.

Theorem 4.2. *Let R, d odd and $m \geq 2$ be as above and suppose*

$$s^* > 3 \cdot 2^{d-1}(d-1)R\max\{Rr+1, \gamma_d^*(m,R)\}.$$

Then one has an asymptotic formula as in (4.1.2) *and the product of the local densities is positive.*

Again, $\gamma_d^*(R,m)$ can be controlled by inserting bounds from the literature. In fact, by inserting the bound from [18, Theorem 1] into [77, Theorem 2.4] one obtains

$$\gamma_d^*(R,m) \leq (R^2d^2 + mR)^{2^{d-2}} d^{2^{d-1}}.$$

In the cubic case one can use the explicit value given in the remark following [18, Theorem 1] instead, and then [77, Theorem 2.4] gives

$$\gamma_3^*(R,m) \leq 10(6R^2 + mR)^2.$$

With these values, the bound of Theorem 4.2 takes the shape

$$s^* > \begin{cases} 3 \cdot 2^{d-1}(d-1)d^{2^{d-1}} R\left(R^2d^2 + Rm\right)^{2^{d-2}} & \text{for } d \geq 5, \\ 24R\max\left\{10(6R^2+mR)^2, Rr+1\right\} & \text{if } d = 3. \end{cases} \tag{4.1.5}$$

The case distinction in the cubic case arises from the somewhat surprising fact that one can ensure the existence of local p-adic solutions for some choices of m and R with much looser conditions on the number of variables than what is needed to establish a Hasse principle, where s is needed to grow at least cubically in m by Theorem 2.1 and (4.1.1). Furthermore, note that due to the generality of the setting, the bound given in (4.1.5) will not be sharp for any typical set of parameters, and in any given special case the numerical constants can be improved just by inserting the best available bounds for the respective situations.

It should be mentioned that the methods used here do not yield honest linear spaces unless we can ensure that the span of the vectors is of the right dimension. Suppose for convenience that m is even. Then it follows from the argument in [42, p. 283] that the set of $\overline{\mathbf{x}}$ in $\mathbb{Z}^{ms}$ determining a subspace of dimension at most $m/2$ forms a linear space of dimension $ms/2 + (m/2)^2$. It follows that the number of such $\overline{\mathbf{x}}$ of height at most P is $O(P^{ms/2+(m/2)^2})$, so if one has

$$ms - Rrd > ms/2 + (m/2)^2,$$

the number of vectors $\overline{\mathbf{x}}$ solving the system

$$F^{(\rho)}(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) = 0, \quad 1 \leq \rho \leq R,$$

exceeds the number of vectors whose span is of dimension at most $m/2$. This implies that we can find distinct vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ contained in the intersection of the R hyperplanes that are of rank at least $m/2$, provided that the number of variables satisfies

$$s > 2Rrd/m + m/2.$$

This requirement is, however, easily implied by the conditions of our theorems.

4.2 The singularity condition for linear spaces

A crucial stepping stone in the proof of Theorem 4.1 is a suitable version of Hensel's Lemma adapted for linear space situations. As is usual with arguments in the spirit of Hensel's Lemma, this involves finding a non-singular approximate solution which can then be lifted to higher moduli and ultimately generate solutions in $\mathbb{Q}_p$. It is, however, not obvious what shape the non-singularity condition takes in our case, as the multi-dimensional situation has a higher degree of intrinsic complexity than the one-dimensional setting. This makes it necessary to spend some time specifying what it means for a solution to be non-singular.

We adopt all the notation introduced in Section 2.2. Recall that one has a correspondence between linear spaces on the complete intersection defined by $\mathbf{F} = \mathbf{0}$ and points on an expanded system of equations which is given as follows. For $1 \leq \rho \leq R$ let $\Phi^{(\rho)}$ be the multilinear form associated to $F^{(\rho)}$, then one has

$$F^{(\rho)}\left(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m\right) = \sum_{\mathbf{j}\in J} A(\mathbf{j}) t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d} \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}), \tag{4.2.1}$$

where J is the set of multi-indices $\{j_1, \ldots, j_d\} \in \{1, \ldots, m\}^d$, $A(\mathbf{j})$ is a combinatorial factor and we abbreviate $\Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}) = \Phi^{(\rho)}(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_d})$.

Let $\Gamma^{(d)}_{s,R,m}(p^l)$ denote the number of solutions $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{Z}/p^l\mathbb{Z}$ of the simultaneous congruences

$$F^{(\rho)}(\mathbf{x}_1 t_1 + \ldots + \mathbf{x}_m t_m) \equiv 0 \pmod{p^l}, \quad 1 \leq \rho \leq R,$$

identically in $t_1, \ldots, t_m$. By (4.2.1) this can be written as

$$\Gamma^{(d)}_{s,R,m}(p^l) = \operatorname{Card}\left\{\overline{\mathbf{a}} \in (\mathbb{Z}/p^l\mathbb{Z})^{ms} \colon \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}}) \equiv 0 \pmod{p^l} \text{ for all } \mathbf{j}, \rho\right\}. \tag{4.2.2}$$

Suppose $\overline{\mathbf{a}} \in (\mathbb{Z}_p)^{ms}$ is a solution to

$$\Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}}) = 0 \quad \text{ for all } \mathbf{j} \in J, 1 \leq \rho \leq R. \tag{4.2.3}$$

Consider a small disturbance $\overline{\mathbf{a}} + p^h\overline{\mathbf{x}}$ of $\overline{\mathbf{a}}$, then an application of Taylor's Theorem yields

$$\begin{aligned}\Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}} + p^h\overline{\mathbf{x}}) &= \Phi^{(\rho)}_{\mathbf{j}}(\mathbf{a}_{j_1} + p^h\mathbf{x}_{j_1}, \mathbf{a}_{j_2} + p^h\mathbf{x}_{j_2}, \ldots, \mathbf{a}_{j_d} + p^h\mathbf{x}_{j_d}) \\ &= \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}}) + p^h L^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}}) \cdot \overline{\mathbf{x}} + R^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}})(p^h\overline{\mathbf{x}}),\end{aligned} \tag{4.2.4}$$

where $\Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}}) = 0$ by assumption, $R^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}})(\overline{\mathbf{x}})$ is a polynomial in $\overline{\mathbf{x}}$ all of whose terms are of degree ≥ 2, and $L^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{a}})$ is defined as follows. If $\mathbf{e}_n$ is the n-th unit vector, we write

$$B^{(\rho)}_n(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d-1)}) = \Phi^{(\rho)}(\mathbf{e}_n, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d-1)}),$$

so that

$$\Phi^{(\rho)}(\mathbf{x}, \mathbf{a}_1, \ldots, \mathbf{a}_{d-1}) = \sum_{n=1}^{s} L_n(\mathbf{a}_1, \ldots, \mathbf{a}_{d-1}) x_n.$$

Then we can define $L_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}})$ via the relation

$$\begin{aligned} L_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) \cdot \overline{\mathbf{x}} &= \sum_{k=1}^{d} \Phi^{(\rho)}(\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_{k-1}}, \mathbf{x}_{j_k}, \mathbf{a}_{j_{k+1}}, \ldots, \mathbf{a}_{j_d}) \\ &= \sum_{k=1}^{d} \sum_{n=1}^{s} B_n^{(\rho)}(\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_{k-1}}, \mathbf{a}_{j_{k+1}}, \ldots, \mathbf{a}_{j_d}) x_{j_k,n} \\ &= \sum_{i=1}^{m} \sum_{n=1}^{s} \left[\sum_{k=1}^{d} B_n^{(\rho)}(\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_{k-1}}, \mathbf{a}_{j_{k+1}}, \ldots, \mathbf{a}_{j_d}) \delta_{j_k,i} \right] x_{i,n}, \end{aligned}$$

where $\delta_{i,j}$ denotes the Kronecker delta. Writing $L_{\mathbf{j},n,i}^{(\rho)}(\overline{\mathbf{a}})$ for the term in the square brackets, one has the $(Rr \times ms)$-matrix

$$\mathscr{L}(\overline{\mathbf{a}}) = \left[L_{\mathbf{j},n,i}^{(\rho)}(\overline{\mathbf{a}}) \right]_{\mathbf{j},\rho;n,i}.$$

A solution $\overline{\mathbf{a}}$ to the system given in (4.2.3) is therefore considered singular if $\operatorname{rank}(\mathscr{L}(\overline{\mathbf{a}})) \leq Rr - 1$, and we can properly define

$$\mathfrak{V}_m = \left\{ \overline{\mathbf{x}} \in \mathbb{A}^{ms}(\mathbb{Q}_p) : \operatorname{rank}(\mathscr{L}(\overline{\mathbf{x}})) \leq Rr - 1 \right\}.$$

As in the introduction, we remark that for this definition we do not require the singular locus $\mathfrak{V}_m$ to be a subset of the variety given by $\mathbf{F} = \mathbf{0}$.

For an integer matrix $A \in \mathbb{Z}^{n \times m}(\mathbf{x})$ with $n \leq m$ let $\operatorname{ord}_p(A)$ denote the least integer h such that p^h does not divide the greatest common divisor of the determinants of the $(n \times n)$-minors of A, if such a number exists, and $\operatorname{ord}_p(A) = \infty$ otherwise. If the matrix A depends on a parameter $\mathbf{x}$ we will write $\operatorname{ord}_p(A) = \min_{\mathbf{x}} \operatorname{ord}_p(A(\mathbf{x}))$. This means that when the system is non-singular in $\mathbb{Q}_p$, it is in some sense non-singular over $\mathbb{Z}/p^h\mathbb{Z}$ for all powers h satisfying $h \geq \operatorname{ord}_p(\mathscr{L}(\overline{\mathbf{x}}))$.

4.3 Hensel's Lemma for linear spaces

Suppose $\mathrm{ord}_p(\mathscr{L}) = \sigma < \infty$, so the system is non-singular. Then for some $\overline{\mathbf{a}}$ the matrix $\mathscr{L}(\overline{\mathbf{a}})$ has an $(Rr \times Rr)$-minor $\mathscr{L}_0(\overline{\mathbf{a}})$ with the property

$$|\det \mathscr{L}_0(\overline{\mathbf{a}})|_p = p^{1-\sigma};$$

we can assume without loss of generality that it is the first minor. Write $M(\sigma,\nu)$ for the number of $\overline{\mathbf{a}} \in (\mathbb{Z}/p^{2\sigma-1+\nu}\mathbb{Z})^{ms}$ such that

$$\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) \equiv 0 \pmod{p^{2\sigma-1+\nu}} \qquad \text{for all } \mathbf{j} \in J, 1 \le \rho \le R, \tag{4.3.1}$$

$$|\det \mathscr{L}_0(\overline{\mathbf{a}})|_p = p^{1-\sigma}. \tag{4.3.2}$$

Then $M(\sigma,\nu)$ can be bounded below.

Lemma 4.1. *Suppose $ms \ge Rr$. For any $\nu \ge 0$ the number of $\overline{\mathbf{a}} \pmod{p^{2\sigma-1+\nu}}$ subject to (4.3.1) and (4.3.2) is at least*

$$M(\sigma,\nu) \ge p^{(ms-Rr)\nu} M(\sigma,0).$$

Proof. The proof is an amalgam of the arguments of [17, Lemma 17.1] and [27, Prop. 5.20]. For $M(\sigma,0) = 0$ the result is immediate, so it suffices to consider the case $M(\sigma,0) > 0$. Also, the lemma is trivially true for $\nu = 0$. We can therefore proceed by induction and investigate $M(\sigma,\nu+1)$ under the assumption that for some given $\nu \ge 0$ the statement is true. Let $\overline{\mathbf{a}}$ be one of the $\ge p^{(ms-Rr)\nu} M(\sigma,0)$ solutions to (4.3.1) and (4.3.2) counted by $M(\sigma,\nu)$. As in (4.2.4), Taylor's Theorem yields

$$\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}} + p^{\sigma+\nu}\overline{\mathbf{x}}) \equiv \Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) + p^{\sigma+\nu} L_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) \cdot \overline{\mathbf{x}} \pmod{p^{2\sigma+2\nu}},$$

and by (4.3.1) we can write $\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) = p^{2\sigma-1+\nu}\varphi_{\mathbf{j}}^{(\rho)}$ for some $\varphi_{\mathbf{j}}^{(\rho)} \in \mathbb{Z}$. Also, by (4.3.2) there is a unimodular matrix $\mathscr{N} \in \mathbb{Z}^{Rr\times Rr}$ such that

$$\mathscr{N}\mathscr{L}_0(\overline{\mathbf{a}}) = \det(\mathscr{L}_0(\overline{\mathbf{a}}))\,\mathrm{Id}_{Rr} = \beta p^{\sigma-1}\,\mathrm{Id}_{Rr}$$

for some β coprime to p, so the system of congruences

$$\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) + p^{\sigma+\nu} L_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{a}}) \cdot \overline{\mathbf{x}} \equiv 0 \pmod{p^{2\sigma+\nu}}, \quad \mathbf{j} \in J, 1 \le \rho \le R,$$

is equivalent to

$$p^{2\sigma-1+\nu}\mathscr{N}\boldsymbol{\varphi}+p^{\sigma+\nu}\mathscr{N}\mathscr{L}(\overline{\mathbf{a}})\cdot\overline{\mathbf{x}}\equiv 0 \pmod{p^{2\sigma+\nu}}. \tag{4.3.3}$$

The matrix $\mathscr{N}\mathscr{L}(\overline{\mathbf{a}})$ can be written in block structure as

$$\mathscr{N}\mathscr{L}(\overline{\mathbf{a}})=\mathscr{N}\left(\mathscr{L}_0(\overline{\mathbf{a}})\middle|\mathscr{L}_1(\overline{\mathbf{a}})\right)=\left(\beta p^{\sigma-1}\,\mathrm{Id}_{Rr}\middle|\mathscr{N}\mathscr{L}_1(\overline{\mathbf{a}})\right)$$

for some matrix $\mathscr{L}_1(\overline{\mathbf{a}})$, so if we restrict ourselves to counting solutions $\overline{\mathbf{x}}$ of the form

$$x_i=\begin{cases} y_i & 1\leq i\leq Rr,\\ p^{\sigma-1}y_i & Rr+1\leq i\leq ms,\end{cases}$$

we can write

$$\mathscr{N}\mathscr{L}(\overline{\mathbf{a}})\cdot\overline{\mathbf{x}}=p^{\sigma-1}\hat{\mathscr{L}}\cdot\overline{\mathbf{y}},$$

where

$$\hat{\mathscr{L}}=\left(\beta\,\mathrm{Id}_{Rr}\middle|\mathscr{N}\mathscr{L}_1(\overline{\mathbf{a}})\right).$$

Thus the number of solutions to (4.3.3) is at least as big as the solution set of

$$\mathscr{N}\boldsymbol{\varphi}+\hat{\mathscr{L}}\overline{\mathbf{y}}\equiv 0 \pmod{p}.$$

Since $\mathscr{N}$, $\boldsymbol{\varphi}$ and $\hat{\mathscr{L}}$ are fixed, this can be considered as a non-singular system of Rr linear equations in ms variables and therefore has p^{ms-Rr} solutions. Thus altogether we have

$$M(\sigma,\nu+1)\geq p^{(ms-Rr)}M(\sigma,\nu)\geq p^{(ms-Rr)(\nu+1)}M(\sigma,0)$$

as claimed. □

Lemma 4.1 shows that as soon as there is one non-singular solution $\overline{\mathbf{a}}$ counted by $M(\sigma,0)$ for some $\sigma<\infty$, it can be lifted to a p-adic solution and thus establishes a form of Hensel's Lemma for linear spaces. Notice also that for $m=1$ one has $r=1$ and hence Theorem 4.4 reproduces standard results such as [17, Lemma 17.1].

Write

$$\mathfrak{X} = \left\{ \overline{\mathbf{x}} \in \mathbb{A}^{ms}(\mathbb{Q}_p) \colon \Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}}) = 0 \quad \text{for all } \mathbf{j} \in J, 1 \leq \rho \leq R \right\}$$

for the space of solutions to the system (4.2.3). It is clear from our previous considerations that if $\mathfrak{X} \backslash \mathfrak{V}_m \neq \emptyset$, there exists an integer σ such that $M(\sigma, 0) > 0$. We can therefore find a non-singular solution and lift it to a p-adic solution by means of Lemma 4.1. This establishes an alternative version of Hensel's Lemma.

Theorem 4.3. *Let $F^{(1)}, \ldots, F^{(R)} \in \mathbb{Z}[x_1, \ldots, x_s]$ be forms of degree d, and m an integer. Furthermore, let $ms \geq rR$. Suppose that one has*

$$\mathfrak{X} \setminus \mathfrak{V}_m \neq \emptyset.$$

Then there exists a constant c_p depending only on the system $\mathbf{F}$ such that for any $\nu \in \mathbb{N}$ one has

$$\Gamma_{s,R,m}^{(d)}(p^\nu) \geq c_p p^{\nu(ms - Rr)}.$$

In particular, the variety defined by $\mathbf{F} = \mathbf{0}$ contains a non-singular p-adic linear space of dimension m.

Note that Lemma 4.1 allows us to take $c_p = p^{(1-2\sigma)(ms-Rr)}$, where $\sigma = \mathrm{ord}_p(\mathscr{L})$ is a finite parameter depending only on p, the dimension m and the system $F^{(1)}, \ldots, F^{(\rho)}$.

In its present shape, our new version of Hensel's Lemma is only of limited use to us, as it is not well adapted to our general state of knowledge. In fact, while there are many results available about the least number of variables required to guarantee the existence of p-adic solutions to polynomial equations, we tend to have little quantitative information regarding the number of solutions for any given modulus or over $\mathbb{Q}_p$. It would therefore be desirable to rephrase the conditions in Theorem 4.3 in terms of existence of p-adic solutions only. This is indeed possible.

Lemma 4.2. *Suppose*

$$ms - \dim \mathfrak{V}_m \geq \gamma_d^p(R, m).$$

Then the system (4.2.3) *possesses a non-singular p-adic solution.*

Proof. Let $b = \dim \mathfrak{V}_m$. Recall that $\mathfrak{V}_m$ really is a projective variety, so it is of projective dimension $b-1$. By a suitable version of Bertini's Theorem we can find a hyperplane $H_1 \subset \mathbb{P}_{\mathbb{Q}_p}^{ms-1}$ such that

$$\dim_{\text{proj}}(\mathfrak{V}_m \cap H_1) \leq \max\{-1, b-2\}.$$

Thus after k iterations one obtains a hyperplane section

$$\mathfrak{H}_k = H_1 \cap \cdots \cap H_k$$

such that the restriction of the expanded singular locus $\mathfrak{V}_m \cap \mathfrak{H}_k$ is of projective dimension at most $b-k-1$. For $k=b$ this equals -1, so we can infer that the set $\mathfrak{V}_m \cap \mathfrak{H}_b$ is empty as a projective variety and thus contains only the origin as an affine variety. Hence all non-zero elements of $\mathfrak{X} \cap \mathfrak{H}_b$ are non-singular. It suffices therefore to show that the set $\mathfrak{X} \cap \mathfrak{H}_b$ contains a non-trivial point.

Since all the H_k can be viewed as hyperplanes contained in $\mathbb{A}_{\mathbb{Q}_p}^{ms}$, the space $\mathfrak{H}_b$ is isomorphic to $\mathbb{A}_{\mathbb{Q}_p}^{ms-b}$. In order to show that $\mathfrak{X} \cap \mathfrak{H}_b$ is non-empty, it is therefore enough to show that the system $\mathbf{F} = \mathbf{0}$ contains an m-dimensional p-adic space contained in $\mathfrak{H}_b$. However, by the definition of $\gamma_d^p(R, m)$ this is guaranteed if

$$ms - b \geq \gamma_d^p(R, m).$$

□

Altogether, on combining Lemma 4.2 with Theorem 4.3 we obtain the following.

Theorem 4.4. *Let $F^{(1)}, \ldots, F^{(R)} \in \mathbb{Z}[x_1, \ldots, x_s]$ be as above, and m an integer. Furthermore, suppose that $ms \geq Rr$, and that $(ms)^* \geq \gamma_d^p(R, m)$. Then for any $\nu \in \mathbb{N}$ one has*

$$\Gamma_{s,R,m}^{(d)}(p^\nu) \geq c_p p^{\nu(ms-Rr)},$$

where the constant is as in Theorem 4.3.

This implies that for any given system of equations which is not too singular the quantity $\Gamma^{(d)}_{s,R,m}(p^h)$ is of the expected order of magnitude essentially as soon as it has a non-trivial p-adic solution.

4.4 The singular series in Birch's Theorem

In order to deduce Theorem 4.1 from Theorem 4.4 it is necessary to recall some of the notation of Chapter 2. In (2.6.8) we showed that the p-adic solution density χ_p can be expressed as

$$\chi_p = \sum_{l=0}^{\infty} p^{-lms} \sum_{\overline{\mathbf{x}}=1}^{p^l} \sum_{\substack{\underline{\mathbf{u}}=1 \\ (\underline{\mathbf{u}},p)=1}}^{p^l} e\left(\mathfrak{F}(\overline{\mathbf{x}}; p^{-l}\underline{\mathbf{u}})\right), \tag{4.4.1}$$

where we followed the notation from Section 2.6 in writing $\underline{\mathbf{u}} = (u^{(\rho)}_{\mathbf{j}})_{\mathbf{j},\rho}$ and

$$\mathfrak{F}\left(\overline{\mathbf{x}}; \underline{\mathbf{u}}\right) = \sum_{\rho=1}^{R} \sum_{\mathbf{j}\in J} u^{(\rho)}_{\mathbf{j}} \Phi^{(\rho)}_{\mathbf{j}}(\overline{\mathbf{x}}).$$

The expression in (4.4.1) can be transformed by what are essentially standard operations into the following, more intuitively accessible version.

Lemma 4.3. *We have*

$$\sum_{i=0}^{l} p^{-ims} \sum_{\overline{\mathbf{x}}=1}^{p^i} \sum_{\substack{\underline{\mathbf{u}}=1 \\ (\underline{\mathbf{u}},p)=1}}^{p^i} e\left(\mathfrak{F}(\overline{\mathbf{x}}; p^{-i}\underline{\mathbf{u}})\right) = p^{l(Rr-ms)} \Gamma^{(d)}_{s,R,m}(p^l),$$

and consequently

$$\chi_p = \lim_{l\to\infty} p^{l(Rr-ms)} \Gamma^{(d)}_{s,R,m}(p^l).$$

Proof. This is like [17, Lemma 5.3]. Notice that the vector $\underline{\mathbf{u}}$ has Rr components, so by (4.2.2) and standard orthogonality relations one has

$$\Gamma^{(d)}_{s,R,m}(p^l) = p^{-lRr} \sum_{\overline{\mathbf{x}}=1}^{p^l} \sum_{\underline{\mathbf{u}}=1}^{p^l} e\left(\mathfrak{F}(\overline{\mathbf{x}}; p^{-l}\underline{\mathbf{u}})\right).$$

Furthermore, we can force a coprimality condition on the sum by writing

$$\begin{aligned}\sum_{\overline{\mathbf{x}}=1}^{p^l}\sum_{\underline{\mathbf{u}}=1}^{p^l} e\left(\mathfrak{F}(\overline{\mathbf{x}};p^{-l}\underline{\mathbf{u}})\right) &= \sum_{i\leq l}\sum_{\substack{\underline{\mathbf{u}}=1\\(\underline{\mathbf{u}},p)=1}}^{p^i}\sum_{\overline{\mathbf{x}}=1}^{p^l} e\left(\mathfrak{F}(\overline{\mathbf{x}};p^{-i}\underline{\mathbf{u}})\right)\\ &= \sum_{i\leq l} p^{(l-i)ms}\sum_{\substack{\underline{\mathbf{u}}=1\\(\underline{\mathbf{u}},p)=1}}^{p^i}\sum_{\overline{\mathbf{x}}=1}^{p^i} e\left(\mathfrak{F}(\overline{\mathbf{x}};p^{-i}\underline{\mathbf{u}})\right).\end{aligned}$$

Combining these two statements and taking the limit $l\to\infty$ completes the proof. □

By Lemma 2.14 the product of the χ_p converges under the conditions of Theorem 4.1, so in order to establish positivity of the singular series it remains to show that each individual χ_p is positive, and by Lemma 4.3 this is the case as soon as there is a constant c_p such that

$$\Gamma^{(d)}_{s,R,m}(p^l)\geq c_p p^{(ms-Rr)l}$$

for all sufficiently large l. This is, however, immediate from Theorem 4.4, provided that one has $(ms)^*\geq\gamma_d^p(R,m)$. It follows that $\chi_p>0$ for all primes p.

In the case of Theorem 4.2, the unwieldy non-singularity condition in the case $m\geq 2$ means that we are prevented from using Theorem 4.4 and have to employ different methods instead.

Lemma 4.4. *We have*

$$\Gamma^{(d)}_{s,R,m}(p^l)\gg p^{l\left(ms-\gamma_d^*(R,m)\right)}.$$

Proof. This is a generalised version of [60, Lemma 2]. The proof counts the number of solutions that are primitive in the sense that they do not vanish modulo p. We consider the additive group $X=(\mathbb{Z}/p^l\mathbb{Z})^{ms}$, and say a subgroup $H\subset X$ has property (P) if it is the sum of $g=\gamma_d^p(R,m)+1$ cyclic groups of order p^l.

We now let

$$\begin{aligned}
\alpha_1 &= \mathrm{Card}\{H \subset X \colon H \text{ has property (P)}\},\\
\alpha_2(\overline{\mathbf{x}}) &= \mathrm{Card}\{H \subset X \colon \overline{\mathbf{x}} \in H \text{ and } H \text{ has property (P)}\},\\
\beta_1 &= \mathrm{Card}\{\overline{\mathbf{x}} \in X \colon \mathrm{ord}(\overline{\mathbf{x}}) = p^l\},\\
\beta_2(H) &= \mathrm{Card}\{\overline{\mathbf{x}} \in H \colon \mathrm{ord}(\overline{\mathbf{x}}) = p^l\}.
\end{aligned}$$

By symmetry considerations one sees that $\alpha_2(\overline{\mathbf{x}})$ is constant for all $\overline{\mathbf{x}}$ of group order p^l, so we denote that value by α_2. Similarly, $\beta_2(H)$ will take the same value β_2 for all H with property (P). Notice that

$$\sum_{\overline{\mathbf{x}} \colon \mathrm{ord}(\overline{\mathbf{x}}) = p^l} \sum_{\substack{H \ni \overline{\mathbf{x}} \\ H \text{ has property (P)}}} 1 = \sum_{H \colon H \text{ has property (P)}} \sum_{\substack{\overline{\mathbf{x}} \in H \\ \mathrm{ord}(\overline{\mathbf{x}}) = p^l}} 1,$$

which in our notation can be simplified to $\beta_1\alpha_2 = \alpha_1\beta_2$. It follows that we have

$$\frac{\alpha_1}{\alpha_2} = \frac{\beta_1}{\beta_2} = \frac{p^{lms} - p^{(l-1)ms}}{p^{lg} - p^{(l-1)g}} = p^{l(ms-g)} \frac{1-p^{-ms}}{1-p^{-g}}.$$

By the definition of $\gamma_d^p(R,m)$ every system of R forms in at least g variables contains a non-trivial m-dimensional linear space in $\mathbb{Q}_p$ and thus by homogeneity in $\mathbb{Z}_p$. In our notation in this thesis, this means that there is an ms-tuple $\overline{\mathbf{x}} \in \mathbb{Z}_p^{ms}$ that solves the system (4.2.3). It follows that every subgroup H with property (P) contains a solution $\overline{\mathbf{x}}$ that solves the system modulo p^l and which is non-vanishing modulo p. By homogeneity again all its multiples are also solutions, and $p^l - p^{l-1}$ of these are also primitive. Hence the number $\widehat{\Gamma}^{(d)}_{s,R,m}(p^l)$ of primitive solutions of the system is

$$\begin{aligned}
\widehat{\Gamma}^{(d)}_{s,R,m}(p^l) &\geq (p^l - p^{l-1}) \frac{\alpha_1}{\alpha_2}\\
&= (p^l - p^{l-1}) p^{l(ms-g)} \frac{1-p^{-ms}}{1-p^{-g}}\\
&= p^{l(ms-g+1)} \frac{(1-p^{-ms})(1-p^{-1})}{1-p^{-g}},
\end{aligned}$$

and the last factor is $\gg 1$. Thus indeed with $g = \gamma_3^p(R,m)+1$ we have

$$\Gamma^{(d)}_{s,R,m}(p^l) \geq \widehat{\Gamma}^{(d)}_{s,R,m}(p^l) \gg p^{l(ms-g+1)} = p^{l\left(ms-\gamma_3^p(R,m)\right)}.$$

□

With this result at our hands we can proceed. Let p be fixed, then we split the factors χ_p and write

$$\chi_p = \sum_{i=0}^{l} \sum_{\substack{|\underline{\mathbf{a}}|<p^i \\ (\underline{\mathbf{a}},p)=1}} p^{-ims} S_{p^i}(\underline{\mathbf{a}}) + \sum_{i=l+1}^{\infty} \sum_{\substack{|\underline{\mathbf{a}}|<p^i \\ (\underline{\mathbf{a}},p)=1}} p^{-ims} S_{p^i}(\underline{\mathbf{a}}) = I_l + I_\infty,$$

where we adopted the notation

$$S_q(\underline{\mathbf{a}}) = \sum_{\overline{\mathbf{x}}=1}^{q} e\left(\frac{\mathfrak{F}(\overline{\mathbf{x}};\underline{\mathbf{a}})}{q}\right)$$

as in (2.6.2). We show that I_l is the dominant term and characterises the number of solutions modulo p^l. Indeed, by Lemmata 4.3 and 4.4 one has

$$I_l = p^{l(Rr-ms)} \Gamma^{(d)}_{s,R,m}(p^l) \gg p^{l(Rr-ms)} p^{l(ms-\gamma_d^p(R,m))} \gg p^{l(Rr-\gamma_d^p(R,m))}.$$

Now consider I_∞. Choosing $W > \max\{\gamma_d^p(R,m), Rr\}$ in Lemma 2.13 yields

$$I_\infty \ll \sum_{i=l+1}^{\infty} \sum_{\substack{|\underline{\mathbf{a}}|<p^i \\ (\underline{\mathbf{a}},p)=1}} p^{-ims} S_{p^i}(\underline{\mathbf{a}}) \ll \sum_{i=l+1}^{\infty} p^{i(Rr-W)} \ll p^{l(Rr-\gamma_d^p(R,m)-\delta)}$$

for some $\delta > 0$. Together, these two estimates yield $\chi_p = I_l + I_\infty \gg 1$ for some suitable l, provided that

$$k > 3(d-1)RW > 3(d-1)R\max\{\gamma_d^p(R,m), Rr\}.$$

Recalling that in order to prove convergence of the singular series and integral in Lemma 2.14 we had the additional requirement $k > 3R(d-1)(Rr+1)$, it follows that one has an asymptotic formula with $\prod_p \chi_p > 0$ as soon as

$$s^* > 3 \cdot 2^{d-1}(d-1)R\max\{\gamma_d^p(R,m), Rr+1\}.$$

This proves the p-adic aspect of Theorem 4.2, and replacing $\gamma_d^p(R,m)$ by $\gamma_d^*(R,m)$ therefore implies that the product over all p-adic densities is positive.

4.5 Schmidt's treatment of the singular integral

This section collects the arguments of Schmidt's treatment in [60] and [61] to show that the singular integral is indeed positive, provided the degree d of the system of equations is odd. The proof consists of three steps.

Lemma 4.5. *Suppose the forms $F^{(\rho)}$ are of odd degree. Then the variety*

$$M = \left\{ \overline{\mathbf{x}} \in \mathbb{R}^{ms} \colon \Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}}) = 0 \quad (\mathbf{j} \in J,\, 1 \leq \rho \leq R) \right\} \subset \mathbb{R}^{ms}$$

is of dimension at least $ms - Rr$.

Proof. This follows by the argument of [60, §2]. Observe that the forms $\Phi_{\mathbf{j}}^{(\rho)}$ define a map $\mathbb{R}^{ms} \to \mathbb{R}^{Rr}$. Every $(Rr+1)$-dimensional subspace of $\mathbb{R}^{ms}$ contains an Rr-sphere S^{Rr}, so by the Borsuk–Ulam Theorem applied to the restriction of the forms $\Phi_{\mathbf{j}}^{(\rho)}$ to S^{Rr} there exists a point $\overline{\mathbf{x}} \in S^{Rr}$ with the property that $\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}}) = \Phi_{\mathbf{j}}^{(\rho)}(-\overline{\mathbf{x}})$ for all $\mathbf{j} \in J$, $1 \leq \rho \leq R$. Since the forms $\Phi_{\mathbf{j}}^{(\rho)}$ are odd functions, this implies that $\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}}) = 0$ for all $\mathbf{j}$ and ρ. It follows that M intersects every $(Rr+1)$-dimensional subspace of $\mathbb{R}^{ms}$ non-trivially.

Now let δ be minimal with the property that M intersects every space A of dimension strictly greater than δ non-trivially. By the assumption of minimality there exists a δ-dimensional space A intersecting M only in the origin. Let B be the orthogonal complement of A in $\mathbb{R}^{ms}$, so $\dim(B) = ms - \delta$, and let $\overline{\mathbf{b}}$ be an arbitrary non-zero element of B. By the hypothesis, the space spanned by A and $\overline{\mathbf{b}}$ intersects M non-trivially, so we can find $\overline{\mathbf{a}} \in A$ and $\lambda \neq 0$ so that $\overline{\mathbf{u}} = \overline{\mathbf{a}} + \lambda \overline{\mathbf{b}}$ lies in the intersection. It follows that $\lambda \overline{\mathbf{b}}$ lies in the orthogonal projection of M onto B, and since $\overline{\mathbf{b}}$ was arbitrary, it follows that the orthogonal projection of M onto B covers the whole of B. Now recall that by the first part of the argument we had $\delta \leq Rr$, so we have

$$\dim(M) \geq \dim(B) = ms - \delta \geq ms - Rr$$

as claimed. □

Following Schmidt [61, §4], we define a weight function w_L dependent on a parameter L as

$$w_L(x) = \begin{cases} L(1 - L\,|x|) & \text{if } |x| \leq L^{-1}, \\ 0 & \text{else.} \end{cases}$$

This allows us to define a family of approximate singular integrals given by

$$\mathfrak{J}_L = \int_{|\overline{\boldsymbol{\xi}}| \leq 1} \prod_{\rho=1}^{R} \prod_{\mathbf{j} \in J} w_L(\Phi_{\mathbf{j}}^{(\rho)}(\overline{\boldsymbol{\xi}})) \mathrm{d}\overline{\boldsymbol{\xi}}.$$

Note that the family of weights $w_L(x)$ converges to a δ-distribution as L tends to infinity, so we expect the approximate singular integrals $\mathfrak{J}_L$ to converge to our standard singular integral $\mathfrak{J}$. This is the content of [61, §11], which we reproduce here in the next lemma.

Lemma 4.6. *Suppose*

$$s^* > 3 \cdot 2^{d-1}(d-1)R(Rr+1).$$

Then the approximate singular integrals $\mathfrak{J}_L$ converge to $\mathfrak{J}$ as L tends to infinity.

Proof. Since

$$w_L(x) = \int_{\mathbb{R}} e(\beta x) \left(\frac{\sin(\pi\beta/L)}{\pi\beta/L} \right)^2 \mathrm{d}\beta,$$

we may write

$$\int_{|\overline{\boldsymbol{\xi}}| \leq 1} \prod_{\rho=1}^{R} \prod_{\mathbf{j} \in J} w_L(\Phi_{\mathbf{j}}^{(\rho)}(\overline{\boldsymbol{\xi}})) \mathrm{d}\overline{\boldsymbol{\xi}} = \int_{\mathbb{R}^{Rr}} \int_{|\overline{\boldsymbol{\xi}}| \leq 1} e(\mathfrak{F}(\overline{\mathbf{x}}; \underline{\boldsymbol{\beta}})) \prod_{\rho, \mathbf{j}} \left(\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} \right)^2 \mathrm{d}\underline{\boldsymbol{\beta}} \mathrm{d}\overline{\boldsymbol{\xi}}.$$

It follows that

$$\begin{aligned} \mathfrak{J} - \mathfrak{J}_L &= \int_{\mathbb{R}^{Rr}} \int_{|\overline{\boldsymbol{\xi}}| \leq 1} e(\mathfrak{F}(\overline{\boldsymbol{\xi}}; \underline{\boldsymbol{\beta}})) \left\{ 1 - \prod_{\rho, \mathbf{j}} \left(\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} \right)^2 \right\} \mathrm{d}\underline{\boldsymbol{\beta}} \mathrm{d}\overline{\boldsymbol{\xi}} \\ &\ll \int_{\mathbb{R}^{Rr}} |\underline{\boldsymbol{\beta}}|^{-Rr-1} \left\{ 1 - \prod_{\rho, \mathbf{j}} \left(\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} \right)^2 \right\} \mathrm{d}\underline{\boldsymbol{\beta}} \end{aligned}$$

by Lemma 2.12 with $W = Rr + 1$.

By the power series expansion of the sine function one has

$$\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} = \frac{\pi\beta_{\mathbf{j}}^{(\rho)}/L + O\left((\pi\beta_{\mathbf{j}}^{(\rho)}/L)^3\right)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} = 1 + O\left((\beta_{\mathbf{j}}^{(\rho)}/L)^2\right)$$

in the range $|\beta_{\mathbf{j}}^{(\rho)}| < L$, so

$$\int_{|\underline{\boldsymbol{\beta}}|<L} |\underline{\boldsymbol{\beta}}|^{-Rr-1} \left\{ 1 - \prod_{\rho,\mathbf{j}} \left(\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} \right)^2 \right\} \mathrm{d}\underline{\boldsymbol{\beta}} \\ \ll L^{-2} \int_{|\underline{\boldsymbol{\beta}}|<L} |\underline{\boldsymbol{\beta}}|^{2-Rr-1} \mathrm{d}\underline{\boldsymbol{\beta}} \ll L^{-1}.$$

On the other hand, if $|\beta_{\mathbf{j}}^{(\rho)}| > L$, bounding the sine trivially yields

$$\int_{|\underline{\boldsymbol{\beta}}|>L} |\underline{\boldsymbol{\beta}}|^{-Rr-1} \left\{ 1 - \prod_{\rho,\mathbf{j}} \left(\frac{\sin(\pi\beta_{\mathbf{j}}^{(\rho)}/L)}{\pi\beta_{\mathbf{j}}^{(\rho)}/L} \right)^2 \right\} \mathrm{d}\underline{\boldsymbol{\beta}} \\ \ll \int_{|\underline{\boldsymbol{\beta}}|>L} |\underline{\boldsymbol{\beta}}|^{-Rr-1} \mathrm{d}\underline{\boldsymbol{\beta}} \ll L^{-1}$$

as well. Hence we have indeed $\mathfrak{J}_L \to \mathfrak{J}$ as L tends to infinity. □

It remains to show that each of the approximate singular integrals is positive. This is indeed the case, provided the variety M is sufficiently large.

Lemma 4.7. *If* $\dim M \geq ms - Rr$*, one has* $\mathfrak{J}_L \gg 1$ *uniformly in* L*.*

Proof. This is [61, Lemma 2] and the proof goes as follows. Under the assumption that $\dim(M) \geq ms - Rr$ one finds submanifolds $M_1 \subseteq M$ with $\dim(M_1) = ms - Rr$ which lies within a positive distance ϵ from the boundary of the unit cube, and $M_2 \subseteq M_1$ that can be parametrised by the first $ms - Rr$ coordinates. Write $\overline{\mathbf{x}} = (\boldsymbol{\zeta}, \boldsymbol{\xi}) \in \mathbb{R}^{ms-Rr} \times \mathbb{R}^{Rr}$. There is an open set $\mathscr{O} \subset \mathbb{R}^{ms-Rr}$ and a continuous map $f : \mathscr{O} \to \mathbb{R}^{Rr}$ such that $(\boldsymbol{\zeta}, f(\boldsymbol{\zeta})) \in M_2$ for all $\boldsymbol{\zeta} \in \mathscr{O}$.

The set

$$S_\epsilon = \{(\boldsymbol{\zeta}, \boldsymbol{\xi}) : \boldsymbol{\zeta} \in \mathscr{O}, |f(\boldsymbol{\zeta}) - \boldsymbol{\xi}| < \epsilon\}$$

lies still in the unit cube. Furthermore, since $M_2 \subseteq M$ one has $\Phi_{\mathbf{j}}^{(\rho)}(\boldsymbol{\zeta}, f(\boldsymbol{\zeta})) = 0$ for all $\mathbf{j}$ and ρ and therefore by Lipschitz continuity $\left|\Phi_{\mathbf{j}}^{(\rho)}(\boldsymbol{\zeta}, \boldsymbol{\xi})\right| < (2L)^{-1}$ for $(\boldsymbol{\zeta}, \boldsymbol{\xi}) \in S_{cL^{-1}}$ for some suitable constant c. This implies that

$$\prod_{\mathbf{j},\rho} w_L(\Phi_{\mathbf{j}}^{(\rho)}(\overline{\mathbf{x}})) \geq (L/2)^{Rr}$$

for all $\overline{\mathbf{x}} \in S_{cL^{-1}}$ and thus

$$\mathfrak{J}_L \gg \mathrm{vol}(S_{cL^{-1}})(L/2)^{Rr} \gg 1,$$

since $\mathrm{vol}(S_{cL^{-1}}) \gg L^{-Rr}$. □

These three lemmata suffice to prove the positivity of the singular integral. Indeed, by Lemma 4.6 one sees that the singular integral $\mathfrak{J}$ is positive as soon as the approximate singular integrals $\mathfrak{J}_L$ are bounded away from zero uniformly in L. Lemma 4.7 proves that this is the case provided the variety M is sufficiently large, which in turn follows by Lemma 4.5 from the fact that we assumed the degree to be odd.

Chapter 5

Unconditional bounds

5.1 Background and history

Up to this point we have been working under the assumption that the systems of equations that are being studied are not too singular. While this condition is certainly necessary in order to obtain quantitative estimates regarding the number of solutions, it is natural to ask whether the non-singularity requirement can be dispensed with if we ask only for the existence of rational points, particularly since the very existence of large singularities is often an indication that the problem in consideration can be reduced to problems of a lower complexity which are usually expected to have a larger number of solutions, so it should be easier to establish solubility.

Problems of this flavour have received considerable attention since the seminal work of Birch [6], in which he applied a diagonalisation method to prove the existence of arbitrarily many hyperplanes of any given dimension on the intersection of an arbitrary number of hypersurfaces, provided only that the number of variables be sufficiently large, and in order to steer clear of obstructions to the real solubility, he requires the degrees of the hypersurfaces to be odd. This method is, however, extremely wasteful in the number of variables (a quantified version has been provided by Wooley [76, Theorem 1]), so the focus shifted to trying to obtain bounds that are closer to the expected values at least in some simpler special cases.

Before embarking on a more detailed discussion of the available results in this area, a remark is in order about what one might expect to be the true lower bound on the number s of variables that ensures the existence of linear spaces. Note that the shape of the exponent $ms - Rrd$ in the main term in Theorem 2.1 will be positive only if $s \gg Rm^{d-1}$. This lower bound is confirmed by a comparison with related problems in $\mathbb{R}$ (see [42, Theorem 4]) and $\mathbb{C}$. In the latter case, apart from results concerning the multilinear Waring's problem such as [1], there is a bound for the existence of linear spaces due to Debarre and Manivel [19, Theorem 2.1(b)], which shows the existence of m-dimensional linear spaces on the intersection of R hypersurfaces of degree d if

$$s > \min \left\{ m + \frac{R}{m+1} \binom{d+m}{m}, 2m + r \right\},$$

and this is sharp.

Meanwhile, in the case $d = 3$ and $R = 1$ an upper bound of the desired quality has been provided by Dietmann and Wooley [24, Theorem 2(a)], which Dietmann [20, Theorem 6] complemented with an explicit lower bound. Together, these results strongly suggest that the true growth rate of s in m really is $\asymp m^{d-1}$. Less is known about the R-aspect, and while no compelling reason is known why the growth rates should not be linear, there might be anomalies that have as yet not been spotted. We notice, however, that the condition on s we have established in Theorem 4.2 falls short of the expected values, so one cannot expect that the derived bounds are anywhere close to the conjecture. This is due mainly to the relatively large contribution arising from the local solubility condition; however, even in Theorem 2.1 we miss the aim by a factor of mR.

We denote by $\gamma_d(R, m)$ the least number of variables that are necessary to ensure the existence of an m-dimensional linear space on the intersection of R hypersurfaces of degree d. In 1982, Schmidt studied systems of cubic forms [57–60] in order to establish a bound on $\gamma_3^*(R,1)$ and then pursue an iterating argument allowing him to exclude the singular cases in Theorem 4.2. We remark here that in our setting the dimension m of the linear space is understood in the affine sense, so one-dimensional linear spaces are really just projective points. This argument enabled him to obtain the long-standing bound

$$\gamma_3(R,1) \le (10R)^5. \tag{5.1.1}$$

The result from (5.1.1) can be extended in a fairly straightforward manner to larger values of m, yielding

$$\gamma_3(R,m) \ll R^5 m^{14} \tag{5.1.2}$$

(see [42, p. 283]). One notices, however, that for large values of m the bound in (5.1.2) is far from optimal, and one would expect to obtain better results by an independent approach that makes direct use of the linear spaces situation. This has first been successfully attempted by Lewis and Schulze-Pillot [42, p. 282], who obtained

$$\gamma_3(R,m) \ll R^{11} m + R^3 m^5.$$

A partial improvement is due to Wooley [74, Corollary of Theorem 1], who was able to show $\gamma_3(R,m) \ll R^{8+\epsilon} m^5$, and in the same work [74, Theorem 3.1] proved $\gamma_3(R,m) \ll_R m^\alpha$ with $\alpha = (5+\sqrt{17})/2 = 4.56155....$ More recently, Dietmann [20, Theorem 2] established a Hasse principle for the number of linear spaces on hypersurfaces which enabled him to show that

$$\gamma_3(R,m) \ll R^6 m^5 + R^4 m^6.$$

Our work in Chapter 2 allows us to refine Dietmann's methods and prove the following.

Theorem 5.1. *One has*

$$\gamma_3(R,m) \ll R^6 + R^3 m^3.$$

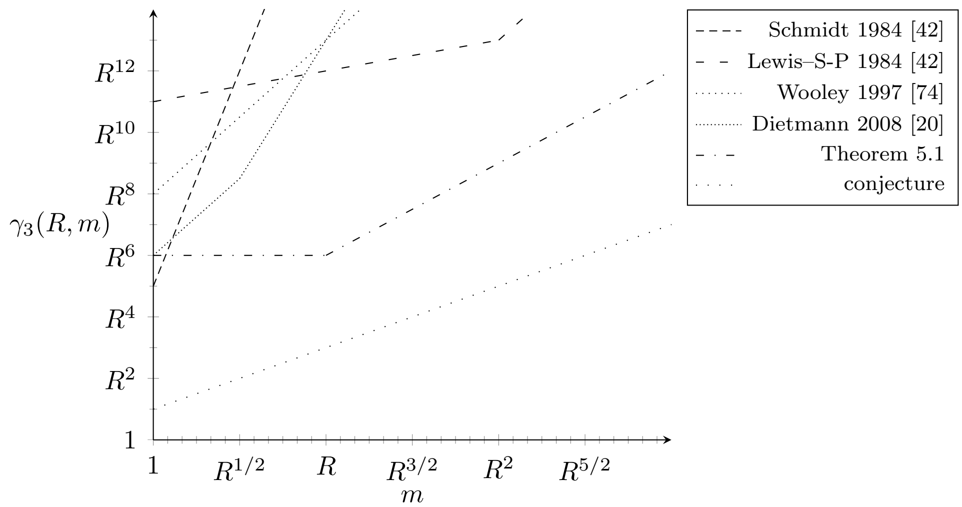

Figure 5.1: History of bounds for $\gamma_3(R, m)$

Note that the bound in Theorem 5.1 is of cubic growth in m, thus superseding all previous results in the m-aspect. Note that in the case $R = 1$ different methods [24, Theorem 2] yield

$$\gamma_3(1, m) \leq \frac{1}{2}(5m^2 + 33m + 32),$$

which beats the bounds we are able to obtain in the respective case even for small values of m. However, Theorem 5.1 is new even in the case $R = 2$, superseding a previous result by Wooley who bounded $\gamma_3(2, m)$ by a quartic polynomial (see [75, Theorem 2(b)] and the discussion of the corollary in [74, p. 3]).

Following the example of Dietmann [21, Theorem 2], we can use Theorem 5.1 and apply an iterating argument to derive an unconditional bound for the number of m-dimensional spaces on quintic hypersurfaces.

Theorem 5.2. *We have*

$$\gamma_5(R, m) \ll_R m^{12(3^{R-1}-1)+48\cdot 3^{R-1}}.$$

In particular, $\gamma_5(1, m) \ll m^{48}$.

This is the first time that polynomial growth in m has been established for the problem of finding m-spaces on systems of quintic forms, thus improving

Dietmann's treatment [21, Theorem 2] of the case $R = 1$, for which he requires at least $\gamma_5(1, m) \ll m^{439}$ variables, both in quality and in generality. In fact, by a more careful analysis it is possible track the dependence on R, and the same methods will yield a bound of the general shape

$$\gamma_5(R, m) \le (AmR)^{B^{CR}}$$

with explicit constants A, B, C. For large R this bound is not very satisfactory, as the advantage stemming from polynomial behaviour in m will soon be nullified by the number R of equations occurring in the second order exponent. However, in the light of the work by Wooley [76], especially the discussion in Section 6 of his paper, this drawback comes as no surprise. In fact, Wooley's bound [74, Theorem 2] of the shape

$$\gamma_5(R, m) \ll (3mR)^{A(mR)^c}$$

with numerical constants A and c will prevail as soon as $R \gg \log m$.

5.2 Preliminaries: The h-invariant

Theorem 2.1 implies that the number of m-dimensional rational linear spaces whose generators lie in a given box of sidelength P is given by

$$N_{s,R,m}^{(d)}(P) = P^{ms-Rrd} \chi_\infty \prod_{p \text{ prime}} \chi_p + o(P^{ms-Rrd}), \tag{5.2.1}$$

provided the relative number of variables s^* is large enough. Often it is useful to express results of a shape similar to that of Theorem 2.1 in a way that avoids explicit mention of the singularities. This section will therefore be used to restate and to some extent reprove [62, Theorem II] and [22, Theorem 2] for multidimensional situations.

A common and useful reformulation of the results in Chapters 2 and 4 is in terms of the h-invariant. This is an arithmetic invariant attached to a system of forms of equal degree and is defined as follows.

Definition 5.1. *(i) The h-invariant $h(F)$ of a form F denotes the least integer h that allows F to be written identically as a decomposition*

$$F(\mathbf{x}) = G_1(\mathbf{x})H_1(\mathbf{x}) + \ldots + G_h(\mathbf{x})H_h(\mathbf{x}) \tag{5.2.2}$$

of rational forms G_i, H_i of degree strictly smaller than $\deg(F)$.

(ii) For a system of forms $\mathbf{F}$ the h-invariant $h(\mathbf{F})$ is given by the minimum of the h-invariants over the forms in the rational pencil of the $F^{(\rho)}$, so one has

$$h(\mathbf{F}) = \min_{\mathbf{c}} h\left(c^{(1)}F^{(1)} + \ldots + c^{(R)}F^{(R)}\right),$$

where $\mathbf{c}$ runs over the non-zero elements of $\mathbb{Q}^R$.

One can show that this is indeed invariant with respect to changes of variables or linear combinations of the forms.

Write

$$H_d(R, m) = \tau_m \cdot 2^{d-1}(d-1)R(Rr+1),$$

for the bound on the variables in Birch's Theorem and Theorem 2.1, where $\tau_1 = 1$ and $\tau_m = 3$ for $m \geq 2$, and let

$$K_d(R, m) = \tau_m \cdot 2^{d-1}(d-1)R\max\{Rr+1, \gamma_d^*(R, m)\}$$

denote the bound established by Theorem 4.2. We now have the following restatement of Theorem 2.1, which is more suitable for some applications.

Theorem 5.3. *Let $F^{(1)}, \ldots, F^{(R)}, R$, d and m be as in Theorem 2.1. Then the number of m-dimensional linear spaces contained in $\mathbf{F} = \mathbf{0}$ satisfies* (5.2.1), *provided that*

$$h(\mathbf{F}) > (\log 2)^{-d} d!\big(H_d(R, m) + (d-1)R(R+1)\big),$$

and the product of positive densities $\chi_\infty \prod_p \chi_p$ is positive if $H_d(R, m)$ is replaced by $K_d(R, m)$.

This is a consequence of Theorem 2.1. Recall the tripartite case distinction from Lemma 2.4, which states that, for some suitable parameters k and θ, either we are in the minor arcs situation with $|T(\underline{\boldsymbol{\alpha}})| \ll P^{ms-k\theta}$, or $\underline{\boldsymbol{\alpha}}$ is contained in the major arcs, i.e. the coefficients $\alpha_{\mathbf{j}}^{(\rho)}$ have good rational approximations, or the system of forms has been singular from the beginning, that is to say, the number of $(d-1)$-tuples $(\mathbf{h}_1, \dots, \mathbf{h}_{d-1}) \leq P^\theta$ that satisfy

$$\operatorname{rank}\left(B_i^{(\rho)}(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})\right)_{i,\rho} \leq R-1 \tag{5.2.3}$$

is asymptotically greater than $\left(P^\theta\right)^{(d-1)s-2^{d-1}k-\epsilon}$. If we write $\mathfrak{V}$ for the variety defined by (5.2.3) and $z_P(\mathfrak{V})$ for the number of integer points of height at most P contained in $\mathfrak{V}$, then the singular case can be characterised by the relation

$$z_{P^\theta}(\mathfrak{V}) \gg (P^\theta)^{(d-1)s-2^{d-1}k} \tag{5.2.4}$$

for suitable values of k and θ. Furthermore, recall that Chapters 2 and 4 establish that, if the singular case is excluded, one obtains an asymptotic formula if $2^{d-1}k > H_d(R,m)$, and the main term will be positive if $2^{d-1}k > K_d(R,m)$. The goal is therefore to understand under what conditions the singular case can be excluded.

Write $g(\mathbf{F})$ for the largest number g such that the number of integer points of height at most P contained in $\mathfrak{V}$ is bounded by $z_P(\mathfrak{V}) \ll P^{(d-1)s-g+\epsilon}$. The Corollary of [62, Proposition III] says that

$$h(\mathbf{F}) \leq \phi(d)([g(\mathbf{F})] + (d-1)R(R+1)), \tag{5.2.5}$$

and $\phi(d)$ is given in the Corollary of Proposition $\mathrm{III}_{\mathbb{C}}$ as

$$\phi(d) < (\log 2)^{-d} d!.$$

One should remark that the arguments employed by Schmidt in order to derive these results remain unchanged by the linear spaces, since the condition of being singular is a property of the system of forms in question and is independent of the dimension of the solution space.

Now suppose that we are in the singular situation as characterised above, so (5.2.4) is true. Then we have the chain of inequalities

$$(P^\theta)^{(d-1)s-2^{d-1}k} \ll z_{P^\theta}(\mathfrak{V}) \ll (P^\theta)^{(d-1)s-g+\epsilon},$$

whence $g(\mathbf{F}) \leq 2^{d-1}k$ and therefore, by (5.2.5),

$$h(\mathbf{F}) \leq \phi(d)(2^{d-1}k + (d-1)R(R+1)).$$

It follows that if

$$h(\mathbf{F}) > \phi(d)(2^{d-1}k + (d-1)R(R+1)),$$

the singular case is excluded, so by fixing a small positive parameter δ and letting $k = 2^{1-d}H_d(R,m) + \delta$ or $k = 2^{1-d}K_d(R,m) + \delta$, we can indeed follow through the proofs of Chapters 2 or 4, respectively, and derive the asymptotic formula under the condition

$$h(\mathbf{F}) > \phi(d)(H_d(R,m) + 2^{d-1}\delta + (d-1)R(R+1))$$

or the respective result with $H_d(R,m)$ replaced by $K_d(R,m)$. Since $\delta > 0$ has been arbitrary, this establishes the theorem.

In the cubic case one can do better than in Theorem 5.3. In fact, Schmidt's work [60, 61] on the subject has recently received an improvement by Dietmann [22], which translates into our case and which we will apply in our proofs of Theorems 5.1 and 5.2.

Theorem 5.4. *Let $F^{(1)}, \ldots, F^{(R)}$ be cubic forms in s variables such that no form in the rational pencil vanishes on a linear space of codimension less than $H_3(R,m)$. Then we have*

$$N^{(3)}_{s,R,m}(P) = P^{ms-3Rr}\chi_\infty \prod_{p \text{ prime}} \chi_p + o\left(P^{ms-3Rr}\right).$$

The product of the local densities $\chi_\infty \prod_p \chi_p$ is positive if $H_3(R,m)$ is replaced by $K_3(R,m)$.

This is essentially [22, Theorem 2], but the setting is different enough to warrant some further justification. Just as in the case of Theorem 5.3, the proof rests on understanding the singular case of Weyl's inequality for points on the intersection on R cubic forms.

For a given R-tuple $(w_1, \dots, w_R)$ let $V(w_1, \dots, w_R; P)$ denote the number of solutions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in [-P, P]^s$ of

$$\sum_{\rho=1}^{R} w_\rho B_i^{(\rho)}\big(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}\big) = 0 \quad (1 \leq i \leq s),$$

where the $B_i^{(\rho)}$ are as in Lemma 2.2. Furthermore, recall that Lemma 2.3 states that if $|T(\underline{\boldsymbol{\alpha}})|$ is large, one has

$$\text{Card}\left\{\mathbf{h}_{j_1}, \mathbf{h}_{j_2} \leq P^\theta \colon \left| M(\mathbf{j}) \sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\mathbf{h}_{j_1}, \mathbf{h}_{j_2}) \right| < P^{-3+2\theta} \right\} \gg (P^\theta)^{2s-4k-\epsilon}. \tag{5.2.6}$$

We have the following alternative version of the singular case of Lemma 2.4.

Lemma 5.1. *Suppose $d = 3$ we are in the situation of Lemma 2.4* (C). *Then there exist integers $w_1, \dots, w_R$, not all of which are zero, such that*

$$V(w_1, \dots, w_R; P^\theta) \gg (P^\theta)^{2s-4k-\epsilon}.$$

Proof. This is [22, Lemma 2] in the case $d = 3$ and the proof is similar to that of Lemma 2.4. Suppose that neither of the first two alternatives in 2.4 is true. Now consider the matrix

$$B_{\mathbf{j}} = \left(B_i^{(\rho)}(\mathbf{h}_{j_1}, \mathbf{h}_{j_2})\right)_{i, \mathbf{h}_{j_1}, \mathbf{h}_{j_2}; \rho}$$

whose rows range from $1 \leq \rho \leq R$ and correspond to the initial system of forms, whereas the columns range over all $1 \leq i \leq s$ and all $\mathbf{h}_{j_1}, \mathbf{h}_{j_2} \leq P^\theta$ counted by (5.2.6), so this matrix is essentially the matrix that is obtained from collecting all the matrices $\left(B_i^{(\rho)}(\mathbf{h}_{j_1}, \mathbf{h}_{j_2})\right)_{i;\rho}$ considered in the proof of Lemma 2.4 for which $\sum_{\rho=1}^{R} \alpha_{\mathbf{j}}^{(\rho)} B_i^{(\rho)}(\mathbf{h}_{j_1}, \mathbf{h}_{j_2})$ is small. If this matrix is of rank R, there is a non-singular $(R \times R)$ submatrix $B_{\mathbf{j}}^0$ which can be used to derive an approximation of $\alpha_{\mathbf{j}}$ literally as in the proof of Lemma 2.4. Since we had assumed that there is no such approximation, this is a contradiction.

It therefore follows that the rank of the matrix is strictly less than R, so one finds a vanishing linear combination of the rows of $B_{\mathbf{j}}$. Let the coefficients of this linear combination be given by $w_1, \ldots, w_R$, then this implies that

$$\sum_{\rho=1}^{R} w_\rho B_i^{(\rho)}(\mathbf{h}_{j_1}, \mathbf{h}_{j_2}) = 0$$

for every $1 \leq i \leq s$ and for every choice of $\mathbf{h}_{j_1}, \mathbf{h}_{j_2}$ counted by (5.2.6). It follows that with this choice of $w_1, \ldots, w_R$ every pair $\mathbf{h}_{j_1}, \mathbf{h}_{j_2}$ that is counted by (5.2.6) is also counted by $V(w_1, \ldots, w_R; P^\theta)$. However, by (5.2.6) the number of these is $\gg (P^\theta)^{2s-4k-\epsilon}$. The lemma follows. □

Lemma 5.2. *Suppose that $d = 3$ and each form of the rational pencil of the $F^{(1)}, \ldots, F^{(R)}$ has h-invariant greater than $4k$. Then the singular case of Lemma 2.4 is excluded.*

Proof. This is [22, Lemma 6]. Lemma 5.1 states that in the singular case (C) of Lemma 2.4 one can find integers $w_1, \ldots, w_R$ such that

$$V(w_1, \ldots, w_R; P^\theta) \gg (P^\theta)^{2s-4k-\epsilon}.$$

For a fixed set of such w_ρ consider the cubic form $C = \sum_\rho w_\rho F^{(\rho)}$ and notice that by Definition 5.1(ii) this implies that $h(C) \geq h(\mathbf{F})$. Recalling that we had

$$F^{(\rho)}(\mathbf{x}) = \sum_{1 \leq i,j,k \leq s} c_{i,j,k}^{(\rho)} x_i x_j x_k,$$

one sees that the trilinear form Φ associated to C is given by

$$\begin{aligned} \Phi(\mathbf{x}, \mathbf{y}, \mathbf{z}) &= \sum_{\rho=1}^{R} w_\rho \sum_{1 \leq i,j,k \leq s} c_{i,j,k}^{(\rho)} x_i y_j z_k \\ &= \sum_{1 \leq i,j,k \leq s} \tilde{c}_{i,j,k} x_i y_j z_k, \end{aligned}$$

where we wrote

$$\tilde{c}_{i,j,k} = \sum_{\rho=1}^{R} w_\rho c_{i,j,k}^{(\rho)}.$$

As before, we have bilinear forms B_i associated to C that are defined by the relation

$$\Phi(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i=1}^{s} x_i B_i(\mathbf{y},\mathbf{z}).$$

In the light of the previous arguments one one sees now that

$$B_i(\mathbf{y},\mathbf{z}) = \sum_{1\le j,k\le s} \tilde{c}_{i,j,k} y_j z_k = \sum_{1\le j,k\le s}\sum_{\rho=1}^{R} w_\rho c_{i,j,k}^{(\rho)} y_j z_k = \sum_{\rho=1}^{R} w_\rho B_i^{(\rho)}(\mathbf{y},\mathbf{z}).$$

One can now appeal to [22, Lemma 5], which implies that

$$\text{Card}\,\{\mathbf{y},\mathbf{z}\in[-P,P]^s\colon B_i(\mathbf{y},\mathbf{z})=0 \quad (1\le i\le s)\} \ll P^{2s-h(C)},$$

whence we have the sequence of inequalities

$$\begin{aligned}(P^\theta)^{2s-4k-\epsilon} &\ll V(w_1,\dots,w_R;P^\theta)\\ &\ll \text{Card}\left\{\mathbf{y},\mathbf{z}\in[-P^\theta,P^\theta]^s : B_i(\mathbf{y},\mathbf{z})=0 \quad \forall i\right\}\\ &\ll (P^\theta)^{2s-h(C)}.\end{aligned}$$

Since $h(C)\ge h(\mathbf{F})$, this contradicts our hypothesis that $h(\mathbf{F})>4k$. □

Theorem 5.4 now follows from the fact that in the cubic case Definition 5.1(ii) implies that the condition $h(\mathbf{F})>4k$ is equivalent to saying that at least one form in the rational pencil of the $F^{(\rho)}$ vanishes on a linear space of codimension less than $4k$. We also remark that, just as in the case of Theorem 5.3, the further analysis of the proof of Theorem 2.1 remains valid if we set $4k = H_3(R,m)$ or $4k = K_3(R,m)$, respectively.

5.3 Linear spaces on the intersection of hypersurfaces

The proof of Theorem 5.1 rests on the following key lemma.

Lemma 5.3. *One has*

$$\gamma_3(R,m) \le \sum_{\rho=1}^{R} K_3(\rho,m) \ll RK_3(R,m).$$

Proof. Firstly, suppose that no form in the rational pencil of the $F^{(1)}, \dots, F^{(R)}$ vanishes on a rational subspace with codimension $\leq K_3(R,m)$. Then Theorem 5.4 gives $\gamma_3(R,m) \leq K_3(R,m)$. We may therefore assume that at least one of the forms in the linear pencil does vanish on a linear space Y with $\dim Y \geq s - K_3(R,m)$; by relabelling and considering linear combinations we may assume without loss of generality that this is the form $F^{(R)}$. If the dimension of this space is large enough and one has $\dim Y \geq \gamma_3(R-1,m)$, we may solve the remaining $R-1$ equations on Y. Hence if $s - K_3(R,m) \geq \gamma_3(R-1,m)$, we will either have an m-dimensional space directly via Theorem 5.4, or else we may reduce the problem to finding an m-dimensional linear space on the intersection of Y with the hypersurfaces associated to $F^{(1)}, \dots, F^{(R-1)}$. Altogether we obtain the recursion formula

$$\gamma_3(\rho,m) \leq \max\{K_3(\rho,m), \gamma_3(\rho-1,m) + K_3(\rho,m)\} = \gamma_3(\rho-1,m) + K_3(\rho,m),$$

and after at most R iterations one recovers the statement. □

We can now prove Theorem 5.1. By Theorem 4.2 and (4.1.5) we have

$$K_3(R,m) = 24R \max\left\{10(Rm + 6R^2)^2, Rr + 1\right\} \asymp R^5 + R^2 m^3,$$

and an application of Lemma 5.3 yields

$$\begin{aligned}\sum_{\rho=1}^{R} 240\rho(\rho m + 6\rho^2)^2 &\leq 60m^2(R+1)^4 m^2 + 576m(R+1)^5 + 1440(R+1)^6 \\ &= 12(R+1)^4\left(5m^2 + 48m(R+1) + 120(R+1)^2\right) \\ &\leq 60(R+1)^4\left(m + 5(R+1)\right)^2 \\ &\leq 1860\left((R+1)^6 + (R+1)^4 m^2\right).\end{aligned}$$

On the other hand, one has

$$\sum_{\rho=1}^{R} 24\rho(\rho r + 1) \leq 8(R+1)^3 r + 12R(R+1),$$

where

$$r = m(m+1)(m+2)/6 \leq (m+1)^3/6.$$

Combining these estimates one sees that

$$\gamma_3(R,m) \leq 1860 \max\big((R+1)^6 + (R+1)^3(m+1)^3\big),$$

which proves Theorem 5.1.

It remains to complete the proof of Theorem 5.2. Again, we proceed by induction. For the case $R = 1$ we imitate Dietmann [21], using Theorem 4.2 instead of the weaker bounds applied by him. Applying Theorem 5.3 with $d = 5$, one sees that any single quintic hypersurface F contains a linear space of dimension m as soon as

$$h(F) \gg (d^2 + m)^{2^{d-2}} \asymp m^8,$$

so we may suppose that $h(F) \leq C_1 m^8$ for some constant C_1. By Definition 5.1(i), this means that one can find forms G_i, H_i $(i = 1, \dots, h)$ of degree less than five such that the form F can be written in the shape given in (5.2.2). We can, without loss of generality, assume the H_i to be of odd degree. Suppose k of these are linear, then we want to find an $(m+k)$-dimensional linear space on the remaining system of $h-k$ cubic forms, and one sees from our bound on $\gamma_3(R,m)$ that the worst case scenario is the case when $k = 0$. It follows that we may assume without loss of generality that all forms H_i are cubic. Theorem 5.1 now implies that the intersection of the H_i contains a linear m-space if $s \gg h^3 m^3 + h^6$. Since we had assumed $h(F) \leq C_1 m^8$, this gives the result.

Now consider $\gamma_5(R,m)$ for $R > 1$. As before, in the case $h(\mathbf{F}) \gg_R m^8$ the claim follows from Theorem 4.2. Let us therefore suppose that $F^{(R)}$ possesses a decomposition as in (5.2.2) with $h(\mathbf{F}) \leq C_2(R) m^8$ for some $C_2(R)$. As above, it suffices to consider the worst case scenario that all forms $H_i^{(R)}$ are cubic, and by Theorem 5.1 we can find that the intersection of the hypersurfaces $H_i^{(R)} = 0$ contains a linear space $L^{(R)}$ of dimension $\lambda^{(R)}$ as long as the number of variables exceeds

$$\gamma_3\big(C_2(R)m^8, \lambda^{(R)}\big) \ll_R m^{48} + m^{24}(\lambda^{(R)})^3.$$

Thus we can reduce the problem to solving the remaining $R-1$ equations on $L^{(R)}$. In order for the residual system to be accessible to our methods, we need $\lambda^{(R)} \geq \gamma_5(R-1,m)$, whence by the induction hypothesis $\gamma_5(R,m)$ is bounded by

$$\begin{aligned}\gamma_5(R,m) &\ll_R m^{48} + m^{24}\big(\gamma_5(R-1,m)\big)^3\\ &\ll_R m^{48} + m^{24}\Big(m^{12(3^{R-2}-1)+48\cdot 3^{R-2}}\Big)^3.\end{aligned}$$

We may therefore conclude that there exists some function $A(R)$ such that

$$\gamma_5(R,m) \leq A(R) m^{12(3^{R-1}-1)+48\cdot 3^{R-1}},$$

as claimed. In fact, one may verify that the statement holds with

$$A(R) \leq (18R)^{43\cdot 3^R}.$$

Bibliography

[1] J. Alexander and A. Hirschowitz. Polynomial interpolation in several variables. *J. Algebraic Geom.*, 4:201–222, 1995.

[2] G. I. Arkhipov and A. A. Karatsuba. A multidimensional analogue of Waring's problem. *Dokl. Akad. Nauk SSSR*, 295:521–523, 1987.

[3] R. C. Baker. Diagonal cubic equations. II. *Acta Arith.*, 53:217–250, 1989.

[4] R. Balasubramanian, J.-M. Deshouillers, and F. Dress. Problème de Waring pour les bicarrés. I. Schéma de la solution. *C. R. Acad. Sci. Paris Sér. I Math.*, 303:85–88, 1986.

[5] R. Balasubramanian, J.-M. Deshouillers, and F. Dress. Problème de Waring pour les bicarrés. II. Résultats auxiliaires pour le théorème asymptotique. *C. R. Acad. Sci. Paris Sér. I Math.*, 303:161–163, 1986.

[6] B. J. Birch. Homogeneous forms of odd degree in a large number of variables. *Mathematika*, 4:102–105, 1957.

[7] B. J. Birch. Waring's problem in algebraic number fields. *Proc. Cambridge Philos. Soc.*, 57:449–459, 1961.

[8] B. J. Birch. Forms in many variables. *Proc. Roy. Soc. Ser. A*, 265:245–263, 1961/1962.

[9] E. Bombieri and H. Iwaniec. On the order of $\zeta(\frac{1}{2}+it)$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 13:449–472, 1986.

[10] J. Brandes. Forms representing forms and linear spaces on hypersurfaces. *Proc. Lond. Math. Soc.*, 108:809–835, 2014.

[11] R. Brauer. A note on systems of homogeneous algebraic equations. *Bull. Amer. Math. Soc.*, 51:749–755, 1945.

[12] H. Davenport. On Waring's problem for fourth powers. *Ann. of Math. (2)*, 40:731–747, 1939.

[13] H. Davenport. Cubic forms in thirty-two variables. *Philos. Trans. Roy. Soc. London. Ser. A*, 251:193–232, 1959.

[14] H. Davenport. Cubic forms in 29 variables. *Proc. Roy. Soc. Ser. A*, 266:287–298, 1962.

[15] H. Davenport. Cubic forms in sixteen variables. *Proc. Roy. Soc. Ser. A*, 272:285–303, 1963.

[16] H. Davenport. *The collected works of Harold Davenport. Vol. III.* Academic Press [Harcourt Brace Jovanovich Publishers], London, 1977. Edited by B. J. Birch, H. Halberstam and C. A. Rogers.

[17] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005.

[18] H. Davenport and D. J. Lewis. Homogeneous additive equations. *Proc. Roy. Soc. Ser. A*, 274:443–460, 1963.

[19] Olivier Debarre and Laurent Manivel. Sur la variété des espaces linéaires contenus dans une intersection complète. *Math. Ann.*, 312:549–574, 1998.

[20] R. Dietmann. Systems of cubic forms. *J. Lond. Math. Soc. (2)*, 77:666–686, 2008.

[21] R. Dietmann. Linear spaces on rational hypersurfaces of odd degree. *Bull. Lond. Math. Soc.*, 42:891–895, 2010.

[22] R. Dietmann. Weyl's inequality and systems of forms. 2012, arXiv:1208.1968. submitted.

[23] R. Dietmann and M. Harvey. On the representation of quadratic forms by quadratic forms. *Michigan Math. J.*, 62:673–896, 2013.

[24] R. Dietmann and T. D. Wooley. Pairs of cubic forms in many variables. *Acta Arith.*, 110:125–140, 2003.

[25] J. S. Ellenberg and A. Venkatesh. Local-global principles for representations of quadratic forms. *Invent. Math.*, 171:257–279, 2008.

[26] C. F. Gauß. *Disquisitiones Arithmeticae.* Lipsiae, 1801.

[27] M. J. Greenberg. *Lectures on forms in many variables.* W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[28] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio numerorum'; I: A new solution of Waring's Problem. *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl.*, 1920:33–54, 1920.

[29] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio Numerorum': IV. The singular series in Waring's Problem and the value of the number $G(k)$. *Math. Z.*, 12:161–188, 1922.

[30] G. H. Hardy and S. Ramanujan. Asymptotic formulae in combinatory analysis. *Proc. London Math. Soc.*, 17:75–115, 1917.

[31] D. R. Heath-Brown. Cubic forms in ten variables. *Proc. London Math. Soc. (3)*, 47:225–257, 1983.

[32] D. R. Heath-Brown. Cubic forms in 14 variables. *Invent. Math.*, 170:199–230, 2007.

[33] D. Hilbert. Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem). *Math. Ann.*, 67:281–300, 1909.

[34] C. Hooley. On nonary cubic forms. *J. Reine Angew. Math.*, 386:32–98, 1988.

[35] C. Hooley. On octonary cubic forms. *Proc. Lond. Math. Soc.*, 2014. in press.

[36] J. S. Hsia, Y. Kitaoka, and M. Kneser. Representations of positive definite quadratic forms. *J. Reine Angew. Math.*, 301:132–141, 1978.

[37] L.-K. Hua. On Waring's problem. *Quart. J. Math. Oxford*, 9:199–202, 1938.

[38] A. Kempner. Bemerkungen zum Waringschen Problem. *Math. Ann.*, 72:387–399, 1912.

[39] E. Landau. Über eine Anwendung der Primzahltheorie auf das Waringsche Problem in der elementaren Zahlentheorie. *Math. Ann.*, 66:102–105, 1908.

[40] D. J. Lewis. Cubic homogeneous polynomials over p-adic number fields. *Ann. of Math. (2)*, 56:473–478, 1952.

[41] D. J. Lewis. Cubic forms over algebraic number fields. *Mathematika*, 4:97–101, 1957.

[42] D. J. Lewis and R. Schulze-Pillot. Linear spaces on the intersection of cubic hypersurfaces. *Monatsh. Math.*, 97:277–285, 1984.

[43] Yu. V. Linnik. On the representation of large numbers as sums of seven cubes. *Rec. Math. [Mat. Sbornik] N. S.*, 12(54):218–224, 1943.

[44] D. T. Loughran. *Manin's conjecture for Del Pezzo surfaces*. PhD thesis, University of Bristol, 2011.

[45] Ju. V. Matijasevič. The Diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR*, 191:279–282, 1970.

[46] R. Miranda. Linear systems of plane curves. *Notices Amer. Math. Soc.*, 46:192–201, 1999.

[47] S. T. Parsell. The density of rational lines on cubic hypersurfaces. *Trans. Amer. Math. Soc.*, 352:5045–5062 (electronic), 2000.

[48] S. T. Parsell. Multiple exponential sums over smooth numbers. *J. Reine Angew. Math.*, 532:47–104, 2001.

[49] S. T. Parsell. A generalization of Vinogradov's mean value theorem. *Proc. London Math. Soc. (3)*, 91:1–32, 2005.

[50] S. T. Parsell. Asymptotic estimates for rational linear spaces on hypersurfaces. *Trans. Amer. Math. Soc.*, 361:2929–2957, 2009.

[51] S. T. Parsell. Hua-type iteration for multidimensional Weyl sums. *Mathematika*, 58:209–224, 2012.

[52] S. T. Parsell, S. M. Prendiville, and T. D. Wooley. Near-optimal mean value estimates for multidimensional Weyl sums. *Geom. Funct. Anal.*, 23:1962–2024, 2013.

[53] W. R. Paton, editor. *Anthologia Graeca*. William Heinemann Ltd, London, 1927.

[54] K. Ranestad and F.-O. Schreyer. Varieties of sums of powers. *J. Reine Angew. Math.*, 525:147–181, 2000.

[55] J. Robinson. Unsolvable diophantine problems. *Proc. Amer. Math. Soc.*, 22:534–538, 1969.

[56] D. Schindler. Bihomogeneous forms in many variables. *J. Théor. Nombres Bordeaux*, 2014, arXiv:1301:6516. in press.

[57] W. M. Schmidt. On cubic polynomials. I. Hua's estimate of exponential sums. *Monatsh. Math.*, 93:63–74, 1982.

[58] W. M. Schmidt. On cubic polynomials. II. Multiple exponential sums. *Monatsh. Math.*, 93:141–168, 1982.

[59] W. M. Schmidt. On cubic polynomials. III. Systems of p-adic equations. *Monatsh. Math.*, 93:211–223, 1982.

[60] W. M. Schmidt. On cubic polynomials. IV. Systems of rational equations. *Monatsh. Math.*, 93:329–348, 1982.

[61] W. M. Schmidt. Simultaneous rational zeros of quadratic forms. In *Seminar on Number Theory, Paris 1980-81 (Paris, 1980/1981)*, volume 22 of *Progr. Math.*, pages 281–307. Birkhäuser Boston, Mass., 1982.

[62] W. M. Schmidt. The density of integer points on homogeneous varieties. *Acta Math.*, 154:243–296, 1985.

[63] R. Schulze-Pillot. Local conditions for global representations of quadratic forms. *Acta Arith.*, 138:289–299, 2009.

[64] E. S. Selmer. The Diophantine equation $ax^3 + by^3 + cz^3 = 0$. *Acta Math.*, 85:203–362, 1951.

[65] C. L. Siegel. Über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36:527–606, 1935.

[66] C. L. Siegel. Über die analytische Theorie der quadratischen Formen. II. *Ann. of Math. (2)*, 37:230–263, 1936.

[67] C. L. Siegel. Über die analytische Theorie der quadratischen Formen. III. *Ann. of Math. (2)*, 38:212–291, 1937.

[68] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997.

[69] R. C. Vaughan and T. D. Wooley. Waring's problem: a survey. In *Number theory for the millennium, III (Urbana, IL, 2000)*, pages 301–340. A K Peters, Natick, MA, 2002.

[70] I. M. Vinogradov. Sur le problème de Waring. *C. R. Acad. Sci. USSR*, pages 393–400, 1928.

[71] E. Waring. *Meditationes Algebraicae*. J. Archdeacon, Cambridge, third edition, 1782.

[72] A. Wieferich. Beweis des Satzes, daß sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen läßt. *Math. Ann.*, 66:95–101, 1908.

[73] T. D. Wooley. New estimates for smooth Weyl sums. *J. London Math. Soc. (2)*, 51:1–13, 1995.

[74] T. D. Wooley. Forms in many variables. In *Analytic number theory (Kyoto, 1996)*, volume 247 of *London Math. Soc. Lecture Note Ser.*, pages 361–376. Cambridge Univ. Press, Cambridge, 1997.

[75] T. D. Wooley. Linear spaces on cubic hypersurfaces, and pairs of homogeneous cubic equations. *Bull. London Math. Soc.*, 29:556–562, 1997.

[76] T. D. Wooley. An explicit version of Birch's theorem. *Acta Arith.*, 85:79–96, 1998.

[77] T. D. Wooley. On the local solubility of Diophantine systems. *Compositio Math.*, 111:149–165, 1998.

[78] T. D. Wooley. Vinogradov's mean value theorem via efficient congruencing. *Ann. of Math. (2)*, 175:1575–1627, 2012.

[79] T. D. Wooley. Vinogradov's mean value theorem via efficient congruencing, II. *Duke Math. J.*, 162:673 – 730, 2013.