Rational points and lines on cubic hypersurfaces

Dissertation

for the award of the degree

"Doctor rerum naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program "Mathematical Sciences"

of the Georg-August University School of Science (GAUSS)

submitted by

Christian Bernert

from Bückeburg

Göttingen, 2023

Thesis advisory committee

Prof. Dr. Jörg Brüdern,
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Harald Helfgott,
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Preda Mihailescu,
Mathematisches Institut, Georg-August-Universität Göttingen

Members of the examination board

Reviewer:
Prof. Dr. Jörg Brüdern,
Mathematisches Institut, Georg-August-Universität Göttingen

Second Reviewer:
Prof. Dr. Damaris Schindler,
Mathematisches Institut, Georg-August-Universität Göttingen

Further members of the examination board

Prof. Dr. Preda Mihailescu,
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Harald Helfgott,
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Thomas Schick,
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Axel Munk,
Institut für Mathematische Stochastik, Georg-August-Universität Göttingen

Date of oral examination: 26 June 2023

Declaration

Chapter 3, which deals with cubic forms over imaginary quadratic number fields was created in joint work with fellow Göttingen PhD student Leonhard Hochfilzer (LH). In the spring of 2022 this topic came up in a conversation between CB and LH motivated by a previous talk by Rainer Dietmann. After a conversation with Tim Browning on this topic, LH and CB started working on the problem. The main ideas and sketches for the proof of the main result were worked out on blackboards in collaborative sessions in the mathematical institute in Göttingen. Once the authors agreed that they now together worked out a line of attack on the problem, they decided to divide the effort of writing up the content and working out the remaining details involved. In particular LH conceived first drafts of Section 3.4, Section 3.5, Section 3.6 and Section 3.7, which is comprised of some technical preliminary lemmas, the treatment of the major arcs and a Weyl differencing estimate, which is needed for a part of the minor arcs. CB conceived first drafts of Section 3.2, Section 3.8 and Section 3.9, which includes the deduction of the applications to our result, the averaged van der Corput estimate and the estimation of the minor arcs. The introductory content of Section 3.1 and Section 3.3 were written up en passant and so both CB and LH contributed equally to the conception of these sections. All sections were proofread by both LH and CB and changes were made afterwards accordingly.

Acknowledgements

First of all, I would like to thank my supervisor Jörg Brüdern for guiding me not only through the almost four years of my Ph.D., but indeed through almost all the eight wonderful years that I spent in Göttingen. Through his series of lectures on analytic number theory, I first got in contact with this beautiful area of mathematics, and indeed many of the projects that I am now working on have their roots in topics discussed during these lectures.

Besides this guidance leading to a Bachelor Thesis, a Master Thesis and now finally a Ph.D. Thesis, I am equally grateful to him for always allowing me to decide independently on which problems I wanted to work on, and for giving me a lot of freedom in how to approach these problems.

I am very grateful to Damaris Schindler who, although not formally part of my Thesis Advisory Committee, has been an extremely important mentor for me, helping out with both mathematical and practical problems surrounding the life of a Ph.D. student.

I would also like to thank Thomas Schick for teaching me so much interesting mathematics during all my time in Göttingen, and even after I went into a slightly different area of mathematics, his advice, experience and passion for mathematics have continued to be very helpful and inspiring.

I would like to thank Leonhard Hochfilzer for a very enjoyable and fruitful time of collaboration, which I certainly hope will not have been our last one.

I would like to thank all my colleagues from the RTG and in particular from the Göttingen number theory group for making life at our university just so much more enjoyable, especially in the last year after the difficult Covid times. Special thanks to Rok for proofreading large parts of the manuscript.

Finally, I would like to thank my family and my friends for their support, without which I would have become a very different person than I now am.

Contents

Chapter 1

Introduction

This thesis is concerned with the arithmetic properties of cubic forms and more generally cubic polynomials. By this, we mean polynomials of degree 3 with integer coefficients in possibly many variables $x_1, \ldots, x_n$. If the polynomial is homogeneous, i.e. does not have any lower order terms, we call it a *cubic form*.

Before we restrict to the cubic case, let us first discuss the case of a general polynomial F with integer coefficients in variables $x_1, \ldots, x_n$.

The zero set $F(x_1, \ldots, x_n) = 0$ defines a certain subset of n-dimensional space, more precisely an *algebraic variety*. Traditionally, such an equation with integer coefficients is known as a *Diophantine equation*.

It has been an immensely fruitful approach in the 20th century to study the arithmetic properties of polynomials with the help of these associated geometric objects. This field of research is now known as *Arithmetic (Algebraic) Geometry*.

For us, at this point the geometric approach merely serves well to illustrate the questions that we will study in this thesis. The fundamental question from the perspective of a number theorist is the following:

Question 1: *Given a polynomial F, can we describe the solutions of the Diophantine equation $F(\mathbf{x}) = 0$ in the set of integers or the rational numbers? Geometrically, can we describe the integral/rational points on the variety $F(\mathbf{x}) = 0$, i.e. the points with integral/rational coordinates?*

This is clearly not a precise mathematical question and indeed there are many different ways, depending on the context, to turn this into a meaningful mathematical problem.

One example with a long history and a very satisfactory answer is given by the quadratic equation $F(x, y) = x^2 + y^2 - 1$ in two variables. Here, the set of (real) zeroes of F defines the unit circle in the plane. Rational solutions of $x^2 + y^2 = 1$ correspond to *Pythagorean triples*, i.e. integer solutions to the equation $x^2 + y^2 = z^2$, after clearing denominators. For instance, the classical solution $3^2 + 4^2 = 5^2$ corresponds to the rational point $\left(\frac{3}{5}, \frac{4}{5}\right)$ on the unit circle.

Already the Babylonians studied these triples and it is a classical result, essentially already known to Euclid, that the rational points (x, y) on the unit circle $x^2 + y^2 = 1$ can be parametrized by the formula

$$(x, y) = \left(\frac{u^2 - v^2}{u^2 + v^2}, \frac{2uv}{u^2 + v^2}\right)$$

where u and v are arbitrary integers. In particular, infinitely many essentially different Pythagorean triples exist and we have a systematic way of generating them.

While this result can be generalized to more general quadratic equations in two variables (defining a *conic*), for a general polynomial F we cannot hope for an explicit description of all its rational solutions.

Instead, a version of Question 1 that in many cases is already very hard, is the following.

Question 2: *Does a (non-trivial) integral/rational solution to $F(\mathbf{x}) = 0$ exist? If yes, are there infinitely many?*

This question still involves the vague notion of *non-trivial* solutions, the exact definition of which will typically depend on the context. To give a very concrete and famous example, Fermat's Last Theorem asks whether the equation $x^n + y^n = z^n$ has a non-trivial solution for some $n \geq 3$, where here 'non-trivial' means precisely that all the variables should be different from 0.

Returning to the case of polynomials in two variables discussed above, the next simplest case after quadratic equations would be to study cubic equations. Generically, a cubic polynomial in two variables defines an *elliptic curve.*

The theory of rational and integral points on elliptic curves has become an extremely rich area of research in the last century, forming an important part of the proof of Fermat's Last Theorem due to Andrew Wiles, but also leading to powerful applications in modern-day cryptography.

A special feature of elliptic curves is that they are examples of *abelian varieties*, which means that the set of rational points forms an (algebraic) group. Very concretely, this allows us to generate new points from old ones. In geometric terms, this can be described by taking the third point of intersection of the tangent line to the curve at a given rational point.

As a very concrete example, for the curve given by $x^3 + y^3 = 9$, we can start with the solution $(1, 2)$ to generate iteratively the solution $\left(-\frac{17}{7}, \frac{20}{7}\right)$ and then the next two solutions are given by

$$\left(\frac{188479}{90391}, \frac{36520}{90391}\right) \text{ and } \left(\frac{1243617733990094836481}{609623835676137297449}, \frac{487267171714352336560}{609623835676137297449}\right).$$

What this example shows quite impressively is that while this procedure has a good chance of producing infinitely many different rational points, the *height* of the solutions, which we can think of as the size of numerator and denominator, grows extremely rapidly.

This line of thought leads us to another refinement of Question 1.

Question 3: Can we quantify the distribution of solutions to $F(\mathbf{x}) = 0$ in the case where infinitely many solutions exist?

One way to do this is to count *points of bounded height* which in the case of integer solutions means that we should study the *counting function*

$$N_F(P) := \#\{\mathbf{x} \in \mathbb{Z}^n : \|\mathbf{x}\|_\infty \leq P, F(\mathbf{x}) = 0\}$$

for a real parameter $P > 0$. In particular, it is an interesting question to determine the asymptotic growth of $N_F(P)$ as $P \to \infty$.

What kind of behaviour can we expect? A first approximation is given by the following probabilistic heuristic: The range $\|\mathbf{x}\| \leq P$ contains roughly P^n many integer points. If F is a polynomial of degree d, then its values on this set are roughly bounded by P^d, i.e. there are roughly P^d many different values that F can take. If we expect the particular 0 to obtain its 'fair share', we could therefore conjecture that $N_F(P)$ should roughly be of the size P^{n-d}.

A different heuristic leading to a more precise prediction for the size of $N_F(P)$ and indeed in many situations providing a viable approach of actually proving it, comes from a *Fourier-analytic perspective* via the *Hardy-Littlewood circle method.* The key players here are the linear characters

$$\alpha \mapsto e(n\alpha) := \exp(2\pi i n\alpha)$$

on the torus $\mathbb{R}/\mathbb{Z}$ and the corresponding orthogonality result

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

which therefore allows us to detect whether an expression is zero by computing a suitable Fourier integral. Applying this elementary observation with the value $n = F(\mathbf{x})$ and summing over all values of $\mathbf{x}$ in consideration, we thus obtain the following fundamental identity

$$N_F(P) = \int_0^1 S(\alpha) d\alpha,$$

where

$$S(\alpha) = \sum_{|\mathbf{x}| \leq P} e(\alpha F(\mathbf{x}))$$

is the *Weyl sum* associated to F.

Note that the size of $S(\alpha)$ is clearly bounded by (a constant times) P^n. However, since we are summing complex numbers on the unit circle, one might expect that for *generic* α, there is a large amount of cancellation, possibly leading to $S(\alpha)$ being roughly as small as $P^{n/2}$, a phenomenon known as *square-root cancellation*, familiar e.g. from the theory of random walks.

On the other hand, clearly $S(0)$ is as large as P^n, and indeed this fact persists for α appreciably smaller than $\frac{1}{P^d}$ since then $\alpha F(\mathbf{x})$ remains small.

We thus conclude that such a small interval for α around the point 0 contributes roughly P^{n-d} to the whole integral and hence to $N(P)$.

Naively, one might now hope to show that for all other α, the sum $S(\alpha)$ is indeed small and so the contribution computed above indeed leads to an asymptotic $N_F(P) \approx P^{n-d}$.

However, this cannot possibly be the case since so far we have not taken into account the arithmetic nature of F. For instance, it could happen that $F(\mathbf{x}) = 2G(\mathbf{x}) - 1$ is always odd and hence no integer solution to $F(\mathbf{x}) = 0$ can possibly exist.

Fortunately, this issue can be located in Fourier space as well: In that case we will have $S\left(\frac{1}{2}\right) = -S(0)$ and a small interval around $\frac{1}{2}$ will also make a contribution of the size P^{n-d} to the whole integral, cancelling out the contribution from the interval around 0.

As this phenomenon is not restricted to the number 2, it then transpires that we should do the following: For each rational number $\frac{a}{q}$, consider a small interval around $\frac{a}{q}$ and show that it contributes roughly

$$\frac{S(q,a)}{q^n} \cdot P^{n-d}$$

to the integral, where

$$S(q,a) := \sum_{\mathbf{x} \pmod q} e\left(\frac{aF(\mathbf{x})}{q}\right)$$

is the *Gauß Sum* associated to F. Then sum up the contributions from all these small intervals to end up with the prediction

$$N(P) \sim \mathfrak{J} \cdot \mathfrak{S} \cdot P^{n-d} \tag{1.0.1}$$

where $\mathfrak{J}$ is the *singular integral*, a correction factor that has to be introduced because $S(\alpha)$ is not quite constant on each of these small intervals (but which will in general be harmless) and

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\gcd(a;q)=1} \frac{S(q,a)}{q^n}$$

is the *singular series*. It is not hard to prove that the Gauß sums $S(q,a)$ enjoy a certain multiplicativity in the modulus q, allowing us to rewrite

$$\mathfrak{S} = \prod_p \chi_p$$

where the product is an *Euler product* over all primes p and

$$\chi_p = \sum_{k=0}^{\infty} \sum_{\gcd(a;p^k)=1} \frac{S(p^k,a)}{p^{kn}}$$

is a *p-adic density* which, as the name suggests, measures the density of solutions of the equation $F(\mathbf{x}) = 0$ in the p-adic numbers.

In very explicit terms, in the example above where F is always odd, we can conclude that no 2-adic solutions exist and we will have $\chi_2 = 0$.

In general, one can show that $\chi_p > 0$ if and only if a suitable p-adic solution exists.

This is very nice because it means that the asymptotic (1.0.1) captures a *Local-Global Principle* or *Hasse Principle*. It is trivial and always the case that (as in the case above where F is always odd) there can be no integer solution if there is a *congruence obstruction*, i.e. if $F(\mathbf{x}) = 0$ is insoluble modulo a certain number q. The converse however is far from trivial and not always true: If there are no congruence obstructions, does $F(\mathbf{x}) = 0$ always have an integer solution? If the prediction (1.0.1) is true, the answer is yes, because then the leading coefficient is positive so that there are indeed many solutions!

The terminology *Local-Global Principle* comes from the fact that $\mathbb{Q}$ is a global field and its completions are given by the local fields $\mathbb{R}$ and $\mathbb{Q}_p$ for p prime[1].

Of course the discussion above leading to (1.0.1) was far from rigorous. By summing over small intervals around all rational numbers, we would obtain significant overlaps. Moreover, the approximation to $S(\alpha)$ on each of these intervals becomes worse with growing denominator q.

In practice, one therefore uses a truncated version of this process. One only sums over intervals around rationals with small denominator, where the threshold is determined by the size of P. These *major arcs* can then be shown rigorously to lead to a contribution of the shape (1.0.1) involving a truncated version of $\mathfrak{S}$.

Note however, that contrary to what the nomenclature might suggest and despite contributing the main term to our asymptotic formula, the major

[1] The alert reader will have noticed that we did not mention local solubility at the infinite place corresponding to $\mathbb{R}$. Indeed, a local obstruction here would mean that F does not have a real solution, for example if it is positive definite. In that case one can show that the singular integral $\mathfrak{I}$ vanishes, so that this aspect is also captured in our asymptotic formula.

arcs in general constitute only a tiny part of the interval $[0, 1]$.
It thus remains to deal with the *minor arcs*, which are defined as the complement of the major arcs, hence consisting of real numbers that are not too close to a rational number with small denominator. Here, we can now hope to realize the '*random walk principle*' of showing that $S(\alpha)$ is small. This however is easier said than done and usually constitutes the hardest part of the method, thus also limiting its applicability significantly.

From this discussion, it transpires that the method will work better when the number of variables is relatively large. In particular, since we cannot hope for a better bound than $P^{n/2}$ on the minor arcs and the expected main term is of the size P^{n-d}, we cannot in general hope to apply the method when n is less than $2d$. In practice the ideal exponent $\frac{n}{2}$ on the minor arcs is hardly ever realized, so that the method usually requires even more variables.
We want to stress, however, that even in the case of many variables where we expect many solutions to exist, it is not even clear a priori that even a single solution exists. The Hardy-Littlewood method should therefore be seen as the primary example of the following fundamental principle of Diophantine Analysis:

Show that solutions exist by showing that there are indeed many!

We now come to the discussion of the main results of this thesis and for this purpose we will soon restrict to the case of cubic polynomials.

Why should we study cubic polynomials? It was already discussed earlier that cubic polynomials in two variables are essentially elliptic curves and indeed results on cubic surfaces or higher-dimensional cubic hypersurfaces can often be related to the rich and highly developed theory of elliptic curves. From another point of view, one can see the theory in the cubic case as the first interesting instance of the general theory of higher-degree polynomials. The case of linear equations is of course trivial and the theory of quadratic forms is classical and well-understood. In contrast, the cubic case exhibits new features and we are still far from being able to answer many natural questions. Most of these questions can then also be raised in the higher-degree case, but it is natural to first study the simplest non-trivial case.

After this motivation, let us now first consider a homogeneous cubic polynomial C with integer coefficients in n variables $x_1, \dots, x_n$.
Since C is homogeneous, the study of rational solutions is equivalent to studying integer solutions, as we may always clear denominators. Moreover, note that $(0, 0, \dots, 0)$ always is a solution. We thus say that an integer solution is non-trivial if it is different from this solution.

While there are examples of cubic forms in up to 9 variables for which there are congruence obstructions to non-trivial solutions, it has been known for a long time that such obstructions cannot occur when $n \geq 10$.
In view of the previous discussion, it is therefore a natural conjecture that the equation $C(\mathbf{x}) = 0$ should always have a non-trivial integer solution when $n \geq 10$.
This remains an open problem, but recent years have seen progress on this question, culminating in the seminal work of Heath-Brown [16] who proved the existence of non-trivial solutions for $n \geq 14$.
His method uses the Hardy-Littlewood method as described above, including a grain of salt. Indeed, let us note that the asymptotic formula

$$N(P) = (1 + o(1))\mathfrak{J} \cdot \mathfrak{S} \cdot P^{n-d}, \quad P \to \infty \tag{1.0.2}$$

cannot possibly hold for all cubic forms, as indeed in certain non-degenerate situations there are *too many solutions.* This happens for instance in the case of a reducible cubic form $C(\mathbf{x}) = L(\mathbf{x})Q(\mathbf{x})$ which clearly has $N(P) \gg P^{n-1}$, and even in cases like

$$C(\mathbf{x}) = L_1(\mathbf{x})Q_1(\mathbf{x}) + L_2(\mathbf{x})Q_2(\mathbf{x}) \tag{1.0.3}$$

for certain linear forms L_1, L_2 and quadratic forms Q_1, Q_2, where we still have $N(P) \gg P^{n-2}$ simply by setting $x_1 = x_2 = 0$.

Following an ingenious idea of Davenport, Heath-Brown introduces a dichotomy to deal with this issue: If C satisfies a certain genericity condition, called *Davenport's Geometric Condition*, we can apply the Hardy-Littlewood method to prove (1.0.2) and in particular deduce the existence of solutions. If the Geometric Condition fails, we cannot use the circle method, but already Davenport managed to show that this failure can be turned into an alternative existence proof for a non-trivial solution.

While this alternative argument would still work for $n \geq 10$, it is the application of the circle method and more precisely our poor understanding of the Weyl sum $S(\alpha)$ on the minor arcs, that leads to the restriction $n \geq 14$.

It is thus natural to conjecture that even the asymptotic formula (1.0.2) should continue to hold for $n \geq 10$, under the assumption of the Geometric Condition.

In Chapter 2, we make progress towards this conjecture with the following result:

Theorem 1.0.1. *Assume that $n \geq 10$ and that C satisfies Davenport's Geometric Condition. Then the singular series $\mathfrak{S}$ is absolutely convergent. In particular, $\mathfrak{S} > 0$.*

This can be seen as a step towards the conjecture in two different ways: Firstly, while a proof of the asymptotic formula (1.0.2) for $n \geq 10$ remains elusive, our result establishes the convergence and positivity of the objects involved.

Indeed, recall that $\mathfrak{S}$ can be written as an infinite Euler product $\prod_p \chi_p$ over all primes. While we know that the individual χ_p are all positive for $n \geq 10$ (since local solutions exist), this alone is not enough to deduce the positivity of the product, unless we establish its absolute convergence.

Secondly, the proof of the theorem naturally involves new bounds for the Gauß sums $S(q, a)$. Recalling that these are complete exponential sums which are closely connected to the more general Weyl sum $S(\alpha)$, bounds for the Gauß sums can be seen as a natural model problem for the harder problem of improving the known bounds for the Weyl sum on the minor arcs, which would be required to improve on Heath-Brown's 14-variable result.

We also mention that all previous bounds on the Weyl sum for a general cubic form were based on *Davenport's Shrinking Lemma*, an elementary result on diophantine inequalities which, however, so far did not have an elementary proof.

In Chapter 2, we also give a new elementary and short proof of the Shrinking Lemma. While this is very satisfatory, it does not immediately lead to a better result. Instead, we reinterpret the previous use of the Shrinking Lemma in terms of the $\mathbb{F}_p$-rank of certain matrices. It is this reinterpretation that eventually allows us to improve the existing bounds and establish the result for $n \geq 10$.

In Chapter 3, which is joint work with Leonhard Hochfilzer, we are interested in the solubility of the equation $C(\mathbf{x}) = 0$ in the prime numbers.

Many of the most famous problems in classical number theory, such as Goldbach's Problem or the Twin Prime Conjecture can be interpreted as asking for the solubility of a certain Diophantine equation in the set of prime numbers. While we cannot quite prove the existence of prime solutions to the cubic equation $C(\mathbf{x}) = 0$, we establish the following:

Theorem 1.0.2. *Let C be a cubic form in $n \geq 33$ variables with rational coefficients. Then there are almost prime solutions to $C(\mathbf{x}) = 0$ in the following sense: There are coprime integers $c_1, \ldots, c_n$ such that the equation*

$$C(c_1p_1, c_2p_2, \ldots, c_np_n) = 0$$

has infinitely many solutions in primes $p_1, \ldots, p_n$, not all equal.

Perhaps surprisingly, this is a consequence of the following geometric result, in conjunction with the Green-Tao Theorem on primes in arithmetic progressions:

Theorem 1.0.3. *Let C be a cubic form in $n \geq 33$ variables with rational coefficients. Then the projective cubic hypersurface defined by $C(\mathbf{x}) = 0$ contains a rational projective line.*

Both theorems improve on results of Wooley [29] by four variables. Interestingly, contrary to the question of existence of rational points, there does not seem to be a clear heuristic of how many variables should be required to ensure the existence of a rational line, but we certainly do not expect the bound 33 to be optimal. Indeed, combining our method with recent ideas of Brandes and Dietmann, one can save another two variables and prove both results for $n \geq 31$.

How does one prove the existence of a rational line? Expanding the condition $C(\mathbf{x} + t\mathbf{y}) = 0$ for the existence of such a line in terms of t, we are led to a system

$$C(\mathbf{x}) = Q_{\mathbf{y}}(\mathbf{x}) = L_{\mathbf{y}}(\mathbf{x}) = C(\mathbf{y}) = 0.$$

A natural strategy is to start by choosing a point $\mathbf{y}$ with $C(\mathbf{y}) = 0$, which we can do by Heath-Brown's result. This leaves us with a system of a linear, quadratic and cubic equation in $\mathbf{x}$. While the linear equation does not constitute a problem, the quadratic equation is an issue as it is hard to control its signature so that it might be definite or nearly so.

The solution, as pioneered by Wooley, is to pass to an imaginary quadratic number field $K/\mathbb{Q}$ in which the definiteness issue disappears. By an observation of Lewis, finding a suitable K-rational line would suffice to deduce the existence of a $\mathbb{Q}$-rational line.

The key problem then becomes to study the existence of solutions to cubic homogeneous equations over number fields. We are able to generalize Heath-Brown's method and obtain the following:

Theorem 1.0.4. *Let $K/\mathbb{Q}$ be an imaginary quadratic number field. If $C(\mathbf{x})$ is a homogeneous cubic form over K in at least 14 variables, then $C(\mathbf{x}) = 0$ has a non-trivial solution.*

We remark that this by far does not constitute the first instance of an application of the Hardy-Littlewood method over number fields. However, Heath-Brown needed to introduce several technical innovations to the method in order to treat the case of 14 variables. In trying to generalize them to the number field case, one encounters serious difficulties, and indeed we are able to surmount all these difficulties only in the case of imaginary quadratic number fields. While one would certainly hope to extend this result to all number fields, this level of generality is sufficient to deduce Theorems 1.0.3 and 1.0.2.

In total, the proof of Theorem 1.0.2 thus consists of a mixture of results from classical prime number theory (Green-Tao Theorem), geometric ideas (Wooley, Lewis) and heavy Fourier-analytic machinery in the form of the Hardy-Littlewood method over number fields.

In Chapter 4, we finally introduce two new aspects to the problems discussed previously. The first is the generalization to *inhomogeneous* cubic equations $\phi(\mathbf{x}) = 0$. While many aspects of the method are similar to the homogeneous case, new phenomena arise in this case.

First of all, it is no longer the case that the assumption of sufficiently many variables ensures that there are no congruence obstructions. We thus always need to assume the *Necessary Congruence Condition* that $\phi(\mathbf{x}) \equiv 0 \pmod{N}$ is solvable for all N. However, as the following example

$$\phi(\mathbf{x}) = (2x_1 - 1)(1 + x_1^2 + x_2^2 + \cdots + x_n^2) + x_1 x_2$$

of Watson shows, the Necessary Congruence Condition is not in general sufficient to ensure the existence of integer solutions, even if the number of variables is large.
Following an approach of Davenport and Lewis, we introduce the h-invariant of a cubic form C to be the least positive integer such that

$$C(\mathbf{x}) = \sum_{i=1}^{h} L_i(\mathbf{x}) Q_i(\mathbf{x})$$

for appropriate linear forms $L_1, \ldots, L_h$ and quadratic forms $Q_1, \ldots, Q_h$. Equivalently, $n - h$ is the largest dimension of a linear subspace contained in the cubic hypersurface defined by the equation $C(\mathbf{x}) = 0$. With a slight variant of the definition of the counting function $N(P)$, we are then able to prove the following:

Theorem 1.0.5. *Assume that $\phi = C + Q + L + N$ is of degree 3, non-degenerate, satisfies the Necessary Congruence Condition and $h(C) \geq 14$. Then*

$$N(P) = (1 + o(1))\mathfrak{S} \cdot \mathfrak{I} \cdot P^{n-3}, \quad P \to \infty$$

where $\mathfrak{S}$ and $\mathfrak{I}$ are the usual singular series and singular integral, respectively. We have $\mathfrak{S} > 0$ and $\mathfrak{I} > 0$. In particular, there is a solution $\mathbf{x} \in \mathbb{Z}^n$ to $\phi(\mathbf{x}) = 0$.

Note that this also shows the asymptotic formula in the homogeneous case under the assumption $h(C) \geq 14$. The relevance of the h-invariant can already be seen from the fact that the counterexample (1.0.3) to the asymptotic from above has $h(C) = 2$.

The second new aspect introduced in Chapter 4 is a different approach to our Question 3 from above, i.e. to quantify the distribution of solutions. Indeed, one may ask for an upper bound on the size of the smallest non-trivial solution in terms of the coefficients of the polynomial ϕ.

In many situations, we expect such a *search bound* to be polynomial in M, where M is the maximal absolute value among the coefficients of ϕ. Improving on results of Browning-Dietmann-Elliott [4], we are able to prove the following:

Theorem 1.0.6. *With the same assumptions as in Theorem 1.0.5, there exists a vector* $\mathbf{x} \in \mathbb{Z}^n$ *with* $\phi(\mathbf{x}) = 0$ *and*

$$\max_i |x_i| \ll M^{6407n^2}.$$

In the homogeneous case, if $n \geq 14$ *and* $C \in \mathbb{Z}[x_1, \ldots, x_n]$ *is a cubic form, there exists a vector* $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ *with* $C(\mathbf{x}) = 0$ *and*

$$\max_i |x_i| \ll M^{132484}.$$

If, additionally, we assume C *to be non-singular, then for* $n = 14$ *we can ensure that*

$$\max_i |x_i| \ll M^{2049}.$$

The proof involves a careful generalization of Heath-Brown's method, together with some new ideas regarding uniform lower bounds for the singular series and the singular integral.

Chapter 2

The singular series of a cubic form in many variables and a new proof of Davenport's Shrinking Lemma

2.1 Introduction

In this chapter, let $C(x_1, \dots, x_n) \in \mathbb{Z}[x_1, \dots, x_n]$ be a cubic form. We are interested in the existence of nontrivial integer solutions, i.e. nonzero vectors $\mathbf{x} \in \mathbb{Z}^n$ with $C(\mathbf{x}) = 0$.

Davenport [10] proved that if $n \geq 16$, such nontrivial solutions always exist. This remained the state of the art for almost half a century until Heath-Brown [16] could extend the admissible range to $n \geq 14$, this has not been improved to date. Given that 10 variables suffice to guarantee local solubility [21], it is generally expected that the same result should hold already when $n \geq 10$.

The Hardy-Littlewood Circle Method aims to prove the existence of solutions by proving that there are indeed many. If it works, it provides us with an asymptotic formula of the shape

$$\#\{\mathbf{x} \in \mathbb{Z}^n, x \in P\mathcal{B}\} = (1 + o(1)) \cdot \mathfrak{I} \cdot \mathfrak{S} \cdot P^{n-3} \tag{2.1.1}$$

as $P \to \infty$. Here $\mathcal{B} \subset \mathbb{R}^n$ is a suitably chosen box and $\mathfrak{I}$ and $\mathfrak{S}$ denote the usual *singular integral* and the *singular series* of the cubic form C, respectively, measuring the local solubility of C over the fields $\mathbb{R}$ and $\mathbb{Q}_p$ for all primes p. The singular integral is rather unimportant for this paper, so we refer the reader to [10] for its precise definition and only mention that it is known to be positive for a suitable choice of $\mathcal{B}$ as soon as $n \geq 4$. The singular series is the key object of the present paper and will be defined and discussed in more detail in the next section.

For now, let us continue discussing the heuristic asymptotic formula (2.1.1) and let us note that it clearly fails in certain degenerate situations. Indeed, when C is reducible, it is easy to see that the count on the left-hand side is already $\gg P^{n-1}$. More generally, if our cubic form is of the shape $C(\mathbf{x}) = x_1Q_1(\mathbf{x}) + x_2Q_2(\mathbf{x})$ for certain quadratic forms Q_1 and Q_2, we still have $\gg P^{n-2}$ solutions and hence too many for (2.1.1) to possibly hold.

The ingenious idea of Davenport to circumvent this fundamental problem was to establish a certain dichotomy: If the circle method fails to produce the asymptotic (2.1.1), then this failure could be turned into an alternative proof of the existence of solutions, though not in such a precise quantitative manner.

To describe Davenport's idea in more detail, we write $C(\mathbf{x}) = \sum_{i,j,k} c_{ijk}x_ix_jx_k$ where we assume the c_{ijk} to be symmetric and integers (as we may by multiplying C by 6 if necessary). We then define the bilinear forms

$$B_i(\mathbf{x}, \mathbf{y}) = \sum_{j,k=1}^{n} c_{ijk}x_jy_k$$

and the matrix $M(\mathbf{x})$ with entries

$$M(\mathbf{x})_{jk} = \sum_{i=1}^{n} c_{ijk} x_i$$

so that $M(\mathbf{x})\mathbf{y}$ is the vector with entries $B_i(\mathbf{x}, \mathbf{y})$. For later use we let $D(\mathbf{x}) = \det M(\mathbf{x})$ and $r(\mathbf{x}) = \mathrm{rk} M(\mathbf{x})$. For a prime p, we will also need to consider the $\mathbb{F}_p$-rank of $M(\mathbf{x})$ which we denote by $r_p(\mathbf{x})$.

Let us now say that C satisfies *Davenport's Geometric Condition* if

$$\#\{\mathbf{x} \in \mathbb{Z}^n : \|x\|_\infty \leq P, r(\mathbf{x}) = r\} \ll P^{r+\varepsilon} \tag{2.1.2}$$

is satisfied for all integers r with $0 \leq r \leq n$.

We can then describe Davenport's result more concisely as follows:

Theorem A. *If C does not satisfy Davenport's Geometric Condition* (2.1.2), *then the equation $C(\mathbf{x}) = 0$ has a non-trivial integer solution.*

Theorem B. *If C satisfies Davenport's Geometric Condition* (2.1.2), *then the asymptotic formula* (2.1.1) *holds with $\mathfrak{J}, \mathfrak{S} > 0$ as soon as $n \geq 16$. In particular, there are non-trivial integer solutions to $C(\mathbf{x}) = 0$.*

Note that Theorem A does not make any assumption on the number of variables n. This means that in trying to improve on the constraint on the number of variables, we are free to assume that the Geometric Condition is satisfied.

Indeed, this is what Heath-Brown did, showing

Theorem C. *If C satisfies Davenport's Geometric Condition* (2.1.2), *then the asymptotic formula* (2.1.1) *holds with $\mathfrak{J}, \mathfrak{S} > 0$ as soon as $n \geq 14$. In particular, there are non-trivial integer solutions to $C(\mathbf{x}) = 0$.*

In view of the above discussion, it is natural to conjecture that this should extend to $n \geq 10$:

Conjecture 2.1.1. *If C satisfies Davenport's Geometric Condition* (2.1.2), *then the asymptotic formula* (2.1.1) *holds with $\mathfrak{I}, \mathfrak{S} > 0$ as soon as $n \geq 10$. In particular, there are non-trivial integer solutions to $C(\mathbf{x}) = 0$.*

2.2 Main results

We now describe our main results. To this end, we need to return to the singular series $\mathfrak{S}$. It is defined in terms of the *Gauß sums*

$$S(q,a) = \sum_{\mathbf{x} \pmod q} e\left(\frac{aC(\mathbf{x})}{q}\right)$$

via

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{(a;q)=1} \frac{S(q,a)}{q^n}.$$

By standard multiplicativity properties of the Gauß sums, this can (at least formally) also be written as an Euler product

$$\mathfrak{S} = \prod_p \chi_p$$

over all primes p where

$$\chi_p = \sum_{k=0}^{\infty} \sum_{(a;p^k)=1} \frac{S(p^k,a)}{p^{kn}}$$

is known as the p-adic density. By classical arguments it follows that $\chi_p > 0$ if and only if $C(\mathbf{x}) = 0$ has a non-trivial solution over $\mathbb{Q}_p$. In particular, from [21] we conclude that $\chi_p > 0$ for all p whenever $n \geq 10$.

So far we have ignored all convergence issues. The rearrangement between the series and the product representation of $\mathfrak{S}$ is only valid when either of the two is known to be absolutely convergent. Proving absolute convergence of $\mathfrak{S}$ is therefore crucial for switching between the two representations and also to conclude its positivity from the positivity of all individual factors χ_p. Only then, the formula (2.1.1) truly captures the expected Local-Global Principle.

Previously, the absolute convergence for $\mathfrak{S}$ under the assumption of Davenport's Geometric Condition (2.1.2) was known for $n \geq 11$ by work of Heath-Brown [16].

We begin by giving a new short and self-contained proof of this result. This new method then allows us to improve on previous work and establish the following.

Theorem 2.2.1. *Assume that $n \geq 10$ and that C satisfies Davenport's geometric condition. Then the singular series $\mathfrak{S}$ is absolutely convergent. In particular, $\mathfrak{S} > 0$.*

This can be seen as giving further evidence to Conjecture 2.1.1. Moreover, the Gauß sums featuring in the definition of the singular series are closely related to the Weyl sums that would appear in a circle method proof of (2.1.1). It is therefore to be hoped that the study of the Gauß sums and hence of the singular series can serve as a good model problem for our understanding of the more difficult Circle Method Problem.

We can also say something about the case $n = 9$. We begin by proving that the only possible obstructions to absolute convergence are the Gauß sums with prime moduli. To deal with them, we then propose the following conjecture:

Conjecture 2.2.2. *Assume that C satisfied the Geometric Condition* (2.1.2)*. Then for all n and uniformly in $1 \leq H \leq R$, we have*

$$\#\{\mathbf{h} \leq H, R < p \leq 2R : r_p(\mathbf{h}) \leq r\} \ll H^r \cdot R^{1+\varepsilon}.$$

We are able to prove the following:

Theorem 2.2.3. *Under the assumption of Conjecture 2.2.2, the singular series is absolutely convergent for $n = 9$.*

In the last section, we return to the work of Davenport and Heath-Brown and give a short and elementary proof of *Davenport's Shrinking Lemma*, which is a crucial ingredient in the circle method approach to the cubic forms problem as pioneered by Davenport. The only previous proof of the Shrinking Lemma is due to Davenport and uses rather intricate tools from the geometry of numbers.

2.2.1 Notation

We use the usual notation $\mathcal{O}(\dots)$ and $\ll$ where the implicit constants are always allowed to depend on the cubic form $C(\mathbf{x})$. Moreover, whenever a bound involves ε, it means that the bound is true for all sufficiently small $\varepsilon > 0$, but the implicit constant is allowed to depend on ε.
Moreover, we use the notation $e(x) = e^{2\pi i x}$ and $\|x\| = \min_{n \in \mathbb{Z}} |x - n|$. Whenever we write something like $\sum_{\mathbf{h}}$, the sum is restricted to integer vectors $\mathbf{h}$ and the given restrictions on the summation are to be read component-wise. Finally, the condition $r \sim R$ denotes a restriction of r to a dyadic interval $(R, 2R]$.

2.3 Review of previous bounds for $S(q, a)$

The following simple lemma is good enough to recover all results previously obtained:

Lemma 2.3.1. *Let q and n be positive integers and let M be a $n \times n$ matrix with integer coefficients. Then the size of the kernel of M viewed as a map from $(\mathbb{Z}/q\mathbb{Z})^n$ to itself divides* $\det M$. *In particular, if q is a prime and M has $\mathbb{F}_q$-rank at most $n - r$, then $p^r \mid \det M$.*

Proof. Without loss of generality (that is, up to multiplication from both sides by invertible matrices), we may assume that M is in Smith Normal Form with diagonal entries $a_1, \ldots, a_n$. Then the kernel has size $\prod_{i=1}^n (a_i; q)$ which divides $\det M = \prod_{i=1}^n a_i$. □

We now recall the classical van der Corput differencing:

Lemma 2.3.2 (Initial van der Corput Bound)**.** *Let $H \geq 1$ be arbitrary. Then, in the above notation, we have*

$$\left(\frac{S(q,a)}{q^n}\right)^2 \ll \frac{1}{H^n} \sum_{1 \leq \mathbf{h} \leq H} \sqrt{\frac{1}{q^n} \#\{\mathbf{y} \pmod q : q \mid B_i(\mathbf{y}, \mathbf{h})\}}. \tag{2.3.1}$$

Proof. We set out by applying Cauchy-Schwarz to the identity

$$S(q,a) = \frac{1}{H^n} \sum_{\mathbf{x} \pmod q} \sum_{1 \leq \mathbf{h} \leq H} e\left(\frac{aC(\mathbf{x}+\mathbf{h})}{q}\right)$$

to obtain after some manipulations

$$|S(q,a)|^2 \ll \frac{q^n}{H^n} \sum_{-H \leq \mathbf{h} \leq H} \left| \sum_{\mathbf{x} \pmod q} e\left(\frac{a\left(C(\mathbf{x}+\mathbf{h}) - C(\mathbf{x})\right)}{q}\right)\right|.$$

The lemma now follows by noting that the square of the absolute value of the inner sum is just

$$\begin{aligned}
&\sum_{\mathbf{x},\mathbf{y}} e\left(\frac{a\left(C(\mathbf{x}+\mathbf{y}+\mathbf{h}) - C(\mathbf{x}+\mathbf{y}) - C(\mathbf{x}+\mathbf{h}) + C(\mathbf{x})\right)}{q}\right) \\
&= \sum_{\mathbf{x},\mathbf{y}} e\left(\frac{a\sum_i x_i B_i(\mathbf{y},\mathbf{h})}{q}\right)
\end{aligned}$$

and using orthogonality. □

Next, from Lemma 2.3.1 we see that $q^{r(\mathbf{h})-n}\#\{\mathbf{y} \pmod q : q \mid B_i(\mathbf{y},\mathbf{h})\}$ divides a non-zero $r(\mathbf{h}) \times r(\mathbf{h})$ minor of M so that in particular

$$\frac{1}{q^n}\#\{\mathbf{y} \pmod q : q \mid B_i(\mathbf{y},\mathbf{h})\} \ll \left(\frac{H}{q}\right)^{r(\mathbf{h})}.$$

Inserting this into Lemma 2.3.2 and using the geometric condition (2.1.2), we find that

$$\begin{aligned}\left(\frac{S(q,a)}{q^n}\right)^2 &\ll \frac{1}{H^n} \sum_{-H\leq \mathbf{h}\leq H} \left(\frac{H}{q}\right)^{r(\mathbf{h})/2} \\ &\ll \frac{q^\varepsilon}{H^n} \sum_{r=0}^{n} \left(\frac{H^3}{q}\right)^{r/2} \\ &\ll q^\varepsilon \left(\frac{1}{H^n} + \frac{H^{n/2}}{q^{n/2}}\right)\end{aligned}$$

and putting $H = q^{1/3}$, we recover Heath-Brown's pointwise bound $S(q,a) \ll q^{5n/6+\varepsilon}$.

Recalling the definition of the p-adic factor in the product expansion of $\mathfrak{S}$, we now find that

$$\begin{aligned}\chi_p &= \sum_{k=0}^{\infty} \sum_{(a;p^k)=1} \frac{S(p^k,a)}{p^{nk}} \\ &= 1 + \sum_{(a;p)=1} \frac{S(p,a)}{p^n} + \mathcal{O}\left(\sum_{k=2}^{\infty} p^{k(1-n/6)+\varepsilon}\right) \\ &= 1 + \sum_{(a;p)=1} \frac{S(p,a)}{p^n} + \mathcal{O}\left(p^{2-n/3+\varepsilon}\right)\end{aligned}$$

so that the estimation of the terms with $k \geq 2$ is satisfactory for the question of absolute convergence of $\mathfrak{S}$ as soon as $n > 9$.

To establish Theorem 2.2.1, it therefore remains to show that

$$\sum_p \sum_{(a;p)=1} \frac{S(p,a)}{p^n}$$

converges absolutely for $n \geq 10$. It would therefore clearly suffice to show that

$$\sum_{p\sim R} \max_{(a;p)=1} \left|\frac{S(p,a)}{p^n}\right| \ll R^{-1-\delta}$$

for which, by Cauchy-Schwarz, it suffices to establish

$$\sum_{p\sim R} \max_{(a;p)=1} \left|\frac{S(p,a)}{p^n}\right|^2 \ll R^{-3-\delta}$$

for all choices of $R \geq 1$.

Using our previous line of argument, the LHS is bounded by

$$\begin{aligned}&\frac{1}{H^n}\sum_{-H\leq \mathbf{h}\leq H}\sum_{R\leq p<2R}\sqrt{\frac{1}{p^n}\#\{\mathbf{y} \pmod p : p \mid B_i(\mathbf{y},\mathbf{h})\}}\\ =&\frac{1}{H^n}\sum_{-H\leq \mathbf{h}\leq H}\sum_{R\leq p<2R} p^{-r_p(\mathbf{h})/2}\end{aligned}$$

where $r_p(\mathbf{h})$ is the $\mathbb{F}_p$-rank of $M(\mathbf{h})$. Heath-Brown's idea is now to distinguish two cases:

Those pairs $(\mathbf{h}, p)$ with $r_p(\mathbf{h}) = r(\mathbf{h})$ give a contribution bounded by

$$\frac{1}{H^n}\sum_{-H\leq \mathbf{h}\leq H}\sum_{R\leq p<2R} p^{-r(\mathbf{h})/2} \ll \frac{1}{H^n}\sum_{r=0}^{n} H^{r+\varepsilon}R^{1-r/2} \ll H^{\varepsilon}\left(\frac{R}{H^n}+\frac{R}{R^{n/2}}\right) \tag{2.3.2}$$

by the Geometric Condition (2.1.2). The last term is satisfactory for $n > 8$. On the other hand, we need to estimate the contribution from those pairs $(\mathbf{h}, p)$ with $r_p(\mathbf{h}) < r(\mathbf{h})$. Here we use the implication from Lemma 2.3.1 that $p^{r(\mathbf{h})-r_p(\mathbf{h})}$ must divide a non-zero $r(\mathbf{h}) \times r(\mathbf{h})$-minor of $M(\mathbf{h})$ and is hence $\mathcal{O}(H^{r(\mathbf{h})})$ so that

$$p^{-r_p(\mathbf{h})} \ll \left(\frac{H}{p}\right)^{r(\mathbf{h})}. \tag{2.3.3}$$

Moreover, p being a divisor of such a minor, there are at most H^{ε} choices of such p for any fixed $\mathbf{h}$. The total contribution of such pairs $(\mathbf{h}, p)$ can therefore be bounded by

$$\frac{H^{\varepsilon}}{H^n}\sum_{\mathbf{h}}\left(\frac{H}{R}\right)^{\frac{r(\mathbf{h})}{2}} \ll H^{\varepsilon-n}\sum_{r=0}^{n}\left(\frac{H}{R}\right)^{\frac{r}{2}} \ll H^{\varepsilon}\left(\frac{1}{H^n}+\left(\frac{H^3}{R}\right)^{\frac{n}{2}}\right) \tag{2.3.4}$$

again using the geometric condition.

Comparing the contributions from (2.3.2) and (2.3.4) we find that the optimal choice is $H = R^{\frac{n+2}{3n}}$ leading to the bound $R^{-(n-1)/3+\varepsilon}$ which is satisfactory when $n > 10$.

2.4 The case of ten variables

When $n = 10$, we observe that $H = R^{2/5+\delta}$ for sufficiently small $\delta > 0$ leads to a satisfatory contribution from (2.3.2) and from all terms in (2.3.4) except when $r = n = 10$. Moreover, even for this term it suffices to save another small power of R, which we do in (2.3.3) unless $r_p(\mathbf{h}) = 6$. It therefore suffices to show that

$$\#\{\mathbf{h} \leq H, p \sim R : r(\mathbf{h}) = 10, r_p(\mathbf{h}) = 6\} \ll H^{10-\delta'}$$

for some $\delta' > 0$ whenever $H = R^{2/5+\delta}$ for sufficiently small $\delta > 0$.

To prove this, we use an argument inspired by a trick of Davenport [10] which he used to go from 17 to 16 variables. However, the presence of the extra averaging over p requires a new idea.

By Lemma 2.3.1, we have $p^4 \mid D(\mathbf{h})$ for all such $\mathbf{h}$. Moreover, there are p^4 vectors $\mathbf{y} \in \{0, 1, 2, \ldots, p-1\}^n$ with $p \mid B_i(\mathbf{y}, \mathbf{h})$ for all i.

By the Pigeonhole principle, two of them differ by $\mathcal{O}(p^{3/5})$ in each component and by linearity of the B_i, this means that for each such $\mathbf{h}$ we get one solution $\mathbf{y} = \mathbf{y}(\mathbf{h}) \neq 0$ with $\|\mathbf{y}\|_\infty \ll p^{3/5}$ and $p \mid B_i(\mathbf{y}, \mathbf{h})$.

Writing $B_i(\mathbf{y}, \mathbf{h}) = pm_i$, we find that $m_i = m_i(\mathbf{h}) \ll R^\delta$. Moreover, not all m_i are zero since we assumed $r(\mathbf{h}) = 10$.

We can now count the number of pairs $(\mathbf{h}, p)$ in question as follows: There are $\ll R^{10\delta}$ possible choices of the m_i. For a fixed choice of $(m_1, \ldots, m_n)$, we then study the number of possible choices of $(\mathbf{h}, p)$. The general solution of the system $B_i(\mathbf{y}, \mathbf{h}) = pm_i$ is given by

$$y_j = p \cdot \frac{\sum_k m_k E_{jk}(\mathbf{h})}{D(\mathbf{h})}$$

where the E_{jk} are certain 9×9 minors of $M(\mathbf{h})$, in particular homogeneous forms of degree 9 in $\mathbf{h}$.

Now certainly, for our given choice of the m_i, there is one j such that the degree-9 form $E(\mathbf{h}) := \sum_k m_k E_{j,k}(\mathbf{h})$ is not identically zero. We conclude that $D(\mathbf{h}) \mid p \cdot E(\mathbf{h})$.

Let G be the greatest common divisor of D and E and write $D = GD'$ and $E = GE'$ so that $D'(\mathbf{h}) \mid p \cdot E'(\mathbf{h})$ and D' is coprime to E'. We thus find by Bézout's Theorem a non-zero linear combination F of D' and E' that depends only on $h_2, \ldots, h_n$. Hence $D'(\mathbf{h}) \mid p \cdot F(h_2, \ldots, h_n)$. Note that the coefficients of all the polynomials depend on the m_i, but are all polynomially bounded in terms of R which is sufficient for our application.

Now there are at most H^9 values of $\mathbf{h}$ where F is zero and then p as a divisor of $D(\mathbf{h})$ is determined up to H^ε many choices, leading to a total bound of $H^{9+\varepsilon}$ for the number of pairs $(\mathbf{h}, p)$ in this case.

On the other hand, if $F(h_2, \ldots, h_{10})$ is non-zero, we see that $p \mid F(h_2, \ldots, h_{10})$ by the following ad-hoc bootstrapping argument: Since $p^4 \mid D(\mathbf{h}) = G(\mathbf{h}) \cdot D'(\mathbf{h})$ and $\deg G \leq 9$ we have $G(\mathbf{h}) \ll H^9 < p^4$ if $\delta > 0$ is sufficiently small. Hence $p \mid D'(\mathbf{h})$. But if δ is small, this forces $\deg D' \geq 3$ and hence $\deg G \leq 7$ so that $G(\mathbf{h}) \ll H^7 < p^3$, again if δ is small. Hence $p^2 \mid D'(\mathbf{h})$ and hence $p \mid F(h_2, \ldots, h_{10})$ as desired.

Finally, for any choice of $h_2, \ldots, h_{10}$ with $F(\mathbf{h}) \neq 0$, this determines p and $D'(\mathbf{h})$ up to H^ε many choices and then also h_1 is determined up to finitely many choices, unless we are in a proper Zariski-closed subset of $h_2, \ldots, h_{10}$. In any case, the total number of pairs $(\mathbf{h}, p)$ can be bounded by $H^{9+\varepsilon}$. Summing up, we have thus shown that

$$\#\{\mathbf{h} \leq H, p \sim R : r(\mathbf{h}) = 10, r_p(\mathbf{h}) = 6\} \ll R^{10\delta} \cdot H^{9+\varepsilon}$$

which is satisfactory for $\delta > 0$ sufficiently small. This finishes the proof of Theorem 2.2.1.

2.5 The case of nine variables

We now set out to discuss the case $n = 9$, aiming for a proof of Theorem 2.2.3. To begin with, we need to discuss the case of higher prime powers. The contribution to χ_p of $S(p^k, a)$ for $k \geq 3$ is seen to be satisfactory even for $n = 9$. For the contribution of the terms with $k = 2$, our pointwise bound $S(p^2, a) \ll p^{5n/3+\varepsilon}$ just fails to be good enough when $n = 9$.
However, we can use the averaging trick introduced in the previous section to also improve on this bound and therefore reduce the problem of absolute convergence of $\mathfrak{S}$ for $n = 9$ to the study of $S(p, a)$ for primes p:

Lemma 2.5.1. *For $n = 9$, the sum*

$$\sum_p \sum_{(a;p^2)=1} \frac{S(p^2, a)}{p^{2n}}$$

is absolutely convergent. In particular, the singular series for $n = 9$ converges absolutely if and only if

$$\sum_p \sum_{(a;p)=1} \frac{S(p, a)}{p^n}$$

is absolutely convergent.

Proof. As before, a dyadic decomposition and an application of Cauchy-Schwarz reduce the problem to showing that

$$\sum_{p \sim R} \max_{(a;p^2)=1} \left| \frac{S(p^2, a)}{p^{2n}} \right|^2 \ll R^{-5-\delta}.$$

From Lemma 2.3.2, we see that the LHS is bounded by

$$\frac{1}{H^n} \sum_{-H \leq \mathbf{h} \leq H} \sum_{p \sim R} \sqrt{\frac{1}{p^{2n}} \#\{\mathbf{y} \pmod{p^2} : p^2 \mid B_i(\mathbf{y}, \mathbf{h})\}}.$$

We continue by separating the cases $r_p(\mathbf{h}) = r(\mathbf{h})$ and $r_p(\mathbf{h}) < r(\mathbf{h})$. In the first case, the expression under the root is $(p^2)^{-r(\mathbf{h})}$ and using the geometric condition (2.1.2) we obtain a contribution bounded by

$$\frac{RH^\varepsilon}{H^n}\sum_{r=0}^{n}\frac{H^r}{(R^2)^{r/2}} \ll R^{1+\varepsilon}\left(\frac{1}{H^n}+\frac{1}{R^n}\right)$$

(compare this with (2.3.2)). In the second case, for each $\mathbf{h}$, there are at most R^ε choices of p and for each such pair the expression under the root is bounded by $\left(\frac{H}{p^2}\right)^{r(\mathbf{h})/2}$ so that the contribution in this case can be bounded by

$$\frac{R^\varepsilon}{H^n}\sum_{r=0}^{n}H^r\left(\frac{H}{R^2}\right)^{r/2} \ll R^\varepsilon\cdot\left(\frac{1}{H^n}+\left(\frac{H}{R^2}\right)^{n/2}\right)$$

(compare this with (2.3.4)) and choosing $H=(R^2)^{\frac{n+1}{3n}}$ we end up with the total contribution of $\ll R^{1-\frac{2(n+1)}{3}}$ from both cases together, which is satisfactory as soon as $n > 8$. □

We are now ready to prove Theorem 2.2.3:

Proof of Theorem 2.2.3. By Lemma 2.5.1 and the arguments from the previous discussion, it suffices to prove that

$$\sum_{p\sim R}\max_{(a;p)=1}\left|\frac{S(p,a)}{p^n}\right|^2 \ll R^{-1-\delta}$$

for all choices of $R \geq 1$. Using Lemma 2.3.2, the LHS is bounded by

$$\frac{1}{H^n}\sum_{\mathbf{h}}\sum_{p\sim R}p^{-\frac{r_p(\mathbf{h})}{2}} \ll \frac{1}{H^n}\sum_{r=0}^{n}R^{-\frac{r}{2}}\#\{\mathbf{h}\leq H, p\sim R : r_p(\mathbf{h})=r\}.$$

Assuming Conjecture 2.2.2, this can be further estimated as

$$\ll \frac{1}{H^n}\sum_{r=0}^{n}R^{-r/2}H^rR^{1+\varepsilon} \ll \frac{R^{1+\varepsilon}}{H^n}+\frac{R^{1+\varepsilon}}{R^{n/2}}.$$

Choosing e.g. $H = R^{1/2}$ we see that this is satisfactory as soon as $n > 8$. □

Indeed, as can be seen from the above proof, only something weaker than Conjecture 2.2.2 is actually required. However, we do believe that this is the 'right' way to put the conjecture, as the proposed upper bound is exactly the contribution that we a priori get from the terms with $r(\mathbf{h}) = r$ and p arbitrary, using the geometric condition (2.1.2).

We close this section by a few more remarks regarding Conjecture 2.2.2. To start with, the cases $r = 0$ and $r = n$ are easy to establish. Moreover, we can also prove the case $r = n - 1$: Those $\mathbf{h}$ with $r(\mathbf{h}) = n - 1$ produce a satisfactory contribution by the geometric condition (2.1.2), as explained above. On the other hand, there can be only $\mathcal{O}(H^{n+\varepsilon})$ pairs $(\mathbf{h}, p)$ with $r(\mathbf{h}) = n$ and $r_p(\mathbf{h}) = n - 1$ as then $p \mid D(\mathbf{h})$ and so p is determined by $\mathbf{h}$ up to at most H^ε many choices.

2.6 A new proof of Davenport's Shrinking Lemma

In previous work on general cubic forms, a crucial tool for dealing with the bilinear counting problems as seen in (2.3.1) as well as more general versions for the Weyl sums was the following result of Davenport, also known as the Shrinking Lemma.

Lemma 2.6.1 (Davenport's Shrinking Lemma)**.** *Let $L = (L_1, \ldots, L_n) \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Let $P \geq 1$ and $0 < Z < 1$ be real numbers. Then*

$$\begin{aligned} &\#\left\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| \leq P, \|L_i(\mathbf{x})\| < \frac{1}{2nP} \forall i\right\} \\ &\leq \left(\frac{4}{Z}\right)^n \cdot \#\left\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| \leq ZP, \|L_i(\mathbf{x})\| < \frac{Z}{2nP} \forall i\right\}. \end{aligned}$$

Here, $\|z\|$ denotes the distance of z to the nearest integer.

In only dealing with the Gauß sums in the above discussion we were able to circumvent the use of the lemma, using Lemma 2.3.1 and the Pigeonhole Principle as a substitute, but for the Weyl sums it remains an essential ingredient. Since the only previous proof uses rather intricate tools from the geometry of numbers, it is therefore desirable to present a short and elementary proof which we do in this section.

Proof of Lemma 2.6.1. We begin by choosing a prime q such that $\frac{2}{Z} \leq q \leq \frac{4}{Z}$ which is always possible. Then it will suffice to prove that

$$\begin{aligned}&\#\left\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| \leq P, \|L_i(\mathbf{x})\| < \frac{1}{2nP} \forall i\right\}\\ \leq &q^n \cdot \#\left\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| \leq \frac{2P}{q}, \|L_i(\mathbf{x})\| < \frac{1}{nqP} \forall i\right\}.\end{aligned}$$

Denote by $[z]$ the nearest integer to z. For each $(\mathbf{a}, \mathbf{b}) \in (\mathbb{Z}/q\mathbb{Z})^2$ let

$$N_{\mathbf{a},\mathbf{b}} = \#\left\{|\mathbf{x}| \leq P, \|L_i(\mathbf{x})\| < \frac{1}{2nP}, \mathbf{x} \equiv \mathbf{a} \pmod{q}, ([L_i(\mathbf{x})])_i \equiv \mathbf{b} \pmod{q}\right\}.$$

Clearly the LHS of our inequality now decomposes as $\sum_{\mathbf{a},\mathbf{b}} N_{\mathbf{a},\mathbf{b}}$. Now observe that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are counted by $N_{\mathbf{a},\mathbf{b}}$, then $\mathbf{x} := \frac{\mathbf{x}_2 - \mathbf{x}_1}{q}$ is counted by the RHS of our inequality. Hence it follows that

$$N_{\mathbf{a},\mathbf{b}} \leq \#\left\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| \leq \frac{2P}{q}, \|L_i(\mathbf{x})\| < \frac{1}{nqP} \forall i\right\}$$

which is already enough to deduce our claimed inequality with a factor of q^{2n} instead of q^n, since there are q^{2n} choices of $(\mathbf{a}, \mathbf{b})$.

To conclude the stronger claim, it will thus suffice to show that $N_{\mathbf{a},\mathbf{b}} \neq 0$ only for at most q^n choices of $(\mathbf{a}, \mathbf{b})$. Indeed, this will follow immediately if we can show that the $2n \times 2n$ matrix with columns $(\mathbf{x}, [L_i(\mathbf{x})])$ for $\mathbf{x}$ counted by the LHS of our inequality has rank at most n.

However, note that by our estimate on $\|L_i(\mathbf{x})\|$ and the symmetry of L we have

$$\mathbf{y} \cdot ([L_i(\mathbf{x})])_i = \mathbf{x} \cdot ([L_i(\mathbf{y})])_i$$

for all $\mathbf{x}$ and $\mathbf{y}$ counted, since both sides are integers and differ by less than $2n \cdot P \cdot \frac{1}{2nP} = 1$.

Hence, if we add to our matrix the columns $(-[L_i(\mathbf{x})], \mathbf{x})$ each column of the new part will be orthogonal to each column of the old part, and since they both have the same rank, both parts can have rank at most n, as desired. $\square$

Chapter 3

Cubic forms over imaginary quadratic number fields and rational lines on cubic hypersurfaces

3.1 Introduction

The study of integer solutions to polynomial equations is one of the most fundamental mathematical problems. Quadratic forms are very well understood but the situation already becomes much more difficult when studying cubic equations. A cubic form $C(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_s]$ is a homogeneous polynomial of degree 3. We say that C represents zero non-trivially if there is a vector $\mathbf{x} \in \mathbb{Z}^s \backslash \{\mathbf{0}\}$ such that $C(\mathbf{x}) = 0$. Lewis [20] and Birch [1] both independently showed that every cubic form in sufficiently many variables represents zero non-trivially.

Using the Hardy-Littlewood circle method, Davenport [11] showed that it suffices to assume $s \geq 32$ in order to show that C represents zero non-trivially, which he then improved to $s \geq 16$ in a series of papers [9, 10]. The current state of the art is due to Heath-Brown [16] who showed that 14 variables suffice.

The best one can hope for is that every cubic forms in at least 10 variables represents zero non-trivially since there exist cubic forms in 9 variables, which do not have non-trivial p-adic solutions and hence also do not represent zero non-trivially over the integers.

More is known when the cubic form is assumed to be non-singular. In this case Heath-Brown [17] showed that if $s \geq 10$ then the cubic form represents zero non-trivially, and Hooley [18] established the Hasse Principle if $s \geq 9$. That is, he showed that if a non-singular cubic form over $\mathbb{Q}$ in at least nine variables has a non-trivial p-adic solution for every p and a non-trivial real solution then it also represents zero non-trivially over the rational numbers.

One may also consider these problems for cubic forms over a number field $K/\mathbb{Q}$. In fact the above mentioned result by Lewis was proved for any number field $K/\mathbb{Q}$. Using the circle method the number of variables required was reduced to 54 by Ramanujam [25], which was subsequently improved to 17 variables by Ryavec [26] and 16 variables by Pleasants [24]. If one assumes the cubic form to be non-singular then recent work by Browning–Vishe [6] shows that ten variables suffice in order to infer the existence of a non-trivial zero, which improves previous work by Skinner [27].

The main result of this paper is the following.

Theorem 3.1.1. *Let $K/\mathbb{Q}$ be an imaginary quadratic number field. If $C(\mathbf{x})$ is a homogeneous cubic form over K in at least 14 variables then $C(\mathbf{x})$ represents zero nontrivially.*

It seems likely that our result should remain true for general number fields, however there are two serious obstructions in generalizing Heath-Brown's ideas to the number field setting, as we discuss in the course of our proof.

We are able to remove these difficulties only in the special case of imaginary quadratic number fields.

Our result has some interesting applications to problems that do not involve, prima facie, any number fields. The first of these concerns rational lines on cubic hypersurfaces.

Theorem 3.1.2. *Let C be a cubic form in $s \geq 33$ variables with rational coefficients. Then the projective cubic hypersurface defined by $C(\mathbf{x}) = 0$ contains a rational projective line.*

This improves on work of Wooley [29] who had the same result under the assumption $s \geq 37$. We note that another two variables can be saved using ideas from forthcoming work by Brandes and Dietmann [3], thus leading to a result for $s \geq 31$ variables. More specifically, while our argument (building on Wooley's) only requires Theorem 3.1.1 for one imaginary quadratic number field (e.g. $\mathbb{Q}(i)$), the full generality of Theorem 3.1.1 is required in the argument of Brandes and Dietmann.

It is also worth mentioning that in a different paper of the same authors [2], the result for $s \geq 31$ variables is already established under the assumption that the underlying hypersurface is nonsingular.

Based on an observation of Brüdern–Dietmann–Liu–Wooley [7], the existence of rational lines can be used in conjunction with the Green–Tao Theorem to produce almost prime solutions to cubic forms as follows:

Theorem 3.1.3. *Let C be a cubic form in $s \geq 33$ variables with rational coefficients. Then there are almost prime solutions to $C(\mathbf{x}) = 0$ in the following sense: There are coprime integers $c_1, \ldots, c_s$ such that the equation*

$$C(c_1p_1, c_2p_2, \ldots, c_sp_s) = 0$$

has infinitely many solutions in primes $p_1, \ldots, p_s$, not all equal.

We note that one can obtain the same result for $s \geq 31$, assuming the corresponding version of Theorem 3.1.2.

For comparison, the existence of prime solutions is only known for non-singular cubic forms and under the assumption of a much larger number of variables, cf. the work of Yamagishi [30] and Liu–Zhao [22]. These authors require 8996 and 9216 variables, respectively, in the case of cubic forms.

Notation

We use $e(\alpha) = e^{2\pi i\alpha}$ and the notation $O(\dots)$ and $\ll$ of Landau and Vinogradov, respectively. All implied constants are allowed to depend on the number field K, a choice of integral basis Ω for K, the cubic form C and a small parameter $\varepsilon > 0$ whenever it appears.

As is convenient in analytic number theory, this parameter ε may change its value finitely many times. In particular, we may write something like $M^{2\varepsilon} \ll M^{\varepsilon}$.

We often use the notation $q \sim R$ to denote the dyadic condition $R < q \leq 2R$.

3.2 Deduction of Theorems 3.1.2 and 3.1.3

In this section, we give the proofs of Theorems 3.1.2 and 3.1.3 assuming Theorem 3.1.1.

We begin with the observation that the existence of a rational line on the cubic hypersurface defined by C is equivalent to the existence of linearly independent vectors $\mathbf{v}$ and $\mathbf{w}$ such that $C(\mathbf{v} + t\mathbf{w}) = 0$ identically in t. Expanding this formally as a cubic polynomial in t, we obtain the equivalent

$$C(\mathbf{v}) + tQ_{\mathbf{w}}(\mathbf{v}) + t^2 L_{\mathbf{w}}(\mathbf{v}) + t^3 C(\mathbf{w}) = 0$$

for certain quadratic resp. linear forms $Q_{\mathbf{w}}$ and $L_{\mathbf{w}}$ depending on $\mathbf{w}$. We therefore need to find linearly independent $\mathbf{v}$ and $\mathbf{w}$ such that

$$C(\mathbf{v}) = Q_{\mathbf{w}}(\mathbf{v}) = L_{\mathbf{w}}(\mathbf{v}) = C(\mathbf{w}) = 0.$$

If we start by choosing a solution $\mathbf{w} \neq 0$ of $C(\mathbf{w}) = 0$, the linear equation $L_{\mathbf{w}}(\mathbf{v}) = 0$ and the linear independence to $\mathbf{w}$ reduce the degrees of freedom for $\mathbf{v}$ by two. We are thus looking for a solution to the system $C(\mathbf{v}) = Q_{\mathbf{w}}(\mathbf{v}) = 0$ of one cubic and one quadratic equation in $s-2$ variables. If we knew that the signature of the quadratic form $Q_{\mathbf{w}}$ was sufficiently indefinite, we could infer the existence of a sufficiently large linear space on which $Q_{\mathbf{w}}$ vanishes, leaving us with a single cubic form in many variables, that can be dealt with by the work of Heath-Brown[16].

The crux however is that it is in general hard to control the signature of $Q_{\mathbf{w}}$. Instead we avoid the indefiniteness issue by passing to an imaginary quadratic number field of $\mathbb{Q}$, thus requiring our Theorem 3.1.1.

We now present the complete argument in order: We begin by choosing a $\mathbf{w} \in \mathbb{Q}^s \backslash \{\mathbf{0}\}$ such that $C(\mathbf{w}) = 0$, which exists by the work of Heath-Brown. Letting $K/\mathbb{Q}$ be any imaginary quadratic number field, we next show the existence of a vector $\mathbf{v} \in K^s$, linearly independent to $\mathbf{w}$ and satisfying

$$C(\mathbf{v}) = Q_{\mathbf{w}}(\mathbf{v}) = L_{\mathbf{w}}(\mathbf{v}) = 0.$$

To this end, we use that a hypersurface $Q(\mathbf{x}) = 0$ defined by a quadratic form Q in s variables contains a $\left\lfloor \frac{s-3}{2} \right\rfloor$-dimensional K-linear subspace, a fact that is easily proved by induction.

The linear space of vectors $\mathbf{v}$ orthogonal to $\mathbf{w}$ and satisfying $L_{\mathbf{w}}(\mathbf{v}) = 0$ is at least $(s-2)$-dimensional. Thus, $Q_{\mathbf{w}}$ vanishes on a linear subspace of dimension at least $\left\lfloor \frac{(s-2)-3}{2} \right\rfloor = \left\lfloor \frac{s-5}{2} \right\rfloor$. Note that by our assumption on s we have $\left\lfloor \frac{s-5}{2} \right\rfloor \geq 14$. We are then left to solve the equation $C(\mathbf{v}) = 0$ on a 14-dimensional linear space which can be done by Theorem 3.1.1.

We have thus proved that $C(\mathbf{v} + t\mathbf{w}) = 0$ identically in t for some linearly independent vectors $\mathbf{v} \in K^s$ and $\mathbf{w} \in \mathbb{Q}^s$.

By an observation of Lewis, this is enough to deduce the existence of a rational line, as we explain now, following an argument of Dietmann-Wooley [13].

Consider the K-rational spaces V spanned by $\mathbf{v}$ and $\mathbf{w}$ as well as V^* spanned by $\mathbf{v}^*$ and $\mathbf{w}$ where * denotes conjugation in K. If $\mathbf{v} \in \mathbb{Q}^s$ we are already done. Else, consider the three-dimensional space W spanned by $\mathbf{v}$, $\mathbf{v}^*$ and $\mathbf{w}$. If C vanishes on W, we are also done as W clearly contains a two-dimensional $\mathbb{Q}$-rational subspace. Else, by intersection theory the hypersurface defined by C must intersect W in a third two-dimensional K-rational subspace L. More precisely, by Theorem I.7.7 in Hartshorne [15] we have

$$i(W, C; V) + i(W, C; V^*) + \sum_j i(W, C; Z_j) \cdot \deg Z_j = (\deg W)(\deg C) = 3$$

where $i(W, C; V)$ denotes the intersection multiplicity and Z_i are the other irreducible components of $C \cap W$. Since W is invariant under conjugation, we must have $i(W, C; V) = i(W, C; V^*)$ and thus both numbers are equal to 1, implying that there is a unique third component $L = Z_1$ which is then necessarily linear. Finally, since W and C are conjugation invariant, the three spaces V, V^* and L are permuted under conjugation and thus L itself is conjugation invariant, i.e. describes the desired rational line. □

We remark that the use of intersection theory in the previous argument can be replaced by an explicit algebraic computation, as shown in Wooley [29].

To deduce Theorem 3.1.3, we follow the strategy in [7]. In particular, we show that the existence of a rational line implies the existence of almost prime solutions, regardless of the number of variables. We thus assume that for some linearly independent vectors $\mathbf{a}, \mathbf{b} \in \mathbb{Z}^s$, we have $C(\mathbf{a}t + \mathbf{b}u) = 0$ identically in t and u. If $a_i = b_i = 0$ for some i, then we can set $c_i = 1$ and continue to work with the other variables. By taking a suitable linear combination, we can then assume that indeed all a_i and b_i are different from

0. Rescaling u by a factor of $a_1a_2\dots a_s$ and then rescaling the variables by a factor of a_i (thereby changing c_i by a factor of a_i), we may even assume that all the a_i are equal to 1, i.e.

$$C(t+b_1u, t+b_2u, \dots, t+b_nu) = 0$$

identically in t and u. By the Green–Tao Theorem [14], the primes contain infinitely many arithmetic progressions of length $2M+1$ where $M = 2\max_i |b_i| + 1$, i.e. there are infinitely many pairs (ℓ, d) such that $\ell + kd$ is prime for all $|k| \leq M$. Choosing $t = \ell$ and $u = k$ then yields the desired result with $c_i = 1$. □

3.3 Algebraic Preliminaries

While our main result is proved only for imaginary quadratic number fields we will introduce the matter in a general fashion and not restrict ourselves to these fields for now. We will aim to highlight whenever phenomena occur that set apart the situation for imaginary quadratic number fields from a general setting. In particular, even when $K/\mathbb{Q}$ is an imaginary quadratic number field we still sometimes prefer to write $n = [K : \mathbb{Q}]$.

Let K be a number field of degree n over $\mathbb{Q}$ and denote by $\mathcal{O}$ its ring of integers.

Define the $\mathbb{R}$-vector space $K_\mathbb{R} \coloneqq K \otimes_\mathbb{Q} \mathbb{R}$ and note that we have natural embeddings $\mathcal{O} \subset K \subset K_\mathbb{R}$. The space $K_\mathbb{R}$ is sometimes referred to as the *Minkowski space* of K. Note that there exist integers n_1 and n_2 with $n_1 + n_2 = n$ such that K admits n_1 real embeddings $\sigma_1, \dots, \sigma_{n_1}$ and $2n_2$ complex embeddings $\sigma_{n_1+1}, \overline{\sigma}_{n_1+1}, \dots, \sigma_{n_1+n_2}, \overline{\sigma}_{n_1+n_2}$ so that $K_\mathbb{R} \cong \mathbb{R}^{n_1} \times \mathbb{C}^{n_2}$.

Denote by π_i the projection from $V_\mathbb{R} \cong \mathbb{R}^{n_1} \times \mathbb{C}^{n_2}$ to the i-th coordinate, which may take real or complex values. We define the trace map $\operatorname{tr}\colon K_\mathbb{R} \to \mathbb{R}$ and norm map Norm: $K_{\mathbb{R}} \to \mathbb{R}$ as

$$\operatorname{tr}(\alpha) = \sum_{i=1}^{n_1} \pi_i(\alpha) + \sum_{i=n_1+1}^{n_2} \operatorname{Re}(\pi_i(\alpha)),$$

and

$$\operatorname{Norm}(\alpha) = \prod_{i=1}^{n_1} |\pi_i(\alpha)| \prod_{i=n_1+1}^{n_2} |\pi_i(\alpha)|^2 ,$$

respectively. If $\alpha \in K$ then these are just the usual norm and trace function from algebraic number theory.

Pick a basis $\Omega = \{\omega_1, \ldots, \omega_n\}$ of $\mathcal{O}$. Any element $\alpha \in K_{\mathbb{R}}$ may be expressed in the form $\alpha = \sum_{j=1}^{n} \alpha_j \omega_j$ for some $\alpha_j \in \mathbb{R}$. For such α we define a height

$$|\alpha| \coloneqq \max_j |\alpha_j|.$$

Note that this depends on the choice of basis Ω for $\mathcal{O}$. Given a vector $\boldsymbol{\alpha} = (\alpha^{(1)}, \ldots, \alpha^{(s)}) \in K_{\mathbb{R}}^s$ we further denote

$$|\boldsymbol{\alpha}| \coloneqq \max_k |\alpha^{(k)}|.$$

We may alternatively define another height on $K_{\mathbb{R}}$ given by

$$|\alpha|_K \coloneqq \max_p |\pi_p(\alpha)| .$$

As noted by Pleasants [24, Section 2] we have

$$|\alpha| \asymp |\alpha|_K,$$

for all $\alpha \in K_{\mathbb{R}}$. If $\alpha, \beta \in K_{\mathbb{R}}$ then it is easy to see that this height satisfies

$$\begin{aligned} |\alpha\beta|_K &\leq |\alpha|_K |\beta|_K, \\ |\alpha + \beta|_K &\leq |\alpha|_K + |\beta|_K \\ |\alpha^{-1}|_K &\leq \frac{|\alpha|_K^{n-1}}{\operatorname{Norm}(\alpha)}. \end{aligned}$$

The same inequalities therefore hold for $|\cdot|$ if we replace the symbols $\leq$ by $\ll_K$. It would be desirable to have the last inequality in the form $|\alpha^{-1}| \asymp |\alpha|^{-1}$ which would result if $\mathrm{Norm}(\alpha) \asymp |\alpha|^n$. However, if α is a unit in $\mathcal{O}$ then $\mathrm{Norm}(\alpha) = 1$ while the height $|\alpha|$ may be unbounded, at least whenever K is not an imaginary quadratic number field. This is one of the points where our argument crucially depends on the latter assumption.

If $K = \mathbb{Q}(\sqrt{-d})$ is an imaginary quadratic number field then, depending on the value of the residue class of d mod 4 we can choose $\{1, \sqrt{-d}\}$ or $\{1, (1+\sqrt{-d})/2\}$ as an integral basis for $\mathcal{O}$. We thus find that

$$\mathrm{Norm}(\alpha) \asymp |\alpha|^2.$$

In particular we find

$$|\alpha^{-1}| \asymp |\alpha|^{-1}.$$

Given an ideal $J \subset \mathcal{O}$ we recall that $\mathcal{O}/J$ is finite and we define as usual the norm of the ideal to be

$$N(J) := \#\left(\mathcal{O}/J\right).$$

For a fractional ideal of K this norm is, as usual, extended multiplicatively using the unique factorization into prime ideals inside K. Given $\gamma \in K$ we further define the *denominator ideal* of γ as

$$\mathfrak{a}_\gamma := \{x \in \mathcal{O} \colon x\gamma \in \mathcal{O}\}.$$

As the name suggests, and this is not very difficult to verify, $\mathfrak{a}_\gamma$ is an ideal inside $\mathcal{O}$, contained in the fractional ideal $(\gamma)^{-1}$. We will need the following fact several times.

Lemma 3.3.1. *Let $J \subset \mathcal{O}$ be an ideal. Then there are at most $N(J)$ different elements $\gamma \in K/\mathcal{O}$ such that $\mathfrak{a}_\gamma = J$.*

Proof. To see this, note first that for any two fractional ideals $\mathfrak{b}, \mathfrak{c} \subset K$ with $\mathfrak{b} \supset \mathfrak{c}$ there exists some $d \in \mathcal{O}$ such that $d\mathfrak{b}, d\mathfrak{c} \subset \mathcal{O}$. Thus

$$[\mathfrak{b} \colon \mathfrak{c}] = [d\mathfrak{b} \colon d\mathfrak{c}] = \frac{[\mathcal{O} \colon d\mathfrak{c}]}{[\mathcal{O} \colon d\mathfrak{b}]} = N(d\mathfrak{c})/N(d\mathfrak{b}) = N(\mathfrak{c})/N(\mathfrak{b}).$$

Now note that if $\mathfrak{a}_\gamma = J$ we must have $\gamma \in J^{-1}\mathcal{O}$, where

$$J^{-1} = \{x \in K \colon xJ \subset \mathcal{O}\}.$$

But now $[J^{-1}\mathcal{O} \colon \mathcal{O}] = N(J)$ and so the result follows. □

We shall further require a version of Dirichlet's approximation theorem.

Lemma 3.3.2. *Let $K/\mathbb{Q}$ be a number field of degree n. Let $\alpha \in K_\mathbb{R}$ and let $Q \geq 1$. Then there exist some $a, q \in \mathcal{O}$ with $1 \leq |q| \leq Q$ such that*

$$|q\alpha - a| \leq \frac{1}{Q}. \tag{3.3.1}$$

Proof. Consider the set $\mathcal{Q}$ of algebraic integers given by

$$\mathcal{Q} = \left\{ \sum_j q_j \omega_j \in \mathcal{O} \colon 0 \leq q_j \leq Q \right\}.$$

For any $q \in \mathcal{Q}$ we may express $q\alpha$ as

$$q\alpha = a_q + x_q,$$

where $a_q \in \mathcal{O}$ and $x_q = \sum_j x_{q,j}\omega_j$ such that $0 \leq x_{q,j} < 1$ for $j = 1, \ldots, n$. We may partition $K_\mathbb{R}/\mathcal{O} = \left\{ \sum_j x_j \omega_j \colon 0 \leq x_j < 1 \right\}$ into Q^n boxes such that the height of the difference of two points in the same box is bounded by $1/Q$. Since $\mathcal{Q}$ has $(Q+1)^n$ elements, by the pigeonhole principle there must be $q_1, q_2 \in \mathcal{Q}$ such that x_{q_1} and x_{q_2} lie in the same box according to the partition above. Therefore we find

$$|(q_1 - q_2)\alpha - (a_{q_1} - a_{q_2})| = |x_{q_1} - x_{q_2}| \leq 1/Q.$$

Taking $q = q_1 - q_2$ and $a = a_{q_1} - a_{q_2}$ delivers the result. □

For the application to the mean-square averaging method introduced by Heath-Brown, we need a fractional form of Dirichlet's theorem. We are only able to obtain a satisfactory version for imaginary quadratic number fields, this being the first of the obstructions regarding possible generalizations mentioned in the introduction. Note that this is special to Heath-Brown's method and hence was not an issue in the work of Ramanujam, Ryavec and Pleasants.

Lemma 3.3.3. *Let $K/\mathbb{Q}$ be an imaginary quadratic number field (in particular $n = 2$). Let $\alpha \in K_{\mathbb{R}}$ and let $Q \geq 1$. Then there exists some $\gamma \in K$ with $N(\mathfrak{a}_\gamma) \leq Q^n$ such that*

$$|\alpha - \gamma| \ll \frac{1}{N(\mathfrak{a}_\gamma)^{\frac{1}{n}} Q}. \tag{3.3.2}$$

Proof. From Lemma 3.3.3 we find that there exist $a, q \in \mathcal{O}$ with $|q| \leq Q$ such that

$$|q\alpha - a| \leq 1/Q.$$

Set $\gamma = a/q \in K$ and note that $(q) \subseteq \mathfrak{a}_\gamma$. In particular from this it follows that

$$N(\mathfrak{a}_\gamma) \leq N((q)) = \mathrm{Norm}(q) \asymp |q|^n,$$

where the last estimate is true since K is an imaginary quadratic number field. Thus

$$|q|^{-1} \ll N(\mathfrak{a}_\gamma)^{-1/n},$$

and so we obtain

$$|\alpha - \gamma| \ll |q|^{-1} |q\alpha - a| \ll \frac{1}{N(\mathfrak{a}_\gamma)^{\frac{1}{n}} Q},$$

as desired. □

We shall sometimes require the following easy lemma.

Lemma 3.3.4. *Let $J \subset \mathcal{O}$ be an ideal. Then there exist constants c_1, c_2 only depending on K such that for any non-zero $g \in J$ we have*

$$c_1 N(J)^{1/n} \leq |g|,$$

and we may always find a non-zero element $a \in J$ such that

$$|a| \leq c_2 N(J)^{1/n}.$$

Proof. First note that if $g \in J$ then $(g) \subset J$ and therefore

$$N(J) \leq N((g)) = \mathrm{Norm}(g) \ll |g|^n.$$

For the second inequality note that there are at least $N(J)+1$ algebraic integers whose height does not exceed $N(J)^{1/n}$. By definition $N(J) = \#(\mathcal{O}/J)$ and hence at least two of these integers must lie in the same residue class modulo J. Their difference is therefore an algebraic integer $a \in J$ with $|a| \leq 2N(J)^{1/n}$. □

Finally we will also need the following.

Lemma 3.3.5. *Let $K/\mathbb{Q}$ be a number field and let Δ be the discriminant of this extension. Let $\alpha \in K_{\mathbb{R}}$ and assume that $\{\omega_i\}_i$ is an integral basis for $\mathcal{O}$. If*

$$\Delta^{-1}\mathrm{tr}(\alpha\omega_i) \in \mathbb{Z}$$

holds for all $i = 1, \ldots, n$ then $\alpha \in \mathcal{O}$.

Proof. Write $\alpha = \sum_{j=1}^{n} \alpha_j \omega_j$, where $\alpha_j \in \mathbb{R}$. Due to the additivity of the trace we have

$$\mathrm{tr}(\alpha\omega_i) = \sum_{j=1}^{n} \alpha_j \mathrm{tr}(\omega_i\omega_j).$$

Denote by $\mathbf{T}$ the trace form, that is, the $n \times n$ matrix with entries $\mathrm{tr}(\omega_i\omega_j)$. Then if we identify $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n$, the assumption of the lemma is equivalent to

$$\Delta^{-1}\mathbf{T}(\alpha) \in \mathbb{Z}^n.$$

By definition $\det \mathbf{T} = \Delta$. Hence $\mathbf{T}' := \Delta \mathbf{T}^{-1}$ has integer entries. Combining this with our previous observation yields

$$\alpha = \mathbf{T}^{-1}\mathbf{T}(\alpha) = \mathbf{T}'(\Delta^{-1}\mathbf{T}(\alpha)) \in \mathbb{Z}^n.$$

Hence $\alpha \in \mathcal{O}$ as required. □

3.4 The Dichotomy

Let $C \in \mathcal{O}[x_1, \ldots, x_s]$ be a homogeneous cubic form. Our goal is to show that there always exists a non-trivial solution to $C = 0$ over K provided $s \geq 14$ and K is an imaginary quadratic number field. We follow the strategy of Davenport that was later refined by Heath-Brown [16]: Either C represents zero non-trivially for 'geometric reasons', or we can establish an asymptotic formula for the number of solutions of bounded height, using the circle method.

3.4.1 Davenport's Geometric Condition

We may express $C(\mathbf{x})$ as

$$C(\mathbf{x}) = \sum_{i,j,k} c_{ijk} x_i x_j x_k,$$

where the coefficients c_{ijk} are fully symmetric in the indices and lie in $\mathcal{O}$, after replacing $C(\mathbf{x})$ by $6C(\mathbf{x})$ if required. For $i = 1, \ldots, s$ further define the bilinear forms $B_i(\mathbf{x}, \mathbf{y})$ by

$$B_i(\mathbf{x}, \mathbf{y}) = \sum_{j,k} c_{ijk} x_j y_k.$$

Finally, we also consider an $s \times s$ matrix $M(\mathbf{x})$, the *Hessian* of $C(\mathbf{x})$, whose entries are defined by

$$M(\mathbf{x})_{jk} = \sum_{i} c_{ijk} x_i,$$

so that

$$(M(\mathbf{x})\mathbf{y})_i = B_i(\mathbf{x}, \mathbf{y}).$$

We note that the entries are linear forms in the variables $\mathbf{x}$. Denote the rank of the matrix by

$$r(\mathbf{x}) = \operatorname{rank}(M(\mathbf{x})).$$

As in Davenport's and Heath-Brown's work we obtain a dichotomy.

Lemma 3.4.1. *One of the following two alternatives holds.*

1. *Davenport's Geometric Condition: For every integer* $0 \leq r \leq s$ *we have*

$$\#\{\mathbf{x} \in \mathcal{O}^s \colon |\mathbf{x}| < H,\ r(\mathbf{x}) = r\} \ll H^{nr}. \tag{3.4.1}$$

2. *The cubic form* $C(\mathbf{x})$ *has a non-trivial zero in* $\mathcal{O}$.

Proof. Consider the least integer $h = h(C)$ such that the cubic form may be written as

$$C(\mathbf{x}) = \sum_{i=1}^{h} L_i(\mathbf{x}) Q_i(\mathbf{x}),$$

where L_i are linear and Q_i are quadratic forms defined over K. This is the h-invariant of C. It is easy to see that $1 \leq h \leq s$ holds, and that $C(\mathbf{x}) = 0$ has a non-trivial solution over K if and only if $h < s$.

We will show that if $h = s$ then alternative (1) holds. In fact, Pleasants [24, Lemma 3.5] showed that the number of points $\mathbf{x} \in \mathcal{O}^s$ such that $|\mathbf{x}| < H$ holds, for which the equations $B_i(\mathbf{x}, \mathbf{y}) = 0$, $j = 1, \dots, s$ have exactly $s - r$ linearly independent solutions $\mathbf{y}$ is bounded by $O(H^{n(s-h+r)})$. Hence taking $h = s$ delivers the desired bound (3.4.1). □

We will henceforth assume that Davenport's Geometric Condition (3.4.1) is satisfied and apply the circle method. In particular as in [16] this condition implies that we have

$$\#\{\mathbf{x}, \mathbf{y} \in \mathcal{O}^s \colon |\mathbf{x}|, |\mathbf{y}| < H, B_i(\mathbf{x}, \mathbf{y}) = 0, \forall i\} \ll H^{ns}, \tag{3.4.2}$$

for any $H \geq 1$.

3.4.2 The Circle Method

Let $\mathcal{B} \subset K_{\mathbb{R}}^s \cong \mathbb{R}^{ns}$ be a box of the form

$$\mathcal{B} = \left\{ \left(\sum_j \alpha_{ij}\omega_j \right)_i \in K_{\mathbb{R}}^s \colon b_{ij}^- \leq \alpha_{ij} \leq b_{ij}^+ \right\},$$

where $b_{ij}^- < b_{ij}^+$ are some real numbers . For $P \geq 1$ consider the counting function

$$N(P;\mathcal{B}) = N(P) = \{\mathbf{x} \in P\mathcal{B} \cap \mathcal{O}^s \colon C(\mathbf{x}) = 0\}.$$

For $\alpha \in K_{\mathbb{R}}$ and $P \geq 1$ we define the exponential sum

$$S(\alpha) = S(\alpha; P) = \sum_{\mathbf{x} \in P\mathcal{B} \cap \mathcal{O}^s} e\left(\operatorname{tr}(\alpha C(\mathbf{x}))\right).$$

Denote by $I \subset K_{\mathbb{R}}$ the set given by

$$I = \left\{ \sum_{j=1}^n \alpha_j \omega_j \colon 0 \leq \alpha_j \leq 1 \right\},$$

which may also be regarded as $K_{\mathbb{R}}/\mathcal{O}$. Due to orthogonality of characters we obtain

$$N(P) = \int_{\alpha \in I} S(\alpha) d\alpha.$$

We are now able to state the main technical theorem of our paper.

Theorem 3.4.2. *Let $K/\mathbb{Q}$ be an imaginary quadratic number field and let $C(\mathbf{x})$ be a cubic form in $s \geq 14$ variables over K. Suppose that $C(\mathbf{x})$ is irreducible over K and that Davenport's Geometric Condition* (3.4.1) *is satisfied. Then we have the asymptotic formula*

$$N(P) = \sigma P^{n(s-3)} + o\left(P^{n(s-3)}\right), \quad \textit{as} \quad P \to \infty,$$

where $\sigma > 0$ is the product of the usual singular integral and singular series.

Thus Theorem 3.1.1 follows directly from Lemma 3.4.1 and Theorem 3.4.2 where we also note that a reducible cubic form always contains a linear factor over K and therefore has a non-trivial solution for obvious reasons.

3.4.3 The major arcs

For this section we do not need to assume that K is an imaginary quadratic number field of $\mathbb{Q}$. As in Pleasants, we choose as center of our box $\mathcal{B} = \mathcal{B}(\mathbf{z})$ a solution $\mathbf{z} \in K_{\mathbb{R}}$ of $C(\mathbf{z}) = 0$ satisfying $\frac{\partial C}{\partial x_1}(\mathbf{z}) \neq 0$ and $z_1, \ldots, z_n \neq 0$. Such a vector $\mathbf{z}$ always exists by [24, Lemma 7.2] provided C is irreducible.
Let $\gamma \in K/\mathcal{O}$ and define

$$\mathfrak{M}_\gamma \coloneqq \left\{ \alpha \in I \colon |\alpha - \gamma| < P^{-3+\nu} \right\},$$

where we regard $I = K_{\mathbb{R}}/\mathcal{O}$. We define the *major arcs* as

$$\mathfrak{M} = \bigcup_{\substack{\gamma \in K/\mathcal{O} \\ N(\mathfrak{a}_\gamma) \leq P^\nu}} \mathfrak{M}_\gamma,$$

and the *minor arcs* as

$$\mathfrak{m} = I \setminus \mathfrak{M}.$$

Further, define the sum S_γ via

$$S_\gamma = \sum_{\mathbf{x} \bmod N(\mathfrak{a}_\gamma)} e(\operatorname{tr}(\gamma C(\mathbf{x})).$$

Given a parameter $R \geq 1$ we define the *truncated singular series* to be

$$\mathfrak{S}(R) \coloneqq \sum_{\substack{\gamma \in K/\mathcal{O} \\ N(\mathfrak{a}_\gamma) \leq R}} N(\mathfrak{a}_\gamma)^{-ns} S_\gamma,$$

and the *truncated singular integral* to be

$$\mathfrak{J}(R) \coloneqq \int_{|\zeta| < R^\nu} \int_{\mathcal{B}} e(\operatorname{tr}(\zeta R^{-3} C(R\boldsymbol{\xi})))d\boldsymbol{\xi} d\zeta.$$

Pleasants [24, Lemma 7.1] shows that if $\nu < \frac{1}{n+4}$ is satisfied then we have

$$\int_{\mathfrak{M}} S(\alpha) d\alpha = \mathfrak{S}(P^\nu)\mathfrak{J}(P)P^{n(s-3)} + o(P^{n(s-3)}).$$

Moreover, if $\mathcal{B} = \mathcal{B}(\mathbf{z})$ is the box as in the beginning of the section, provided that the sidelengths of the boxes are sufficiently small, and if $C(\mathbf{x})$ is irreducible over K then Pleasants [24, Lemma 7.2] further shows that $\mathfrak{I}(R)$ converges absolutely to a positive number $\mathfrak{I}$.

We remark that Lemma 7.2 in [24] was originally stated under the weaker assumption that $C(\mathbf{x})$ is not a rational multiple of a cube of a linear form. His proof relies on a result by Davenport [11, Lemma 6.2], which assumes the existence of a non-singular, real solution $\boldsymbol{\xi} \in \mathbb{R}^n$ of a rational cubic form G such that

$$\frac{\partial G}{\partial x_i}(\boldsymbol{\xi}) \neq 0, \quad \xi_i \neq 0,$$

holds for some i. In particular Pleasants writes that *"this hypothesis is not used in the proof of the lemma, however, and in any case the argument that follows could easily be adapted to provide it"*. While one can always find $\boldsymbol{\xi} \in \mathbb{R}^s$ with $\frac{\partial G}{\partial x_i}(\boldsymbol{\xi}) \neq 0$ unless G is a rational multiple of a cube of a linear form, one can not necessarily ensure that $\xi_i \neq 0$ for the same index i. Consider for example $G(x_1, \dots, x_n) = x_1(x_2^2 + \cdots + x_n^2)$. It is possible that Davenport's result [11, Lemma 6.2] holds nevertheless in this generality but at least the standard method of establishing bounded variation of the auxiliary function involved in the proof by showing the existence of right and left derivatives, see for example [8, Lemma 16.1], fails in general.

The singular series $\mathfrak{S}(R)$ may or may not converge absolutely as $R \to \infty$. If it does converge, then provided non-singular $\mathfrak{p}$-adic solutions of $C(\mathbf{x}) = 0$ exist for all primes $\mathfrak{p}$, by standard arguments it follows that $\mathfrak{S} > 0$. See for example the proof of Lemma 7.4 in [24], where this argumentation is carried out in our setting. Finally, Lewis [21] showed that these non-singular $\mathfrak{p}$-adic solutions always exist whenever $s \geq 10$. Therefore we obtain the following.

Theorem 3.4.3. *Let $C \in \mathcal{O}[x_1, \dots, x_s]$ be an irreducible cubic form. Assume that $s \geq 10$. If the singular series $\mathfrak{S}(R)$ converges absolutely as $R \to \infty$ then*

$$\int_{\mathfrak{M}} S(\alpha) d\alpha = \sigma P^{n(s-3)} + o(P^{n(s-3)}),$$

for some $\sigma > 0$ as $P \to \infty$.

In particular, in Section 3.7 we will establish the following.

Theorem 3.4.4. *Assume that $s \geq 13$ and that Davenport's Geometric Condition* (3.4.1) *is satisfied then the singular series converges absolutely. Therefore if $C(\boldsymbol{x})$ is irreducible we have*

$$\int_{\mathfrak{M}} S(\alpha) d\alpha = \sigma P^{n(s-3)} + o\left(P^{n(s-3)}\right),$$

for some $\sigma > 0$ as $P \to \infty$.

We remark that we show this result for any number field K.

3.5 Auxiliary Diophantine Inequalities

To bound the Weyl sum $S(\alpha)$ of a general cubic form, classical Weyl differencing leaves us with the task of examining the number of solutions to certain auxiliary Diophantine inequalities. Davenport's crucial idea was to bootstrap these inequalities using his *Shrinking Lemma*, combined with the observation that sufficiently strong Diophantine inequalities already imply divisibility or even equality.

In this section, we prepare these arguments by providing a version of this observation adapted to our setting. We are only able to show a satisfactory version of this lemma if $K/\mathbb{Q}$ is an imaginary quadratic number field, this being the second of the obstructions mentioned in the introduction.

Lemma 3.5.1. *Assume that $K/\mathbb{Q}$ is a number field and denote by Δ the discriminant of this extension. There exists a real positive constant $A > 0$ depending only on K and the choice of integral basis Ω for K such that the following statement holds.*
Let $M \geq 0$ be a real number and let $\alpha \in K_{\mathbb{R}}$. Suppose that $\alpha = \gamma + \theta$ with $\gamma \in K$ and $M|\theta|N(\mathfrak{a}_\gamma)^{1/n} \leq A$. If $m \in \mathcal{O}$ is such that $|m| \leq M$ and $\|\Delta^{-1}\mathrm{tr}(\alpha m \omega_j)\| < P_0^{-1}$ holds for all $j = 1, \ldots, n$ where $AP_0 \geq N(\mathfrak{a}_\gamma)^{1/n}$ then $m \in \mathfrak{a}_\gamma$. In particular if either of the conditions

1. *$M \leq AN(\mathfrak{a}_\gamma)^{1/n}$, or*

2. *K is an imaginary quadratic number field and $A|\theta| \geq N(\mathfrak{a}_\gamma)^{-1/n}P_0^{-1}$*

is satisfied, then we must have $m = 0$.

Proof. Note first that

$$\left\|\Delta^{-1}\mathrm{tr}(\gamma m \omega_j)\right\| \leq \left\|\Delta^{-1}\mathrm{tr}(\alpha m \omega_j)\right\| + \left\|\Delta^{-1}\mathrm{tr}(\theta m \omega_j)\right\|.$$

Now due to our assumption we have $\|\Delta^{-1}\mathrm{tr}(\alpha m \omega_j)\| < P_0^{-1}$. Further it is easy to see that

$$\Delta^{-1}|\mathrm{tr}(\theta m \omega_j)| \ll |\theta| M.$$

Therefore choosing A sufficiently small we find

$$\left\|\Delta^{-1}\mathrm{tr}(\gamma m \omega_j)\right\| < \frac{A^{1/2}}{N(\mathfrak{a}_\gamma)^{1/n}}, \tag{3.5.1}$$

for all $j = 1, \ldots, n$. As before write $\mathbf{T} = (\mathrm{tr}(\omega_i \omega_j))_{i,j}$ for the trace form. Write $\mathbf{x} \in \mathbb{R}^n$ for the real vector obtained from γm under the isomorphism $K_{\mathbb{R}} \cong \mathbb{R}^n$. Then (3.5.1) is equivalent to saying that there exist $\mathbf{a} \in \mathbb{Z}^n$ and $\mathbf{r} \in \mathbb{R}^n$ with $|\mathbf{r}| < \frac{A^{1/2}}{N(\mathfrak{a}_\gamma)^{1/n}}$ such that

$$\mathbf{T}(\Delta^{-1}\mathbf{x}) = \mathbf{a} + \mathbf{r}.$$

Recall that $\Delta\mathbf{T}^{-1}$ is an integral matrix whose entries are bounded in terms of K. Therefore

$$\mathbf{x} = \Delta\mathbf{T}^{-1}(\mathbf{a}) + \Delta\mathbf{T}^{-1}(\mathbf{r}).$$

Now $\mathbf{T}^{-1}(\mathbf{a}) \in \mathbb{Z}^n$ and

$$|\Delta\mathbf{T}^{-1}(\mathbf{r})| < \frac{A^{1/3}}{N(\mathfrak{a}_\gamma)^{1/n}},$$

after decreasing A if necessary. We thus find that

$$\gamma m = a + \rho,$$

where $a \in \mathcal{O}$ and $|\rho| < \frac{A^{1/3}}{N(\mathfrak{a}_\gamma)^{1/n}}$. By Lemma 3.3.4 there exists $g \in \mathfrak{a}_\gamma$ with $|g| \asymp N(\mathfrak{a}_\gamma)^{1/n}$. From the above equation we see that $g\rho \in \mathcal{O}$, and so, unless $\rho = 0$ we have

$$1 \leq |g\rho| < A^{1/4},$$

after decreasing A if necessary. Choosing A suitably small therefore leads to a contradiction whence we must have $\rho = 0$, and so $m \in \mathfrak{a}_\gamma$. This finishes the first part of the proof.

If we now assume that $M \leq AN(\mathfrak{a}_\gamma)^{1/n}$ is satisfied then by choosing A suitably small this implies that $m = 0$ via Lemma 3.3.4.

Finally, assume that $A|\theta| > (N(\mathfrak{a}_\gamma)^{1/n}P_0)^{-1}$ is satisfied and that K is an imaginary quadratic number field. Upon choosing A even smaller if necessary, we find that

$$\Delta^{-1}|\mathrm{tr}(\theta m\omega_j)| \leq \frac{1}{2},$$

for all $j = 1, \ldots, n$ and thus

$$\Delta^{-1}|\mathrm{tr}(\theta m\omega_j)| = \left\|\Delta^{-1}\mathrm{tr}(\theta m\omega_j)\right\| = \left\|\Delta^{-1}\mathrm{tr}(\alpha m\omega_j)\right\| < P_0^{-1},$$

for all $j = 1, \ldots, n$. Write $\mathbf{y} = (y_1, \ldots, y_n)$ for the image of θm under the isomorphism $K_\mathbb{R} \cong \mathbb{R}^n$ and let $\mathbf{T}$ be the trace form as above. The above inequality is equivalent to saying that there exists some $\mathbf{t} \in \mathbb{R}^n$ with $|\mathbf{t}| < P_0^{-1}$ such that

$$\mathbf{T}(\Delta^{-1}\mathbf{y}) = \mathbf{t}.$$

As before the inverse of $\mathbf{T}$ is a matrix with rational entries, whose absolute value is bounded by $O(1)$. Hence

$$|\mathbf{y}| = \Delta|\mathbf{T}^{-1}(\mathbf{t})| \ll |\mathbf{t}| < P_0^{-1}.$$

Further $|\mathbf{y}| = |\theta m|$, and since K is an imaginary quadratic number field we have $|\theta^{-1}| \asymp |\theta|^{-1}$ and so

$$|m| \ll (P_0|\theta|)^{-1}.$$

Hence for sufficiently small A we obtain

$$|m| < A^{1/2} N(\mathfrak{a}_\gamma)^{1/n}.$$

Choosing A to be suitably small implies $m = 0$ by Lemma 3.3.4. □

We now recall Davenport's shrinking lemma [8, Lemma 12.6].

Lemma 3.5.2. *Let $L\colon \mathbb{R}^m \to \mathbb{R}^m$ be a linear map. Let $a > 0$ be a real number and for a real number $Z > 0$ consider*

$$N(Z) = \left\{\mathbf{u} \in \mathbb{Z}^m \colon |\mathbf{u}| < aZ,\ \|(L(\mathbf{u}))_i\| < a^{-1}Z, \textit{for all } i\right\}.$$

Then if $0 < Z \leq 1$ we have

$$N(1) \ll_m Z^{-m} N(Z).$$

As noted in [16] the lemma was originally only stated when $a \geq 1$ but we may extend the range of a to all positive real numbers since the result holds trivially if $0 < a < 1$.

3.6 Weyl Differencing

One of the main innovations in [16] is to introduce an averaged van der Corput differencing approach in order to bound the contribution from the minor arcs. Since this cannot handle the entire range of minor arcs we need to supplement it with an estimate coming from conventional Weyl differencing. Let $\alpha \in K_{\mathbb{R}}$. Throughout this section we will write

$$\alpha = \gamma + \theta,$$

where $\gamma \in K$ and $\theta \in K_{\mathbb{R}}$. Note as in [24, Lemma 2.1] we find

$$|S(\alpha)|^4 \ll P^{ns} \sum_{|\mathbf{x}|,|\mathbf{y}|<P} \prod_{i=1}^{s} \prod_{j=1}^{n} \min\left(P, \|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x},\mathbf{y}))\|^{-1}\right). \qquad (3.6.1)$$

This estimate is proved using a classical Weyl differencing procedure adjusted to this context. Following standard arguments as in Davenport [8, Chapter 13] we now transform this into a counting problem.
Given $\alpha \in \mathbb{R}$ and $P \geq 1$ define

$$N(\alpha, P) := \#\left\{(\mathbf{x},\mathbf{y}) \in \mathcal{O}^{2s} \colon |\mathbf{x}| < P,\ |\mathbf{y}| < P,\ \|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x},\mathbf{y}))\| < P^{-1}\right\}.$$

For a fixed $\mathbf{x} \in \mathcal{O}^s$ write further

$$N(\mathbf{x}) := \#\left\{\mathbf{y} \in \mathcal{O}^s \colon |\mathbf{y}| < P,\ \|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x},\mathbf{y}))\| < P^{-1},\ \forall i,j\right\},$$

so that

$$N(\alpha, P) = \sum_{|\mathbf{x}|<P} N(\mathbf{x}).$$

Let r_{ij} be integers such that $0 \leq r_{ij} < P$ for $i = 1, \ldots, s$, $j = 1, \ldots, n$. We claim that there exist no more than $N(\mathbf{x})$ integer tuples $\mathbf{y} \in \mathcal{O}^s$, which lie in a box whose edges have sidelengths at most P such that

$$\frac{r_{ij}}{P} \leq \{\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x},\mathbf{y}))\} < \frac{r_{ij}+1}{P}$$

is satisfied for all $i = 1, \dots, s$ and $j = 1, \dots, n$, where $\{x\}$ denotes the fractional part of a real number x. Indeed, if $\mathbf{y}_1$ and $\mathbf{y}_2$ are two such integer tuples satisfying the above system of inequalities then $|\mathbf{y}_1 - \mathbf{y}_2| < P$ and

$$\|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x}, \mathbf{y}_1 - \mathbf{y}_2))\| < P^{-1}$$

holds for all i, j. Hence, since $\mathbf{y} = \mathbf{0}$ is a possible solution, there are no more than $N(\mathbf{x})$ possible solutions to the system of inequalities above. Dividing the box $P\mathcal{B}$ into 2^{ns} boxes whose edges have sidelength at most P we find

$$\begin{aligned}
&\sum_{|\mathbf{y}|<P} \prod_{i=1}^{s} \prod_{j=1}^{i} \min\left(P, \|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x}, \mathbf{y}))\|^{-1}\right) \\
\ll &N(\mathbf{x}) \prod_{i,j} \sum_{r_{ij}=0}^{P} \min\left(P, \frac{P}{r_{ij}}, \frac{P}{P - r_{ij} - 1}\right) \\
\ll &N(\mathbf{x})(P \log P)^{ns}.
\end{aligned}$$

Upon summing this estimate over $|\mathbf{x}| < P$ and using (3.6.1) we obtain

$$|S(\alpha)|^4 \ll P^{2ns}(\log P)^{ns} N(\alpha, P). \tag{3.6.2}$$

We now proceed to estimate $N(\alpha, P)$ using the results from the previous section.

For fixed $\mathbf{x} \in \mathcal{O}^s$ identifying $\mathcal{O}^s \cong \mathbb{R}^{ns}$ and given $\mathbf{y} \in \mathcal{O}^s$ one may view the map

$$\mathbf{y} \mapsto (\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x}, \mathbf{y})))_{i,j}$$

as a linear map $\mathbb{R}^{ns} \to \mathbb{R}^{ns}$. Hence we can apply Lemma 3.5.2 where $N(\mathbf{x}) = N(1)$ in the notation of the lemma where Z is to be determined in due course. Summing over the $|\mathbf{x}| < P$ then yields

$$\begin{aligned}
N(\alpha, P) \ll Z^{-ns} \# \big\{(\mathbf{x}, \mathbf{y}) \in \mathcal{O}^{2s} \colon |\mathbf{x}| < P,\ &|\mathbf{y}| < ZP, \\
&\|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x}, \mathbf{y}))\| < ZP^{-1}, \forall i, j\big\}.
\end{aligned} \tag{3.6.3}$$

If we apply the same procedure to the quantity on the right hand side of (3.6.3), but now with the roles of $\mathbf{x}$ and $\mathbf{y}$ reversed we obtain

$$N(\alpha, P) \ll Z^{-2ns} \# \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{O}^{2s} \colon |\mathbf{x}| < ZP,\, |\mathbf{y}| < ZP, \right. \\ \left. \|\mathrm{tr}(6\alpha\omega_j B_i(\mathbf{x}, \mathbf{y}))\| < Z^2 P^{-1},\, \forall i, j \right\}. \quad (3.6.4)$$

At this point we will employ Lemma 3.5.1. We wish to choose Z such that the bilinear forms appearing in the right hand side of (3.6.4) are forced to vanish. To this end, in the notation of the lemma we take $m = 6\Delta B_i(\mathbf{x}, \mathbf{y})$, $M \asymp 6Z^2P^2$ and $P_0^{-1} = Z^2P^{-1}$. Choose the parameter Z so that it satisfies

$$0 < Z < 1, \quad Z^2 \ll (P^2|\theta|N(\mathfrak{a}_\gamma)^{1/n})^{-1}, \quad Z^2 \ll \frac{P}{N(\mathfrak{a}_\gamma)^{1/n}},$$

as well as

$$Z^2 \ll \max\left(\frac{N(\mathfrak{a}_\gamma)^{1/n}}{P^2}, N(\mathfrak{a}_\gamma)^{1/n}|\theta|P \right),$$

where the implicit constants involved are sufficiently small such that the assumptions of Lemma 3.5.1 are satisfied. Provided K is an imaginary quadratic number field, Lemma 3.5.1 and (3.6.4) give

$$N(\alpha, P) \ll Z^{-2ns} \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{O}^{2s} \colon |\mathbf{x}| < ZP,\, |\mathbf{y}| < ZP,\, B_i(\mathbf{x}, \mathbf{y}) = 0,\, \forall i \right\},$$

where we note that clearly $6\Delta B_i(\mathbf{x}, \mathbf{y}) = 0$ if and only if $B_i(\mathbf{x}, \mathbf{y}) = 0$.
Since we assume that Davenport's Geometric Condition (3.4.1) is satisfied it follows from the simple observation (3.4.2) that

$$N(\alpha, P) \ll Z^{-2ns}(ZP)^{ns}.$$

From (3.6.2) for permissible Z as described above we therefore have

$$|S(\alpha)|^4 \ll P^{3ns+\varepsilon} Z^{-ns}. \quad (3.6.5)$$

The estimate is optimised when Z is as large as possible. Hence if we take

$$Z^2 \asymp \min\left\{ 1, \frac{1}{P^2|\theta|N(\mathfrak{a}_\gamma)^{1/n}}, \frac{P}{N(\mathfrak{a}_\gamma)^{1/n}}, \max\left(\frac{N(\mathfrak{a}_\gamma)^{1/n}}{P^2}, N(\mathfrak{a}_\gamma)^{1/n}|\theta|P \right) \right\}$$

then Z is clearly in the permissible range, and we deduce

$$|S(\alpha)|^4 \ll P^{3ns+\varepsilon}\left(1+P^2|\theta|N(\mathfrak{a}_\gamma)^{1/n}+P^{-1}N(\mathfrak{a}_\gamma)^{1/n}\right.$$
$$\left.+\min\Big(PN(\mathfrak{a}_\gamma)^{-1/n},\big(N(\mathfrak{a}_\gamma)^{1/n}|\theta|P\big)^{-1}\Big)\right)^{\frac{ns}{2}}.$$

In particular, if $N(\mathfrak{a}_\gamma)^{1/n}\le P^{3/2}$ then $P^{-1}N(\mathfrak{a}_\gamma)^{1/n}\le P^{1/2}$ and so we find

$$|S(\alpha)|\ll P^{ns+\varepsilon}\left(N(\mathfrak{a}_\gamma)^{1/n}|\theta|+(N(\mathfrak{a}_\gamma)^{1/n}|\theta|P^3)^{-1}+P^{-3/2}\right)^{\frac{ns}{8}}$$

in this case. Finally since $X^{1/2}\le X/Y+Y$ for any two positive real numbers X and Y we see that the last term of the right hand side above is dominated by the other two summands. We summarise the main result of this section.

Lemma 3.6.1. *Let $K/\mathbb{Q}$ be an imaginary quadratic number field. Let $\alpha\in K_\mathbb{R}$ and write $\alpha=\gamma+\theta$ where $\gamma\in K$ and $\theta\in K_\mathbb{R}$. If $N(\mathfrak{a}_\gamma)^{1/n}\le P^{3/2}$ then we have*

$$S(\alpha)\ll P^{ns+\varepsilon}\left(N(\mathfrak{a}_\gamma)^{1/n}|\theta|+(N(\mathfrak{a}_\gamma)^{1/n}|\theta|P^3)^{-1}\right)^{\frac{ns}{8}}. \qquad (3.6.6)$$

This bound will be useful for the range in the minor arcs when the parameter θ is small.

3.7 Pointwise van der Corput Differencing and the singular series

In this section we will perform a pointwise van der Corput differencing argument, in order to show that the singular series converges absolutely. This argument works over a general number field. We start by considering the exponential sum $S(\gamma)$, where $\gamma\in K$ and we set $P=N(\mathfrak{a}_\gamma)$. Further in this section we take the box $\mathcal{B}=\mathcal{B}_\mathfrak{S}=\{(\sum_j x_{ij}\omega_j)_i\in K_\mathbb{R}^s\colon 0\le x_{ij}<1\}$

so that the goal of this section is to study the sum S_γ as it was defined in Section 3.4.3. To be completely explicit with our choice of box we then have

$$S_\gamma = S(\gamma) = \sum_{0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma)} e\left(\mathrm{tr}(\gamma C(\mathbf{x}))\right),$$

where the condition $0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma)$ denotes the sum over elements $\mathbf{x} = \left(\sum_j x_{ij}\omega_j\right)_i \in \mathcal{O}^s$ such that $0 \leq x_{ij} < N(\mathfrak{a}_\gamma)$ holds. The main goal of this section is to establish the bound

$$S_\gamma \ll N(\mathfrak{a}_\gamma)^{s(n-1/6)+\varepsilon}. \tag{3.7.1}$$

Let H be a positive integer that satisfies $H \leq N(\mathfrak{a}_\gamma)$. Clearly we have

$$H^{ns} S(\gamma) = \sum_{0 \leq \mathbf{h} < H} \sum_{\substack{0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma) \\ 0 \leq \mathbf{x}+\mathbf{h} < N(\mathfrak{a}_\gamma)}} e\left(\mathrm{tr}(\gamma C(\mathbf{x}+\mathbf{h}))\right).$$

Interchanging the order of summation gives

$$H^{ns} S(\gamma) = \sum_{0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma)} \sum_{\substack{0 \leq \mathbf{h} < H \\ 0 \leq \mathbf{x}+\mathbf{h} < N(\mathfrak{a}_\gamma)}} e\left(\mathrm{tr}(\gamma C(\mathbf{x}+\mathbf{h}))\right).$$

Since $H \leq N(\mathfrak{a}_\gamma)$ the number of non-zero summands of the inner sum is bounded by $O(N(\mathfrak{a}_\gamma)^{ns})$. Therefore, an application of Cauchy-Schwarz yields

$$H^{2ns} |S(\gamma)|^2 \ll N(\mathfrak{a}_\gamma)^{ns} \sum_{0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma)} \left| \sum_{\substack{0 \leq \mathbf{h} < H \\ 0 \leq \mathbf{x}+\mathbf{h} < N(\mathfrak{a}_\gamma)}} e\left(\mathrm{tr}(\gamma C(\mathbf{x}+\mathbf{h}))\right) \right|^2.$$

Expanding the square one obtains that $H^{2ns}|S(\gamma)|^2$ is

$$\ll N(\mathfrak{a}_\gamma)^{ns} \sum_{0 \leq \mathbf{x} < N(\mathfrak{a}_\gamma)} \sum_{\substack{0 \leq \mathbf{h}_1, \mathbf{h}_2 < H \\ 0 \leq \mathbf{x}+\mathbf{h}_1, \mathbf{x}+\mathbf{h}_2 < N(\mathfrak{a}_\gamma)}} e\left(\mathrm{tr}(\gamma C(\mathbf{x}+\mathbf{h}_1) - C(\mathbf{x}+\mathbf{h}_2))\right).$$

Set $\mathbf{y} = \mathbf{x} + \mathbf{h}_2$ and $\mathbf{h} = \mathbf{h}_1 - \mathbf{h}_2$. Note that after this change of coordinates each value of $\mathbf{h}$ in the sum above appears at most H^{ns} times. Therefore the previous display gives

$$H^{ns}|S(\gamma)|^2 \ll N(\mathfrak{a}_\gamma)^{ns} \sum_{|\mathbf{h}| \leq H} |T(\mathbf{h}, \gamma)| \,, \tag{3.7.2}$$

where

$$T(\mathbf{h}, \gamma) = \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} e\left(\operatorname{tr}(\gamma(C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})))\right),$$

and where $\mathcal{R}(\mathbf{h})$ is a box whose sidelengths are $O(N(\mathfrak{a}_\gamma))$. We take the square of the absolute value of this expression, and expand the resulting sum in order to obtain

$$|T(\mathbf{h}, \gamma)|^2 = \sum_{\mathbf{y}, \mathbf{z} \in \mathcal{R}(\mathbf{h})} e\left(\operatorname{tr}(\gamma(C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y}) - C(\mathbf{z}+\mathbf{h}) + C(\mathbf{z})))\right).$$

Making the change of variables $\mathbf{y} = \mathbf{z} + \mathbf{w}$ we find

$$|T(\mathbf{h}, \gamma)|^2 = \sum_{|\mathbf{w}| < N(\mathfrak{a}_\gamma)} \sum_{\mathbf{z}} e\left(\operatorname{tr}(\gamma C(\mathbf{w}, \mathbf{h}, \mathbf{z}))\right),$$

where the inner sum ranges over a (potentially empty) box $\mathcal{S}(\mathbf{h}, \mathbf{w})$ whose sidelengths are $O(N(\mathfrak{a}_\gamma))$ and where we write $C(\mathbf{w}, \mathbf{h}, \mathbf{z})$ for the multilinear form given by

$$C(\mathbf{w}, \mathbf{h}, \mathbf{z}) = C(\mathbf{w} + \mathbf{h} + \mathbf{z}) - C(\mathbf{w} + \mathbf{z}) - C(\mathbf{h} + \mathbf{z}) + C(\mathbf{z}).$$

In particular we have

$$C(\mathbf{w}, \mathbf{h}, \mathbf{z}) = 6 \sum_{i=1}^{s} z_i B_i(\mathbf{w}, \mathbf{h}) + \Psi(\mathbf{w}, \mathbf{h}),$$

where B_i are the bilinear forms associated to C, and where Ψ is a certain polynomial whose precise shape is of no importance to us. Therefore we find

$$|T(\mathbf{h}, \gamma)|^2 = \sum_{\mathbf{w}} \sum_{\mathbf{z}} e\left(\operatorname{tr}\left(6\gamma \sum_{i=1}^{s} z_i B_i(\mathbf{w}, \mathbf{h}) + \gamma \Psi(\mathbf{w}, \mathbf{h})\right)\right).$$

Writing $z_i = \sum_j z_{ij}\omega_j$ we may regard the inner sum as an exponential sum over integer variables z_{ij}. This is a linear exponential sum and the coefficient of z_{ij} is given by $6\mathrm{tr}(\gamma\omega_j B_i(\mathbf{w},\mathbf{h}))$. A standard argument regarding geometric sums now yields

$$|T(\mathbf{h},\gamma)|^2 \ll \sum_{\mathbf{w}} \prod_{i=1}^{s} \prod_{j=1}^{n} \min\left(N(\mathfrak{a}_\gamma), \|6\mathrm{tr}(\gamma\omega_j B_i(\mathbf{w},\mathbf{h}))\|^{-1}\right).$$

In particular the same argument that led to (3.6.2) shows that

$$|T(\mathbf{h},\gamma)|^2 \ll N(\mathfrak{a}_\gamma)^{ns+\varepsilon} N(\gamma, N(\mathfrak{a}_\gamma), \mathbf{h}), \tag{3.7.3}$$

where

$$N(\gamma, N(\mathfrak{a}_\gamma), \mathbf{h}) = \#\left\{\mathbf{w} \in \mathcal{O}^s \colon |\mathbf{w}| < N(\mathfrak{a}_\gamma),\ \|6\mathrm{tr}(\gamma\omega_j B_i(\mathbf{w},\mathbf{h}))\| < N(\mathfrak{a}_\gamma)^{-1}\right\}.$$

Note that the condition in the sum already implies that $6\Delta B_i(\mathbf{x},\mathbf{y}) \in \mathfrak{a}_\gamma$ holds for all i, but we prefer to write it in the above shape in order to highlight the similarities with the argument in the previous section.

As in Section 3.6 we may regard $\mathbf{w} \mapsto \mathrm{tr}(\gamma\omega_j B_i(\mathbf{w},\mathbf{h}))$ as a linear map $\mathbb{R}^{ns} \to \mathbb{R}^{ns}$. Hence we can apply Lemma 3.5.2 so that for any $Z \in (0,1]$ we have

$$N(\gamma, N(\mathfrak{a}_\gamma), \mathbf{h}) \ll Z^{-ns} \#\left\{|\mathbf{w}| < ZN(\mathfrak{a}_\gamma),\ \|6\mathrm{tr}(\gamma\omega_j B_i(\mathbf{w},\mathbf{h}))\| < ZN(\mathfrak{a}_\gamma)^{-1}\right\}.$$

We now wish to choose Z in such a way that we can apply Lemma 3.5.1. In the notation of this lemma we have $m = \Delta\omega_j B_i(\mathbf{w},\mathbf{h})$ and $\theta = 0$. We take $Z \in (0,1]$ such that $Z \asymp H^{-1}N(\mathfrak{a}_\gamma)^{\frac{1}{n}-1}$ for a suitable implied constant. Then Lemma 3.5.1 implies

$$N(\gamma, P, \mathbf{h}) \ll H^{ns} N(\mathfrak{a}_\gamma)^{ns-s} \#\left\{\mathbf{w} \in \mathcal{O}^s \colon |\mathbf{w}| < H^{-1}N(\mathfrak{a}_\gamma)^{1/n},\ B_i(\mathbf{w},\mathbf{h}) = 0\right\}.$$

Recalling that $r(\mathbf{h})$ is the rank of $B_i(\mathbf{h},\cdot)\colon K_{\mathbb{R}}^s \to K_{\mathbb{R}}^s$, using (3.7.3) we find

$$T(\mathbf{h},\gamma) \ll N(\mathfrak{a}_\gamma)^{ns-\frac{r(\mathbf{h})}{2}+\varepsilon} H^{\frac{nr(\mathbf{h})}{2}}.$$

Hence (3.7.2) delivers

$$|S(\gamma)|^2 \ll H^{-ns} N(\mathfrak{a}_\gamma)^{2ns+\varepsilon} \sum_{|\mathbf{h}|\leq H} \left(H^n N(\mathfrak{a}_\gamma)^{-1}\right)^{\frac{r(\mathbf{h})}{2}}.$$

By (3.4.1), for any r the number of $\mathbf{h}$ with $r(\mathbf{h}) = r$ is $O(H^{nr})$. Therefore we find

$$|S(\gamma)|^2 \ll H^{-ns} N(\mathfrak{a}_\gamma)^{2ns+\varepsilon} \sum_{r=0}^{s} \left(H^{3n} N(\mathfrak{a}_\gamma)^{-1}\right)^{\frac{r}{2}}.$$

The sum is maximal either when $r = 0$ or when $r = s$, and thus

$$|S(\gamma)|^2 \ll H^{-ns} N(\mathfrak{a}_\gamma)^{2ns+\varepsilon} \left(1 + H^{3ns/2} N(\mathfrak{a}_\gamma)^{-s/2}\right).$$

Choosing $H = \lfloor N(\mathfrak{a}_\gamma)\rfloor^{1/3n}$ this finally yields

$$S(\gamma) \ll N(\mathfrak{a}_\gamma)^{s(n-1/6)+\varepsilon}.$$

3.7.1 Proof of Theorem 3.4.4

By Theorem 3.4.3 it suffices to show that $\mathfrak{S}(R)$ converges absolutely as $R \to \infty$.

Given a positive integer k the number of ideals of $\mathcal{O}$ of norm k is $O(k^\varepsilon)$ using the divisor bound. Hence together with Lemma 3.3.1 we obtain that the number of $\gamma \in K/\mathcal{O}$ such that $N(\mathfrak{a}_\gamma) = k$ is bounded by $O(k^{1+\varepsilon})$. Thus, using (3.7.1) we find

$$\mathfrak{S}(R) \ll \sum_{k=0}^{R} k^{-ns+1+\varepsilon} k^{ns-s/6} = \sum_{k=0}^{R} k^{1-s/6+\varepsilon}.$$

Therefore $\mathfrak{S}(R)$ converges absolutely to some real number $\mathfrak{S}$ as $R \to \infty$ provided $s \geq 13$. □

We remark that using the ideas of Heath-Brown [16, Section 7] it would be possible to establish the absolute convergence of $\mathfrak{S}(R)$ already for $s \geq 11$.

3.8 Van der Corput on average

In this section, we work towards a bound for the Weyl sum $S(\alpha)$ on the minor arcs. As observed by Heath-Brown, the simple pointwise van der Corput differencing is not sufficient to improve on Davenport's result for $s \geq 16$.

It is however possible to exploit the fact that we are averaging both over the modulus $\mathfrak{a}_\gamma$ as well as the integration variable β in the minor arcs, thus leading to a version of van der Corput differencing on average.

From now on we continue to work with the box $\mathcal{B} = \mathcal{B}(\mathbf{z})$ as defined in the beginning of Section 3.4.3. Instead of a pointwise bound for $S(\alpha)$, we will seek to bound the mean-square average

$$M(\alpha, \kappa) = \int_{|\beta-\alpha|<\kappa} |S(\beta)|^2 d\beta$$

for $\alpha \in K_\mathbb{R}$ and a small parameter $\kappa \in (0,1)$, where we remind the reader that the integration is over a region of $K_\mathbb{R}$.

In conjunction with the Cauchy-Schwarz inequality and an appropriate dyadic dissection of the minor arcs, a satisfactory bound for $M(\alpha,\kappa)$ will be sufficient to control the total minor arc contribution.

The idea now is that the mean square integral automatically shortens all the n coordinates of h_1 in the van der Corput differencing, allowing us to effectively save a factor $\frac{H^n}{P^n}$ over the pointwise bound. Here and throughout we denote $\mathbf{h} = (h_i)_i = \left(\sum_j h_{ij}\omega_j\right)_i \in \mathcal{O}^s$.

To this end, we initiate the van der Corput differencing with parameters $1 \leq H_{ij} \leq P$ to be determined, obtaining

$$\begin{aligned}\prod_{i,j} H_{ij} S(\beta) &= \sum_{0 \leq h_{ij} < H_{ij}} \sum_{\mathbf{x}+\mathbf{h} \in P\mathcal{B}} e\left(\operatorname{tr}(\beta C(\mathbf{x}+\mathbf{h}))\right)\\ &= \sum_{\mathbf{x} \in \mathcal{O}^s} \sum_{\mathbf{x}+\mathbf{h} \in P\mathcal{B}} e\left(\operatorname{tr}(\beta C(\mathbf{x}+\mathbf{h}))\right),\end{aligned}$$

where implicitly we still restrict to $\mathbf{h}$ such that $0 \leq h_{ij} < H_{ij}$ is satisfied. Note that the condition $H_{ij} \leq P$ ensures that the sum over $\mathbf{x}$ is restricted to

$O(P^{ns})$ many summands. An application of Cauchy-Schwarz thus yields

$$\prod_{i,j} H_{ij}^2 |S(\beta)|^2 \ll P^{ns} \sum_{\mathbf{x}\in\mathcal{O}^s} \left| \sum_{\mathbf{x}+\mathbf{h}\in P\mathcal{B}} e\left(\operatorname{tr}(\beta C(\mathbf{x}+\mathbf{h}))\right) \right|^2 .$$

Opening the square on the RHS, we obtain

$$\prod_{i,j} H_{ij}^2 |S(\beta)|^2 \ll P^{ns} \sum_{\mathbf{x}\in\mathcal{O}^s} \sum_{\mathbf{x}+\mathbf{h_1},\mathbf{x}+\mathbf{h_2}\in P\mathcal{B}} e\left(\operatorname{tr}(\beta\left[C(\mathbf{x}+\mathbf{h_1})-C(\mathbf{x}+\mathbf{h_2})\right])\right).$$

On substituting $\mathbf{y}=\mathbf{x}+\mathbf{h_2}$ and $\mathbf{h}=\mathbf{h_1}-\mathbf{h_2}$, this becomes

$$\prod_{i,j} H_{ij}^2 |S(\beta)|^2 \ll P^{ns} \sum_{|h_{ij}|\le H_{ij}} w(\mathbf{h}) \sum_{\mathbf{y}\in\mathcal{R}(\mathbf{h})} e\left(\operatorname{tr}(\beta\left[C(\mathbf{y}+\mathbf{h})-C(\mathbf{y})\right])\right)$$

where $w(\mathbf{h}) = \#\{\mathbf{h_1},\mathbf{h_2} : \mathbf{h}=\mathbf{h_1}-\mathbf{h_2}\} \le \prod_{i,j} H_{ij}$ and $\mathcal{R}(\mathbf{h})$ is a certain box depending only on $\mathbf{h}$.

Instead of taking absolute values, we now first integrate over $\beta = \sum_j \beta_j\omega_j$ with a smooth cutoff function to obtain

$$\begin{aligned} M(\alpha,\kappa) &\le e^n \int_{K_\mathbb{R}} \exp\left(-\frac{\sum_j(\beta_j-\alpha_j)^2}{\kappa^2}\right)\cdot |S(\beta)|^2 d\beta \\ &\ll \frac{P^{ns}}{\prod_{i,j} H_{ij}^2} \sum_{|h_{ij}|\le H_{ij}} w(\mathbf{h}) \sum_{\mathbf{y}\in\mathcal{R}(\mathbf{h})} I(\mathbf{h},\mathbf{y}) \\ &\ll \frac{P^{ns}}{\prod_{i,j} H_{ij}} \sum_{|h_{ij}|\le H_{ij}} \left| \sum_{\mathbf{y}\in\mathcal{R}(\mathbf{h})} I(\mathbf{h},\mathbf{y}) \right|, \end{aligned}$$

where

$$\begin{aligned} I(\mathbf{h},\mathbf{y}) &= \int_{K_\mathbb{R}} \exp\left(-\frac{\sum_j(\beta_j-\alpha_j)^2}{\kappa^2}\right)\cdot e\left(\operatorname{tr}(\beta\left[C(\mathbf{y}+\mathbf{h})-C(\mathbf{y})\right])\right) d\beta \\ &= \pi^{n/2}\kappa^n \prod_{j=1}^n \exp\left(-\pi^2\kappa^2 \operatorname{tr}(\omega_j\left[C(\mathbf{y}+\mathbf{h})-C(\mathbf{y})\right])^2\right)\cdot e(\operatorname{tr}(\alpha\left[C(\mathbf{y}+\mathbf{h})-C(\mathbf{y})\right])). \end{aligned}$$

Heuristically, for large $h_1\in\mathcal{O}$, we should have $C(\mathbf{y}+\mathbf{h})-C(\mathbf{y}) \approx h_1\cdot \frac{\partial C(\mathbf{y})}{\partial x_1}$ so that by our choice of the box $\mathcal{B}(\mathbf{z})$, this difference is large. But then for some j, the trace of this number multiplied with ω_j must be large as well, leading to a negligible contribution to $M(\alpha, \kappa)$ from those terms, thus effectively cutting down the range to small h_1.

We now fix the choice $H_{ij} = H$ for $i \neq 1$ and $H_{1j} = cP$ for a sufficiently small constant c and make the above heuristic discussion precise. For $\mathbf{y} \in \mathcal{R}(\mathbf{h})$ we have

$$C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y}) = h_1 \cdot \frac{\partial C(\mathbf{y})}{\partial x_1} + O(HP^2 + |h_1|^2|\mathbf{y}|).$$

If the width of the box $\mathcal{B}(\boldsymbol{z})$ and c are sufficiently small, then $\frac{\partial C(\mathbf{z})}{\partial x_1} \neq 0$ implies that

$$|C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})| \gg |h_1| \cdot P^2$$

unless $|h_1| \ll H$. Additionally, unless $|h_1| \ll \frac{(\log P)^2}{\kappa P^2}$, we even have that

$$|C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})| \gg \frac{(\log P)^2}{\kappa}$$

so that for some j we must have

$$|\mathrm{tr}\,(\omega_j\,[C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})])| \gg \frac{(\log P)^2}{\kappa}$$

and we infer from our previous calculations that the contribution of such $\mathbf{h}$ to $M(\alpha, \kappa)$ is $O(1)$. Hence,

$$M(\alpha, \kappa) \ll 1 + \frac{P^{ns-n}}{H^{ns-n}} \sum_{|h_i| \ll H} \left| \sum_{\mathbf{y}} I(\mathbf{h}, \mathbf{y}) \right|$$

if we choose $\kappa \asymp \frac{(\log P)^2}{HP^2}$.

Moreover, the range $|\beta - \alpha| \geq \kappa \log P$ in the definition of $I(\mathbf{h}, \mathbf{y})$ clearly gives a total contribution of $O(1)$ to $M(\alpha, \kappa)$ so that we end up with the estimate

$$M(\alpha, \kappa) \ll 1 + \frac{P^{ns-n}}{H^{ns-n}} \sum_{|h_i| \ll H} \int_{|\beta - \alpha| < \kappa \log P} |T(\mathbf{h}, \beta)| d\beta$$

with

$$T(\mathbf{h},\beta) = \sum_{\mathbf{y}\in\mathcal{R}(\mathbf{h})} e\left(\mathrm{tr}(\beta\left[C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})\right])\right).$$

As in Section 3.7, we obtain

$$|T(\mathbf{h},\beta)|^2 \ll P^{ns+\varepsilon} N(\beta,P,\mathbf{h})$$

where

$$N(\beta,P,\mathbf{h}) = \#\{\mathbf{w}\in\mathcal{O}^s : |\mathbf{w}| < P, \|6\mathrm{tr}(\beta\omega_j B_i(\mathbf{w},\mathbf{h}))\| < P^{-1}, \forall i,j\}$$

so that

$$M(\alpha,\kappa) \ll 1 + \frac{\kappa^n P^{\frac{3ns}{2}-n+\varepsilon}}{H^{ns-n}} \sum_{|h_i|\ll H} \max_{\beta\in\mathcal{I}} N(\beta,P,\mathbf{h})^{\frac{1}{2}} \tag{3.8.1}$$

for $\mathcal{I} = \{\beta : |\beta-\alpha| \le \kappa\log P\}$.
We next claim that

$$\max_{\beta\in\mathcal{I}} N(\beta,P,\mathbf{h}) \ll P^{\varepsilon} N(\alpha,P,\mathbf{h}).$$

Indeed, consider a vector $\mathbf{w}$ counted by $N(\beta,P,\mathbf{h})$. It thus satisfies $|\mathbf{w}| < P$ as well as $\|6\mathrm{tr}(\beta\omega_j B_i(\mathbf{w},\mathbf{h}))\| < P^{-1}$ so that

$$\|6\mathrm{tr}(\alpha\omega_j B_i(\mathbf{w},\mathbf{h}))\| \ll \frac{1}{P} + |\beta-\alpha|\cdot|B_i(\mathbf{w},\mathbf{h})| \ll \frac{1}{P} + \kappa\log P\cdot HP \ll \frac{(\log P)^3}{P}.$$

We thus obtain

$$\begin{aligned} N(\beta,P,\mathbf{h}) &\ll \#\{\mathbf{w}\in\mathcal{O}^s : |\mathbf{w}| < P, \|6\mathrm{tr}(\alpha\omega_j B_i(\mathbf{w},\mathbf{h}))\| \ll \frac{(\log P)^3}{P}, \forall i,j\} \\ &\ll P^{\varepsilon} N(\alpha,P,\mathbf{h}) \end{aligned}$$

where the last estimate is a consequence of Lemma 3.5.2 upon choosing suitable $Z \asymp (\log P)^{-3}$.

We conclude that

$$M(\alpha,\kappa) \ll 1 + \frac{\kappa^n P^{\frac{3ns}{2}-n+\varepsilon}}{H^{ns-n}} \sum_{|h_i|\ll H} N(\alpha,P,\mathbf{h})^{\frac{1}{2}}.$$

Let $\alpha = \gamma + \theta$ with $\gamma \in K$ and $\theta \in K_{\mathbb{R}}$ (which we think of as being small). We are now prepared for an application of Lemmas 3.5.2 and 3.5.1. Indeed, Lemma 3.5.2 implies that

$$N(\alpha,P,\mathbf{h}) \ll Z^{-ns}\#\{\mathbf{w} \in \mathcal{O}^s : |\mathbf{w}| < ZP, \|6\mathrm{tr}(\alpha\omega_j B_i(\mathbf{w},\mathbf{h}))\| < ZP^{-1}\}.$$

Following Heath-Brown, we will make two different choices of Z: In the first one, we will choose $Z = Z_1$ sufficiently small so that Lemma 3.5.1 implies that $B_i(\mathbf{w},\mathbf{h}) = 0$. In the second choice $Z = Z_2$, we will only force $6\Delta B_i(\mathbf{w},\mathbf{h}) \in \mathfrak{a}_\gamma$, a consequence followed by a study of how often such a divisibility property can occur, crucially using an average over γ.
By Lemma 3.5.1, if we choose $Z \leq 1$ satisfying

$$Z \ll \frac{P}{N(\mathfrak{a}_\gamma)^{1/n}}$$

and

$$Z \ll \frac{1}{PH|\theta|N(\mathfrak{a}_\gamma)^{1/n}}$$

we can conclude that $6\Delta B_i(\mathbf{w},\mathbf{h}) \in \mathfrak{a}_\gamma$. If, moreover

$$Z \ll \frac{N(\mathfrak{a}_\gamma)^{1/n}}{PH}$$

or

$$Z \ll |\theta|PN(\mathfrak{a}_\gamma)^{1/n}$$

we obtain that $B_i(\mathbf{w},\mathbf{h}) = 0$. Here, all the implicit constants need to be sufficiently small in order to satisfy the conditions in Lemma 3.5.1.
Writing

$$\eta = |\theta| + \frac{1}{P^2H} \tag{3.8.2}$$

we should therefore choose

$$Z_1 \asymp \min\left(N(\mathfrak{a}_\gamma)^{1/n}P\eta, \frac{1}{PH\eta N(\mathfrak{a}_\gamma)^{1/n}}\right),$$

noting that this automatically implies that $Z_1 \leq 1$. Similarly we should choose

$$Z_2 \asymp \min\left(1, \frac{1}{PH\eta N(\mathfrak{a}_\gamma)^{1/n}}\right).$$

In the application with $Z = Z_1$, we thus obtain

$$\begin{aligned}N(\alpha, P, \mathbf{h}) &\ll Z_1^{-ns}\#\{\mathbf{w} \in \mathcal{O}^s : |\mathbf{w}| < Z_1 P, B_i(\mathbf{w}, \mathbf{h}) = 0, \forall i\}\\ &\ll Z_1^{-ns} \cdot (Z_1 P)^{n(s-r)}\\ &\ll P^{ns} \cdot \left(\frac{1}{N(\mathfrak{a}_\gamma)^{1/n}P^2\eta} + H\eta N(\mathfrak{a}_\gamma)^{1/n}\right)^{nr}\end{aligned}$$

with $r = r(\mathbf{h})$. Instead, in the application with $Z = Z_2$, we end up with the bound

$$N(\alpha, P, \mathbf{h}) \ll Z_2^{-ns}\#\{\mathbf{w} \in \mathcal{O}^s : |\mathbf{w}| < Z_2 P, 6\Delta B_i(\mathbf{w}, \mathbf{h}) \in \mathfrak{a}_\gamma, \forall i\}. \tag{3.8.3}$$

We thus need to count vectors $\mathbf{w}$ with $6\Delta B_i(\mathbf{w}, \mathbf{h}) \in \mathfrak{a}_\gamma$. For any prime ideal $\mathfrak{p}$, let $r_\mathfrak{p}(\mathbf{h})$ be the rank of $M(\mathbf{h})$ modulo $\mathfrak{p}$. Clearly, $r_\mathfrak{p}(\mathbf{h}) \leq r(\mathbf{h}) = r$ with strict inequality if and only if $\mathfrak{p}$ divides all $r \times r$ minors of $M(\mathbf{h})$. This means that there are only relatively few such 'bad' primes, which we will exploit later.

We now decompose $\mathfrak{a}_\gamma = \mathfrak{q}_1 \cdot \mathfrak{q}_2$ where $\mathfrak{q}_1$ contains all the primes $\mathfrak{p}$ dividing $\mathfrak{a}_\gamma$ with $r_\mathfrak{p}(\mathbf{h}) < r$ and $\mathfrak{q}_2$ consists of those with $r_\mathfrak{p}(\mathbf{h}) = r$.

As we are looking for an upper bound, we can replace $\mathfrak{a}_\gamma$ by the larger $\mathfrak{q}_2$ in (3.8.3).

For fixed $\mathbf{h}$ with $r(\mathbf{h}) = r$, the condition $6\Delta B_i(\mathbf{h}, \mathbf{w}) \in \mathfrak{q}_2, \forall i$ defines a lattice $\Lambda(\mathbf{h})$ for $\mathbf{w} \in \mathcal{O}^s$ which we view as a lattice in $\mathbb{R}^{ns}$.

To estimate the number of integer points in such a lattice we use [16, Lemma 5.1] implying that

$$\#\{\mathbf{x} \in \Lambda(\mathbf{h}) : |\mathbf{x}| \le B\} \ll \prod_i \left(1 + \frac{B}{\lambda_i}\right) \tag{3.8.4}$$

where $\lambda_1, \dots, \lambda_{ns}$ are the successive minima of $\Lambda(\mathbf{h})$.

In order to make this estimate useful, we need a bound on the determinant/covolume $d(\Lambda)$ which is proportional to $\prod_i \lambda_i$ as well as a bound on the skewness of the measure, i.e. upper and lower bounds for the λ_i.

For the determinant, we note that for $\mathfrak{p}^e \mid \mathfrak{q}_2$, the matrix $M(\mathbf{h})$ has rank r modulo $\mathfrak{p}$ (hence also modulo $\mathfrak{p}^e$) and therefore $B_i(\mathbf{h}, \mathbf{w})$ has $N(\mathfrak{p}^e)^{s-r}$ solutions modulo $\mathfrak{p}^e$ so that $N(\mathfrak{p}^e)^r$ divides $d(\Lambda)$. It thus follows that $N(\mathfrak{q}_2)^r \mid d(\Lambda)$ and hence $d(\Lambda) \ge N(\mathfrak{q}_2)^r$.

Regarding the skewness, we clearly have $\lambda_i \gg 1$ for all i, while in the other direction we have $\mathfrak{q}_2\mathcal{O}^s \subset \Lambda(\mathbf{h})$ so that Lemma 3.3.4 implies $\lambda_i \ll N(\mathfrak{q}_2)^{1/n}$. Optimizing the RHS of (3.8.4) with these constraints shows that the maximum is obtained when rn of the λ_i are of order $N(\mathfrak{q}_2)^{1/n}$ while the others are of order 1.

This shows that

$$N(\alpha, P, \mathbf{h}) \ll Z_2^{-ns} \left(1 + \frac{Z_2 P}{N(\mathfrak{q}_2)^{1/n}}\right)^{rn} \cdot (Z_2 P)^{(s-r)n} = P^{ns} \left(\frac{1}{Z_2 P} + \frac{1}{N(\mathfrak{q}_2)^{1/n}}\right)^{rn}$$

if $Z_2 P \gg 1$ but we note that the bound is trivially true for $Z_2 P \ll 1$.

Recalling our choice of Z_2, this bound becomes

$$N(\alpha, P, \mathbf{h}) \ll P^{ns} \left(\frac{1}{P} + \frac{1}{N(\mathfrak{q}_2)^{1/n}} + H\eta N(\mathfrak{a}_\gamma)^{1/n}\right)^{rn}.$$

Combining our two estimates, we obtain

$$N(\alpha, P, \mathbf{h}) \ll P^{ns} \left(\frac{1}{P} + H\eta N(\mathfrak{a}_\gamma)^{1/n} + \min\left(\frac{1}{N(\mathfrak{a}_\gamma)^{1/n} P^2 \eta}, \frac{1}{N(\mathfrak{q}_2)^{1/n}}\right)\right)^{rn}.$$

We now need to insert this into our expression for $M(\alpha, \kappa)$ which already involves the average over $\mathbf{h}$. Additionally, we want to average over $\mathfrak{a}_\gamma$ allowing us to use that $N(\mathfrak{q}_2)$ is almost as large as $N(\mathfrak{a}_\gamma)$ most of the time.
Our object of study thus becomes

$$A(\theta, R, H, P) := \sum_{\gamma : N(\mathfrak{a}_\gamma)^{1/n} \sim R} \sum_{|h_i| \ll H} N(\alpha, P, \mathbf{h})^{1/2} \tag{3.8.5}$$

where we continue to write $\alpha = \gamma + \theta$ and we remind the reader of the notation $q \sim R$ for the dyadic condition $R < q \leq 2R$. We then obtain that $A(\theta, R, H, P)$ is

$$\ll R^n P^{ns/2} \sum_{|h_i| \ll H} \sum_{N(\mathfrak{a})^{1/n} \sim R} \left(\frac{1}{P} + H\eta R + \min\left(\frac{1}{RP^2\eta}, \frac{1}{N(\mathfrak{q}_2)^{1/n}} \right) \right)^{\frac{r(\mathbf{h})n}{2}}$$

where we used that there are at most $N(\mathfrak{a})$ choices of γ with $\mathfrak{a}_\gamma = \mathfrak{a}$ by Lemma 3.3.1 and we remind the reader that $\mathfrak{q}_2$ depends on $\mathfrak{a}$ and $\mathbf{h}$.
We thus proceed to estimate

$$V(\mathbf{h}, R, \eta) := \sum_{N(\mathfrak{a})^{1/n} \sim R} \min\left(\frac{1}{RP^2\eta}, \frac{1}{N(\mathfrak{q}_2)^{1/n}} \right)^{\frac{rn}{2}}$$

for $r = r(\mathbf{h})$ via a dyadic decomposition as follows:

$$\begin{aligned} V(\mathbf{h}, R, \eta) &\ll P^\varepsilon \max_{S \leq R} \sum_{N(\mathfrak{q}_1)^{1/n} \sim S} \sum_{N(\mathfrak{q}_2)^{1/n} \sim \frac{R}{S}} \min\left(\frac{1}{RP^2\eta}, \frac{S}{R} \right)^{\frac{rn}{2}} \\ &\ll P^\varepsilon \max_{S \leq R} \frac{R^n}{S^n} \min\left(\frac{1}{RP^2\eta}, \frac{S}{R} \right)^{\frac{rn}{2}} \#\{\mathfrak{q}_1 : N(\mathfrak{q}_1)^{1/n} \leq 2S\}. \end{aligned}$$

Now recall that $\mathfrak{q}_1$ only contains prime ideals dividing a certain non-zero $r \times r$ determinant M_0 of $M(\mathbf{h})$. In particular, we have $M_0 \ll H^r$. Applying Rankin's trick, we then obtain

$$\#\{\mathfrak{q}_1 : N(\mathfrak{q}_1)^{1/n} \leq 2S\} \ll S^\varepsilon \sum_{\mathfrak{q}_1} N(\mathfrak{q}_1)^{-\varepsilon} = S^\varepsilon \prod_{\mathfrak{p} | M_0} \frac{1}{1 - N(\mathfrak{p})^{-\varepsilon}} \ll P^\varepsilon$$

and thus

$$V(\mathbf{h}, R, \eta) \ll P^{\varepsilon} \max_{S \leq R} \frac{R^n}{S^n} \min\left(\frac{1}{RP^2\eta}, \frac{S}{R}\right)^{\frac{rn}{2}}.$$

Maximizing for S we find that

$$V(\mathbf{h}, R, \eta) \ll P^{\varepsilon} \frac{R^n}{(RP^2\eta)^{rn/2}} \min(1, P^2\eta)^{ne(r)}$$

with $e(0) = 0$, $e(1) = \frac{1}{2}$ and $e(r) = 1$ for $r \geq 2$.

Putting everything together, we obtain the estimate

$$\begin{aligned} A(\theta, R, H, P) &\ll R^{2n} P^{\frac{ns}{2}} \sum_{|h_i| \ll H} \left[\left(\frac{1}{P} + H\eta R\right)^{\frac{nr(\mathbf{h})}{2}} + \frac{1}{R^n} V(\mathbf{h}, R, \eta)\right] \\ &\ll R^{2n} P^{\frac{ns}{2}+\varepsilon} \sum_{|h_i| \ll H} \left[\left(\frac{1}{P} + H\eta R\right)^{\frac{nr(\mathbf{h})}{2}} + \frac{1}{(RP^2\eta)^{\frac{r(\mathbf{h})n}{2}}} \min(1, P^2\eta)^{ne(r(\mathbf{h}))}\right] \\ &\ll R^{2n} P^{\frac{ns}{2}+\varepsilon} \sum_{r=0}^{s} H^{nr} \left[\left(\frac{1}{P} + H\eta R\right)^{\frac{nr}{2}} + \frac{1}{(RP^2\eta)^{\frac{rn}{2}}} \min(1, P^2\eta)^{ne(r)}\right] \\ &\ll \left[R^2 P^{s/2+\varepsilon} \left(1 + (RH^3\eta)^{s/2} + \frac{H^s}{P^{s/2}} + \frac{H^s}{(RP^2\eta)^{s/2}} \min(1, P^2\eta)\right)\right]^n. \end{aligned}$$

Finally, we argue that the third term $\frac{H^s}{P^{s/2}}$ is negligible.

Indeed, if $HRP\eta \geq 1$, then it is smaller than the second term. Otherwise, if $HRP\eta \leq 1$, we have $(RP\eta)^{s/2} \leq RP\eta \leq \frac{1}{H} \leq \min(1, \eta P^2)$ on recalling that $\eta \geq \frac{1}{P^2 H}$ and hence the term $\frac{H^s}{P^{s/2}}$ is dominated by the fourth term in that case.

In any case, it now follows that

$$A(\theta, R, H, P) \ll \left[R^2 P^{s/2+\varepsilon} \left(1 + (RH^3\eta)^{s/2} + \frac{H^s}{(RP^2\eta)^{s/2}} \min(1, P^2\eta)\right)\right]^n. \tag{3.8.6}$$

3.9 The minor arcs

Finally, we synthesize the bounds obtained by Weyl and van der Corput differencing to estimate the total minor arc contribution $\int_{\mathfrak{m}} S(\alpha) d\alpha$.

We dissect $\mathfrak{m}$ with the help of the version of Dirichlet's Approximation Theorem provided by Lemma 3.3.3, applied for some parameter $1 \leq Q \leq P^{3/2}$ to be determined. Thus, every $\alpha \in K_{\mathbb{R}}$ has an approximation $\alpha = \gamma + \theta$ with $\gamma \in K$ and $N(\mathfrak{a}_\gamma) \leq Q^n$ as well as $|\theta| \ll \frac{1}{N(\mathfrak{a}_\gamma)^{1/n}Q}$.
The assumption $\alpha \in \mathfrak{m}$ then implies that $N(\mathfrak{a}_\gamma) > P^\nu$ or $|\theta| > P^{-3+\nu}$. Note that as the contribution to the minor arcs coming from $|\theta| \leq \frac{1}{P^s}$ is $O(Q^{n+1})$, we may assume that $|\theta| \geq P^{-s}$.
By a double dyadic decomposition with respect to $|\theta|$ and $N(\mathfrak{a}_\gamma)^{1/n}$, we then obtain that

$$\int_{\mathfrak{m}} S(\alpha) d\alpha \ll Q^{n+1} + P^\varepsilon \max_{R \leq Q, \phi \leq \frac{1}{RQ}} \Sigma(R, \phi)$$

where

$$\Sigma(R, \phi) := \sum_{\gamma : N(\mathfrak{a}_\gamma)^{1/n} \sim R} \int_{|\theta| \sim \phi} |S(\gamma + \theta)| \, d\theta$$

and we note that the region of integration is given by a rectangular annulus. To establish Theorem 3.4.2, it thus suffices to prove $\Sigma(R, \phi) \ll P^{n(s-3)-\varepsilon}$. To employ the mean-value estimates from the previous section, we use Cauchy-Schwarz to obtain

$$\Sigma(R, \phi) \ll R^n \phi^{n/2} \left(\sum_{\gamma : N(\mathfrak{a}_\gamma)^{1/n} \sim R} \int_{|\theta| \sim \phi} |S(\gamma + \theta)|^2 \, d\theta \right)^{1/2} .$$

We next cover the annulus $|\theta| \sim \phi$ with $O\left(\left(1 + \frac{\phi}{\kappa}\right)^n\right)$ boxes of size κ, all centered at values of $\alpha = \gamma + \theta$ with $|\theta| \sim \phi$, so that we obtain

$$\Sigma(R, \phi) \ll R^n \phi^{n/2} \left(1 + \frac{\phi}{\kappa}\right)^{n/2} \max_{|\theta| \sim \phi} \left(\sum_{\gamma : N(\mathfrak{a}_\gamma)^{1/n} \sim R} M(\gamma + \theta, \kappa) \right)^{1/2}$$

and using (3.8.1) and (3.8.5) we obtain

$$\Sigma(R, \phi) \ll R^n \phi^{n/2} \left(1 + \frac{\phi}{\kappa}\right)^{n/2} \max_{|\theta| \sim \phi} \left(R^{2n+\varepsilon} + \frac{\kappa^n P^{\frac{3ns}{2} - n + \varepsilon}}{H^{ns-n}} A(\theta, R, H, P) \right)^{1/2}$$

so that (3.8.6) implies that

$$\Sigma(R,\phi) \ll \left[P^{\varepsilon} R^2 \phi^{1/2} \left(1+\frac{\phi}{\kappa}\right)^{1/2} \left(1+\frac{\kappa P^{2s-1}}{H^{s-1}} E\right)^{1/2}\right]^n \tag{3.9.1}$$

where $E = 1 + (RH^3\eta)^{s/2} + \frac{H^s}{(RP^2\eta)^{s/2}} P^2\eta$. Here we used $\min(1, P^2\eta) \leq P^2\eta$ which turns out to be sufficient.

Suppose we can show that $E \ll 1$. Recall that $\kappa \asymp \frac{(\log P)^2}{HP^2}$ so that

$$1 + \frac{\phi}{\kappa} \ll \frac{P^{\varepsilon}\eta}{\kappa}$$

from the definition (3.8.2) of η.

Since $\kappa \gg \frac{1}{P^s}$, both summands in the last bracket of (3.9.1) are bounded by $\frac{\kappa P^{2s-1}}{H^{s-1}}$. Still assuming $E \ll 1$, we then obtain

$$\Sigma(R,\phi) \ll \left[P^{\varepsilon} R^2 \phi^{1/2} \eta^{1/2} \frac{P^{s-\frac{1}{2}}}{H^{\frac{s-1}{2}}}\right]^n.$$

Recalling our desired bound $\Sigma(R,\phi) \ll P^{n(s-3)-\varepsilon}$, it now suffices to prove that

$$H^{s-1} \gg R^4 \phi \eta P^{5+\varepsilon},$$

still under the assumption $E \ll 1$. Putting $s = 14$ for convenience (as we may without loss of generality) and recalling the definition (3.8.2) of η, it suffices to have

$$H^{13} \gg R^4 \phi^2 P^{5+\varepsilon}$$

as well as

$$H^{14} \gg R^4 \phi P^{3+\varepsilon}.$$

We thus choose

$$H \asymp P^{\varepsilon} \max\left\{\left(R^4\phi^2P^5\right)^{1/13}, \left(R^4\phi P^3\right)^{1/14}, 1\right\}.$$

In order for this choice to satisfy $H \leq P$, we require $R^4\phi^2 \ll P^{8-\varepsilon}$ as well as $R^4\phi \ll P^{11-\varepsilon}$.

Recalling $\phi \leq \frac{1}{QR} \leq \frac{1}{R^2}$, both conditions are satisfied for any $Q \leq P^{3/2}$.
We have thus found an admissible choice of H, leading to a satisfactory estimate for $\Sigma(R,\phi)$ under the assumption of $E \ll 1$.
We now enquire under which circumstances this assumption is justified.
For convenience, denote $\phi_0 = (R^4P^{31})^{-\frac{1}{15}}$. The relevance of this parameter comes from the observation that for $\phi \leq \phi_0$, one has

$$H \asymp P^{\varepsilon} \max\left\{\left(R^4\phi P^3\right)^{1/14}, 1\right\}$$

and $\eta \asymp \frac{1}{HP^2}$ whereas for $\phi \geq \phi_0$, one has

$$H \asymp P^{\varepsilon} \max\left\{\left(R^4\phi^2 P^5\right)^{1/13}, 1\right\}$$

and $\eta \asymp \phi$.
To prove $E \ll 1$, we need $RH^3\eta \ll P^{-\varepsilon}$ as well as $\left(\frac{H^2}{RP^2\eta}\right)^7 P^2\eta \ll P^{-\varepsilon}$.
We begin by checking that $RH^3\eta \ll P^{-\varepsilon}$. First, if $\phi \leq \phi_0$, we have

$$\begin{aligned} RH^3\eta &\ll P^{\varepsilon}\frac{QH^2}{P^2} \\ &\ll P^{\varepsilon}\frac{Q}{P^2}\left(1 + (R^4\phi P^3)^{1/14}\right) \\ &\ll P^{\varepsilon} \cdot \left(\frac{Q}{P^2} + \frac{Q^{9/7}}{P^{11/7}}\right). \end{aligned}$$

This bound is satisfactory if $Q \ll P^{11/9-\varepsilon}$.
Next, if $\phi \geq \phi_0$, we have

$$\begin{aligned} RH^3\eta &\ll P^{\varepsilon}RH^3\phi \\ &\ll P^{\varepsilon} \cdot \frac{1}{Q} \cdot \left(1 + \left(R^4\phi^2P^5\right)^{3/13}\right) \\ &\ll P^{\varepsilon} \cdot \frac{P^{15/13}}{Q} \end{aligned}$$

which is satisfactory if $Q \gg P^{15/13+\varepsilon}$.

We thus choose $Q = P^{13/11}$, ensuring that $RH^3\eta \ll P^{-\varepsilon}$ in both cases, and noting that this also satisfies our earlier rough assumption $Q \leq P^{3/2}$.

Finally, we need to enquire whether $\left(\frac{H^2}{RP^2\eta}\right)^7 P^2\eta \ll P^{-\varepsilon}$.

For $\phi \leq \phi_0$, we have $\eta \asymp \frac{1}{HP^2}$ so that

$$\left(\frac{H^2}{RP^2\eta}\right)^7 P^2\eta \ll P^{\varepsilon}\frac{H^{20}}{R^7}$$

so that it suffices to have $H \ll R^{7/20-\varepsilon}$.

Recalling our choice of H in this case, it is thus sufficient to have $R \gg P^{\varepsilon}$ as well as additionally $\phi \leq \phi_1$ where

$$\phi_1 = R^{9/10}P^{-3-\varepsilon}.$$

Similarly, if $\phi \geq \phi_0$ we have $\eta \asymp \phi$ so that

$$\left(\frac{H^2}{RP^2\eta}\right)^7 P^2\eta \ll \frac{H^{14}}{R^7P^{12}\phi^6}$$

and hence by our definition of H, it suffices to have $R \gg P^{\varepsilon}$ as well as additionally $\phi \geq \phi_2$ where

$$\phi_2 = \frac{1}{P^{\frac{43}{25}-\varepsilon}R^{7/10}}.$$

Summarizing, we have obtained a satisfactory bound for $\Sigma(R, \phi)$ if $R \gg P^{\varepsilon}$ and $\phi \leq \min(\phi_0, \phi_1)$ or $\phi \geq \max(\phi_0, \phi_2)$.

Letting $R_0 = P^{4/5+\varepsilon}$, a quick computation shows that $\phi_2 \leq \phi_0 \leq \phi_1$ if $R \geq R_0$ whereas $P^{-\varepsilon}\phi_1 \leq \phi_0 \leq \phi_2 P^{\varepsilon}$ if $R \leq R_0$.

In the first case, our argument already covers all possible values of ϕ. We are thus left with the case where $R \leq R_0$ and $P^{-\varepsilon}\phi_1 \leq \phi \leq \phi_2 P^{\varepsilon}$ or $R \leq P^{\varepsilon}$.

It is here that we require the bound obtained by Weyl differencing. Indeed, applying Lemma 3.6.1 with $s = 14$ and noting that the assumption $Q \leq P^{3/2}$ is satisfied, we obtain

$$\Sigma(R, \phi) \ll P^{\varepsilon}\left[R^2\phi P^{14}\left(R\phi + \frac{1}{R\phi P^3}\right)^{7/4}\right]^n.$$

Recalling our goal $\Sigma(R,\phi) \ll P^{11n-\varepsilon}$, it then suffices to have

$$R^2\phi P^3 \left(R\phi + \frac{1}{R\phi P^3}\right)^{7/4} \ll P^{-\varepsilon}.$$

But this will be satisfied if

$$\frac{R^{1/3}}{P^{3-\varepsilon}} \ll \phi \ll \frac{1}{P^{12/11+\varepsilon}R^{15/11}}. \tag{3.9.2}$$

Under the assumption $R \leq R_0$ and $P^{-\varepsilon}\phi_1 \leq \phi \leq \phi_2 P^{\varepsilon}$, this will thus be true as soon as

$$\phi_1 \gg \frac{R^{1/3+\varepsilon}}{P^3}$$

as well as

$$\phi_2 \ll \frac{1}{P^{12/11+\varepsilon}R^{15/11}}.$$

The first condition is always satisfied for $R \gg P^{\varepsilon}$ while the second one is satisfied for $R \ll P^{\frac{346}{365}-\varepsilon}$ which is indeed true under the assumption $R \leq R_0$. Finally, we need to treat the cases where $R \leq P^{\varepsilon}$. Here of course, we need to use that we are on the minor arcs so that $\phi \geq P^{-3+\nu}$. But it is easy to see that in that case (3.9.2) is also satisfied, thus finishing our proof of Theorem 3.4.2. □

Chapter 4

Small solutions to homogeneous and inhomogeneous cubic equations

4.1 Introduction

In this chapter, we study the existence of integer solutions to equations $\phi(\mathbf{x}) = 0$ where $\phi \in \mathbb{Z}[x_1, \ldots, x_n]$ is a (not necessarily homogeneous) integer polynomial of degree 3. We denote the homogeneous parts of degrees $3, 2, 1$ and 0 by C, Q, L and N, respectively, so that we can write

$$\phi(\mathbf{x}) = C(\mathbf{x}) + Q(\mathbf{x}) + L(\mathbf{x}) + N.$$

Much work has been done in the case of homogeneous equations where Heath-Brown [16] proved that for $n \geq 14$, non-trivial solutions $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ always exist. Conjecturally, this should be true already in the wider range $n \geq 10$. It is known that no congruence obstructions appear for $n \geq 10$ and that the bound is sharp in this respect.

In the inhomogeneous case, the situation is more complicated. First of all, it is easy to produce examples of cubic polynomials in any number of variables which have a congruence obstruction, e.g.

$$\phi(\mathbf{x}) = 2C_1(\mathbf{x}) + 1$$

where $C_1(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$ is an arbitrary cubic form. It is therefore clearly necessary to stipulate the following *Necessary Congruence Condition*:
(NCC) For any prime power p^k, the equation $\phi(\mathbf{x}) \equiv 0 \pmod{p^k}$ has a solution.
Slightly more subtly, this necessary condition is still not sufficient in the inhomogeneous case, even in the case of many variables, as the following example of Watson shows: The equation

$$\phi(\mathbf{x}) = (2x_1 - 1)(1 + x_1^2 + x_2^2 + \cdots + x_n^2) + x_1x_2$$

is easily seen to satisfy the NCC as soon as $n \geq 5$, but clearly does not admit any integral solution as the absolute value of the first term is always at least $1 + x_1^2 + x_2^2$ and hence larger than the absolute value of the second term.
It is therefore necessary to make further assumptions on the polynomial ϕ. Hitherto, this has been realized in two variations:
Browning and Heath-Brown [5], following work of Heath-Brown [17] in the homogeneous case, have shown that a solution to $\phi(\mathbf{x}) = 0$ exists if ϕ satisfies the NCC, $n \geq 10$ and the cubic part C is non-singular.
Instead, we will focus on an older approach by Davenport and Lewis [12] who introduced the h-invariant of a cubic form C to be the least positive integer such that

$$C(\mathbf{x}) = \sum_{i=1}^{h} L_i(\mathbf{x})Q_i(\mathbf{x})$$

for appropriate linear forms $L_1, \dots, L_h$ and quadratic forms $Q_1, \dots, Q_h$. Equivalently, $n - h$ is the largest dimension of a linear subspace contained in the cubic hypersurface defined by the equation $C(\mathbf{x}) = 0$.

Davenport and Lewis then managed to show, following work of Davenport in the homogeneous case, that a solution to $\phi(\mathbf{x}) = 0$ exists if ϕ satisfies the NCC and $h(C) \geq 17$.

Pleasants [24] extended this to cover the range $h(C) \geq 16$.

Given the progress on Davenport's work [10] in the homogenous case made by Heath-Brown [16], it is natural to ask whether his method can be applied to extend the range to $h(C) \geq 14$.

We achieve this goal, indeed proving an asymptotic formula for the number of solutions in a bounded region. To this end, let us define the counting function

$$N(P) = N(P, \mathcal{B}) = \#\{\mathbf{x} \in \mathbb{Z}^n \cap P\mathcal{B} : \phi(\mathbf{x}) = 0\},$$

where $\mathcal{B} \subset \mathbb{R}^n$ is a box of the shape

$$\mathcal{B} = \prod_{1 \leq i \leq n} [z_i - 1, z_i + 1]$$

for some vector $\mathbf{z} \in \mathbb{R}^n$ with $|\mathbf{z}| \geq 2$ (so that $\mathcal{B}$ does not contain the origin) and $P\mathcal{B} = \{P\mathbf{x} : \mathbf{x} \in \mathcal{B}\}$. We then prove the following:

Theorem 4.1.1. *Assume that $\phi = C + Q + L + N$ is of degree 3, non-degenerate, satisfies the NCC and that $h(C) \geq 14$. If the centre $\mathbf{z}$ of the box $\mathcal{B}$ is a suitable non-singular point of the hypersurface $C(\mathbf{x}) = 0$, then*

$$N(P) = (1 + o(1))\mathfrak{S} \cdot \mathfrak{I} \cdot P^{n-3}, \quad P \to \infty$$

where $\mathfrak{S}$ and $\mathfrak{I}$ are the usual singular series and singular integral, respectively. We have $\mathfrak{S} > 0$ and $\mathfrak{I} > 0$. In particular, there is a solution $\mathbf{x} \in \mathbb{Z}^n$ to $\phi(\mathbf{x}) = 0$.

Note that we do not assume that C is non-degenerate, but that C does not vanish identically. Also note that the restriction $h(C) \geq 14$ automatically implies $n \geq 15$ since Heath-Brown's result is equivalent to saying that a cubic form in 14 variables has $h(C) \leq 13$.

In the culmination of a long series of papers, Watson [28] has established the existence of integer solutions to $\phi(\mathbf{x}) = 0$ under the assumptions $4 \leq h(C) \leq n - 3$ and $n \geq 15$. Combining this with Theorem 4.1.1, we obtain the following:

Theorem 4.1.2. *Assume that $\phi = C + Q + L + N$ satisfies the NCC and that $h(C) \geq 4$ and $n \geq 16$. Then there is a solution $\mathbf{x} \in \mathbb{Z}^n$ to the equation $\phi(\mathbf{x}) = 0$.*

Indeed, under the assumption $n \geq 16$, one of the conditions $h(C) \leq n - 3$ and $h(C) \geq 14$ is always satisfied, allowing us to deduce Theorem 4.1.2. Following a line of investigation revitalized by the work of Browning, Dietmann and Elliott [4], we also provide a bound on the smallest solution $\mathbf{x}$ in Theorem 4.1.1 uniform in the coefficients of the polynomial ϕ:

Theorem 4.1.3. *With the same assumptions as in Theorem 4.1.1, there exists a vector $\mathbf{x} \in \mathbb{Z}^n$ with $\phi(\mathbf{x}) = 0$ and*

$$\max_i |x_i| \ll M^{6407n^2},$$

where M is the maximum of the absolute values of the coefficients of ϕ. The implicit constant is absolute.

It is natural to ask whether Watson's method can be made uniform in the coefficients as well, allowing us to obtain a uniform version of Theorem 4.1.2. However, due to the intricate structure of Watson's argument, it is not immediately transparent to the author whether such an extension is possible. Our method also delivers new results on the smallest solution in the homogeneous case:

Theorem 4.1.4. *Let $n \geq 14$ and let $C \in \mathbb{Z}[x_1, \ldots, x_n]$ be a cubic form. Then there exists a vector $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ with $C(\mathbf{x}) = 0$ and*

$$\max_i |x_i| \ll M^{132484}.$$

If additionally we assume C to be non-singular, then for $n = 14$ we can ensure that

$$\max_i |x_i| \ll M^{2049}.$$

This improves on the work of Browning-Dietmann-Elliott [4], who had the same results for $n \geq 17$ and with the worse exponents 360000 in the general and 1071 in the non-singular case. However, as we will explain in more detail later, the result in [4], in particular the lower bound for the singular series, relies on a computational mistake which, if corrected, would yield an exponent larger than our exponent 2049 and also larger than what our method would yield for $n = 17$.

It also improves on work in the author's Master Thesis (unpublished) in which the result for general cubic forms with $n \geq 14$ was established with the slightly larger exponent 141718. The improvement here comes from a slight variation in the van der Corput differencing argument as well as an improved treatment of the singular integral.

We already record at this point for later convenience that it suffices to prove Theorem 4.1.4 for $n = 14$, as we may always set variables to zero. We may also assume that C is non-degenerate since otherwise the result is trivial.

In the result for non-singular forms, we have restricted to the case $n = 14$ for convenience. It is clear that the same method would also give similar results for all $n \geq 14$, but here the general case does not immediately reduce to the special case $n = 14$, as setting variables to zero could leave us a with a singular form.

We also remark that using the Kloosterman refinement of the circle method, the existence of non-trivial solutions to non-singular cubic forms is known already for $n \geq 10$ due to work of Heath-Brown [17] and – under the assumption of local solubility – even for $n = 9$ due to work of Hooley [18]. It would be interesting to work out a bound for the smallest solution using their method and compare it with our result.

Notation

We use $e(\alpha) = e^{2\pi i\alpha}$ and $e_q(x) = e\left(\frac{x}{q}\right)$ and the notations $\mathcal{O}$ and $\ll$ due to Landau and Vinogradov, respectively. For a subset $\mathcal{A} \subset \mathbb{R}^n$, we use the summation condition $\sum_{\mathbf{x}\in\mathcal{A}}$ as a shorthand for $\sum_{\mathbf{x}\in\mathcal{A}\cap\mathbb{Z}^n}$ i.e. we sum over all the lattice points in the set. For such a vector $\mathbf{x}$ we write $|\mathbf{x}| := \max_i |x_i|$. For a positive number $X > 0$, we shall also use the notation $\sum_{\mathbf{x}\le X}$ to mean $\sum_{0<x_1,\dots,x_n\le X}$ and similarly $\sum_{\mathbf{x}(q)}$ to mean $\sum_{x_1,\dots,x_n \pmod q}$.

The letter ε stands for a sufficiently small positive real number, which by convention may change its value finitely many times. In particular, we may write something like $M^{2\varepsilon} \ll M^{\varepsilon}$. Implicit constants are usually allowed to depend on ε. If we want to stress that it may depend on a parameter d, we write $\mathcal{O}_d$ instead of $\mathcal{O}$.

Setup

The general strategy to estimate the counting function $N(P)$ (and hence to prove the existence of integer solutions to $\phi(\mathbf{x}) = 0$) is to use the Hardy-Littlewood Circle Method. To this end, let

$$S(\alpha) = \sum_{\mathbf{x}\in P\mathcal{B}} e(\alpha\phi(\mathbf{x})).$$

It is then clear by orthogonality that

$$N(P) = \int_0^1 S(\alpha)d\alpha.$$

The next step is to dissect the interval $[0,1]$ of integration into two sets $\mathfrak{M}$ and $\mathfrak{m}$, the major and minor arcs, respectively. In our setup, the major arcs will be defined by taking

$$\mathfrak{M}(a,q) = \left[\frac{a}{q} - \frac{u}{P^3}, \frac{a}{q} + \frac{u}{P^3}\right]$$

for coprime integers a and q with $0 \leq a < q \leq P_0$ and then defining

$$\mathfrak{M} = \bigcup_{q \leq P_0} \bigcup_{(a;q)=1} \mathfrak{M}(a,q)$$

as the major arcs, and the complement $\mathfrak{m} := [0,1] \backslash \mathfrak{M}$ as the minor arcs. Here, u and P_0 are certain parameters which will be chosen as fixed powers of P eventually. In the proofs of Theorem 4.1.3 and 4.1.4 we will also choose P as a fixed power of M allowing us to deduce $N(P) > 0$ and thus establishing the existence of a small solution. We note that these results are certainly true for bounded M (by the above cited results) and so we may assume that M is sufficiently large, if necessary. In the case of Theorem 4.1.1, we will fix M and let $P \to \infty$.
We note at this point that under the harmless assumption $\frac{2u}{P^3} < \frac{1}{P_0^2}$, the arcs $\mathfrak{M}(a,q)$ will be disjoint. We denote the equivalent assumption $2P_0^2 u < P^3$ by $(\mathfrak{M}_1)$ for future reference. Since several more such assumptions will be added in the course of our work, a list of all of them is maintained at the end of the chapter for the convenience of the reader.

We denote the cubic form by $C(\mathbf{x}) = \sum_{i,j,k} c_{ijk} x_i x_j x_k$ where we assume the c_{ijk} to be symmetric and integral (as this can be achieved by rescaling the equation with a factor of 6 if necessary).
For further use let us denote the Hessian matrix of C by $M(\mathbf{x})$, where

$$M(\mathbf{x})_{ij} = \sum_k c_{ijk} x_k,$$

so that the entries of $M(\mathbf{x})\mathbf{y}$ are given by the bilinear forms

$$B_i(\mathbf{x}, \mathbf{y}) = \sum_{j,k} c_{ijk} x_j y_k.$$

We also denote by $r(\mathbf{x})$ the rank of $M(\mathbf{x})$ and by $r_p(\mathbf{x})$ the $\mathbb{F}_p$-rank of $M(\mathbf{x})$ for a prime p.

It will be convenient to assume that the coefficient c_{111} is positive and $\gg M$ and that none of the second partial derivatives of ϕ vanish identically. Both of this can always be achieved by a suitable change of coordinates with bounded coefficients.

4.2 Davenport's Geometric Condition

For general cubic forms, an asymptotic formula of the shape $N(P) \asymp P^{n-3}$ cannot always be true. On the technical side, to successfully bound the contribution from the minor arcs $\mathfrak{m}$, we need to be able to bound the number of solutions to a certain auxiliary system of bilinear equations.

To this end, slightly varying a definition of [4], let us say that a cubic form C is *ψ-good* if the assumption

$$\#\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| < H, r(\mathbf{x}) = r\} \ll H^{n-14+r+\varepsilon} \tag{4.2.1}$$

for all $0 \leq r \leq n$, holds uniformly in the range $1 \leq H \leq M^{\psi}$. Note that this estimate is trivially true for $r \geq 14$.

To relate this to the assumptions in our results, we require the following observations: The first is Lemma 28 in [19]:

Lemma 4.2.1. *If C is non-singular, then*

$$\dim\{\mathbf{x} : r(\mathbf{x}) \leq r\} \leq r.$$

Hence,

$$\#\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| < H, r(\mathbf{x}) = r\} \ll H^{r+\varepsilon}$$

uniformly over all H. In particular, C is ψ-good whenever $n \geq 14$.

Lemma 4.2.2. *If C is not ψ-good, then it vanishes on a $(n-13)$-dimensional subspace containing a non-zero element of size $\ll M^{97+91\psi}$. In particular, $h(C) \leq 13$.*

Proof. We follow the proof of Lemma 3 in [12]. By assumption, for some $H \leq M^{\psi}$ and for some r, there are more than $H^{n-14+r+\varepsilon}$ points $\mathbf{x}$ with $|\mathbf{x}| < H$ and $r(\mathbf{x}) = r$. Note that this already implies $r \leq 13$. This means that for some particular $r \times r$ minor of M, the number of points $\mathbf{x}$ with $r(\mathbf{x}) = r$ for which this particular minor does not vanish, is more than $H^{r+\varepsilon}$. We can then find $n-r$ linearly independent solutions $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(n-r)}$ for each such $\mathbf{x}$, which indeed depend polynomially on $\mathbf{x}$, since they can be expressed as certain $r \times r$ minors of $M(\mathbf{x})$.

These polynomial solutions satisfy

$$\sum_i \sum_k c_{ijk} x_i y_k^{(p)}(\mathbf{x}) = \Delta_{j,p}(\mathbf{x})$$

identically in $\mathbf{x}$ where $\Delta_{j,p}$ is a certain $(r+1) \times (r+1)$-minor of $M(\mathbf{x})$.

Differentiating this identity with respect to x_ν, then multiplying by $y_j^{(q)}$ and summing over j we obtain

$$\sum_j \sum_k c_{\nu jk} y_k^{(p)} y_j^{(q)} + \sum_k \Delta_{k,q} \frac{\partial y_k^{(p)}}{\partial x_\nu} = \sum_j y_j^{(q)} \frac{\partial \Delta_{j,p}}{\partial x_\nu} \tag{4.2.2}$$

for all $1 \leq \nu \leq n$ and $1 \leq p, q \leq n-r$.

Now, since all $\Delta_{k,p}$ vanish on more than $H^{n-14+r+\varepsilon}$ points $\mathbf{x}$ with $|\mathbf{x}| < H$ it follows that the dimension of the variety described by the vanishing of all these determinants must be at least $n - 13 + r$. In particular, there must be a point $\mathbf{x}$ where the rank of the Jacobian matrix of these derivatives is at most $13 - r$.

We now choose $\mathbf{x}$ as such a point. This means there are numbers $W_{j,p,\tau}$ and $U_{\tau,v}$ such that

$$\frac{\partial \Delta_{j,p}}{\partial x_\nu} = \sum_{\tau=1}^{13-r} W_{j,p,\tau} U_{\tau,\nu}.$$

The equation (4.2.2) now becomes

$$\sum_j \sum_k c_{\nu jk} y_k^{(p)} y_j^{(q)} = \sum_j y_j^{(q)} \sum_{\tau=1}^{13-r} W_{j,p,\tau} U_{\tau,\nu}.$$

Writing

$$\mathbf{Y} = T_1\mathbf{y}^{(1)} + \cdots + T_{n-r}\mathbf{y}^{(n-r)}$$

for indeterminates $T_1, \ldots, T_{n-r}$, multiplying the previous display by T_pT_q and summing over q we end up with

$$\sum_{j,k} c_{\nu jk} Y_j Y_k = \sum_{j} \sum_{p=1}^{n-r} \sum_{q=1}^{n-r} T_p T_q y_j^{(q)} \sum_{\tau=1}^{13-r} W_{j,p,\tau} U_{\tau,\nu} = \sum_{\tau=1}^{13-r} V_\tau U_{\tau,\nu}$$

for certain numbers V_τ. Multiplying by Y_ν and summing over ν we find that

$$C(\mathbf{Y}) = \sum_{\tau=1}^{13-r} V_\tau \sum_{p=1}^{n-r} T_p \sum_{\nu=1}^{n} y_\nu^{(p)} U_{\tau,\nu}.$$

Note that the interior double sum is a linear form in the T_p for each τ. If all of these $13-r$ linear forms vanish, we see that $C(\mathbf{Y}) = 0$. But this means that C vanishes on a linear subspace of dimension $(n-r)-(13-r) = n-13$ as desired and hence $h(C) \leq 13$.

Finally, we estimate the size of the smallest solution in this subspace. From their definition as $r \times r$-minors we have $|\mathbf{y}^{(p)}| \ll H^r M^r$. Moreover, the values $U_{\tau,\nu}$ can be chosen as values of $\frac{\partial \Delta_{j,p}}{\partial x_\nu}$ and hence are bounded by $M^{r+1}H^r$. The coefficients of the linear system for the T_i are therefore bounded by $M^{2r+1}H^{2r}$. By an application of Siegel's Lemma, such a system has a non-trivial solution with $T_i \ll (M^{2r+1}H^{2r})^{13-r}$. Finally, this means that there is a non-trivial solution $\mathbf{Y}$ of $C(\mathbf{Y}) = 0$ satisfying

$$Y \ll M^r H^r (M^{2r+1}H^{2r})^{13-r}.$$

It is now readily checked that this is bounded by $M^{97}H^{91}$ for all choices of r. Since $H \leq M^\psi$ by assumption, the result follows. □

Summarizing, we may therefore assume that C is ∞-good for the purpose of Theorems 4.1.1 and 4.1.3 as well as the non-singular case of 4.1.4. In the general case of Theorem 4.1.4, we may suppose that C is ψ-good for some suitable ψ as otherwise Lemma 4.2.2 allows us to deduce the existence of a relatively small solution.

4.3 The Major Arcs

We begin with the contribution from the major arcs. As remarked in [4], using Poisson's Summation Formula instead of the more elementary Euler Summation formula yields a better error term. For the usual problem of establishing an asymptotic formula, this improvement is irrelevant, but for the uniform version it changes the resulting exponent significantly.

The result of that approximation is the following lemma which is proved during the proof of Lemma 5 in [4].

Lemma 4.3.1. *Suppose that $f \in \mathbb{Z}[X_1, \dots, X_n]$ is a polynomial of degree $d \geq 3$ such that none of the partial derivatives $\frac{\partial^2 f}{\partial X_i^2}$ vanish identically.*
Let $\mathscr{C} = \prod_{i=1}^n [a_i, b_i]$ be a box and put $R_{\mathscr{C}} = \max_i |b_i - a_i|$. Let $\lambda \in \mathbb{R}$ and $\psi \in (0, 1]$ be chosen so that $|\lambda \nabla f(\mathbf{x})| \leq 1 - \psi$ for all $\mathbf{x} \in \mathscr{C}$. Then

$$\sum_{\mathbf{x} \in \mathscr{C}} e(\lambda f(\mathbf{x})) = \int_{\mathscr{C}} e(\lambda f(\mathbf{t})) \mathrm{d}\mathbf{t} + \mathcal{O}_d \left(\psi^{-1} R_{\mathscr{C}}^{n-1} \right).$$

We now wish to estimate the Weyl sum $S(\alpha)$, where $\alpha = \frac{a}{q} + \beta$ for some $q \leq P$. Sorting the initial sum by congruence classes modulo q, we get

$$\begin{aligned} S\left(\frac{a}{q} + \beta\right) &= \sum_{\mathbf{x} \in P\mathcal{B} \cap \mathbb{Z}^n} e\left(\left(\frac{a}{q} + \beta\right) \phi(\mathbf{x})\right) \\ &= \sum_{\mathbf{r}(q)} e\left(\frac{a\phi(\mathbf{r})}{q}\right) \sum_{\mathbf{y} \in \mathbb{Z}^n : \mathbf{r} + q\mathbf{y} \in P\mathcal{B}} e\left(\beta \phi(\mathbf{r} + q\mathbf{y})\right). \end{aligned}$$

We continue by applying Lemma 4.3.1 with $f(\mathbf{y}) = \phi(\mathbf{r} + q\mathbf{y})$, $\lambda = \beta$ and $\mathscr{C} = \{\mathbf{y} : \mathbf{r} + q\mathbf{y} \in P\mathcal{B}\}$, so that $R_{\mathscr{C}} = \frac{2P}{q}$. For $|\beta| \leq u$, the bound on the derivative will be satisfied with $\psi = \frac{1}{2}$ as soon as

$$u \cdot P_0 \cdot M \cdot |\mathbf{z}|^2 \ll P \qquad (\mathfrak{M}_2)$$

with a sufficiently small implicit constant. This yields

$$\begin{aligned} S\left(\frac{a}{q}+\beta\right) &= \sum_{\mathbf{r}(q)} e\left(\frac{a\phi(\mathbf{r})}{q}\right)\left(\int_{t:\mathbf{r}+q\mathbf{t}\in P\mathcal{B}} e\left(\beta\phi(\mathbf{r}+q\mathbf{t})\right)dt + \mathcal{O}\left(\frac{P^{n-1}}{q^{n-1}}\right)\right) \\ &= \frac{S(q,a)}{q^n}\int_{P\mathcal{B}} e(\beta\phi(\mathbf{t}))dt + \mathcal{O}\left(P^{n-1}q\right), \end{aligned}$$

where we write

$$S(q,a) = \sum_{\mathbf{r}(q)} e\left(\frac{a\phi(\mathbf{r})}{q}\right).$$

Integrating this approximation over $|\beta| \leq \frac{u}{P^3}$ and summing over a and q, we obtain

$$\begin{aligned} \int_{\mathfrak{M}} S(\alpha)d\alpha &= \mathfrak{S}(P_0)\int_{|\beta|\leq\frac{u}{P^3}}\int_{P\mathcal{B}} e(\beta\phi(\mathbf{t}))d\mathbf{t} + \mathcal{O}\left(P^{n-4}P_0^3u\right) \\ &= \mathfrak{S}(P_0)\cdot\int_{|\beta|\leq u}\int_{\mathcal{B}} e\left(\beta\frac{\phi(P\mathbf{t})}{P^3}\right)d\mathbf{t}\cdot P^{n-3} + \mathcal{O}\left(P^{n-4}P_0^3u\right) \\ &= \mathfrak{S}(P_0)\cdot\int_{|\beta|\leq u}\int_{\mathcal{B}} e\left(\beta C(\mathbf{t}) + \mathcal{O}\left(\frac{uM|\mathbf{z}|^2}{P}\right)\right)d\mathbf{t}\cdot P^{n-3} + \mathcal{O}\left(P^{n-4}P_0^3u\right) \\ &= \mathfrak{S}(P_0)\cdot\left(\mathfrak{I}(u) + \mathcal{O}\left(\frac{u^2M|\mathbf{z}|^2}{P}\right)\right)\cdot P^{n-3} + \mathcal{O}\left(P^{n-4}P_0^3u\right), \end{aligned}$$

with

$$\mathfrak{S}(P_0) := \sum_{q\leq P_0}\sum_{(a;q)=1}\frac{S(q,a)}{q^n}$$

and

$$\mathfrak{I}(u) := \int_{|\beta|\leq u}\int_{\mathcal{B}} e(\beta C(\mathbf{t}))d\mathbf{t}.$$

4.4 The singular integral

We now need to estimate the singular integral $\mathfrak{I}(u)$. As usual, this is done by first choosing the center of the box $\mathcal{B}$ to be a suitable non-singular point and then using Fourier's Inversion Theorem to show that $\mathfrak{I}(u)$ converges to a positive number $\mathfrak{I}$ as $u \to \infty$.

However, to obtain the desired uniform result, we also require a bound on the difference $|\mathfrak{I} - \mathfrak{I}(u)|$. In [4], a uniform version of Fourier's Inversion Theorem was cited from the thesis of Lloyd [23]. Since Lloyd's Thesis is not publically available and there actually was a small mistake in the application of his result, we decided to include a self-contained treatment of the singular integral, closely following the argument of Lloyd.
The key technical result is the following:

Lemma 4.4.1. *Suppose that $n \geq 2$ and that $\mathcal{B}$ is a box with center $\mathbf{z}$ and of width ρ and suppose that C satisfies*

$$\frac{\partial C}{\partial x_i} \geq \partial_i$$

for $i = 1, 2$ on all of $\mathcal{B}$ for some positive constants ∂_1, ∂_2. Then, for

$$\mathfrak{I}(Z) := \int_{|\beta| \leq Z} \int_{\mathcal{B}} e(\beta C(\mathbf{x})) d\mathbf{x} dt,$$

one has

$$\mathfrak{I}(Z) = V(0) \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sigma Z}\right)\right) + \mathcal{O}\left(\frac{\rho^{n-2}}{Z}\left(\frac{1}{\partial_1 \partial_2} + \frac{\rho|\mathbf{z}|M}{\partial_1^3} \log(\sigma Z)\right)\right)$$

whenever $\sigma Z \geq 2$, where $\sigma = \max_{\mathbf{x} \in \mathcal{B}} |C(\mathbf{x})|$.
We have the bound

$$V(0) \gg \frac{\rho^{n-1}}{M|\mathbf{z}|^2}. \tag{4.4.1}$$

Some remarks are in order: Compared to Lemma 7 in [4], our last error term is better. Indeed, as we will explain, it is the precise outcome of Lloyd's argument. It has the additional virtue of being scaling invariant: If we replace $\mathcal{B}$ by $T\mathcal{B}$ and Z by $\frac{Z}{T^3}$, then $\mathfrak{I}(Z)$ is multiplied by T^{n-3} and the same is true for all error terms (in contrast to the version in [4]).
Note that we have also refrained from writing the first term as $V(0) + \mathcal{O}\left(\frac{V(0)}{\sigma Z}\right)$. This is convenient because it means that we do not require an upper bound for $V(0)$ in order to deduce a lower bound for $\mathfrak{I}(Z)$.

Finally, note that the condition $\frac{\partial C}{\partial x_i} \geq \partial_i$ on all of $\mathcal{B}$ leads to a restriction on the size of ρ that was overlooked (and indeed not satisfied by the choice made) in [4].

Proof. We begin by interchanging the order of integration to write

$$\mathfrak{I}(Z) = \int_{\mathcal{B}} \frac{\sin 2\pi Z C(\mathbf{x})}{\pi C(\mathbf{x})} d\mathbf{x}.$$

The next step is to make the change of variables $(x_1, \ldots, x_n) \mapsto (t, x_2, \ldots, x_n)$ with $t = C(\mathbf{x})$. The Jacobian of this change of variable is given by $\frac{\partial C}{\partial x_1} \geq \partial_1 > 0$ and so this change of variables is invertible. If we denote by $g(t, x_2, \ldots, x_n)$ the coordinate x_1 of the inverse, then $0 < \frac{\partial g}{\partial t} = \frac{1}{\frac{\partial C}{\partial x_1}} \leq \frac{1}{\partial_1}$ and we can write

$$\mathfrak{I}(Z) = \int_{\mathcal{R}} \frac{\sin 2\pi Z t}{\pi t} \frac{\partial g}{\partial t}(t, x_2, \ldots, x_n) dt dx_2 \ldots dx_n = \int_{-\sigma}^{\sigma} \frac{\sin 2\pi Z t}{\pi t} V(t) dt$$

where

$$V(t) = \int_{(t, x_2, \ldots, x_n) \in \mathcal{R}} \frac{\partial g}{\partial t}(t, x_2, \ldots, x_n) dx_2 \ldots dx_n \tag{4.4.2}$$

and $\mathcal{R}$ is the image of $\mathcal{B}$ under the change of variables.

For later use, we also record the lower bound

$$\frac{\partial g}{\partial t} = \frac{1}{\frac{\partial C}{\partial x_1}} \gg \frac{1}{M|\mathbf{z}|^2}. \tag{4.4.3}$$

To make use of Fourier's Inversion Theorem, we need to estimate right and left derivatives of the function $V(t)$.

To this end, we write

$$V(t) = \int_{\prod_{i=3}^{n} [z_i - \rho, z_i + \rho]} \int_{x_2 \in \mathcal{R}_{t, x_3, \ldots, x_n}} \frac{\partial g}{\partial t}(t, x_2, \ldots, x_n) dx_2 dx_3 \ldots dx_n$$

where

$$\begin{aligned} &\mathcal{R}_{t, x_3, \ldots, x_n} = \{x_2 \in (z_2 - \rho, z_2 + \rho) : \exists x_1 \in (z_1 - \rho, z_1 + \rho) : t = C(\mathbf{x})\} \\ &= \{x_2 \in (z_2 - \rho, z_2 + \rho) : C(z_1 - \rho, x_2, \ldots, x_n) < t < C(z_1 + \rho, x_2, \ldots, x_n)\} \\ &= \{x_2 \in (z_2 - \rho, z_2 + \rho) : b^{(1)}_{x_3, \ldots, x_n}(x_2) < t < b^{(2)}_{x_3, \ldots, x_n}(x_2)\} \end{aligned}$$

with $b^{(1)}$ and $b^{(2)}$ defined by the last equation so that we have $\frac{\partial b^{(i)}}{\partial x_2} \geq \partial_2$ by assumption. In particular, we can write

$$\mathcal{R}_{t,x_3,\ldots,x_n} = (\ell^{(1)}_{x_3,\ldots,x_n}, \ell^{(2)}_{x_3,\ldots,x_n})$$

with $\ell^{(i)}_{x_3,\ldots,x_n}$ continuous everywhere and continuously differentiable with the exception of at most two points. At these two points, left and right derivatives exist and all of these derivatives satisfy $0 < \frac{\partial \ell^{(i)}}{\partial t} \leq \frac{1}{\partial_2}$. We now obtain

$$V(t) = \int_{\prod_{i=3}^{n}[z_i-\rho, z_i+\rho]} \int_{\ell^{(1)}}^{\ell^{(2)}} \frac{\partial g}{\partial t}(t, x_2, \ldots, x_n) dx_2 dx_3 \ldots dx_n.$$

Using Leibniz's rule, we now obtain that V has right and left derivatives everywhere with them disagreeing only at finitely many points. More precisely, the left and right derivatives of the inner integral are given by a linear combination of expressions of the form $\frac{\partial g}{\partial t} \cdot \frac{\partial^{\pm} \ell^{(i)}}{\partial t}$ as well as

$$\int_{\ell^{(1)}}^{\ell^{(2)}} \frac{\partial^2 g}{\partial t^2} dt.$$

The first type of expressions is bounded by $\mathcal{O}\left(\frac{1}{\partial_1 \partial_2}\right)$ and the second one by $\frac{\rho M|\mathbf{z}|}{\partial_1^3}$ on noting that

$$\frac{\partial^2 g}{\partial t^2} = -\frac{\frac{\partial^2 C}{\partial x_1}}{\left(\frac{\partial C}{\partial x_1}\right)^3} \ll \frac{M|\mathbf{z}|}{\partial_1^3}$$

(where we used that $\rho \ll |\mathbf{z}|$ and hence $|\mathbf{z}| + \rho \ll |\mathbf{z}|$ since the box clearly can't contain the origin). It now follows that

$$\frac{\partial^{\pm} V}{\partial t} \ll \rho^{n-2} \cdot \left(\frac{1}{\partial_1 \partial_2} + \frac{\rho M|\mathbf{z}|}{\partial_1^3}\right) =: A.$$

Letting

$$\Phi(t) = \frac{V(t) + V(-t) - 2V(0)}{t},$$

we now see that $\Phi(t) \ll A$ and $\Phi'(t) \ll \frac{A}{t}$ for all $t > 0$.

Finally, we are ready to estimate

$$\mathfrak{I}(Z) = 2V(0) \int_0^\sigma \frac{\sin 2\pi Zt}{\pi t} dt + \frac{1}{\pi} \int_0^\sigma \Phi(t) \sin 2\pi Z t dt.$$

The first integral is easily evaluated to $\frac{1}{2} + \mathcal{O}\left(\frac{1}{\sigma Z}\right)$. For the second integral, we split the range of integration into $t \leq \tau$ and $t \geq \tau$ for a suitable parameter $0 \leq \tau \leq \sigma$. The range $t \leq \tau$ contributes $\ll \tau A$. On the range $t \geq \tau$ we can integrate by parts to obtain

$$\int_\tau^\sigma \Phi(t) \sin 2\pi Z t dt = \left[-\Phi(t) \frac{\cos 2\pi Z t}{2\pi t} \right]_\tau^\sigma + \int_\tau^\sigma \Phi'(t) \frac{\cos 2\pi Z t}{2\pi t} dt \ll \frac{A}{Z} \left(1 + \log \frac{\sigma}{\tau} \right).$$

The main result of the lemma now follows upon choosing $\tau = \frac{1}{Z}$.
Finally, we note that (4.4.1) follows immediately from (4.4.2) and (4.4.3). □

To apply Lemma 4.4.1, we now need to choose a suitable non-singular point $\mathbf{z}$ as the center of our box. The strategy is similar to the one in the proof of Lemma 6 in [4], but we need a variant for the inhomogeneous case.

Lemma 4.4.2. *a) If $h = h(C)$, there is a solution $\widetilde{\mathbf{z}} = (\xi, \mathbf{y}) \in \mathbb{R}^n$ to $C(\mathbf{z}) = 0$ satisfying $|\mathbf{z}| \ll M^{\frac{1}{h-2}}$ and such that (possibly after relabeling the coordinates and changing signs)*

$$\frac{\partial C(\widetilde{\mathbf{z}})}{\partial x_1} \gg M^{-1-\frac{4}{h-2}}$$

and

$$\frac{\partial C(\widetilde{\mathbf{z}})}{\partial x_2} \gg M^{-2-\frac{7}{h-2}}.$$

b) Similarly, unless $C(\mathbf{x}) = 0$ has a non-trivial integer solution with $|\mathbf{x}| \ll M^{\frac{1}{n-2}}$, there is a solution $\widetilde{\mathbf{z}} = (\xi, \mathbf{y}) \in \mathbb{R}^n$ to $C(\widetilde{\mathbf{z}}) = 0$ satisfying $|\mathbf{z}| \ll M^{\frac{1}{n-2}}$ and such that (possibly after relabeling the coordinates and changing signs)

$$\frac{\partial C(\widetilde{\mathbf{z}})}{\partial x_1} \gg M^{-1-\frac{4}{n-2}}$$

and

$$\frac{\partial C(\widetilde{\mathbf{z}})}{\partial x_2} \gg M^{-2-\frac{7}{n-2}}.$$

Proof. We write the cubic form as

$$C(\mathbf{x}) = ax_1^3 + F_1 x_1^2 + F_2 x_1 + F_3$$

where as explained in the introduction we may assume that $a > 0$ and $a \gg M$. In the setting of b), by Siegel's Lemma, we can find a non-trivial integer solution $\mathbf{y}$ to $F_1(\mathbf{y}) = 0$ with $|\mathbf{y}| \ll M^{\frac{1}{n-2}}$. If $F_3(\mathbf{y}) = 0$, we have found the desired small integer solution $(0, \mathbf{y})$. Otherwise we may assume that $|F_3(\mathbf{y})| \geq 1$.

In the setting of a), we argue instead that we can find linearly independent integer solutions $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(n-h+1)}$ of $F_1(\mathbf{y}) = 0$, all of them satisfying $\mathbf{y}^{(i)} \ll M^{\frac{1}{h-2}}$. By definition of $h = h(C)$, one of them must have $F_3(\mathbf{y}) \neq 0$ and hence $|F_3(\mathbf{y})| \geq 1$.

From here on, the argument is identical in both cases, so we only treat a).

Flipping signs if necessary, we may assume that $F_3 = F_3(\mathbf{y}) \leq -1$. We may thus find a real zero $\xi > 0$ of $C(\xi, \mathbf{y}) = a\xi^3 + F_2\xi + F_3 = 0$ where we have written $F_2 = F_2(\mathbf{y})$. The next step is to establish bounds on ξ.

If $F_2 \geq 0$, we can use that $a\xi^3 \leq |F_3| \ll M^{1+\frac{3}{h-2}}$, so that $\xi \ll M^{\frac{1}{h-2}}$ and then

$$\xi \geq \frac{1}{a\xi^2 + F_2} \gg M^{-1-\frac{2}{h-2}}.$$

If $F_2 < 0$, we instead argue that $a\xi^3 \geq 1$, so that $\xi \gg M^{-1/3}$, $a\xi^3 = |F_2|\xi + |F_3|$ and

$$\xi \ll \left|\frac{F_2}{a}\right|^{1/2} + \left|\frac{F_3}{a}\right|^{1/3} \ll M^{\frac{1}{h-2}}.$$

In any case, we have thus established that

$$M^{-1-\frac{2}{h-2}} \ll \xi \ll M^{\frac{1}{h-2}}.$$

Finally, we need to bound the partial derivatives. We have

$$\frac{\partial C(\xi, \mathbf{y})}{\partial x_1} = 3a\xi^2 + F_2 = 2a\xi^2 - \frac{F_3}{\xi} \geq 2a\xi^2 \gg M^{-1-\frac{4}{h-2}}.$$

Using Euler's identity, we now find that

$$\left|y_2\frac{\partial C}{\partial x_2}+\cdots+y_n\frac{\partial C}{\partial x_n}\right|\geq\left|\xi\frac{\partial C}{\partial x_1}\right|-3|C(\xi,\mathbf{y})|\gg M^{-2-\frac{6}{h-2}}$$

and hence w.l.o.g. $\left|y_2\frac{\partial C}{\partial x_2}\right|\gg M^{-2-\frac{6}{h-2}}$, so that $\left|\frac{\partial C}{\partial x_2}\right|\gg M^{-2-\frac{7}{h-2}}$ as desired.

□

We now choose our box $\mathcal{B}$ with center $\mathbf{z}$ and width $\rho=1$ making sure that the assumptions in Lemma 4.4.1 are satisfied. To this end, we choose $\mathbf{z}=AM^{3+\frac{8}{n-2}}\widetilde{\mathbf{z}}$ with $\widetilde{\mathbf{z}}$ as in Lemma 4.4.2 and a sufficiently large constant $A>0$. With the choice of $h=14$ or $n=14$, respectively, we record the properties of this choice in the following lemma.

Lemma 4.4.3. *The point $\mathbf{z}\in\mathbb{R}^n$ is a solution of $C(\mathbf{z})=0$ satisfying $|\mathbf{z}|\ll M^{3.75}$,*

$$\frac{\partial C}{\partial x_1}\gg M^6$$

and

$$\frac{\partial C}{\partial x_2}\gg M^{4.75}$$

on all of $\mathcal{B}$.

For the proof, we only need to note that the bounds for the derivatives at the point $\mathbf{z}$ (which are obtained directly from the bounds for $\widetilde{\mathbf{z}}$ by scaling) extend over all of $\mathcal{B}$ as

$$\frac{\partial C}{\partial x_i}=\frac{\partial C(\mathbf{z})}{\partial x_i}+\mathcal{O}(M|\mathbf{z}|).$$

It is here that we make use of the fact that A is sufficiently large.
It is clear that we may assume that $\frac{\partial C}{\partial x_i}\ll\frac{\partial C}{\partial x_1}$ for all i at $\mathbf{z}$ and then on all of $\mathcal{B}$ as otherwise we can simply permute the variables.
Finally, we can collect the results of this section in the following lemma.

Lemma 4.4.4. *With the box $\mathcal{B} = \mathcal{B}(\mathbf{z})$ chosen as above, we have*

$$\mathfrak{I}(u) = \mathfrak{I} \cdot \left(1 + \mathcal{O}\left(\frac{1}{M^{12.25}u}\right)\right) + \mathcal{O}\left(\frac{1}{uM^{10.75}}\right)$$

for some number $\mathfrak{I} > 0$ with

$$\mathfrak{I} \gg \frac{1}{M^{8.5}}.$$

In particular, we have

$$\mathfrak{I}(u) \gg \frac{1}{M^{8.5}}$$

for any $u \geq 1$.
Under the assumption of $(\mathfrak{M}_1)$ and $(\mathfrak{M}_2)$ as well as

$$u^2 M^{17+\varepsilon} \ll P, \tag{$\mathfrak{I}_1$}$$

it follows that

$$\int_{\mathfrak{M}} S(\alpha) d\alpha = (1 + o(1))\mathfrak{S}(P_0) \cdot \mathfrak{I}(u) \cdot P^{n-3} + \mathcal{O}\left(P^{n-4}P_0^3 u\right).$$

To discuss the singular series $\mathfrak{S}(P_0)$, we need estimates on the Gauß Sums $S(q,a)$, a problem which is related to bounding the Weyl Sum $S(\alpha)$ on the minor arcs. We therefore continue with the discussion of the minor arcs and return to discuss the singular series at the appropriate point.

4.5 The minor arcs

As in Heath-Brown's work [16], we make use of three different methods to bound $S(\alpha)$ on the minor arcs. First of all, there is the classical Weyl differencing method, that was used by Davenport to obtain his result for 16 variables. While alone it is therefore insufficient for our purposes, it still outperforms the other methods in certain regimes of the minor arcs and therefore remains a crucial ingredient.

The second method is the pointwise van der Corput differencing. This improves on the Weyl differencing, but again is not good enough to save more variables. We shall only require it to bound the Gauß Sums $S(q, a)$ and thus for the convergence of the singular series.

Finally, the third method is the mean-square average version of van der Corput differencing which is the key innovation of Heath-Brown in allowing us to obtain results for 14 variables.

Since in all methods, lower order terms of ϕ disappear in the course of differencing, the results of this section are essentially identical with those from the homogeneous case.

4.5.1 Preliminaries

We begin by recalling the general strategy implicit already in Davenport's work:

By an appropriate combination of squaring and Cauchy-Schwarz, one reduces the cubic exponential sum to one over a linear form. While in the classical case of a diagonal cubic form, the resulting sum is easy to handle, the general shape of a cubic form begins to cause problems.

In general, this step allows us to reduce a bound for $S(\alpha)$ to one for the number of solutions to a system of certain auxiliary diophantine inequalities involving the bilinear forms $B_i(\mathbf{x}, \mathbf{y})$.

These diophantine inequalities are dealt with by an application of Davenport's Shrinking Lemma. Roughly speaking, the Shrinking Lemma allows us to bootstrap the diophantine inequalities in a way that forces equality.

Finally, the number of solutions to the resulting system of auxililary diophantine equations involving the bilinear forms $B_i(\mathbf{x}, \mathbf{y})$ can be estimated using Davenport's Geometric Condition which for us is captured by the assumption (4.2.1) that C is ψ-good.

At this point, we record two of the mentioned key ingredients. The first one is Davenport's Shrinking Lemma (see e.g. [16], Lemma 2.2).

Lemma 4.5.1. *Let $L \in M_n(\mathbb{R})$ be a real symmetric $n \times n$ matrix. Let $a > 0$ be real, and let*

$$N(Z) := \#\{\mathbf{u} \in \mathbb{Z}^n : |\mathbf{u}| \leq aZ, \|(L\mathbf{u})_i\| < a^{-1}Z, 1 \leq i \leq n\}.$$

Then if $0 < Z \leq 1$, we have

$$N(1) \ll_n Z^{-n} N(Z).$$

The second one is Lemma 2.3 from [16] and will allow us to deduce that a sufficiently strong diophantine inequality already forces equality or at least a divisibility condition:

Lemma 4.5.2. *Let a real number $X \geq 0$ be given and let $\alpha = \frac{a}{q} + \theta$ with $(a;q) = 1$ and $2qX|\theta| \leq 1$. Suppose that $m \in \mathbb{Z}$ is such that $|m| \leq X$ and $\|\alpha m\| \leq \frac{1}{P_1}$ for some $P_1 \geq 2q$. Then $q \mid m$. In particular we will have $m = 0$ if in addition $X < q$ or $|\theta| > \frac{1}{qP_1}$.*

4.5.2 Weyl Differencing

Recall the definition

$$S(\alpha) = \sum_{\mathbf{x} \in P\mathscr{B}} e(\alpha\phi(\mathbf{x})).$$

This leads to the identity

$$|S(\alpha)|^2 = \sum_{\mathbf{x},\mathbf{y} \in P\mathscr{B}} e(\alpha(\phi(\mathbf{y}) - \phi(\mathbf{x}))).$$

Writing $\mathbf{y} = \mathbf{x} + \mathbf{d_1}$ we can rewrite this as

$$|S(\alpha)|^2 = \sum_{\mathbf{d_1}} \sum_{\mathbf{x} \in \mathcal{R}(\mathbf{d_1})} e(\alpha(\phi(\mathbf{x}+\mathbf{d_1}) - \phi(\mathbf{x}))),$$

where $\mathcal{R}(\mathbf{d_1}) = P\mathscr{B} \cap (P\mathscr{B} - \mathbf{d_1})$. In particular, the inner sum is empty unless $|\mathbf{d_1}| \leq 2P$. Squaring again and applying Cauchy-Schwarz, we then find that

$$|S(\alpha)|^4 \ll P^n \sum_{\mathbf{d_1}} \sum_{\mathbf{x},\mathbf{z} \in \mathcal{R}(\mathbf{d_1})} e(\alpha(\phi(\mathbf{z}+\mathbf{d_1}) - \phi(\mathbf{z}) - \phi(\mathbf{x}+\mathbf{d_1}) + \phi(\mathbf{x}))).$$

Writing $\mathbf{z} = \mathbf{x} + \mathbf{d_2}$ this can be rewritten as

$$|S(\alpha)|^4 \ll P^n \sum_{\mathbf{d_1},\mathbf{d_2}} \sum_{\mathbf{x} \in \mathcal{S}(\mathbf{d_1},\mathbf{d_2})} e(\alpha C(\mathbf{d_1},\mathbf{d_2},\mathbf{x})) \qquad (4.5.1)$$

where $\mathcal{S}(\mathbf{d_1},\mathbf{d_2}) = \mathcal{R}(\mathbf{d_1}) \cap (\mathcal{R}(\mathbf{d_1}) - \mathbf{d_2})$ and

$$C(\mathbf{d_1},\mathbf{d_2},\mathbf{x}) = \phi(\mathbf{x}+\mathbf{d_1}+\mathbf{d_2}) - \phi(\mathbf{x}+\mathbf{d_1}) - \phi(\mathbf{x}+\mathbf{d_2}) + \phi(\mathbf{x}).$$

Note that the notation $C(\mathbf{d_1},\mathbf{d_2},\mathbf{x})$ is appropriate as the lower order terms of ϕ have disappeared at this point.

All we need to know about $\mathcal{S}(\mathbf{d_1},\mathbf{d_2})$ is that it is a certain box inside $P\mathscr{B}$. Further, note that

$$C(\mathbf{d_1},\mathbf{d_2},\mathbf{x}) = 6\sum_{i=1}^{n} x_i B_i(\mathbf{d_1},\mathbf{d_2}) + \psi(\mathbf{d_1},\mathbf{d_2}),$$

where $\psi(\mathbf{d_1},\mathbf{d_2})$ is independent of $\mathbf{x}$. We now recall the standard bound for linear exponential sums

$$\sum_{x \in I} e(\alpha x) \ll \min(|I|, \|\alpha\|^{-1}),$$

where $I \subset \mathbb{R}$ is any interval and $\|\alpha\| = \min_{n\in\mathbb{Z}} |\alpha - n|$. From this and the previous discussion, it now follows that

$$|S(\alpha)|^4 \ll P^n \sum_{|\mathbf{d_1}|,|\mathbf{d_2}| \leq 2P} \prod_{i=1}^{n} \min\left(P, \|6\alpha B_i(\mathbf{d_1},\mathbf{d_2})\|^{-1}\right). \qquad (4.5.2)$$

The next step is to compute the sum over $\mathbf{d_1}$ and $\mathbf{d_2}$ or rather relate it to the previously mentioned number of solutions to a certain system of inequalities. To this end, let

$$N(\mathbf{d}) = \#\{\mathbf{x} \in \mathbf{Z}^n : |\mathbf{d}| \leq 4P, \|6\alpha B_i(\mathbf{d}, \mathbf{x})\| < \frac{1}{4P}\}.$$

It then follows that for fixed $\mathbf{d_1}$ and integers $r_1, \ldots, r_n$ with $0 \leq r_i < 4P$, there are at most $N(\mathbf{d_1})$ values of $\mathbf{d_2}$ with $|\mathbf{d_2}| \leq 2P$ satisfying

$$\frac{r_i}{4P} \leq \{6\alpha B_i(\mathbf{d_1}, \mathbf{d_2})\} < \frac{r_i + 1}{4P},$$

because for any two such vectors $\mathbf{d_2}$ and $\mathbf{d_2}'$ their difference $\mathbf{d} = \mathbf{d_2} - \mathbf{d_2'}$ must be in the set counted by $N(\mathbf{d_1})$. This yields the estimate

$$\begin{aligned} |S(\alpha)|^4 &\ll P^n \sum_{\mathbf{d_1}} N(\mathbf{d_1}) \sum_{r_1=0}^{4P} \cdots \sum_{r_n=0}^{4P} \prod_{i=1}^{n} \min\left(P, \frac{4P}{r_i}\right) \\ &\ll P^{2n} (\log P)^n \sum_{\mathbf{d_1}} N(\mathbf{d_1}) \end{aligned}$$

so that

$$|S(\alpha)|^4 \ll P^{2n+\varepsilon} \# \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{Z}^{2n} : |\mathbf{x}|, |\mathbf{y}| \leq 4P, \|6\alpha B_i(\mathbf{x}, \mathbf{y})\| < \frac{1}{4P} \right\}.$$

An application of the Shrinking Lemma 4.5.1 now leads to the estimate

$$|S(\alpha)|^4 \ll Z^{-n} P^{2n+\varepsilon} \# \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{Z}^{2n} : |\mathbf{x}| \leq 4P, |\mathbf{y}| \leq 4PZ, \|6\alpha B_i(\mathbf{x}, \mathbf{y})\| < \frac{Z}{4P} \right\}.$$

Reversing the rôles of $\mathbf{x}$ and $\mathbf{y}$ and applying the argument again with slightly different parameters, we arrive at

$$|S(\alpha)|^4 \ll Z^{-2n} P^{2n+\varepsilon} \# \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{Z}^{2n} : |\mathbf{x}|, |\mathbf{y}| \leq 4PZ, \|6\alpha B_i(\mathbf{x}, \mathbf{y})\| < \frac{Z^2}{4P} \right\}. \tag{4.5.3}$$

We now need to choose Z sufficiently small so that Lemma 4.5.2 allows us to conclude $B_i(\mathbf{x}, \mathbf{y}) = 0$.

Here we choose $m = 6B_i(\mathbf{x}, \mathbf{y})$ so that $X \asymp M(PZ)^2$ and $P_1 \asymp \frac{P}{Z^2}$. Thus any choice of Z satisfying $Z \leq 1$ as well as

$$2q|\theta|M(PZ)^2 \ll 1 \quad \text{ and } \quad Z^2 q \ll P$$

as well as

$$M(PZ)^2 \ll q \quad \text{ or } \quad Z^2 \ll |\theta|qP$$

with sufficiently small implicit constants allows us to conclude that

$$|S(\alpha)|^4 \ll Z^{-2n} P^{2n+\varepsilon} \# \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{Z}^{2n} : |\mathbf{x}|, |\mathbf{y}| \leq 4PZ, B_i(\mathbf{x}, \mathbf{y}) = 0 \right\}. \tag{4.5.4}$$

In the end we will choose Z as big as possible, subject to the conditions above, but before making this choice let us see how to estimate the RHS in (4.5.4). At this point we need the assumption that C is ψ-good. Recall that this means that the estimate

$$\#\{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}| < H, r(\mathbf{x}) = r\} \ll H^{n-14+r+\varepsilon} \tag{4.5.5}$$

holds uniformly in $1 \leq H \leq M^{\psi}$ where $r(\mathbf{x})$ is the rank of the matrix $M(\mathbf{x})$. This is clearly related to the system of equations we are studying by the fact that the condition $B_i(\mathbf{x}, \mathbf{y}) = 0$ for all i is equivalent to $M(\mathbf{y})\mathbf{x} = \mathbf{0}$.

But if $\mathbf{y}$ is fixed, the number of $\mathbf{x}$ with $|\mathbf{x}| \leq 4PZ$ and $M(\mathbf{y})\mathbf{x} = 0$ is $\mathcal{O}\left((ZP)^{n-r}\right)$. Hence with $H \asymp PZ$ we find that

$$\begin{aligned} |S(\alpha)|^4 &\ll Z^{-2n} P^{2n+\varepsilon} \sum_{r=0}^{n} \sum_{|\mathbf{w}| \ll H, r(\mathbf{w}) = r} (PZ)^{n-r} \\ &\ll Z^{-2n} P^{2n+\varepsilon} \sum_{r=0}^{14} H^{n-14+r+\varepsilon} (PZ)^{n-r} \\ &\ll P^{4n+\varepsilon} (PZ)^{-14}, \end{aligned}$$

assuming $PZ \ll M^{\psi}$. Here we needed to assume that $PZ \gg 1$ but the final result is trivially true if this assumption fails to be correct. Finally, from this result it is clear that the optimal choice for Z is indeed the maximal one subject to the above conditions. This choice is

$$Z \asymp \min\left(1, \frac{1}{(q|\theta|M)^{\frac{1}{2}}P}, \left(\frac{P}{q}\right)^{\frac{1}{2}}, \frac{M^{\psi}}{P}, \max\left(\frac{q^{\frac{1}{2}}}{M^{\frac{1}{2}}P}, (qP|\theta|)^{\frac{1}{2}}\right)\right)$$

and leads to the following final result.

Lemma 4.5.3. *Assume that C is ψ-good. If $\alpha = \frac{a}{q} + \theta$ for coprime integers $0 \leq a < q$, then*

$$S(\alpha) \ll P^{n+\varepsilon}\left(\frac{1}{P^2} + Mq|\theta| + \frac{q}{P^3} + \frac{1}{q}\min\left(M, \frac{1}{|\theta|P^3}\right) + M^{-2\psi}\right)^{\frac{7}{4}}.$$

4.5.3 A pointwise bound via van der Corput

In this section, we derive a bound for $S(q,a)$ using the version of van der Corput's method from [16] instead of Weyl differencing. To this end, we will temporarily put $\mathscr{B} = (0,1]^n, P = q$ and $\alpha = \frac{a}{q}$ with $(a;q) = 1$, so that $S(\alpha) = S(q,a)$. Of course the arguments in this section can be developed in a much broader context (see [16] for more details), but since in their rôle as bounds for $S(\alpha)$ on the minor arcs they will be insufficient and in fact superseded by the results from the next section, we content ourselves with the treatment of this special case. The basic idea is to write

$$S(q,a) = H^{-n} \sum_{\mathbf{h} \leq H} \sum_{\mathbf{x}:\mathbf{x}+\mathbf{h} \leq q} e_q(a\phi(\mathbf{x}+\mathbf{h})),$$

where $1 \leq H \leq q$ is a suitable parameter. Interchanging the order of summation, this yields

$$S(q,a) = H^{-n} \sum_{\mathbf{x} \in \mathbb{Z}^n} \sum_{\substack{\mathbf{h} \leq H: \\ \mathbf{x}+\mathbf{h} \leq q}} e_q(a\phi(\mathbf{x}+\mathbf{h})).$$

Note that the inner sum is non-empty only for $\mathcal{O}(q^n)$ vectors $\mathbf{x}$, due to the condition $H \leq q$. Hence an application of Cauchy-Schwarz leads to

$$|S(q,a)|^2 \ll H^{-2n} q^n \sum_{\mathbf{x} \in \mathbb{Z}^n} \left| \sum_{\substack{\mathbf{h} \leq H: \\ \mathbf{x}+\mathbf{h} \leq q}} e_q(a\phi(\mathbf{x}+\mathbf{h})) \right|^2 .$$

Opening the square, this yields

$$|S(q,a)|^2 \ll H^{-2n} q^n \sum_{\mathbf{x}} \sum_{\substack{\mathbf{h_1},\mathbf{h_2} \leq H: \\ \mathbf{x}+\mathbf{h_1},\mathbf{x}+\mathbf{h_2} \leq q}} e_q\left(a\left(\phi(\mathbf{x}+\mathbf{h_1}) - \phi(\mathbf{x}+\mathbf{h_2})\right)\right).$$

Writing $\mathbf{y} = \mathbf{x} + \mathbf{h_2}$ and $\mathbf{h} = \mathbf{h_1} - \mathbf{h_2}$, this is equivalent to

$$|S(q,a)|^2 \ll H^{-2n} q^n \sum_{|\mathbf{h}| \leq H} w(\mathbf{h}) \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} e_q(a(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}))),$$

where $w(\mathbf{h}) = \#\{\mathbf{h_1},\mathbf{h_2} : \mathbf{h} = \mathbf{h_1} - \mathbf{h_2}\} \leq H^n$ and $\mathcal{R}(\mathbf{h})$ is a box as before. We have therefore shown that

$$|S(q,a)|^2 \ll H^{-n} q^n \sum_{|\mathbf{h}| \leq H} |T(\mathbf{h},a,q)|,$$

where

$$T(\mathbf{h},a,q) = \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} e_q(a(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}))).$$

Again we reduce the degree of the form once more by squaring and expanding this expression to obtain

$$|T(\mathbf{h},a,q)|^2 = \sum_{\mathbf{x},\mathbf{y} \in \mathcal{R}(\mathbf{h})} e_q(a(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}) - \phi(\mathbf{x}+\mathbf{h}) + \phi(\mathbf{x}))).$$

Writing $\mathbf{y} = \mathbf{x} + \mathbf{d}$ as before, this equals

$$|T(\mathbf{h},a,q)|^2 = \sum_{\mathbf{d}} \sum_{\mathbf{x} \in \mathcal{S}(\mathbf{h},\mathbf{d})} e_q(aC(\mathbf{h},\mathbf{d},\mathbf{x}))$$

with $\mathcal{S}(\mathbf{h},\mathbf{d})$ the box and $C(\mathbf{x},\mathbf{y},\mathbf{z})$ the multilinear form defined before. Again we note that the inner sum is empty unless $|\mathbf{d}| \leq 2q$. So we are in a situation very similar to the one in the previous section and the same argument developed there now shows that

$$|T(\mathbf{h},a,q)|^2 \ll q^{n+\varepsilon} N(a,q,\mathbf{h}),$$

where

$$N(a,q,\mathbf{h}) = \# \left\{ \mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| \leq 2q, \left\| 6\frac{a}{q} B_i(\mathbf{h},\mathbf{d}) \right\| < \frac{1}{q} \right\}.$$

Again applying Lemma 4.5.1, we find that

$$N(a,q,\mathbf{h}) \ll Z^{-n} \# \left\{ \mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| \leq 2qZ, \left\| 6\frac{a}{q} B_i(\mathbf{h},\mathbf{d}) \right\| < \frac{Z}{q} \right\}$$

whenever $0 < Z \leq 1$.

Note that of course the condition $\left\| 6\frac{a}{q} B_i(\mathbf{h},\mathbf{d}) \right\| < \frac{1}{q}$ already implies that $q \mid 6B_i(\mathbf{h},\mathbf{d})$ but we have written it in this form so that we can recognize the condition to be of the same shape as in the earlier argument.

The next step is to apply Lemma 4.5.2 to turn the inequality into the equality $B_i(\mathbf{h},\mathbf{d}) = 0$. Here we choose $m = 6B_i(\mathbf{h},\mathbf{d})$ so that $X \asymp MHqZ$ and $P_1 \asymp \frac{q}{Z}$. Thus any choice of Z satisfying $MHZ \ll 1$ for a sufficiently small implicit constant allows us to conclude that

$$N(a,q,\mathbf{h}) \ll Z^{-n} \# \left\{ \mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| \leq 2qZ, B_i(\mathbf{h},\mathbf{d}) = 0 \right\} \ll Z^{-n} (qZ)^{n-r(\mathbf{h})}$$

and hence

$$|S(q,a)|^2 \ll \frac{q^{\frac{3n}{2}+\varepsilon}}{H^n Z^{\frac{n}{2}}} \sum_{|\mathbf{h}| \leq H} (qZ)^{\frac{n-r(\mathbf{h})}{2}}.$$

If we assume in addition that C is ψ-good and $H \leq M^{\psi}$, we obtain the bound

$$\begin{aligned} |S(q,a)|^2 &\ll \frac{q^{\frac{3n}{2}+\varepsilon}}{H^n Z^{\frac{n}{2}}} \sum_{r=0}^{14} H^{n-14+r+\varepsilon} (qZ)^{\frac{n-r}{2}} \\ &\ll \frac{q^{2n+\varepsilon}}{H^{14}} \sum_{r=0}^{14} \frac{H^r}{(qZ)^{\frac{r}{2}}} \\ &\ll q^{2n+\varepsilon} \left(\frac{1}{H} + \frac{1}{qZ} \right)^{14}. \end{aligned}$$

Again we assumed that $qZ \geq 1$, but the final result is trivial otherwise. So again it will be optimal to choose Z as large as possible i.e. $Z \asymp \frac{1}{MH}$. Inserting this into our result we obtain the bound

$$|S(q,a)| \ll q^{n+\varepsilon} \left(\frac{1}{H^2} + \frac{HM}{q} \right)^{\frac{7}{2}}. \tag{4.5.6}$$

The final step is to choose the value of H minimizing the RHS of (4.5.6). We are given the conditions $1 \leq H \leq q$ and $H \leq M^{\psi}$. Putting $\gamma = \left(\frac{q}{M}\right)^{\frac{1}{3}}$, we see that the RHS of (4.5.6) is decreasing for $H \ll \gamma$ and increasing for $H \gg \gamma$ so that the optimal choice is $H \asymp \min(M^{\psi}, \gamma)$. Note that $H \leq q$ is then automatically satisfied. We have thus proved the following result.

Lemma 4.5.4. *Let a and q be coprime integers with $0 \leq a < q$. Assume that C is ψ-good. Then*

$$\frac{S(q,a)}{q^n} \ll \left(\frac{M}{q} \right)^{\frac{7}{3}+\varepsilon} + M^{-7\psi+\varepsilon}.$$

4.5.4 A mean square average via van der Corput

In this section we finally apply the improved version of van der Corput's method developed in [16] to obtain a satisfying bound for the minor arc contribution.

The idea is to exploit that the minor arc contribution already involves an average over both the modulus q and the integration variable β.

From now on let our box be again $\mathscr{B} = \mathscr{B}(\mathbf{z})$ with the center $\mathbf{z}$ as chosen in Lemma 4.4.3. Instead of a pointwise bound for $S(\alpha)$, we now seek to estimate the mean square average

$$M(\alpha, \kappa) := \int_{\alpha-\kappa}^{\alpha+\kappa} |S(\beta)|^2 d\beta,$$

where $\kappa \in (0,1)$ is a small parameter to be determined. By an appropriate dissection of the minor arcs and an application of Cauchy-Schwarz, a satisfactory estimate for $M(\alpha, H)$ will allow us to bound the minor arc contribution $\int_{\mathfrak{m}} S(\alpha) d\alpha$.

We proceed as in the previous section, only now we consider the more general (but still trivial) identity

$$H_1 H_2 \dots H_n S(\beta) = \sum_{\mathbf{h}: h_i \leq H_i} \sum_{\mathbf{x}+\mathbf{h} \in P\mathscr{B}} e(\beta\phi(\mathbf{x}+\mathbf{h})) = \sum_{\mathbf{x} \in \mathbb{Z}^n} \sum_{\mathbf{h}:\mathbf{x}+\mathbf{h} \in P\mathscr{B}} e(\beta\phi(\mathbf{x}+\mathbf{h}))$$

where $H_1, H_2, \dots, H_n \geq 1$ are certain parameters. We choose $H_1 = P$ and $H_2 = \dots = H_n = H$ for a certain parameter $H \leq P$. Here the special rôle of the first variable comes from its special rôle in the construction of $\mathbf{z}$.

Note that the condition $H_i \leq P$ ensures that the sum over $\mathbf{h}$ is non-empty only for $\mathcal{O}(P^n)$ values of $\mathbf{x}$. Squaring and applying Cauchy-Schwarz, we then find that

$$(H_1^2 \dots H_n^2)|S(\beta)|^2 \ll P^n \sum_{\mathbf{x} \in \mathbb{Z}^n} \left| \sum_{\mathbf{h}:\mathbf{x}+\mathbf{h} \in P\mathscr{B}} e(\beta\phi(\mathbf{x}+\mathbf{h})) \right|^2 .$$

Opening the square, this yields

$$(H_1^2 \dots H_n^2)|S(\beta)|^2 \ll P^n \sum_{\mathbf{x} \in \mathbb{Z}^n} \sum_{\substack{\mathbf{h_1},\mathbf{h_2}: \\ \mathbf{x}+\mathbf{h_1}, \mathbf{x}+\mathbf{h_2} \in P\mathscr{B}}} e\left(\beta\left(\phi(\mathbf{x}+\mathbf{h_1}) - \phi(\mathbf{x}+\mathbf{h_2})\right)\right).$$

Writing $\mathbf{y} = \mathbf{x} + \mathbf{h_2}$ and $\mathbf{h} = \mathbf{h_1} - \mathbf{h_2}$, this is equivalent to

$$(H_1^2 \dots H_n^2)|S(\beta)|^2 \ll P^n \sum_{|h_i| \leq H_i} w(\mathbf{h}) \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} e(\beta(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y})))$$

where $w(\mathbf{h}) = \#\{\mathbf{h_1}, \mathbf{h_2} : \mathbf{h} = \mathbf{h_1} - \mathbf{h_2}\} \leq H_1 H_2 \dots H_n$. Instead of taking absolute values inside as before, we now first integrate over β. Here we use a smooth cutoff function to find that

$$\begin{aligned} M(\alpha, \kappa) &\leq e \int_{\mathbb{R}} \exp\left(-\frac{(\beta-\alpha)^2}{\kappa^2}\right) |S(\beta)|^2 \mathrm{d}\beta \\ &\ll \frac{P^n}{(H_1 \dots H_n)^2} \sum_{\mathbf{h}} w(\mathbf{h}) \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} I(\mathbf{h}, \mathbf{y}) \end{aligned}$$

and hence

$$M(\alpha, \kappa) \ll \frac{P^n}{H_1 \dots H_n} \sum_{\mathbf{h}} \left| \sum_{\mathbf{y} \in \mathcal{R}(\mathbf{h})} I(\mathbf{h}, \mathbf{y}) \right|, \tag{4.5.7}$$

where

$$I(\mathbf{h}, \mathbf{y}) = \int_{\mathbb{R}} \exp\left(-\frac{(\beta-\alpha)^2}{\kappa^2}\right) e\left(\beta(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}))\right) \mathrm{d}\beta$$

which can also be written as

$$I(\mathbf{h}, \mathbf{y}) = \sqrt{\pi}\kappa \exp\left(-\pi^2\kappa^2 \left(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y})\right)^2\right) e(\alpha(\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}))). \tag{4.5.8}$$

Our goal is to bound the contribution of the terms where h_1 is large so that we can effectively bound h_1 to a shorter interval. The point is that by our choice of the box $\mathscr{B}(\mathbf{z})$ we have a good lower bound for the partial derivative $\frac{\partial C}{\partial x_1}$ inside $\mathscr{B}$. But we expect $\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}) \approx h_1 \frac{\partial \phi(\mathbf{y})}{\partial x_1} \approx h_1 \frac{\partial C(\mathbf{y})}{\partial x_1}$ so that this difference should be large if h_1 is large which means that $I(\mathbf{h}, \mathbf{y})$ will be small.

Let us make this precise. By Lemma 4.4.3 we have the bound

$$\frac{\partial C}{\partial x_1} \gg M^6 \tag{4.5.9}$$

on all of $\mathcal{B}$.

By homogeneity this implies that

$$\frac{\partial C}{\partial x_1} \gg P^2 M^6$$

on $P\mathscr{B}$.

We thus find that

$$\frac{\partial \phi}{\partial x_1} = \frac{\partial C}{\partial X_1} + \mathcal{O}\left(PM|\mathbf{z}|\right) \gg P^2M^6$$

on all of $P\mathcal{B}$ as well.

Using $|\mathbf{y}| \ll PM^{3.75}$ for $\mathbf{y} \in P\mathscr{B}$, we now have the approximation

$$\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}) = h_1 \cdot \frac{\partial \phi}{\partial x_1}(\mathbf{y}) + \mathcal{O}\left(HP^2 \max_i \left|\frac{\partial \phi}{\partial x_i}\right| + h_1^2PM^{4.75}\right).$$

Note that $h_1 \leq P$ implies that

$$h_1^2PM^{4.75} \ll h_1P^2M^{4.75}$$

and so we will have

$$\phi(\mathbf{y}+\mathbf{h}) - \phi(\mathbf{y}) \gg |h_1| \cdot P^2M^6$$

unless $|h_1| \ll H$. Indeed, unless also $|h_1| \ll \frac{(\log P)^2}{\kappa P^2M^6}$, this means that

$$|C(\mathbf{y}+\mathbf{h}) - C(\mathbf{y})| \geq \frac{(\log P)^2}{\kappa}$$

and it is easy to see from (4.5.7) and (4.5.8) that the contribution to $M(\alpha,\kappa)$ of such $\mathbf{h}$ is $\mathcal{O}(1)$. Hence we have shown that

$$M(\alpha,\kappa) \ll 1 + \frac{P^{n-1}}{H^{n-1}} \sum_{|h_i| \ll H} \left|\sum_{\mathbf{y}} I(\mathbf{h},\mathbf{y})\right|$$

if we choose $\kappa \asymp \frac{(\log P)^2}{HP^2M^6}$.

Moreover, the range $|\beta - \alpha| \geq \kappa \log P$ in the definition of $I(\mathbf{h},\mathbf{y})$ clearly has a total contribution of $\mathcal{O}(1)$ to $M(\alpha,\kappa)$ so that we end up with the estimate

$$M(\alpha,\kappa) \ll 1 + \frac{P^{n-1}}{H^{n-1}} \sum_{|h_i| \ll H} \int_{\alpha-\kappa\log P}^{\alpha+\kappa\log P} |T(\mathbf{h},\beta)|d\beta,$$

where

$$T(\mathbf{h},\beta)=\sum_{\mathbf{y}\in\mathcal{R}(\mathbf{h})} e\left(\beta(\phi(\mathbf{y}+\mathbf{h})-\phi(\mathbf{y}))\right).$$

The same argument employed already several times now yields

$$|T(\mathbf{h},\beta)|^2 \ll P^{n+\varepsilon}N(\beta,P,\mathbf{h})$$

where

$$N(\beta,P,\mathbf{h})=\#\left\{\mathbf{d}\in\mathbb{Z}^n : |\mathbf{d}|\le 2P, \|6\beta B_i(\mathbf{h},\mathbf{d})\|<\frac{1}{2P}\right\},$$

so that

$$M(\alpha,\kappa)\ll 1+\frac{\kappa P^{\frac{3n}{2}-1+\varepsilon}}{H^{n-1}}\sum_{|h_i|\ll H}\max_{\beta\in I}N(\beta,P,\mathbf{h})^{\frac{1}{2}} \tag{4.5.10}$$

where $I=\{\beta : |\beta-\alpha|\le\kappa\log P\}$.

We next claim that

$$\max_{\beta\in I}N(\beta,P,\mathbf{h})\ll P^{\varepsilon}N(\alpha,P,\mathbf{h}).$$

Indeed, consider a vector $\mathbf{d}$ counted by $N(\beta,P,\mathbf{h})$. By our assumption, it satisfies $|\mathbf{d}|\ll P$ as well as $\|6\beta B_i(\mathbf{h},\mathbf{d})\|\ll\frac{1}{P}$ so that

$$\|6\alpha B_i(\mathbf{h},\mathbf{d})\|\ll\frac{1}{P}+|\beta-\alpha||B_i(\mathbf{h},\mathbf{d})|\ll\frac{1}{P}+\kappa(\log P)MHP\ll\frac{1}{P}$$

by our choice of κ.

We conclude that

$$\max_{\beta}N(\beta,P,\mathbf{h})\le\#\left\{\mathbf{d}\in\mathbb{Z}^n : |\mathbf{d}|\ll P, \|6\alpha B_i(\mathbf{h},\mathbf{d})\|\ll\frac{1}{P}\right\}\ll N(\alpha,P,\mathbf{h}),$$

where the last estimate is a consequence of Lemma 4.5.1 with $Z\asymp 1$ sufficiently small.

We conclude that

$$M(\alpha,\kappa) \ll 1 + \frac{\kappa P^{\frac{3n}{2}-1+\varepsilon}}{H^{n-1}} \sum_{|h_i| \ll H} N(\alpha, P, \mathbf{h})^{\frac{1}{2}}. \tag{4.5.11}$$

We now write $\alpha = \frac{a}{q} + \theta$ in a preparation for applying Lemmas 4.5.1 and 4.5.2. Indeed, Lemma 4.5.1 implies that

$$N(\alpha, P, \mathbf{h}) \ll Z^{-n} \#\{\mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| < ZP, \|6\alpha B_i(\mathbf{h}, \mathbf{d})\| \ll \frac{Z}{P}\}.$$

Following Heath-Brown, we apply this with two different choices of Z. In the first application we choose $Z = Z_1$ sufficiently small to ensure $B_i(\mathbf{h}, \mathbf{d}) = 0$ as before. In the second application however, we make a larger choice of $Z = Z_2$ which only forces that $q \mid B_i(\mathbf{h}, \mathbf{d})$. We then have to consider the implications of this weaker result. It was observed in [16] that only this new trick allows us to handle the case of 14 variables.

To apply Lemma 4.5.2 with $m = 6B_i(\mathbf{h}, \mathbf{d})$ we have to choose $X \asymp MHPZ$ and $P_1 \asymp \frac{P}{Z}$ so that in our application of Lemma 4.5.1 we need to choose $Z \leq 1$ satisfying

$$|\theta| \ll \frac{1}{MHPZq} \quad \text{and} \quad Z \ll \frac{P}{q}$$

with sufficiently small implicit constants. In the first application we should also have

$$MHPZ \ll q \quad \text{or} \quad Z \ll q|\theta|P.$$

Writing

$$\eta = |\theta| + \frac{1}{P^2 HM} \tag{4.5.12}$$

for convenience and assuming $q \sim R$, this means that we need to choose

$$Z_1 \asymp \min\left(R\eta P, \frac{1}{RHMP\eta}\right),$$

noting that this automatically implies $Z_1 \leq 1$. Similarly, we should choose

$$Z_2 \asymp \min\left(1, \frac{1}{RHMP\eta}\right).$$

In the application with $Z = Z_1$, we thus find that

$$\begin{aligned} N(\alpha, P, \mathbf{h}) &\ll Z_1^{-n} \#\{\mathbf{w} \in \mathbb{Z}^n : |\mathbf{w}| \ll P, B_i(\mathbf{h}, \mathbf{d}) = 0\} \\ &\ll Z_1^{-n} (Z_1 P)^{n-r} \\ &\ll P^n \left((R\eta P^2)^{-r} + (RHM\eta)^r \right) \end{aligned}$$

where $r = r(\mathbf{h})$. The argument only works for $Z_1 P \geq 1$ but the estimate is true in any case because of the trivial bound $N(\beta, P, \mathbf{h}) \ll P^n$.
On the other hand, in the application with $Z = Z_2$, we obtain

$$N(\alpha, P, \mathbf{h}) \ll Z_2^{-n} \#\{\mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| \ll Z_2 P, q \mid B_i(\mathbf{h}, \mathbf{d})\}.$$

To make this a useful estimate, we need to count vectors $\mathbf{d}$ with $q \mid B_i(\mathbf{h}, \mathbf{d})$. Here we can copy the results from [16], but we need to introduce some notation. Recall that we have fixed a vector $\mathbf{h}$ with $r(\mathbf{h}) = r$, meaning that the matrix $M(\mathbf{h})$ has rank r. We now distinguish primes p according to whether p divides all the $r \times r$ minors of $M(\mathbf{h})$. If it does, we say that p is of type I and if it does not, we say that it is of type II. We then decompose $q = q_1 q_2$ such that q_1 is a product of type I primes and q_2 a product of type II primes. The argument in [16, p. 218 f.] now shows that

$$\#\{\mathbf{d} \in \mathbb{Z}^n : |\mathbf{d}| \ll Z_2 P, q \mid B_i(\mathbf{h}, \mathbf{d})\} \ll \left(1 + \frac{B}{q_2}\right)^r B^{n-r}$$

with $B = 1 + Z_2 P$. If $Z_2 P \gg 1$ so that $B \asymp Z_2 P$, this shows that

$$\begin{aligned} N(\alpha, P, \mathbf{h}) &\ll Z_2^{-n} \left(1 + \frac{Z_2 P}{q_2}\right)^r (Z_2 P)^{n-r} \\ &= P^n \left(\frac{1}{q_2} + \frac{1}{Z_2 P}\right)^r \\ &\ll P^n \left(\frac{1}{q_2^r} + \frac{1}{P^r} + (RHM\eta)^r\right), \end{aligned}$$

but the intermediate and hence the final result is trivially true if $Z_2 P \ll 1$.
Combining the results of the two applications, we obtain that

$$N(\alpha, P, \mathbf{h}) \ll P^n \left(\frac{1}{P^r} + (RHM\eta)^r + \min\left(\frac{1}{(R\eta P^2)^r}, \frac{1}{q_2^r}\right)\right).$$

We now need to insert this into our estimate for $M(\alpha, \kappa)$ but we also want to introduce an average over q to make use of the fact that q_2 is almost as large as q most of the time.

Our object of study thus becomes

$$A(\theta, R, H, P) := \sum_{q \sim R} \sum_{(a;q)=1} \sum_{|h_i| \ll H} N(\alpha, P, \mathbf{h})^{\frac{1}{2}}, \tag{4.5.13}$$

where we continue to write $\alpha = \frac{a}{q} + \theta$ and we remind the reader of our notation $q \sim R$ for the dyadic condition $R < q \leq 2R$.

The argument above now leads to the bound

$$A(\theta, R, H, P) \ll RP^{\frac{n}{2}} \sum_{|h_i| \ll H} \sum_{q \sim R} \left[\frac{1}{P} + RHM\eta + \min\left(\frac{1}{R\eta P^2}, \frac{1}{q_2} \right) \right]^{\frac{r(\mathbf{h})}{2}}.$$

We then need to estimate

$$V(\mathbf{h}, R, \eta) := \sum_{q \sim R} \min\left(\frac{1}{R\eta P^2}, \frac{1}{q_2} \right)^{r/2}$$

for $r = r(\mathbf{h})$. A double dyadic decomposition leads to

$$\begin{aligned} V(\mathbf{h}, R, \eta) &\ll P^{\varepsilon} \max_{S \leq R} \sum_{q_1 \sim S} \sum_{q_2 \sim \frac{R}{S}} \min\left(\frac{1}{R\eta P^2}, \frac{S}{R} \right)^{r/2} \\ &\ll P^{\varepsilon} \max_{S \leq R} \frac{R}{S} \min\left(\frac{1}{R\eta P^2}, \frac{S}{R} \right)^{r/2} \#\{q_1 \leq 2S\}. \end{aligned}$$

Now recall that q_1 only contains prime factors dividing a certain non-zero $r \times r$ determinant M_0 of $M(\mathbf{h})$. In particular, $M_0 \ll M^r H^r$. Applying Rankin's trick it now follows that

$$\#\{q_1 \leq 2S\} \ll S^{\varepsilon} \sum_{q_1} q_1^{-\varepsilon} = S^{\varepsilon} \prod_{p | M_0} \frac{1}{1 - p^{-\varepsilon}} \ll S^{\varepsilon} M_0^{\varepsilon} \ll M^{\varepsilon}$$

and hence

$$V(\mathbf{h}, R, \eta) \ll M^{\varepsilon} \frac{R}{S} \min\left(\frac{1}{(R\eta P^2)^{\frac{r}{2}}}, \left(\frac{S}{R} \right)^{\frac{r}{2}} \right).$$

Maximizing this function for S we find that

$$V(\mathbf{h}, R, \eta) \ll M^{\varepsilon} \frac{R}{(R\eta P^2)^{\frac{r}{2}}} \cdot \min(1, \eta P^2)^{e(r)},$$

where $e(0) = 0, e(1) = \frac{1}{2}$ and $e(r) = 1$ for $r \geq 2$. Assuming that C is ψ-good and $H \leq M^{\psi}$ it now follows that

$$\begin{aligned} A(\theta, R, H, P) &\ll R^2 P^{\frac{n}{2}} \sum_{|h_i| \ll H} \left[\frac{1}{P^{\frac{r(\mathbf{h})}{2}}} + (RHM\eta)^{\frac{r(\mathbf{h})}{2}} + R^{-1} V(\mathbf{h}, R, \eta) \right] \\ &\ll R^2 P^{\frac{n}{2}} M^{\varepsilon} \sum_{|h_i| \ll H} \left[\frac{1}{P^{\frac{r(\mathbf{h})}{2}}} + (RHM\eta)^{\frac{r(\mathbf{h})}{2}} + \frac{1}{(R\eta P^2)^{\frac{r(\mathbf{h})}{2}}} \cdot \min(1, \eta P^2)^{e(r(\mathbf{h}))} \right] \\ &\ll R^2 P^{\frac{n}{2}} M^{\varepsilon} \sum_{r=0}^{14} H^{n-14+r} \left[\frac{1}{P^{\frac{r}{2}}} + (RHM\eta)^{\frac{r}{2}} + \frac{1}{(R\eta P^2)^{\frac{r}{2}}} \cdot \min(1, \eta P^2)^{e(r)}) \right] \\ &\ll R^2 P^{\frac{n}{2}} H^n M^{\varepsilon} \left(\frac{1}{H^{14}} + \frac{1}{P^7} + (RHM\eta)^7 + \frac{1}{(R\eta P^2)^7} \cdot \min(1, \eta P^2) \right). \end{aligned}$$

Finally, let us show that the term $\frac{1}{P^7}$ is negligble. Indeed, if $HRMP\eta \geq 1$, it is dominated by the third summand.. Otherwise, if $HRMP\eta \leq 1$, we have $(R\eta P)^7 \leq R\eta P \leq \frac{1}{HM} \leq \min(1, \eta P^2)$ on recalling that $\eta \geq \frac{1}{P^2 HM}$ and hence $\frac{1}{P^7}$ is dominated by the last summand in that case. In any case, we now conclude that

$$A(\theta, R, H, P) \ll R^2 P^{\frac{n}{2}} H^n M^{\varepsilon} \left(\frac{1}{H^{14}} + (RHM\eta)^7 + \frac{1}{(R\eta P^2)^7} \cdot \min(1, \eta P^2) \right). \tag{4.5.14}$$

4.6 Intermezzo: The singular series

We are now ready to use the bounds for $S(q, a)$ in order to bound the singular series $\mathfrak{S}(P_0)$. However, we first require another ingredient which is a lower bound on the individual p-adic densities. The non-vanishing of these is a consequence of the existence of a non-singular p-adic solution and Hensel's

Lemma. To get a uniform lower bound, we require a quantitative result on how singular such a solution is.

This is typically described in terms of the invariant $\Delta(C)$ which is defined as the greatest common factor of all $n \times n$-minors of the $n \times \binom{n+1}{2}$-matrix $(c_{ijk})_{i,(j,k)}$. In particular, $\Delta(C) \neq 0$ whenever C is non-degenerate. We also note that if C is degenerate modulo q, then $q \mid \Delta(C)$.

We note that sometimes the invariant we called $\Delta(C)$ is also denoted as $h(C)$, but this already has a different meaning in our work.

We also define $\Delta(\phi)$ to be $\Delta(\widetilde{\phi})$ where $\widetilde{\phi}$ is the homogenized version of ϕ, i.e. a cubic form in $n+1$ variables. Note that $\Delta(C) \mid \Delta(\phi)$.

We will throughout work with the following consequence of Hensel's Lemma:

Lemma 4.6.1. *If $\phi(\mathbf{x}) \equiv 0 \pmod{p^{2\ell-1}}$ and $p^\ell \nmid \nabla\phi(\mathbf{x})$, then x lifts to a non-singular p-adic solution.*

In the homogeneous case, Davenport [8, Lemma 18.7] established the following:

Lemma 4.6.2. *If $C \in \mathbb{Z}[x_1, \ldots, x_n]$ is a non-degenerate cubic form in $n \geq 10$ variables, then for each prime p there is a p-adic solution $\mathbf{x}$ to $C(\mathbf{x}) = 0$ such that $p^\ell \nmid \nabla C(\mathbf{x})$ for*

$$\ell = \ell_C(p) := 3 \cdot \left\lfloor \frac{v_p(\Delta(C))}{n-9} \right\rfloor + 3.$$

In the inhomogeneous case, as observed by Davenport and Lewis [12], the Necessary Congruence Condition is in general not enough to deduce the existence of a non-singular p-adic solution to $\phi(\mathbf{x}) = 0$. A counterexample in the 14 variables x_i for $1 \leq i \leq 4$ and $x_{i,j}$ for $1 \leq i \leq j \leq 4$ is given by

$$\phi(\mathbf{x}) = x_1^2 - Nx_2^2 + p(x_3^2 - Nx_4^2) + p^2 \left(\sum_{1 \leq i \leq j \leq 4} x_{i,j} x_i x_j \right)$$

for a quadratic non-residue N modulo p. However, they managed to prove that the desired implication holds true if $n \geq 15$ and ϕ is non-degenerate. A quantitative version of their argument leads to the following result:

Lemma 4.6.3. *If $\phi \in \mathbb{Z}[x_1, \ldots, x_n]$ is a non-degenerate cubic polynomial in $n \geq 15$ variables, then for each prime p there is a p-adic solution $\mathbf{x}$ to $\phi(\mathbf{x}) = 0$ such that $p^\ell \nmid \nabla\phi(\mathbf{x})$ for*

$$\ell = \ell_\phi(p) := \begin{cases} 98 & p \nmid \Delta(\phi) \\ 144 v_p(\Delta(\phi)) + 2 & p \mid \Delta(\phi) \end{cases}.$$

While the result is probably not optimal and the proof has a certain ad-hoc structure, we note that the counterexample for 14 variables naturally arises from the structure of the proof. We will momentarily see that one could prove a superficially stronger bound, but in the critical case $v_p(\Delta(\phi)) = 1$ we do not lose anything.

Proof. Let $k = v_p(\Delta(\phi))$ and $m = \max(6 \left\lfloor \frac{k}{n-9} \right\rfloor + 5, k+1)$. We will prove the result for $\ell = 48k + 16m + 18$. Note that for $k = 0$, we have $m = 5$ so that $\ell = 98$. For $k \geq 1$, we have $m \leq 6k - 1$ and hence $\ell \leq 144k + 2$ as desired. Here, we used that $\left\lfloor \frac{k}{n-9} \right\rfloor \leq k - 1$ for $k \geq 1$ and $n \geq 11$.

First of all, using the congruence condition, we can find a solution modulo $p^{96k+32m+35}$ and after translating the variables appropriately, we may assume that this solution is given by $(0, 0, \ldots, 0)$ so that

$$\phi(\mathbf{x}) \equiv C(\mathbf{x}) + Q(\mathbf{x}) + L(\mathbf{x}) \pmod{p^{96k+32m+35}}.$$

If L does not vanish identically modulo $p^{48k+16m+18}$, this solution $\mathbf{x} = (0, 0, \ldots, 0)$ satisfies $p^{48k+16m+18} \nmid \nabla\phi(\mathbf{x})$ and $p^{2(48k+16m+18)-1} \mid \phi(\mathbf{x})$ so that it lifts to a p-adic solution of the desired shape by Hensel's Lemma.

Otherwise, we may assume that

$$\phi(\mathbf{x}) \equiv C(\mathbf{x}) + Q(\mathbf{x}) \pmod{p^{48k+16m+18}}.$$

If Q vanishes identically modulo p^m, then C is non-degenerate modulo p^m since otherwise $p^{k+1} \mid \Delta(C)$ by construction of m, contradicting the definition of k.

But if C is non-degenerate modulo p^m, by Lemma 4.6.2, there is a solution of $C(\mathbf{x}) = 0$ with $v_p(\nabla C(\mathbf{x})) \leq 3 \left\lfloor \frac{k}{n-9} \right\rfloor + 3$ and since $m \geq 6 \left\lfloor \frac{k}{n-9} \right\rfloor + 5$, this lifts to a non-singular solution by Hensel's Lemma.

From now on we assume that Q does not vanish identically modulo p^m.

First case: Q has rank at least five modulo p^M where $M = 12k + 4m + 5$. Then we can find a non-singular solution β of $Q(\beta) = 0$ with $p^M \nmid \nabla Q(\beta)$. Rescaling by p^{2M-1}, we have $p^{6M-3} \mid \phi(p^{2M-1}\beta)$ and $p^{3M-1} \nmid \nabla\phi(p^{2m-1}\beta)$ and we obtain a non-singular solution with $p^{3M-1} = p^{36k+12m+15} \nmid \nabla\phi(\mathbf{x})$. Note that here we used that $4M - 2 = 48k + 16m + 18$ so that $p^{6M-3} \mid L(p^{2M-1}\beta)$.

Second case: Q has rank $1 \leq r \leq 4$ modulo p^M. Hence $\phi(\mathbf{x})$ is equivalent to a form $\psi(\mathbf{y})$ with

$$\psi(\mathbf{y}) \equiv y_1 R_1(\mathbf{y}) + \cdots + y_r R_r(\mathbf{y}) + R(y_1, \ldots, y_r) + \Gamma(y_{r+1}, \ldots, y_n) \pmod{p^M}$$

for some cubic form Γ and some quadratic forms $R, R_1, \ldots, R_r$.

First subcase: Γ does not vanish identically modulo p^{2k+1}. We then choose $\delta_1, \ldots, \delta_r$ such that $p^m \nmid R(\delta_1, \ldots, \delta_r)$ and $\delta_{r+1}, \ldots, \delta_n$ such that $p^{2k+1} \nmid \Gamma(\delta_{r+1}, \ldots, \delta_n)$. Let $\rho \leq 2k$ be chosen so that $p^\rho \| \Gamma(\delta_{r+1}, \ldots, \delta_n)$.

We then choose $\varepsilon = (p^{\rho+1}\delta_1, \ldots, p^{\rho+1}\delta_r, \delta_{r+1}, \ldots, \delta_n)$ so that $p^\rho \| y_1 R_1 + \cdots + y_r R_r + \Gamma$ if we insert ε.

We can thus find a p-adic integer μ such that

$$\mu(y_1 R_1 + \cdots + y_r R_r + \Gamma) + R = 0,$$

still everything evaluated at ε. Choosing $y = \mu\varepsilon$, we then have $p^M \mid \psi(\mathbf{y})$. Moreover, by Euler's identity we have that

$$\sum_j \varepsilon_j \psi^{(j)}(\mu\varepsilon) \equiv -\mu \cdot R(\varepsilon) \pmod{p^M}$$

and since

$$v_p(\mu \cdot R(\varepsilon)) \leq 2v_p(R(\varepsilon)) - \rho \leq 3\rho + 4 + 2v_p(R(\delta_1, \dots, \delta_r))$$

$$\leq 2m + 3\rho + 2 \leq 2m + 6k + 2,$$

we have that one of the $\psi^{(j)}$ is divisible at most by $2m + 6k + 2$ so that we have a solution with $\ell = 2m + 6k + 3$ modulo p^M lifting by Hensel since $M = 2\ell - 1$.

Second subcase: Γ vanishes identically modulo p^{2k+1}. Then we can write

$$\psi(\mathbf{y}) \equiv y_1 R_1(\mathbf{y}) + \cdots + y_r R_r(\mathbf{y}) + R(y_1, \dots, y_r) \pmod{p^{2k+1}}.$$

Then we find a solution of the shape $(0, \dots, 0, y_{r+1}, \dots, y_n)$ with $p^{k+1} \nmid \nabla\psi$ (which then again lifts by Hensel) unless all the R_i vanish modulo p^{k+1} at these points. But this means that the variables $y_{r+1}, \dots, y_n$ appear at most linearly in all of the R_i so that we can write

$$\psi(\mathbf{y}) \equiv y_{r+1} S_{r+1}(y_1, \dots, y_r) + \cdots + y_n S_n(y_1, \dots, y_r) + R(y_1, \dots, y_r) \pmod{p^{k+1}}.$$

But then, finally, as there are only at most $\binom{r+1}{2} \leq 10$ linearly independent quadratic monomials in $y_1, \dots, y_r$ and $n - r \geq 11$, the S_i can not be linearly independent and so ψ and hence ϕ must be degenerate modulo p^{k+1} so that $p^{k+1} \mid \Delta(\phi)$ contradicting the definition of k. □

We are now equipped with everything needed for a lower bound of the singular series.

To this end, for each prime p let

$$k_C(p) = \begin{cases} \max_{t \in \mathbb{N}: p^t \leq P_0} \{t\}, & p \nmid \Delta(C), \\ \max_{t \in \mathbb{N}: p^t \leq P_0} \{t, 2\ell_C(p) - 1\}, & p \mid \Delta(C) \end{cases}$$

in the homogeneous case and

$$k_\phi(p) = \begin{cases} \max_{t \in \mathbb{N}: p^t \leq P_0} \{t\}, & p \nmid \Delta(\phi), \\ \max_{t \in \mathbb{N}: p^t \leq P_0} \{t, 2\ell_\phi(p) - 1\}, & p \mid \Delta(\phi) \end{cases}$$

in the inhomogeneous case. We then define the truncated Euler product

$$S_\phi(P_0) = \prod_{p \leq P_0} \sum_{i=0}^{k_\phi(p)} A(p^i),$$

where

$$A(q) = \sum_{(a;q)=1} \frac{S(q,a)}{q^n}$$

and we recall the classical fact that

$$\sum_{i=0}^{k} A(p^i) = p^{-k(n-1)} \rho(p^k),$$

where $\rho(p^k)$ denotes the number of solutions of $\phi(\mathbf{x}) \equiv 0 \pmod{p^k}$.
Similarly, we define

$$S_C(P_0) = \prod_{p \leq P_0} \sum_{i=0}^{k_C(p)} A(p^i).$$

We will first establish a lower bound for the truncated Euler product $S(P_0)$ and then estimate the difference to the truncated singular series $\mathfrak{S}(P_0)$.
We first deal with the primes not dividing Δ. The key ingredient here is the following bound which is Lemma 9 in [4].

Lemma 4.6.4. *Let C be a cubic form in $n \geq 10$ variables. Then for any $p \gg 1$ with $p \nmid \Delta(C)$ and any $k \geq 1$, we have*

$$\rho^*(p^k) \geq p^{k(n-1)} \left(1 + \mathcal{O}\left(\frac{1}{p}\right)\right),$$

where $\rho^*(p^k)$ denotes the number of non-singular solutions of $\phi(\mathbf{x}) \equiv 0 \pmod{p^k}$.

In the homogeneous case, this immediately shows that the contribution from the primes not dividing $\Delta(C)$ to $S_C(P_0)$ is $\gg P_0^{-\varepsilon}$.

In the inhomogeneous case, we note that

$$\rho^*_\phi(p) = \frac{\rho^*_{\widetilde{\phi}}(p) - \rho^*_C(p)}{p-1} \geq \frac{p^n + \mathcal{O}(p^{n-1})}{p-1} \geq p^{(n-1)}\left(1 + \mathcal{O}\left(\frac{1}{p}\right)\right)$$

whenever $p \nmid \Delta(\phi)$ by applying Lemma 4.6.4 to the cubic form $\widetilde{\phi}$. Here, we used that $\rho^*_C(p) \ll p^{n-1}$ for any cubic form C. This follows from $\rho^*_C(p) \leq \rho_C(p) \ll p^{n-1}$ unless C vanishes modulo p, but in that latter case we have $\rho^*_C(p) = 0$.

We thus obtain the same bound as in Lemma 4.6.4 for $\rho^*(p^k)$ by Hensel's Lemma and thus deduce that the contribution from the primes not dividing $\Delta(\phi)$ to $S_\phi(P_0)$ is also $\gg P_0^{-\varepsilon}$.

Note that in both cases the contribution from the primes $p \ll 1$ not dividing Δ is clearly $\gg 1$.

We now need to deal with the primes $p \mid \Delta$.

In the inhomogeneous case, we conclude from Lemma 4.6.3 that

$$\rho(p^{k(p)}) \gg p^{k(p)(n-1)-\ell(p)(n-1)},$$

so that the contribution to $S(P_0)$ from the primes dividing $\Delta(\phi)$ is

$$\gg M^{-\varepsilon} \prod_{p \mid \Delta(\phi)} p^{-(n-1)\ell_\phi(p)} \gg M^{-\varepsilon} \prod_{p \mid \Delta(\phi)} p^{-292(n-1)v_p(\Delta(\phi))} \gg M^{-292(n^2-1)-\varepsilon},$$

using that $\Delta(\phi) \ll M^{n+1}$.

In the homogeneous case, we can do a bit better using the ideas from [4]. We first consider the case that modulo p the form C is not equivalent to one in less than four variables.

In that case, by Lemma 10 from [4], we have

$$\rho^*_C(p^k) \geq p^{k(n-1)}\left(1 + \mathcal{O}\left(p^{-1/2}\right)\right).$$

We thus conclude that primes $p \mid \Delta$ such that the order of $\widetilde{\phi}$ resp. C modulo p is at least four, have

$$\rho(p^{k(p)}) \gg p^{k(p)(n-1)},$$

at least for $p \gg 1$.

Finally, we need to deal with the case where the order modulo p is $t \leq 3$. Hence, C is equivalent to a cubic form $C_1(x_1, \dots, x_t)$ modulo p. If C_1 has a non-singular zero, then immediately $\rho_C^*(p) \geq p^{n-t} \geq p^{n-3}$ as we can vary $x_{t+1}, \dots, x_n$ arbitrarily. Otherwise, if w.l.o.g. $(1, 0, \dots, 0)$ is a singular zero modulo p, then

$$C_1(x_1, \dots, x_t) \equiv x_1 Q(x_2, \dots, x_t) + C_2(x_2, \dots, x_t) \pmod{p}.$$

Here, Q cannot vanish identically modulo p as otherwise the order would be at most $t-1$. We can thus choose $x_2, \dots, x_t$ such that $p \nmid Q$ and then solve the congruence for x_1 to obtain a non-singular solution and conclude as above.

Finally, we need to consider the case where C_1 does not have a non-trivial zero modulo p, i.e all the roots have all variables divisible by p.

But then if we define

$$C'(X_1, \dots, X_n) = p^{-1} C(pX_1, \dots, pX_t, X_{t+1}, \dots, X_n),$$

we have $\rho_C(p^k) = p^{n-t} \rho_{C'}(p^{k-1})$. We can iterate this argument and after at most $\ell_C(p) - 1$ steps we will end up with a cubic form that has a non-singular solution modulo p which we can then treat as above.

Since we lose a factor of at most p^2 in each step (p^{n-3} instead of p^{n-1}), we eventually end up with the bound

$$\rho_C(p^{k(p)}) \gg p^{k(p)(n-1) - 2\ell(p)}$$

and therefore the total contribution to $S(P_0)$ of these primes is

$$\gg M^{-\varepsilon} \prod_{p | \Delta(C)} p^{-2\ell(p)} \gg M^{-\varepsilon} \prod_{p | \Delta(C)} p^{-6 v_p(\Delta(C))} \gg M^{-\varepsilon} \Delta(C)^{-6} \gg M^{-6n-\varepsilon},$$

using $\Delta(C) \ll M^n$.

Here we again used that $\ell(p) = 3\left\lfloor \frac{v_p(\Delta(C))}{n-9} \right\rfloor + 3 \leq 3v_p(\Delta(C))$ for $p \mid \Delta(C)$ which is sharp in the critical case $v_p(\Delta(C)) = 1$. The argument in [4] is not correct and fails in precisely that critical case.
We summarize our results as follows:

Lemma 4.6.5. *We have*

$$S_C(P_0) \gg M^{-6n-\varepsilon}$$

and

$$S_\phi(P_0) \gg M^{-292(n^2-1)-\varepsilon}.$$

Finally, we need to estimate the difference between the truncated singular series $\mathfrak{S}(P_0)$ and $S(P_0)$. To this end, define

$$\mathscr{Q}(P_0) = \{q \in \mathbb{N} : q > P_0, p^i \mid q \Rightarrow p \leq P_0 \text{ and } i \leq k(p)\}$$

where $k(p)$ is either $k_C(p)$ or $k_\phi(p)$, depending on the context.
It is then clear that we have

$$R(P_0) := |\mathfrak{S}(P_0) - S(P_0)| \leq \sum_{q \in \mathscr{Q}(P_0)} |A(q)|.$$

For the case of non-singular forms and the inhomogeneous case, we then have the following result:

Lemma 4.6.6. *If C is ∞-good, then*

$$R(P_0) \ll M^{\frac{7}{3}} P_0^{-\frac{1}{3}+\varepsilon}.$$

Proof. From Lemma 4.5.4 we have $A(q) \ll M^{\frac{7}{3}} q^{-\frac{4}{3}+\varepsilon}$ and the claim follows immediately by summing over $q > P_0$. □

In the general case of a cubic form in 14 variables, we need to work slightly harder. The result is the following.

Lemma 4.6.7. *If C is ψ-good and $p^{k(p)} \leq M^{1+3\psi}$ for all $p \leq P_0$ and $\delta > 2$ satisfies*

$$0 < \frac{14}{14 - 6\delta} < 1 + 3\psi, \tag{$\mathfrak{S}_1$}$$

then

$$R(P_0) \ll M^{\frac{14\delta}{14-6\delta}} P_0^{2-\delta+\varepsilon}.$$

Proof. This is similar to the proof of Lemma 13 in [4] which seems to miss the factor δ in the exponent of M.

We first note that the bound

$$A(q) \ll M^{\frac{7}{3}} q^{-\frac{4}{3}+\varepsilon} + q^{1+\varepsilon} M^{-7\psi+\varepsilon}$$

from Lemma 4.5.4 implies that

$$A(q) \ll M^{\frac{7}{3}} q^{-\frac{4}{3}+\varepsilon} \ll q^{1-\delta}$$

holds uniformly for

$$M^{\frac{14}{14-6\delta}+\varepsilon} \ll q \ll M^{1+3\psi}. \tag{4.6.1}$$

Moreover, for q sufficiently large, we actually have the strict bound $A(q) \leq q^{1-\delta}$ in that range.

We now decompose a general $q \in \mathscr{Q}(P_0)$ into factors of the correct size. Writing $A = \frac{n}{n-6\delta} + \varepsilon$ and $B = 1 + 3\psi$, we find as in [4] a decomposition of the form

$$q = q_1 q_2 \dots q_{t+1}$$

for each $q \in \mathscr{Q}(P_0)$ with q_i pairwise coprime and so that $q_1, \dots, q_t$ are all in the range (4.6.1) and $q_{t+1} < M^A$. Indeed, the only assumption needed for this iterative decomposition is that $2A \leq B$ and $p_i^{k_i} \leq M^B$, which is true by our assumptions.

It then follows that

$$A(q) = A(q_1)A(q_2)\dots A(q_t)A(q_{t+1}) \ll (q_1 \dots q_t)^{1-\delta} q_{t+1} \ll M^A (q_1 \dots q_t)^{1-\delta},$$

using our result from above and the trivial bound $|A(q_{t+1})| \leq q_{t+1}$. With $q_0 = q_1 q_2 \dots q_t$ it now follows that

$$R(P_0) \ll M^A \sum_{q_{t+1} < M^A} \sum_{q_0 > \frac{P_0}{q_{t+1}}} q_0^{1-\delta} \ll M^A \sum_{q < M^A} \left(\frac{P_0}{q}\right)^{2-\delta} \ll M^{A\delta} P_0^{2-\delta}$$

as desired. □

We note that the condition $p^{k(p)} \leq M^{1+3\psi}$ will be satisfied if $P_0 \leq M^{1+3\psi}$ as well as $5n = 70 \leq 1 + 3\psi$, which we denote by $(\mathfrak{S}_2)$ and $(\mathfrak{S}_3)$.

4.7 Synthesis

4.7.1 The major arc contribution

Together with Lemma 4.4.4 we can summarize the major arc contribution as follows:

Lemma 4.7.1. *In the case of a non-singular cubic form C in 14 variables, we have*

$$\mathfrak{S}(P_0) \gg M^{-84-\varepsilon}$$

as soon as $P_0 \gg M^{259+\varepsilon}$.
In the case of an inhomogeneous cubic polynomial ϕ with $h \geq 14$, we have

$$\mathfrak{S}(P_0) \gg M^{-292(n^2-1)-\varepsilon}$$

as soon as $P_0 \gg M^{876(n^2-1)+7+\varepsilon}$.
Finally, in the case of a general cubic form in 14 variables, we have

$$\mathfrak{S}(P_0) \gg M^{-84-\varepsilon}$$

if we assume $(\mathfrak{S}_2)$ and $(\mathfrak{S}_3)$ as well as

$$P_0 \gg M^{\frac{1}{\delta-2}\cdot(84+\frac{14\delta}{14-6\delta})+\varepsilon}. \qquad (\mathfrak{S}_4)$$

Assuming additionally $(\mathfrak{M}_1)$, $(\mathfrak{M}_2)$ *and* $(\mathfrak{I}_1)$ *as well as*

$$P_0^3 u \ll \frac{P}{M^{8.5+T+\varepsilon}}, \tag{$\mathfrak{M}_3$}$$

we have in all three cases

$$\int_{\mathfrak{M}} S(\alpha) d\alpha \gg \frac{P^{n-3}}{M^{8.5+T+\varepsilon}}$$

where $T = 84$ *in the homogeneous case and* $T = 292(n^2-1)$ *in the inhomogeneous case.*

4.7.2 The minor arc contribution

We now use the different bounds obtained in Section 4.5 to bound the total minor arc contribution.

We dissect $\mathfrak{m}$ by an application of Dirichlet's Approximation Theorem for some parameter Q to be determined. For every $\alpha \in \mathbb{R}$, this yields an approximation

$$\alpha = \frac{a}{q} + \theta \quad \text{with} \quad q \leq Q, |\theta| \leq \frac{1}{qQ}.$$

The assumption $\alpha \in \mathfrak{m}$ then implies that $q > P_0$ or $\theta > \frac{u}{P^3}$. Note that since the contribution to the minor arcs from the range $|\theta| \leq \frac{1}{P^n}$ is clearly $\mathcal{O}(Q^2)$, we may also assume that $|\theta| \geq \frac{1}{P^n}$ for $q > P_0$. This allows us to apply a double dyadic decomposition with respect to both $|\theta|$ and q which yields

$$\int_{\mathfrak{m}} S(\alpha) d\alpha \ll Q^2 + P^{\varepsilon} \max_{R \leq Q, \phi \leq \frac{1}{RQ}} \Sigma(R, \phi),$$

where

$$\Sigma(R,\phi) := \sum_{q \sim R} \sum_{(a;q)=1} \int_{|\theta| \sim \phi} \left| S\left(\frac{a}{q} + \theta\right) \right| d\theta.$$

We note that the range of integration is a disjoint union of two intervals.

In view of the major arc contribution estimate from Lemma 4.7.1 it then suffices to show that

$$\Sigma(R,\phi) \ll \frac{P^{n-3}}{M^{8.5+T+\varepsilon}} \tag{4.7.1}$$

if we add the harmless assumption

$$Q^2 \ll \frac{P^{n-3}}{M^{8.5+T+\varepsilon}}. \tag{$\mathfrak{m}_1$}$$

To employ the mean-value estimates developed in Section 4.5, we apply the Cauchy-Schwarz inequality to obtain

$$\Sigma(R,\phi) \ll R\phi^{1/2} \left(\sum_{q\sim R} \sum_{(a;q)=1} \int_{|\theta|\sim\phi} \left| S\left(\frac{a}{q}+\theta\right)\right|^2 d\theta \right)^{1/2}.$$

We next cover the region $|\theta| \sim \phi$ by $\mathcal{O}\left(1+\frac{\phi}{\kappa}\right)$ intervals of size κ centered at values $\alpha = \frac{a}{q}+\theta$ with $|\theta|\sim\phi$. We conclude that

$$\Sigma(R,\phi) \ll R\phi^{1/2} \left(1+\frac{\phi}{\kappa}\right)^{1/2} \left(\sum_{q\sim R} \sum_{(a;q)=1} M\left(\frac{a}{q}+\theta,\kappa\right)\right)^{1/2}$$

for some $\theta \sim \phi$.

Using (4.5.11) and (4.5.13) we thus obtain

$$\Sigma(R,\phi) \ll R\phi^{1/2} \left(1+\frac{\phi}{\kappa}\right) \left(R^2 + \frac{\kappa P^{\frac{3n}{2}-1+\varepsilon}}{H^{n-1}} A(\theta,R,H,P)\right)^{1/2}.$$

Using the bound (4.5.14) for $A(\theta,R,H,P)$ we then find that

$$\Sigma(R,\phi) \ll R^2\phi^{1/2} \left(1+\frac{\phi}{\kappa}\right) \left(1+\kappa P^{2n-1+\varepsilon} HE\right)^{1/2} \tag{4.7.2}$$

with

$$E = \frac{1}{H^{14}} + (RHM\eta)^7 + \frac{\eta P^2}{(R\eta P^2)^7},$$

where we used the bound $\min(1,\eta P^2) \leq \eta P^2$ which turns out to be sufficient.

Suppose that we can show that $E \ll \frac{1}{H^{14}}$. Recalling that $\kappa \asymp \frac{(\log P)^2}{HP^2M^6}$, we have

$$1 + \frac{\phi}{\kappa} \ll \frac{P^\varepsilon \eta}{\kappa}$$

from the definition (4.5.12) of η. As $\kappa \gg \frac{1}{P^n}$, we then obtain that both summands in the last bracket of (4.7.2) are bounded by $\kappa P^{2n-1+\varepsilon} H^{-13}$. Still assuming $E \ll \frac{1}{H^{14}}$, we therefore find that

$$\Sigma(R, \phi) \ll R^2 \phi^{1/2} \eta^{1/2} P^{n-\frac{1}{2}+\varepsilon} H^{-13/2}.$$

Recalling our goal (4.7.1), it thus suffices to have

$$H^{13} \gg M^{2T+17} R^4 \phi^2 P^{5+\varepsilon}$$

as well as

$$H^{14} \gg M^{2T+16} R^4 \phi P^{3+\varepsilon}$$

in view of the definition of η. We hence take

$$H \asymp P^\varepsilon \cdot \max\left((M^{2T+17} R^4 \phi^2 P^5)^{1/13}, (M^{2T+16} R^4 \phi P^3)^{1/14}, 1 \right).$$

We need to check whether this choice satisfies $H \leq P$ and $H \leq M^\psi$. The condition $H \leq P$ will be satisfied if

$$M^{2T+17+\varepsilon} R^4 \phi^2 \ll P^{8-\varepsilon} \quad \text{and} \quad M^{2T+16} R^4 \phi \ll P^{11-\varepsilon}.$$

In view of $\phi R \leq \frac{1}{Q}$ and $R \leq Q$, it will therefore be satisfied if

$$M^{2T+17} \ll P^{8-\varepsilon} \tag{$\mathfrak{m}_2$}$$

and

$$M^{2T+16} Q^2 \ll P^{11-\varepsilon}. \tag{$\mathfrak{m}_3$}$$

Similarly, if $\psi < \infty$, the condition $H \leq M^\psi$ will be satisfied when

$$P^{5+\varepsilon} \ll M^{13\psi - 2T - 17} \tag{$\mathfrak{m}_4$}$$

and

$$P^{3+\varepsilon}Q^2 \ll M^{14\psi-2T-16}. \tag{$\mathfrak{m}_5$}$$

Summarizing, we have found an admissible choice for H that yields a satisfactory bound for the minor arc contribution under the assumption of $E \ll \frac{1}{H^{14}}$. We now need to enquire whether this condition is satisfied.
To this end, it is convenient to introduce the parameter

$$\phi_0 := (R^4 P^{31} M^{2T+30})^{-1/15}.$$

One then readily checks that for $\phi \leq \phi_0$, we have

$$H \asymp P^{\varepsilon} \max((M^{2T+16}R^4\phi P^3)^{1/14}, 1)$$

and

$$\eta \ll \frac{P^{\varepsilon}}{P^2 H M}$$

while for $\phi \geq \phi_0$, we have

$$H \asymp P^{\varepsilon} \max((M^{2T+17}R^4\phi^2 P^5)^{1/13}, 1)$$

and

$$\eta \ll \phi.$$

To prove $E \ll \frac{1}{H^{14}}$ we need to check whether $RH^3M\eta \ll 1$ and $\left(\frac{H^2}{R\eta P^2}\right)^7 \eta P^2 \ll 1$. We begin with the first condition.
If $\phi \leq \phi_0$, we have

$$\begin{aligned} RH^3M\eta &\ll P^{\varepsilon}\frac{QH^2}{P^2} \\ &\ll \frac{Q}{P^{2-\varepsilon}}\left(1 + \left(M^{2T+16}R^4\phi P^3\right)^{1/7}\right) \\ &\ll \frac{Q}{P^{2-\varepsilon}} + \left(\frac{M^{2T+16}Q^9}{P^{11-\varepsilon}}\right)^{1/7}. \end{aligned}$$

This is $\mathcal{O}(1)$ if we assume that

$$Q \ll \frac{P^{\frac{11}{9}-\varepsilon}}{M^{\frac{2T+16}{9}}}. \tag{$\mathfrak{m}_6$}$$

If, conversely, $\phi \geq \phi_0$, we have

$$\begin{aligned} RH^3M\eta &\ll RH^3M^{1+\varepsilon}\phi \\ &\ll \frac{M^{1+\varepsilon}}{Q} \cdot \left(1 + (M^{2T+17}R^4\phi^2P^5)^{3/13}\right) \\ &\ll \frac{M^{1+\varepsilon}}{Q} + \left(\frac{M^{6T+64}P^{15}}{Q^{13}}\right)^{1/13} \end{aligned}$$

which will be $\mathcal{O}(1)$ if

$$Q \gg P^{\frac{15}{13}+\varepsilon}M^{\frac{6T+64}{13}}. \tag{$\mathfrak{m}_7$}$$

We next turn to the question whether or not we have

$$\left(\frac{H^2}{R\eta P^2}\right)^7 \eta P^2 \ll 1. \tag{4.7.3}$$

Again, let us first suppose that $\phi \leq \phi_0$. Then $\eta \gg \frac{1}{P^2HM}$ so that

$$\left(\frac{H^2}{R\eta P^2}\right)^7 \eta P^2 \ll \left(\frac{H^3M}{R}\right)^7 \cdot \frac{1}{HM} = \frac{H^{20}M^6}{R^7}.$$

With our choice of H, this will be $\mathcal{O}(1)$ if $R^7 \gg M^6P^\varepsilon$ as well as $\phi \leq \phi_1$, where

$$\phi_1 := \frac{R^{\frac{9}{10}}}{P^{3+\varepsilon}M^{2T+20.2}}.$$

If, conversely, $\phi \geq \phi_0$, we have $\eta \asymp \phi$ so that (4.7.3) is equivalent to $H^{14} \ll R^7\phi^6P^{12}$.

With our choice of H this will be satisfied as soon as $\phi \geq \phi_2$, where

$$\phi_2 := \frac{M^{\frac{7(2T+17)}{25}}}{R^{\frac{7}{10}}P^{\frac{43}{25}-\varepsilon}}.$$

To summarize, we have obtained a satisfactory bound for the minor arc contribution as soon as $R \gg M^{6/7}P^{\varepsilon}$ as well as $\phi \leq \min(\phi_0, \phi_1)$ or $\phi \geq \max(\phi_0, \phi_2)$.

A quick computation shows that we will have $\phi_2 \leq \phi_0 \leq \phi_1$ as soon as $R \geq R_0$, where

$$R_0 := P^{\frac{4}{5}+\varepsilon}M^{\frac{4}{5}(2T+17)+2}$$

and so in that case the assumptions are always satisfied, while for $R \leq R_0$ we have $P^{-\varepsilon}\phi_1 \leq \phi_0 \leq P^{\varepsilon}\phi_2$.

We are thus left to treat the case where $R \leq R_0$ and $P^{-\varepsilon}\phi_1 \leq \phi_0 \leq P^{\varepsilon}\phi_2$ or $R \ll M^{6/7}P^{\varepsilon}$.

It is here that we use the bound obtained from the Weyl differencing. Noting that $(\mathfrak{m}_6)$ certainly implies $Q \leq P^{3/2}$, applying Lemma 4.5.3 we find that

$$\Sigma(R,\phi) \ll R^2\phi P^{n+\varepsilon}\left(MR\phi + \frac{1}{R\phi P^3} + M^{-2\psi}\right)^{\frac{7}{4}}.$$

Recalling our goal (4.7.1), it will suffice to show that

$$R^2\phi P^3\left(MR\phi + \frac{1}{R\phi P^3} + M^{-2\psi}\right)^{\frac{7}{4}} \ll M^{-8.5-T-\varepsilon}.$$

This will be satisfied as soon as

$$\frac{R^{\frac{1}{3}}M^{\frac{34+4T}{3}}}{P^{3-\varepsilon}} \ll \phi \ll \min\left\{\frac{1}{M^{\frac{41+4T}{11}}P^{\frac{12}{11}}R^{\frac{15}{11}}}, \frac{M^{\frac{7\psi}{2}-8.5-T+\varepsilon}}{R^2P^3}\right\}. \tag{4.7.4}$$

Our bound is therefore satisfactory as soon as

$$\phi_1 \gg \frac{R^{\frac{1}{3}}M^{\frac{34+4T}{3}}}{P^{3-\varepsilon}}$$

as well as

$$\phi_2 \ll \min\left\{\frac{1}{M^{\frac{41+4T}{11}}P^{\frac{12}{11}}R^{\frac{15}{11}}}, \frac{M^{\frac{7\psi}{2}-8.5-T+\varepsilon}}{R^2P^3}\right\}.$$

The first condition will be satisfied as soon as $R \geq R_1$, where

$$R_1 \asymp M^{\frac{50}{17}(2T+17)+\frac{96}{17}},$$

whereas the second one will be satisfied if

$$R \ll \frac{P^{\frac{346}{365}-\varepsilon}}{M^{\frac{254(2T+17)+350}{365}}}$$

and

$$R \ll \frac{M^{\frac{35\psi}{13}-\frac{3}{5}(2T+17)}}{P^{\frac{64}{65}-\varepsilon}}.$$

The latter two conditions are satisfied for $R \leq R_0$ if we assume

$$P \gg M^{\frac{91}{9}(2T+17)+\frac{440}{27}+\varepsilon} \tag{$\mathfrak{m}_8$}$$

and

$$P \ll M^{\frac{175\psi}{116}-\frac{91(2T+17)}{116}-\frac{130}{116}-\varepsilon}. \tag{$\mathfrak{m}_9$}$$

Finally, we are left to deal with the case where $R \leq R_1$. Here, we need to use that we are on the minor arcs. Assuming

$$P_0 \gg M^{\frac{50}{17}(2T+17)+\frac{96}{17}+\varepsilon}, \tag{$\mathfrak{m}_{10}$}$$

we may now assume that $R \leq R_1 \leq P_0$ and hence $\phi \geq \frac{u}{P^3}$.
It then suffices to check that again (4.7.4) is satisfied. Using $\phi R \leq \frac{1}{Q}$, this will be the case if

$$Q \gg P^{\frac{12}{11}+\varepsilon} M^{\frac{234(2T+17)+503}{187}} \tag{$\mathfrak{m}_{11}$}$$

and

$$Q \gg \frac{P^3}{M^{\frac{7\psi}{2}-\frac{117(2T+17)+192}{17}}} \tag{$\mathfrak{m}_{12}$}$$

as well as

$$u \gg M^{\frac{28(2T+17)+32}{17}}. \tag{$\mathfrak{m}_{13}$}$$

4.7.3 The endgame

We have now successfully bounded the minor arc contribution under all the assumptions $\mathfrak{m}_1$ up to $\mathfrak{m}_{13}$ and hence shown that $N(P) > 0$, so that there is a solution $|\mathbf{x}| < P$.

Finally, we are ready to choose the parameters and deduce our theorems.

For Theorem 4.1.1, we may assume that $\psi = \infty$. We then take M fixed and u and P_0 to be a small power of P, we choose $Q = P^{\frac{7}{6}}$ and then let $P \to \infty$. It is then readily checked that all assumptions are indeed satisfied with this choice. Moreover, we have shown that the minor arcs contribute $\mathcal{O}(P^{n-3-\varepsilon})$ and the major arcs contribute the main term from the claimed asymptotic formula.

For Theorems 4.1.3 and the non-singular case of Theorem 4.1.4, we may still assume that $\psi = \infty$ so that all conditions involving ψ are empty.

We choose $P_0 = M^{\frac{50(2T+17)+96}{17}+\varepsilon}$ to satisfy $(\mathfrak{m}_{10})$. We also choose $u = M^{\frac{28(2T+17)+32}{17}+\varepsilon}$ to satisfy $(\mathfrak{m}_{13})$.

We then choose

$$P = M^{\frac{373(2T+17)+640}{34}+\varepsilon}$$

to satisfy $(\mathfrak{M}_{13})$. One now checks that all other conditions are also satisfied with this choice after choosing e.g. $Q = \frac{P^{11/9}}{M^{\frac{2T+16}{9}}}$.

We then plug in the values $T = 292(n^2 - 1)$ and $T = 84$, respectively, to deduce Theorem 4.1.3 and the non-singular case of Theorem 4.1.4.

Finally, we deduce Theorem 4.1.4 in the case of a general cubic form. We choose u, P_0, P and Q as above where now $T = 84$ so that our choice is

$$u \approx M^{306.58}, P_0 \approx M^{549.76}, P \approx M^{2048.38}.$$

Condition $(\mathfrak{m}_9)$ now requires $\psi \geq 1454.8$.

We choose $\psi = 1454.8$ and note that with $\delta = 2.23$ all other conditions are now also satisfied.

If C is ψ-good, we therefore have a solution $|\mathbf{x}| \leq P \ll M^{2049}$. On the other hand, if C is not ψ-good, by Lemma 4.2.2 we have a solution $\mathbf{x} \ll M^{97+91\psi} \ll M^{132484}$ as desired.

4.8 List of assumptions

$$2P_0^2 u < P^3 \tag{$\mathfrak{M}_1$}$$

$$u \cdot P_0 \cdot M \cdot |\mathbf{z}|^2 \ll P \tag{$\mathfrak{M}_2$}$$

$$P_0^3 u \ll \frac{P}{M^{8.5+T+\varepsilon}} \tag{$\mathfrak{M}_3$}$$

$$0 < \frac{14}{14-6\delta} < 1+3\psi \tag{$\mathfrak{S}_1$}$$

$$P_0 \leq M^{1+3\psi} \tag{$\mathfrak{S}_2$}$$

$$\psi \geq 23 \tag{$\mathfrak{S}_3$}$$

$$P_0 \gg M^{\frac{1}{\delta-2}\cdot(84+\frac{14\delta}{14-6\delta})+\varepsilon} \tag{$\mathfrak{S}_4$}$$

$$u^2 M^{17+\varepsilon} \ll P \tag{$\mathfrak{I}_1$}$$

$$Q^2 \ll \frac{P^{n-3}}{M^{8.5+T+\varepsilon}} \tag{$\mathfrak{m}_1$}$$

$$M^{2T+17} \ll P^{8-\varepsilon} \tag{$\mathfrak{m}_2$}$$

$$M^{2T+16} Q^2 \ll P^{11-\varepsilon} \tag{$\mathfrak{m}_3$}$$

$$P^{5+\varepsilon} \ll M^{13\psi-2T-17} \tag{$\mathfrak{m}_4$}$$

$$P^{3+\varepsilon}Q^2 \ll M^{14\psi-2T-16} \tag{$\mathfrak{m}_5$}$$

$$Q \ll \frac{P^{\frac{11}{9}-\varepsilon}}{M^{\frac{2T+16}{9}}} \tag{$\mathfrak{m}_6$}$$

$$Q \gg P^{\frac{15}{13}+\varepsilon} M^{\frac{6T+64}{13}} \tag{$\mathfrak{m}_7$}$$

$$P \gg M^{\frac{91}{9}(2T+17)+\frac{440}{27}+\varepsilon} \tag{$\mathfrak{m}_8$}$$

$$P \ll M^{\frac{175\psi}{116}-\frac{91(2T+17)}{116}-\frac{130}{116}-\varepsilon} \tag{$\mathfrak{m}_9$}$$

$$P_0 \gg M^{\frac{50}{17}(2T+17)+\frac{96}{17}+\varepsilon} \tag{$\mathfrak{m}_{10}$}$$

$$Q \gg P^{\frac{12}{11}+\varepsilon} M^{\frac{234(2T+17)+503}{187}} \tag{$\mathfrak{m}_{11}$}$$

$$Q \gg \frac{P^3}{M^{\frac{7\psi}{2}-\frac{117(2T+17)+192}{17}}} \tag{$\mathfrak{m}_{12}$}$$

$$u \gg M^{\frac{28(2T+17)+32}{17}} \tag{$\mathfrak{m}_{13}$}$$

Bibliography

[1] B. J. Birch. “Homogeneous forms of odd degree in a large number of variables”. In: *Mathematika* 4 (1957), pp. 102–105.

[2] J. Brandes and R. Dietmann. “Rational lines on cubic hypersurfaces”. In: *Math. Proc. Cambridge Philos. Soc.* 171.1 (2021), pp. 99–112.

[3] J. Brandes and R. Dietmann. “Rational lines on cubic hypersurfaces II”. In preparation.

[4] T. D. Browning, R. Dietmann, and P. D. T. A. Elliott. “Least zero of a cubic form”. In: *Math. Ann.* 352.3 (2012), pp. 745–778.

[5] T. D. Browning and D. R. Heath-Brown. “Integral points on cubic hypersurfaces”. In: *Analytic number theory.* Cambridge Univ. Press, Cambridge, 2009, pp. 75–90.

[6] T. D. Browning and P. Vishe. “Cubic hypersurfaces and a version of the circle method for number fields”. In: *Duke Math. J.* 163.10 (2014), pp. 1825–1883.

[7] J. Brüdern, R. Dietmann, J. Y. Liu, and T. D. Wooley. “A Birch-Goldbach theorem”. In: *Arch. Math. (Basel)* 94.1 (2010), pp. 53–58.

[8] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities.* Second. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2005.

[9] H. Davenport. “Cubic forms in 29 variables”. In: *Proc. Roy. Soc. London Ser. A* 266 (1962), pp. 287–298.

[10] H. Davenport. “Cubic forms in sixteen variables”. In: *Proc. Roy. Soc. London Ser. A* 272 (1963), pp. 285–303.

[11] H. Davenport. “Cubic forms in thirty-two variables”. In: *Philos. Trans. Roy. Soc. London Ser. A* 251 (1959), pp. 193–232.

[12] H. Davenport and D. J. Lewis. “Non-homogeneous cubic equations”. In: *J. London Math. Soc.* 39 (1964), pp. 657–671.

[13] R. Dietmann and T. D. Wooley. “Pairs of cubic forms in many variables”. In: *Acta Arith.* 110.2 (2003), pp. 125–140.

[14] B. Green and T. Tao. “The primes contain arbitrarily long arithmetic progressions”. In: *Ann. of Math. (2)* 167.2 (2008), pp. 481–547.

[15] R. Hartshorne. *Algebraic geometry.* Vol. 52. Springer Science & Business Media, 2013.

[16] D. R. Heath-Brown. “Cubic forms in 14 variables”. In: *Invent. Math.* 170.1 (2007), pp. 199–230.

[17] D. R. Heath-Brown. “Cubic forms in ten variables”. In: *Proc. London Math. Soc. (3)* 47.2 (1983), pp. 225–257.

[18] C. Hooley. “On nonary cubic forms.” In: *Journal für die reine und angewandte Mathematik* 386 (1988), pp. 32–98.

[19] C. Hooley. “On nonary cubic forms. II”. In: *J. Reine Angew. Math.* 415 (1991), pp. 95–165.

[20] D. J. Lewis. “Cubic forms over algebraic number fields”. In: *Mathematika* 4 (1957), pp. 97–101.

[21] D. J. Lewis. “Cubic homogeneous polynomials over p-adic number fields”. In: *Ann. of Math. (2)* 56 (1952), pp. 473–478.

[22] J. Liu and L. Zhao. *On forms in prime variables*. 2021. arXiv: `2105.12956 [math.NT]`.

[23] D. Lloyd. "Bounds for solutions of Diophantine equations". In: *University of Adelaide Ph.D. thesis* (1975).

[24] P. A. B. Pleasants. "Cubic polynomials over algebraic number fields". In: *J. Number Theory* 7.3 (1975), pp. 310–344.

[25] C. P. Ramanujam. "Cubic forms over algebraic number fields". In: *Proc. Cambridge Philos. Soc.* 59 (1963), pp. 683–705.

[26] C. Ryavec. "Cubic forms over algebraic number fields". In: *Proc. Cambridge Philos. Soc.* 66 (1969), pp. 323–333.

[27] C. M. Skinner. "Rational points on nonsingular cubic hypersurfaces". In: *Duke Math. J.* 75.2 (1994), pp. 409–466.

[28] G. L. Watson. "Non-homogeneous cubic equations". In: *Proc. London Math. Soc. (3)* 17 (1967), pp. 271–295.

[29] T. D. Wooley. "Linear spaces on cubic hypersurfaces, and pairs of homogeneous cubic equations". In: *Bull. London Math. Soc.* 29.5 (1997), pp. 556–562.

[30] S. Yamagishi. *Diophantine equations in primes: density of prime points on affine hypersurfaces II*. 2021. arXiv: `2111.06122 [math.NT]`.