ROOTS OF POLYNOMIAL CONGRUENCES

By

MATTHEW C. WELSH

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Henryk Iwaniec

And approved by

New Brunswick, New Jersey

May, 2019

ABSTRACT OF THE DISSERTATION

Roots of Polynomial Congruences

By MATTHEW C. WELSH

Dissertation Director: Henryk Iwaniec

In this dissertation we derive and then investigate some consequences of a parametrization of the roots of polynomial congruences. To motivate the later chapters, we begin in chapter two by reviewing known results, presented in our own style, about the roots of the quadratic congruence $\mu^2 \equiv -1 \pmod{m}$. We also review in chapter two applications of the parametrization to the equidistribution and well-spacing of these roots. In chapter three we generalize this classical parametrization of the roots of a quadratic congruence to the cubic congruence $\mu^3 \equiv 2 \pmod{m}$. Several new phenomena are revealed in our derivation, but a special case of the cubic parametrization is seen to be roughly analogous to the quadratic case. We use this case to prove a spacing property analogous to the well-spacing of the quadratic roots, but unfortunately between the points $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ instead of the $\frac{\mu}{m}$ themselves. In chapter four we consider this special case for an arbitrary polynomial congruence of any degree, deriving a parametrization for the roots of these congruences. And just as in chapter three, we are able to prove a spacing property for certain points related to the roots. Finally, in chapter five we return to the congruence $\mu^3 \equiv 2 \pmod{m}$ to explore some of the new phenomena mentioned above with a view towards obtaining equidistribution and well-spacing results. We unfortunately do not prove any concrete results in these directions.

Acknowledgements

I would first like to thank my advisor, Professor Henryk Iwaniec, for his encouragement, support, and guidance. I feel very fortunate to have been able to spend so many lunches with him discussing not only mathematics but life in general. My committee, Professors Steve Miller, Alex Kontorovich, and Nigel Pitt, deserve many thanks as well. Steve provided me much appreciated encouragement when I first began the project that would become this dissertation, and I would also like to thank him for sharing techniques that have made many calculations less of a headache. Alex has made a great effort to build and maintain a community of number theorists among the graduate students and postdocs here at Rutgers, from which I gratefully benefit, and I would also like to thank him for his willingness to look for connections between topics in math. Nigel, who I have had the pleasure to get to know during the past few months, has given me invaluable career advice and much needed motivatation, and I would also like to thank him for his close reading of this dissertation and helpful edits.

Graduate school would be a lonely experience if not for other graduate students and postdocs. In particular I would like to thank all those who participated in the many number theory seminars at Rutgers, and especially the organizers of these seminars over the years. Special thanks go to my friend and colleague Surya Teja Gavva, from whom I have learned a tremendous amount. Outside of the math department, I would like to thank my comrades on the TA/GA steering committee of the Rutgers AAUP-AFT for their work towards winning a democratic university. In particular I would like to thank Anna Barcy for the many hours she has given to the graduate student community.

Finally, and most importantly, I would like to thank my partner and closest friend, Maria Isabel Espinoza Paredes, and my family, Janet, Kevin, Brenda, and Colin, for their love and support.

Table of Contents

Chapter 1

Introduction

The study of the distribution of roots of quadratic congruences provides one of the best applications of the spectral theory of automorphic forms on $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ to questions in arithmetic. The first result on the distribution of these roots came with no reference to automorphic forms in [Hoo63], where it was incidentally proved that the sequence

$$\left\{\frac{\nu}{m} \in \mathbb{R}/\mathbb{Z} \;:\; \nu^2 + D \equiv 0 \pmod{m}\right\}$$

is equidistributed modulo 1. This equidistribution is of course equivalent to finding cancellation in the following Weyl sum:

$$\sum_{m \leq x} \sum_{f(v) \equiv 0 (m)} e\left(\frac{h\nu}{m}\right), \quad h \in \mathbb{Z},$$

with f a fixed quadratic polynomial. In fact strong bounds for this Weyl sum is really what one needs for applications. For an example of such an application see [Iwa78], where bounds for the above sum, and similar sums with m restricted to be divisible by an integer d, are used to show that n^2+1 is infinitely often a prime or product of two primes.

In [Hoo63] the bound $\ll_h x^{3/4}(\log x)^2$ for the Weyl sum above (with $h \neq 0$, of course) was proved using the Weil bound for Kloosterman sums. In retrospect one could have already seen the relevance of the spectral theory of automorphic forms in light of the appearance of a sum of Kloosterman sums. However the introduction of this theory in [Byk87] came in a different package. Only for $D > 0$, [Byk87] proved the bound $\ll_h x^{2/3}\log x$ for the Weyl sum, or $\ll_h x^{1/2}(\log x)^2$ for a smooth version, which is best possible. The proof proceeds by relating the smoothed version to a Poincaré series on $SL_2(\mathbb{Z})\backslash\mathbb{H}$, which was then estimated by its spectral expansion and, in the end, bounds on the Fourier coefficients of automorphic forms. Only a little later[1] bounds of the same strength where produced by a slightly different method in [Hej86] for specific examples of D, all negative.

The strategy of [Byk87] matured in [DFI95], where the method was extended to apply to the Weyl sum with m restricted to be divisible by an integer d. This restricted Weyl sum was also related to a Poincaré series, but now for a congruence subgroup of $SL_2(\mathbb{Z})$, and estimates were produced with enough uniformity in h and d to use in a sieve idea coupled with bilinear forms techniques. The final result was a proof of the equidistribution of the sequence

$$\left\{\frac{\nu}{p} \in \mathbb{R}/\mathbb{Z} \ : \ \nu^2 + D \equiv 0 \pmod{p}\right\},$$

where p is a prime number and $D > 0$.

In contrast to the previously mentioned results [Hoo63] and [Iwa78] that did not use the spectral theory of automorphic forms, the method originating in [Byk87] did not directly transform to Kloosterman sums. It was in [T0́0] that the restriction to $D > 0$ was removed from the equidistribution of roots of quadratic congruences to prime moduli in [DFI95] by doing exactly that: transforming the Weyl sum and then using the spectral theory of automorphic forms to bound the resulting sum of Kloosterman sums.

More recent years have seen significant development of the spectral theory of automorphic forms on $SL_3(\mathbb{Z})\backslash SL_3(\mathbb{R})$. For just a taste, one can see the Kuznetsov-like trace formulae of [Li10], [But12], and [Blo13]. And while this spectral theory has seen some great applications, one can see in addition [BBM17], to the author's knowledge direct applications to arithmetic have been limited. Although there is hope, see for example the introduction to [But12] and the introduction to section 4.1 of [Ter88], that the future will see applications of this theory to questions of a cubic nature in much the same way that the spectral theory of $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ has been so fruitful to questions

[1] Although our reference to [Byk87] dates to 1987, the original version was published in 1981 in Russian.

of a quadratic nature, the distribution of quadratic congruences being perhaps the best example.

It was in this spirit that the author undertook the study of the distribution of cubic congruences with the hope that we might see the application of the spectral theory of $SL_3(\mathbb{Z})\backslash SL_3(\mathbb{R})$ to this question. But despite some encouraging results from our investigations it seems to the author that this hope may have been premature.

Before continuing, we should remark that the equidistribution of the sequence

$$\left\{\frac{\nu}{m} \in \mathbb{R}/\mathbb{Z} \;:\; f(\nu) \equiv 0 \pmod{m}\right\},$$

with f any irreducible, integral polynomial has been proven in [Hoo64]. But as mentioned before, the real interest in the equidistribution lies in strong bounds for the Weyl sum, and the estimate obtained in [Hoo64], saving only a fraction of a power of a logarithm, is far too weak for any applications along the lines of those in [Hoo63] or [Iwa78]. For these applications, we imagine one would need at least a power savings.

We also remark that, for $f(X) = X^3 - 2$, the setting in which we will concern ourselves in chapters 3 and 5, [Hoo78] has touched on the Weyl sums

$$\sum_{m \leq x} \sum_{\mu^3 \equiv 2(m)} e\left(\frac{h\mu}{m}\right), \quad h \in \mathbb{Z},$$

which we consider our ultimate goal, even though we do not make any concrete progress towards a bound here. Indeed in [Hoo78] a parametrization of m and $\mu \pmod{m}$ with $\mu^3 \equiv 2 \pmod{m}$ is given, and we obtain this same parametrization in corollary 3.5 by different means in section 3.2. Our method however reveals something more, namely theorems 3.1 and 3.4.

Finally, we remark that some of the content of chapter 3, and also some of this introduction, appeared in an earlier work by the author, [Wel18]. To be specific, corollary 3.3, corollary 3.5, theorem 3.8, theorem 3.10, and theorem 3.11 were all proved previously in [Wel18]. However, we do take a slightly different approach here, perhaps most importantly by considering corollary 3.3 as a consequence of the more general theorem 3.1. This theorem, theorem 3.1, takes a central role in chapter 5.

We start in chapter 2 by proving in our own way the parametrization of the roots of a quadratic congruence, specifically $\mu^2 + 1 \equiv 0 \pmod{m}$. This parametrization, theorem 2.2, forms the foundation for all the results on the distribution of these roots mentioned above. Differing only superficially from the classical derivation of the parametrization, we first prove in section 2.1 a correspondence, theorem 2.1, between the roots, m and μ (mod m), and ideals in $\mathbb{Z}[i]$. This can be seen to be the same as the classical connection between the roots of the congruence and binary quadratic forms with discriminant -4. We then finish our derivation of the parametrization, theorem 2.2, in section 2.2 using the fact that $\mathbb{Z}[i]$ has class number one, just as in the classical derivation that uses the fact that all binary quadratic forms with discriminant -4 are $SL_2(\mathbb{Z})$ equivalent to $X^2 + Y^2$.

This parametrization of quadratic roots then leads to an approximation, given in theorem 2.3, to $\frac{\mu}{m}$ within $O\left(\frac{1}{m}\right)$ by a fraction with much smaller denominator, size $\sqrt{m}$. This approximation was a key step towards bounding the Weyl sum for the roots in [Hoo63], [Iwa78], and [T0́0]: all the works where the Weyl sum was transformed directly to a sum of Kloosterman sums. We outline this transformation in section 2.5.

In a different direction, the approximation given in theorem 2.3 is also used to provide upper bounds for the number of the $\frac{\mu}{m}$ contained in short intervals, theorem 2.6. In a sense this is about the same as proving a lower bound for the space between different $\frac{\mu}{m}$, only a little weaker. This well-spacing point of view however is not essential and leads to unnecessary difficulties, so in what follows we avoid it in preference to upper bounding the number in short intervals. And we will, somewhat confusingly we admit, refer to these upper bounds in short intervals as spacing results. We present a consequence of this spacing property in theorem 2.5, namely a large sieve inequality for the roots. Although the spacing property is not as refined as the equidistribution, this large sieve has applications beyond what can be produced via bounds for the Weyl sum, see for example [FI97] and [FI98]. This should be encouraging to us because although we do not prove an equidistribution result for the roots of a cubic congruence, we do obtain a somewhat analogous large sieve inequality for these cubic roots in chapter 3. However we unfortunately do not have any applications at this time.

We emphasize that none of the results in chapter 2 are new, only our presentation is our own. The purpose of this chapter is rather to provide motivation for what follows in chapter 3. As mentioned above, generalizing the situation of chapter 2, in particular section 2.5 where the spectral theory of automorphic forms comes into play, to roots of higher degree congruences was a primary motivation for the investigations here, a motivation that is unfortunately still unrealized.

We begin our this generalization in chapter 3 by considering roots of the specific cubic congruence $\mu^3 \equiv 2 \pmod{m}$ in order to facilitate the exposition. The structure of chapter 3 is similar to chapter 2. Just as in the quadratic case, we begin in section 3.1 by proving a correspondence between the roots and ideals in the cubic ring $\mathbb{Z}[2^{1/3}]$, theorem 3.1. This correspondence differs from the quadratic correspondence in that ideals no longer correspond to a single m and $\mu \pmod{m}$ satisfying $\mu^3 \equiv 2 \pmod{m}$, but rather a pair $\mu \pmod{m}$ and $\nu \pmod{n}$ of genuinely distinct roots of the congruence. However, one can specialize to the case with $n = 1$, giving a correspondence between ideals and roots $\mu \pmod{m}$ and, as it turns out, $\mu^2 \pmod{m}$. This is the content of corollary 3.3, and the associated corollary 3.2 gives a characterization of the special ideals corresponding to the case $n = 1$: they are the ideals I such that $\mathbb{Z}[2^{1/3}]/I$ is additively cyclic, or alternatively the ideals with only degree one prime factors, none of which are conjugate.

These correspondences, theorem 3.1 and corollary 3.3, respectively give rise to the parametrizations in theorem 3.4 and corollary 3.5, which are proved in section 3.2. Theorem 3.4 gives a parametrization for both of the roots $\mu \pmod{m}$ and $\nu \pmod{n}$ from the correspondence in theorem 3.1, while corollary 3.5 gives the parametrization of $\mu \pmod{m}$ and $\mu^2 \pmod{m}$. As mentioned previously, this latter parametrization has been proved previously using ternary cubic forms in [Hoo78], although the author found the derivation presented here without knowledge of this work. On the other hand, the more general parametrization in theorem 3.4 is to the author's knowledge new.

We continue in section 3.3 just as in chapter 2 by using the parametrization of the roots to derive, via an LU decomposition, an approximation to the $\frac{\mu}{m}$. The LU decomposition also gives an expression for $\frac{\nu}{n}$ that typically leads to an approximation when it is applied in the context of theorem 3.4, and it gives an approximation to $\frac{\mu^2}{m}$

in the context of corollary 3.5. These approximations are given in theorems 3.7 and 3.8. In theorem 3.7 it is shown that the approximations to $\frac{\mu}{m}$ are all within $O\left(\frac{1}{m}\right)$ and have denominator of size $m^{2/3}n^{1/3}$, perhaps suggesting that it is best to assume $n = 1$. For the remainder of chapter 3 we only consider this case, and we return to general n in chapter 5.

In this special case with $n = 1$, there is another way to derive the approximations to $\frac{\mu}{m}$ and $\frac{\mu^2}{m}$ that is equivalent to the LU decomposition apart from the point of view. We will find this point of view useful in the later sections of chapter 3. The alternative derivation presented in section 3.3.2 manifests the approximation to the point $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ as the intersection of two lines in $\mathbb{R}^2$. The utility of this point of view is revealed in the following section, section 3.4, where we prove proposition 3.9 on the spacing between torsion points in $\mathbb{R}^2/\mathbb{Z}^2$. When this proposition is applied to the approximations of theorem 3.8, it yields theorem 3.10, a spacing property for the points $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$. We then close chapter 3 by deriving a 2-dimensional large sieve inequality, theorem 3.11, almost as a corollary from theorem 3.10.

In both chapters 2 and 3 we used particular examples to facilitate the exposition, the examples being $\mu^2 \equiv -1 \pmod{m}$ and $\mu^3 \equiv 2 \pmod{m}$. Apart from being concrete, these examples have three key properties that aid our proofs: the associated rings $\mathbb{Z}[i]$ and $\mathbb{Z}[2^{1/3}]$ are maximal, these rings are both principle ideal domains, and the units of both rings are convenient to work with. In chapter 4 we both address these concerns and generalize to higher degree, considering a general polynomial congruence, say of degree d. However, we only generalize the aspects of chapter 3 that pertain to the special case with $n = 1$ referred to above. We recall that corollary 3.2 says that the ideals with $n = 1$ in theorem 3.1 are exactly the ideals I with $\mathbb{Z}[2^{1/3}]/I$ additively cyclic. For the general congruence, we prove proposition 4.1 in section 4.1, which characterizes in a similar way the ideals with cyclic quotient, and we prove a correspondence, theorem 4.2, between these ideals and the roots $\mu \pmod{m}$ of the congruence.

As in chapter 3, the general correspondence of theorem 4.2 is used to derive a parametrization of the roots, with a few caveats. First, we remark that the proof of theorem 4.2, in particular the proof of proposition 4.1, seems to depend crucially on the ambient ring being monogenic, $\mathbb{Z}[\alpha]$ for a root α of the polynomial congruence under consideration. This causes us no problems in principle since we are taking the point of view that the polynomial, rather than the order, is the object of interest. However, we do then run into the issue of ideals possibly being not invertible. While it would be interesting to the author to characterize in simpler terms exactly the roots that correspond via theorem 4.2 to non-invertible ideals, here we brush away the issue by considering in the later sections of chapter 4 only those roots that do correspond to invertible ideals.

A second issue is that the number of ideal classes of $\mathbb{Z}[\alpha]$ will typically be greater than one, but this also does not cause too much difficulty since we can simply consider each ideal class separately when it comes to the question of spacing. We execute this splitting in section 4.2, resulting in a parametrization of not only the root $\mu \pmod{m}$, but also all of the $\mu^j \pmod{m}$, $1 \leq j \leq d-1$. This parametrization is given in theorem 4.3.

We continue in section 4.3 as we did in chapters 2 and 3, finding an approximation to the point $\left(\frac{\mu}{m}, \ldots \frac{\mu^{d-1}}{m}\right)$ within $O\left(\frac{1}{m}\right)$, see theorem 4.4. Here we encounter another a difficulty of the general setting. In chapters 2 and 3 we chose a specific fundamental domain for the action of the units in the relevant ring in order to ensure a good error, see proposition 3.6. This can be replicated easily as long as $\mathbb{Q}(\alpha)$ has at least one real embedding, but the author does not know how to make this work when $\mathbb{Q}(\alpha)$ has no real embeddings. Instead we sacrifice some concreteness in the approximation of theorem 4.4 in order to obtain the error bound.

The approximation, as it is in chapter 3, is a torsion point in $\mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$ with torsion $\asymp m^{1-1/d}$. Since again the spacing between torsion points can be controlled by the size of the coefficients of the integral lines containing the point, see proposition 4.5, which is proved in section 4.4, we derive the approximation as the intersection of $d-1$ co-dimension 1 hyper-planes in $\mathbb{R}^{d-1}$. The lines passing through the point are then realized as the intersections of subsets of $d-2$ of these hyper-planes, which leads to the well-spacing of the approximations and then to a spacing property, theorem 4.6, of the points $\left(\frac{\mu}{m}, \ldots, \frac{\mu^{d-1}}{m}\right)$ themselves. We remark that this spacing property can be used to obtain a $(d-1)$-dimensional large sieve inequality, quoted in theorem 4.7, but we do not present a derivation here because the proof is so similar to that of section 3.5.

After achieving some success in generalizing the results on the roots of a quadratic congruence outlined in chapter 2, we should appraise our results. And unfortunately we must admit that they fall short in several ways. For example, the large sieve for the roots of the congruence $\mu^2 \equiv -1 \pmod{m}$, theorem 2.5, plays a key role in [FI97], where it is used to establish level of distribution results for the sequence of integers of the form $a^2 + b^2$ where b is restricted to be prime, for example. A main observation in this application is that one needs an optimal level of distribution in order for sieve/bilinear forms techniques to be able to capture primes, as is done in [FI97]. In turn, this optimal level of distribution depends on the large sieve of theorem 2.5 exhibiting square root cancellation on average. One might hope that the large sieve for roots of the cubic congruence $\mu^3 \equiv 2 \pmod{m}$, theorem 3.11, might be used to establish level of distribution results for sequences of integers with the form $a^3 + 2b^3 + 4c^3 - 6abc$ with some restrictions on b and c. Indeed one can obtain some results in this direction, but to the author's ability these results will not be optimal because the large sieve of theorem 3.11 does not exhibit square-root cancellation on average. On top of this weakness is that level of distribution results for sequences of the form $a^3 + 2b^3$ with restrictions on b would be far more attractive, and it would seem that one needs a 1-dimensional large sieve for this. Even further, we have not yet even considered the question of equidistribution of the roots!

We recall that in section 2.5 we outline a proof for the equidistribution of the $\frac{\mu}{m}$ for the congruence $\mu^2 \equiv -1 \pmod{m}$. The key in this proof is the transformation of the Weyl sum for the $\frac{\mu}{m}$ into a sum of Kloosterman sums, which is achieved by first relating the Weyl sum to a sort of Poincaré series for $SL_2(\mathbb{Z})$ via the parametrization and approximation of theorems 2.2 and 2.3. We remark that the fact that the parametrization uses every coset in the quotient of $SL_2(\mathbb{Z})$ by lower triangular matrices is crucial for making this translation. On the other hand, the parametrization of the cubic roots, theorem 3.4, the cosets used form a small subset; indeed the Plücker coordinates of the used cosets satisfy an additional quadratic constraint in addition to the generic one.

In chapter 5, we take the step suggested by this observation: we parametrize those cosets used in theorem 3.4 in a way that works naturally with the approximation in theorem 3.7. The main tool for realizing this is the arithmetic of binary cubic forms, discussed in section 5.1. The entrance of binary cubic forms is perhaps not surprising, but unfortunately the situation is not nearly so simple as in the quadratic case. In fact this point of view does not lead us to us to any concrete results here. Nevertheless, we do discuss in section 5.3 an attempt at applying these ideas to transform the Weyl sum for the roots of the congruence $\mu^3 \equiv 2 \pmod{m}$. Before this we make a short digression to estimate the number of ideals ordered by the corresponding m and n in theorem 3.1. What we need for the transformations in 5.3 is an estimate, proposition 5.4, for the number of ideals corresponding to a fixed $\mu \pmod{m}$ with control on the size of the corresponding n. And although it is not entirely relevant to our purposes, we end up computing the co-type zeta function for the ring $\mathbb{Z}[2^{1/3}]$, proposition 5.3. The co-type zeta function here is built out of the invariant factors of the ideals in $\mathbb{Z}[2^{1/3}]$, and in general it has been used to study the growth of subgroups, see for example [Pet07] and [CKK17].

Chapter 2

Roots of a quadratic congruence

The classic derivation of the parametrization of the roots of quadratic congruence $\mu^2 \equiv -1 \pmod{m}$, theorem 2.2 below, proceeds by considering binary quadratic forms with discriminant -4. It happens that all these forms are equivalent under the action of $SL_2(\mathbb{Z})$, and in particular equivalent to the form $X^2 + Y^2$. On the other hand, if we have a solution to $\mu^2 \equiv -1 \pmod{m}$, then the integral binary quadratic form

$$mX^2 + 2\mu XY + \frac{\mu^2+1}{m}Y^2 \tag{2.1}$$

has discriminant -4, and we note that every form with discriminant -4 can be written this way. By the $SL_2(\mathbb{Z})$-equivalence of this form to $X^2 + Y^2$, there exists integers a, b, u, and v with $au + bv = 1$, or in other words there is a matrix

$$\begin{pmatrix} a & v \\ -b & u \end{pmatrix} \in SL_2(\mathbb{Z}), \tag{2.2}$$

such that

$$(aX + vY)^2 + (-bX + uY)^2 = mX^2 + 2\mu XY + \frac{\mu^2+1}{m}Y^2. \tag{2.3}$$

Expanding the left side of (2.3) and equating coefficients, we obtain the parametrization given in (2.9).

We remark that the matrix (2.2) used in this parametrization is not unique. For one, we can multiply on the right by the matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.4}$$

which all stabilize the form $X^2 + Y^2$, and the corresponding substitutions would still satisfy (2.3). We can resolve this ambiguity by insisting that $a > 0$ and $-a < b \leq a$.

Secondly, we observe that multiplication of (2.2) on the left by matrices of the form

$$\begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{2.5}$$

shifts the resulting μ in (2.3) by a multiple of m – this can be seen from the parametrization (2.9) or from the action of the matrices (2.5) on the form (2.1).

This argument using binary quadratic forms can be generalized to derive the parametrization of the cubic congruence stated in corollary 3.5 using ternary cubic forms; this is done in [Hoo78]. However, it is simpler and perhaps more illuminating to use ideals in a cubic ring, in our case $\mathbb{Z}[2^{1/3}]$, to prove theorem 3.4, and so in this chapter we will present an alternative proof of the quadratic parametrization, theorem 2.2, using ideals in the quadratic ring $\mathbb{Z}[i]$. Nevertheless the parallels between the two approaches should be quite apparent in this quadratic setting.

We begin in section 2.1 with the proof of the following theorem:

Theorem 2.1. *An ideal $I \subset \mathbb{Z}[i]$ has a unique basis $\{\beta_1, \beta_2\}$ of the form*

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} mn & 0 \\ \mu n & n \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \tag{2.6}$$

where $m, n > 0$ and $\mu^2 \equiv -1 \pmod{m}$, and, for uniqueness to hold, μ is considered as a residue class $\pmod{m}$.

Conversely, given any such m, n, and μ, the sublattice I of $\mathbb{Z}[i]$ with basis $\{\beta_1, \beta_2\}$ given by (2.6) is an ideal.

This theorem is the analogue of the classical connection between roots of the congruence $\mu^2 \equiv -1 \pmod{m}$ and binary quadratic forms of discriminant -4. In fact, one can compute quite easily that the binary quadratic form (2.1) is, at least up to the sign of μ, the norm-form associated to the ideal as stated in theorem 2.1. We remark that n in the statement of this theorem plays no essential role as it simply acts as an overall scaling of the ideal. That is n is simply the largest rational integer divisor of the ideal I. In what follows we lose nothing in assuming that $n = 1$.

Just as in the classical proof, where the fact that all binary quadratic forms of discriminant -4 are $SL_2(\mathbb{Z})$ equivalent to $X^2 + Y^2$ lead to a parametrization of the roots of the quadratic congruence, the fact that all ideals in $\mathbb{Z}[i]$ are principle leads to the following theorem:

Theorem 2.2. *Let a and b be integers such that*

$$\gcd(a,b) = 1, \tag{2.7}$$

so that there are integers u and v so that

$$au + bv = 1. \tag{2.8}$$

Then

$$\begin{aligned} m &= a^2 + b^2 \\ \mu &= av - bu \end{aligned} \tag{2.9}$$

satisfy

$$\mu^2 \equiv -1 \pmod{m}. \tag{2.10}$$

Moreover, every m and μ satisfying (2.10) arises from (2.9), and in fact appears uniquely if we assume, in addition to (2.7), that $a > 0$ and $-a < b \leq a$.

This theorem is proved in the language of ideals in $\mathbb{Z}[i]$ in section 2.2. We remark that, just as in the classical proof using binary quadratic forms, there is an ambiguity here in the definition of u and v via the equation (2.8). But, as is elaborated in section 2.2, we will see that different choices of u and v satisfying (2.8) will simply give different representatives of $\mu \pmod{m}$.

Our next theorem uses the parametrization to derive some statistical information about the roots, it is as follows:

Theorem 2.3. *Let m, μ, a, b, u, and v be as in theorem 2.2. Then*

$$\frac{\mu}{m} = \frac{v}{a} + O\left(\frac{1}{m}\right). \tag{2.11}$$

In this theorem we have an approximation to the fraction $\frac{\mu}{m}$ by another fraction with denominator of size $\sqrt{m}$, and to an error of size $\frac{1}{m}$. We remark that Dirichlet's theorem on Diophantine approximation guarantees that most real numbers will have such an approximation, so the content of theorem 2.3 is really to give a formula for the approximation in this situation. The utility of this approximation can be seen in two applications that we turn to in sections 2.5 and 2.4.

Our first application concerns the distribution of the roots of the quadratic congruence. The Weyl criterion for equidistribution modulo 1 tells us that the sequence of $\frac{\mu}{m}$ (mod 1) with $\mu^2 \equiv -1 \pmod{m}$ and ordered by the size of m is equidistributed if for all non-zero integers h

$$\frac{1}{M} \sum_{m \leq M} \sum_{\mu^2 \equiv -1 (m)} e\left(\frac{h\mu}{m}\right) \to 0 \tag{2.12}$$

as $M \to \infty$. We remark that the factor $\frac{1}{M}$ is the correct scaling because there are asymptotically $\frac{3}{2\pi}M$ of the roots $\mu \pmod{m}$ with $m \leq M$. Denoting by $\rho(m)$ the number of roots modulo m, we can see this fact quite easily by the calculation

$$\sum_{m=1}^{\infty} \frac{\rho(m)}{m^s} = \left(1 + \frac{1}{2^s}\right) \prod_{p \equiv 1(4)} \left(1 + \frac{2p^{-s}}{1 - p^{-s}}\right) = \frac{L(s, \chi_4)}{\zeta(2s)} \zeta(s), \tag{2.13}$$

where χ_4 is the primitive Dirichlet character modulo 4.

Ignoring the dependence on h for exposition, we state the following theorem of Bykovskii [Byk87][1]:

Theorem 2.4. *Let f be a fixed smooth, compactly supported function on the positive real numbers. We have for any real $M > 0$ and any nonzero integer h,*

$$\sum_{m} f\left(\frac{m}{M}\right) \sum_{\mu^2 \equiv -1 (m)} e\left(\frac{h\mu}{m}\right) \ll M^{1/2} (\log M)^2. \tag{2.14}$$

In section 2.5 we outline the first steps of the proof of this theorem, following the method of Toth, [T0́0]. Of particular relevance to us in this dissertation is the important step in the proof in which the sum over the roots of the congruence is approximated by a sum over cosets in $SL_2(\mathbb{Z})$ modulo triangular matrices, see (2.42) and (2.46), which are consequences of theorems 2.2 and 2.3. In more concrete terms, we notice that in the approximation given by theorem 2.3, a and v are more or less independent of each other, that is $\frac{v}{a}$ is an arbitrary fraction with the given size of denominator. Although

[1] Bykovskii actually proved the bound for an arbitrary quadratic congruence with negative discriminant, and with uniformity in h.

in the end quite sophisticated machinery is used to bound the resulting sum optimally, specifically the spectral theory of automorphic forms, this simple observation at least gives one optimism that a power-savings over the trivial bound might be found. That the analogue of this observation does not seem to hold in quite the same way will haunt us in the next chapters.

The other application we discuss is a large sieve inequality proved in [FI97]:

Theorem 2.5. *For an arbitrary sequence of complex numbers α_l,*

$$\sum_{M<m\leq 2M}\sum_{\mu^2\equiv -1(m)}\left|\sum_{|k|\leq K}\alpha_k e\left(\frac{\mu l}{m}\right)\right|^2 \ll (M+K)\sum_{|k|\leq K}|\alpha_k|^2 \tag{2.15}$$

with absolute implied constant.

This large sieve inequality in [FI97] for a crucial step towards proving that there are infinitely many primes of the form n^2+p^2. In addition we should mention that this large sieve, and the approximation given by theorem 2.3, was also crucially used in [FI98] where it is proved that there are infinitely many primes of the form n^2+l^4.

The large sieve inequality of theorem 2.5 is equivalent, at least morally, to the following upper bound for the number of the fraction $\frac{\mu}{m}$ contained in a short interval:

Theorem 2.6. *For any positive real number M, the number of fractions $\frac{\mu}{m}$ $(\text{mod } 1)$ with $\mu^2\equiv -1 \pmod m$ and $M<m\leq 2M$ in any interval of length $\frac{1}{M}$ is bounded by an absolute constant.*

While this upper bound is a less refined statistic than equidistribution, its power comes from its validity at much smaller scales than we could expect equidistribution to hold. Indeed the upper bound gives the correct result all the way down to scales of size $\frac{1}{M}$, the smallest possible because there are asymptotically $\frac{3}{2\pi}M$ roots with modulus $m\leq M$. On the other hand, the bound given in theorem 2.4 suggests that one can only expect equidistribution of the fractions in intervals of length at least $M^{-1/2}$.

2.1 Correspondence between roots and ideals

In this section we prove theorem 2.1. We start by noticing that given a sublattice $I\subset\mathbb{Z}[i]$, we can find a unique $\mathbb{Z}$-basis $\{\beta_1,\beta_2\}$ of I in Hermite normal form with respect to the basis $\{1, i\}$ of $\mathbb{Z}[i]$, by which we mean

$$\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} = \begin{pmatrix}a_{11} & 0\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}1\\ i\end{pmatrix} \tag{2.16}$$

where $a_{11}, a_{22} > 0$ and $0 \leq a_{21} < a_{11}$. The sublattice I will be an ideal of $\mathbb{Z}[i]$ if an only if i, as the generator of the ring, maps I into itself, $iI \subset I$. In coordinates this means that i acts by an integral matrix with respect to the basis (2.16). Explicitly, since

$$\begin{pmatrix}i\beta_1\\ i\beta_2\end{pmatrix} = \begin{pmatrix}a_{11} & 0\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}a_{11} & 0\\ a_{21} & a_{22}\end{pmatrix}^{-1}\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}, \tag{2.17}$$

I will be an ideal of $\mathbb{Z}[i]$ if and only if

$$\frac{1}{a_{11}a_{22}}\begin{pmatrix}a_{11} & 0\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}a_{22} & 0\\ -a_{21} & a_{11}\end{pmatrix} = \begin{pmatrix}-\frac{a_{21}}{a_{22}} & \frac{a_{11}}{a_{22}}\\ -\frac{a_{21}^2+a_{22}^2}{a_{11}a_{22}} & \frac{a_{21}}{a_{22}}\end{pmatrix} \tag{2.18}$$

is an integral matrix.

From this integrality condition on the matrix (2.18), we see that for I to be an ideal, it is necessary that a_{22} divides both a_{11} and a_{21}. Writing $a_{11} = ma_{22}$ and $a_{21} = \mu a_{22}$, the only remaining necessary condition for I to be an ideal is that

$$\mu^2 + 1 \equiv 0 \pmod{m}. \tag{2.19}$$

We close this short section by noticing first that these necessary conditions are also sufficient. Indeed, given m and μ for which (2.19) holds, the above calculations shows that the sublattice with basis given by (2.16) with $a_{11} = mn$, $a_{21} = \mu n$ and $a_{22} = n$, where n is an arbitrary integer, will be an ideal in $\mathbb{Z}[i]$. Finally, we note that it is necessary for the uniqueness of the Hermite normal form that $0 \leq a_{21} < a_{11}$. The same ambiguity occurs in picking a representative of μ modulo m.

2.2 Parametrization of the roots

For proving theorem 2.2, we will restrict our attention to the ideals I of theorem 2.1 with $n = 1$. We note that these ideals are exactly the ideals that are not divisible by any rational integers.

The key observation for the proof is that since $\mathbb{Z}[i]$ has class number one, every ideal I has a generator $\alpha = a+bi$. This generator is unique up to multiplication by the units, $\{1, i, -1, -i\}$, so we choose $a+bi$ to be in a fundamental domain of the action of these units, for example $a > 0$ and $-a < b \leq a^2$.

From the generator α, we get a natural basis for the ideal I, namely $\{\alpha, \alpha i\}$, which is expressed as

$$\begin{pmatrix} \alpha \\ \alpha i \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix}. \tag{2.20}$$

Now the parametrization is powered by the observation that the two bases (2.20) and (2.6) are related by a matrix $SL_2(\mathbb{Z})^3$. That is, there exists a matrix $\gamma \in SL_2(\mathbb{Z})$ such that

$$\begin{pmatrix} m & 0 \\ \mu & 1 \end{pmatrix} = \gamma \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{2.21}$$

From the second column of this matrix equation, we see that gamma must have the form

$$\gamma = \begin{pmatrix} a & -b \\ v & u \end{pmatrix} \tag{2.22}$$

for some integers u and v, which must satisfy $au + bv = 1$ for γ to be in $SL_2(\mathbb{Z})$. The existence of such u and v are guaranteed by $\gcd(a,b) = 1$, which we note is the same as the ideal $I = (a+bi)$ having no rational integer divisors.

The the integers u, v, and hence γ, are determined up to multiplication on the left by matrices of the form

$$\begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}. \tag{2.23}$$

This change in γ, via (2.21), corresponds to multiplying the matrix

$$\begin{pmatrix} m & 0 \\ \mu & 1 \end{pmatrix} \tag{2.24}$$

[2] The exact choice of fundamental domain does not matter so much here, but in section 2.3 we will see why this is a convenient choice.

[3] $SL_2(\mathbb{Z})$ rather than just $GL_2(\mathbb{Z})$ because we are assuming that m is positive in 2.6)

on the left by matrices of the same form, (2.23). Clearly this changes μ by a multiple of m, leaving $\mu \pmod{m}$ unchanged. Hence different choices of u and v satisfying $au + bv = 1$ will simply correspond to different representatives of the residue class $\mu \pmod{m}$.

2.3 Approximation and spacing of the roots

We now use the parametrization of theorem 2.2 to approximate $\frac{\mu}{m}$ by a fraction with much smaller denominator, resulting in theorem 2.3. And since it is an easy consequence of this theorem, at the end of this section we also prove the spacing property of the roots, theorem 2.6.

Using (2.9) we write

$$\frac{\mu}{m} = \frac{av - bu}{a^2 + b^2} = \frac{v}{a} - \frac{b}{a(a^2 + b^2)} \tag{2.25}$$

by (2.8). Now, if we assume $a > 0$ and $-a < b \leq a$, as we do in theorem 2.2, we see that since $a \leq \sqrt{a^2 + b^2} = \sqrt{m}$ and $b \leq a$,[4]

$$\frac{\mu}{m} = \frac{v}{a} + O\left(\frac{1}{m}\right). \tag{2.26}$$

So $\frac{v}{a}$ is indeed an approximation to $\frac{\mu}{m}$ with a much smaller denominator. We note that since all the v satisfying (2.8) are all $\equiv -\overline{b} \pmod{a}$, we have

$$\frac{\mu}{m} \equiv -\frac{\overline{b}}{a} + O\left(\frac{1}{m}\right) \pmod{1}. \tag{2.27}$$

Since we will want to generalize the expression (2.25) in the following chapters, let us give another way of deriving the approximation that comes naturally from the way we proved theorem 2.2. Recall that we have the relation

$$\begin{pmatrix} m & 0 \\ \mu & 1 \end{pmatrix} = \begin{pmatrix} a & -b \\ v & u \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{2.28}$$

[4] For this approximation to hold is in fact the reason why we made this choice of fundamental domain in section 2.2.

Now performing an LU-decomposition on the first matrix on the right of (2.28) gives the equation

$$\begin{pmatrix} m \\ \mu \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{v}{a} & 1 \end{pmatrix} \begin{pmatrix} a & -b \\ 0 & \frac{1}{a} \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{v}{a} & 1 \end{pmatrix} \begin{pmatrix} a^2+b^2 \\ -\frac{b}{a} \end{pmatrix}, \tag{2.29}$$

which gives the expression (2.25).

Having the approximation (2.11), the proof of theorem 2.6 is almost immediate. Indeed, for all the $\frac{\mu}{m}$ in an interval of length $\frac{1}{M}$, the corresponding $\frac{v}{a}$ will be all be forced in an interval of length $\ll \frac{1}{M}$. However, since $a \leq \sqrt{m}$ and fractions with denominator $\ll M^{1/2}$ are spaced by $\ll \frac{1}{M}$, there can be at most $O(1)$ of the $\frac{v}{a}$ in such an interval. We see that the theorem then follows if we can appropriately bound the multiplicity of $\frac{v}{a}$, which is to say that we need to rule out the possibility that a lot of $\frac{\mu}{m}$ correspond to the same $\frac{v}{a}$.

To rule out this possibility, we recall first the fact that $\gcd(a,v)=1$ and $a>0$, so we can recover a and v from the fraction $\frac{v}{a}$. Second, we recall that $b \equiv -\overline{v} \pmod{a}$, and so b, being restricted to $-a < b \leq a$, is determined up to 2 possibilities. Now since having a and b determines both m and $\mu \pmod{m}$, we see that a given $\frac{v}{a}$ corresponds to at most 2 of the $\frac{\mu}{m}$, thus finishing the proof of theorem 2.6.

2.4 Large sieve inequality

The upper bound

$$\sum_{M<m\leq 2M} \sum_{\mu^2\equiv -1(m)} \left| \sum_{|k|\leq K} \alpha_k e\left(\frac{k\mu}{m}\right) \right|^2 \ll (M+K) \sum_{|k|\leq K} |\alpha|^2 \tag{2.30}$$

for an arbitrary sequence of complex numbers α_k is equivalent by the duality principle, see [IK04], to proving

$$\sum_{|k|\leq K} \left| \sum_{M<m\leq 2M} \sum_{\mu^2\equiv -1(m)} \beta_{m,\mu} e\left(\frac{k\mu}{m}\right) \right|^2 \ll (M+K) \sum_{M<m\leq 2M} \sum_{\mu^2\equiv -1(m)} |\beta_{m,\mu}|^2, \tag{2.31}$$

also for an arbitrary sequence of complex numbers $\beta_{m,\mu}$. Letting f be a smooth function with compactly supported Fourier transform, $\hat{f}$,

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e(\xi x)\, dx, \tag{2.32}$$

and such that $f(x) \geq 1$ if $-1 \leq x \leq 1$ and $f(x) \geq 0$[5], the left side of (2.31) is

$$\begin{aligned} &\leq \sum_k f\left(\frac{k}{K}\right) \left| \sum_{M<m\leq 2M} \sum_{\mu^2 \equiv -1(m)} \beta_{m,\mu} e\left(\frac{k\mu}{m}\right) \right|^2 \\ &\qquad = \sum\sum_{M<m,m_1<2M} \sum\sum_{\substack{\mu^2\equiv -1(m) \\ \mu_1^2 \equiv -1(m_1)}} \beta_{m,\mu} \overline{\beta}_{m_1,\mu_1} \sum_k f\left(\frac{k}{K}\right) e\left(k\left(\frac{\mu}{m} - \frac{\mu_1}{m_1}\right)\right). \end{aligned} \tag{2.33}$$

We apply Poisson summation to the inner sum over k in (2.33) to obtain

$$\sum_k f\left(\frac{k}{K}\right) e\left(k\left(\frac{\mu}{m} - \frac{\mu_1}{m_1}\right)\right) = K \sum_k \hat{f}\left(K\left(k - \left(\frac{\mu}{m} - \frac{\mu_1}{m_1}\right)\right)\right). \tag{2.34}$$

Since $\hat{f}$ has compact support, only those k satisfying

$$\left| k - \left(\frac{\mu}{m} - \frac{\mu_1}{m_1}\right)\right| \ll \frac{1}{K}, \tag{2.35}$$

so if K is sufficiently large at most one k will contribute, and even then only if

$$\left|\left| \frac{\mu}{m} - \frac{\mu_1}{m_1} \right|\right| \ll \frac{1}{K}, \tag{2.36}$$

where $||\cdot||$ denotes the distance to the nearest integer. Applying this we see that when K is sufficiently large,

$$\sum_k f\left(\frac{k}{K}\right) e\left(k\left(\frac{\mu}{m} - \frac{\mu_1}{m_1}\right)\right) \ll K \mathbb{1}_{\left|\left|\frac{\mu}{m} - \frac{\mu_1}{m_1}\right|\right| \ll \frac{1}{K}}. \tag{2.37}$$

In fact, by adjusting the implied constants if necessary, this bound holds for all K.

Estimating (2.33) by (2.37), we see that the left side of (2.31) is

$$\ll K \sum\sum_{\substack{M<m\leq 2M \\ \mu^2\equiv -1(m)}} \sum\sum_{\substack{M<m_1\leq 2M \\ \mu_1^2 \equiv -1(m_1) \\ \frac{\mu_1}{m_1} \in I_{m,\mu}}} |\beta_{m,\mu}\beta_{m_1,\mu_1}|, \tag{2.38}$$

where $I_{m,\mu}$ is the interval of x in $\mathbb{R}/\mathbb{Z}$ such that

$$\left|\left| \frac{\mu}{m} - x \right|\right| \ll \frac{1}{K}, \tag{2.39}$$

with the same implied constant as in (2.37).

[5] A re-scaling of $f(x) = \left(\frac{sinx}{x}\right)^2$ will do.

Using the inequality $2|\beta_{m,\mu}\beta_{m_1,\mu_1}| \le |\beta_{m,\mu}|^2 + |\beta_{\mu_1,m_1}|^2$ and exploiting the symmetry, the right side of (2.38) is seen to be

$$\le K \sum\sum_{\substack{M<m\le 2M \\ \mu^2\equiv -1(m)}} |\beta_{m,\mu}|^2 \sum\sum_{\substack{M<m_1\le 2M \\ \mu_1^2\equiv -1(m_1) \\ \frac{\mu_1}{m_1}\in I_{m,\mu}}} 1. \tag{2.40}$$

Since $I_{m,\mu}$ has length $O\left(\frac{1}{K}\right)$, it can be covered by $\ll \frac{M}{K}+1$ intervals of length $\frac{1}{M}$. Since by theorem 2.6 there can be at most $O(1)$ fractions $\frac{\mu_1}{m_1}$ in each of these intervals, we find that (2.40) is

$$\ll (M+K) \sum_{M<m\le 2M} \sum_{\mu^2\equiv -1(m)} |\beta_{m,\mu}|^2, \tag{2.41}$$

which establishes the bound (2.31) and proves theorem 2.5.

2.5 Equidistribution

The parametrization of theorem 2.2 transforms the Weyl sum for the roots $\frac{\mu}{m}$ into

$$\sum_m f\left(\frac{m}{M}\right) \sum_{\mu^2\equiv -1(m)} e\left(\frac{h\mu}{m}\right) = \sum\sum_{\gcd(a,b)=1} g(a,b) f\left(\frac{a^2+b^2}{M}\right) e\left(h\frac{av-bu}{a^2+b^2}\right), \tag{2.42}$$

where u and v are some choice of integers satisfying $au+bv=1$ and $g(a,b)$ is a smooth fundamental domain for the action of the units in $\mathbb{Z}[i]$ supported in $a>0$, $|b|\le 2a$. To be more specific, we mean that $g(a,b)=0$ if a is non-positive or if $|b|>2a$, and

$$g(a,b)+g(-b,a)+g(-a,-b)+g(b,-a)=1. \tag{2.43}$$

We note that such a g can be chosen so that away from the origin

$$\left(\frac{\partial}{\partial y}\right)^j g(x,y) \ll \frac{1}{x^j}. \tag{2.44}$$

We remark that in the language of $SL_2(\mathbb{Z})$, the sum (2.42) can be written naturally as a sum over cosets

$$\begin{pmatrix} a & b \\ -v & u \end{pmatrix} \in \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \backslash SL_2(\mathbb{Z}). \tag{2.45}$$

Now, applying the approximation, theorem 2.3, shows the Weyl sum (2.42) is

$$\begin{aligned} &= \sum\sum_{\gcd(a,b)=1} g(a,b) f\left(\frac{a^2+b^2}{M}\right)\left(e\left(\frac{h\bar{b}}{a}\right)+O\left(\frac{|h|}{M}\right)\right) \\ &= \sum\sum_{\gcd(a,b)=1} g(a,b) f\left(\frac{a^2+b^2}{M}\right) e\left(\frac{h\bar{b}}{m}\right)+O(|h|). \end{aligned} \tag{2.46}$$

This approximation shows that the oscillating part of the Weyl sum only depends on the double coset,

$$\begin{pmatrix} a & b \\ -\bar{b} & * \end{pmatrix} \in \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \backslash SL_2(\mathbb{Z}) / \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}. \tag{2.47}$$

So, rearranging the terms in (2.46) by grouping them according the the double coset gives

$$\sum_{a>0} \sum_{\beta(a)^*} e\left(\frac{h\beta}{a}\right) \sum_{b \equiv \beta(a)} g(a,b) f\left(\frac{a^2+b^2}{M}\right). \tag{2.48}$$

This arrangement is quite similar to how one computes the Fourier coefficients of a Poincaré series, and just as one does in that setting we apply Poisson summation to the b sum to obtain

$$= \sum_{a>0} \frac{1}{a} \sum_{\beta(a)^*} e\left(\frac{h\overline{\beta}}{a}\right) \sum_k e\left(\frac{k\beta}{a}\right) \int\limits_{-\infty}^{\infty} g(a,x) f\left(\frac{a^2+x^2}{M}\right) e\left(\frac{kx}{a}\right) dx. \tag{2.49}$$

Again just as when one computes the Fourier coefficients of a Poincaré series, a sum of Kloosterman sums is the result of the Poisson summation, explicitly

$$= \sum_k \sum_{a>0} \frac{1}{a} S(h,k;a) G(a,k,M), \tag{2.50}$$

where

$$S(h,k;a) = \sum_{\beta(a)^*} e\left(\frac{h\beta + k\overline{\beta}}{a}\right) \tag{2.51}$$

is a Kloosterman sum and

$$G(a,k,M) = \int\limits_{-\infty}^{\infty} g(a,x) f\left(\frac{a^2+x^2}{M}\right) e\left(\frac{kx}{a}\right) dx. \tag{2.52}$$

With little effort one can use the Weil bound for the Kloosterman sums, noticing that only $k \ll M^{\epsilon}$, $a \ll M^{1/2}$ contribute and $G(a,k,M) \ll M^{1/2}$, to estimate this sum of Kloosterman sums by $\ll M^{3/4+\epsilon}$.[6] If we want to do better, one can use the spectral theory of automorphic forms to find cancellation between the Kloosterman sums themselves. The result is quite strong, both in terms of M, obtaining $M^{1/2}(\log M)^2$, but also in terms of uniformity in h.

[6] Of course one would want uniformity in h as well, but this is not our main point, so we do not discuss this here.

Chapter 3

Roots of a cubic congruence

Our attempt to generalize the results of chapter 2 begins with the following analogue of theorem 2.1:

Theorem 3.1. *Let $I \subset \mathbb{Z}[2^{1/3}]$ be an ideal that is not divisible by any rational integers. Then the unique basis $\{\beta_1, \beta_2, \beta_3\}$ of I in Hermite normal form can be written as*

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} mn & 0 & 0 \\ -\mu n & n & 0 \\ \lambda & \nu & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix} \tag{3.1}$$

where m and n are positive integers, $\mu \pmod{m}$ and $\nu \pmod{n}$ satisfy

$$\begin{aligned} \mu^3 &\equiv 2 \pmod{m} \\ \nu^3 &\equiv 2 \pmod{n}. \end{aligned} \tag{3.2}$$

In addition $\gcd(m, n, 6) = 1$ and

$$\gcd(m, n, \mu - \nu) = 1. \tag{3.3}$$

Conversely, given positive integers m and n, such that $\gcd(m, n, 6) = 1$, with residue classes $\mu \pmod{m}$ and $\nu \pmod{n}$ satisfying (3.2) together with (3.3), then there is a unique value of $\lambda \pmod{mn}$ so that the sublattice of $\mathbb{Z}[2^{1/3}]$ defined by (3.1) is an ideal of $\mathbb{Z}[2^{1/3}]$.

Before discussing this theorem in more depth, we state corollaries 3.2 and 3.3. Corollary 3.3 is simply theorem 3.1 in the special case when $n = 1$. Corollary 3.2 classifies the ideals appearing in this special case: they are the ideals I such that $\mathbb{Z}[2^{1/3}]$ is additively cyclic, which can also be stated as the ideals I that have only degree one prime factors, none of which are conjugate.

Corollary 3.2. *Let I be an ideal in $\mathbb{Z}[2^{1/3}]$ and let $l \in \mathbb{Z}$ be so that $\frac{1}{l}I$ has no integer divisors. Further, let m and n be the integers corresponding to $\frac{1}{l}I$ as in theorem 3.1. Then mnl, nl, and l are the invariant factors of $\mathbb{Z}[2^{1/3}]/I$.*

Corollary 3.3. *Let I be an ideal in $\mathbb{Z}[2^{1/3}]$ such that $\mathbb{Z}[2^{1/3}]/I$ is additively cyclic. Then I has a unique basis $\{\beta_1, \beta_2, \beta_3\}$ such that*

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \begin{pmatrix} m & 0 & 0 \\ -\mu & 1 & 0 \\ -\mu^2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix} \tag{3.4}$$

where $m > 0$ and $\mu^3 \equiv 2 \pmod{m}$, and, for uniqueness to hold, μ is considered as a residue class modulo m. Conversely, given given an $m > 0$ and $\mu \pmod{m}$ such that $\mu^3 \equiv 2 \pmod{m}$, the lattice spanned by the basis $\{\beta_1, \beta_2, \beta_3\}$ of (3.4) is an ideal I of $\mathbb{Z}[2^{1/3}]$ such that $\mathbb{Z}[2^{1/3}]/I$ is cyclic.

There are some readily apparent differences between this theorem and the quadratic theorem 2.1. Perhaps most obviously, ideals in this cubic setting no longer correspond roots of a cubic congruence, but rather pairs of roots satisfying a coprimality condition (3.3). We interpret this condition as requiring μ and ν to be genuinely different since it is equivalent to μ and ν being incongruent modulo all primes dividing both m and n.

We can illustrate this condition by considering some examples. A rational prime p different from 2 and 3 either splits completely, factors as a degree one prime times a degree two prime, or stays inert in $\mathbb{Z}[2^{1/3}]$. As implied by either theorem 3.1 or the Dedekind-Kummer theorems, these cases correspond exactly to whether $\mu^3 \equiv 2 \pmod{p}$ has three, one, or zero solutions.

In the first case, we have three degree one prime ideals $\mathfrak{p}_1$, $\mathfrak{p}_2$ and $\mathfrak{p}_3$ lying above p, each one corresponding to a root μ_j of the congruence $\mu^3 \equiv 2 \pmod{p}$. Indeed a basis for $\mathfrak{p}_j$ is given by

$$\begin{pmatrix} p & 0 & 0 \\ -\mu_j & 1 & 0 \\ -\mu_j^2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}. \tag{3.5}$$

If on the other hand $p = \mathfrak{p}_1\mathfrak{p}_2$, with $\mathfrak{p}_1$ having degree one and $\mathfrak{p}_2$ having degree two, there is only one root of the congruence $\mu^3 \equiv 2 \pmod{p}$. A basis for $\mathfrak{p}_1$ has the same form as in (3.5), while a basis for $\mathfrak{p}_2$ is given by

$$\begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ \mu^2 & \mu & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}. \tag{3.6}$$

Finally, returning to the first case where $\mathfrak{p}_j$ are the three degree one prime ideals lying above p, and μ_j are the corresponding roots, then a basis for $\mathfrak{p}_1\mathfrak{p}_2$ is given by

$$\begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ \mu_3^2 & \mu_3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}. \tag{3.7}$$

The fact that we used μ_3 here is exactly because of the coprimality condition (3.3). Indeed, we note that if we multiply $\mathfrak{p}_1\mathfrak{p}_2$ by $\mathfrak{p}_3$, we obtain the rational integer p, which does not appear in theorem 3.1. On the other hand, $\mathfrak{p}_1^2\mathfrak{p}_2$ does appear in theorem 3.1, a basis for this ideal is

$$\begin{pmatrix} p^2 & 0 & 0 \\ -\mu_1 & p & 0 \\ * & \mu_3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}. \tag{3.8}$$

This last example, the basis for $\mathfrak{p}_1^2\mathfrak{p}_2$, also sheds light on the coprimality condition on m and n, that $\gcd(m, n, 6) = 1$. On the one hand, it is not surprising that this condition is related to ramification in $\mathbb{Z}[2^{1/3}]$, and on the other hand, the above examples are enough to convince oneself that products of this form, $\mathfrak{p}_1^2\mathfrak{p}_2$ and their higher-power relatives are entirely responsible for common factors between m and n. Taking these observations together, we note that there is only one prime lying above 3, which is a unit multiple of the cube of this prime, and the same holds for 2. This shows that the products of the above form made with primes of 2 and 3 will necessarily be divisible by a rational integer, 2 or 3, and so does not appear in the correspondence of theorem 3.1.

These differences and complications aside, we can use theorem 3.1 to prove an analogue of the parametrization of roots of a quadratic congruence, theorem 2.2 in a very similar way that we used theorem 2.1. Indeed, the main idea is to relate two bases of the same ideal by a matrix in $SL_3(\mathbb{Z})$, one basis given by theorem 3.1 and the other given by a generator of the ideal, using that $\mathbb{Z}[2^{1/3}]$ has class number one. The result is the following:

Theorem 3.4. *Let $\alpha = a + b2^{1/3} + c2^{2/3}$ and $\alpha' = A + B2^{1/3} + C^{2/3}$ be integers in $\mathbb{Z}[2^{1/3}]$ such that*

$$\gcd(a, b, c) = \gcd(A, B, C) = 1, \tag{3.9}$$

$$\alpha\alpha' \in \mathbb{Z}_{>0}, \tag{3.10}$$

and α in a fundamental domain for the action of the units on $\mathbb{Z}[2^{1/3}]$. Let $\gamma \in SL_3(\mathbb{Z})$ be a representative of the coset in

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} \backslash SL_3(\mathbb{Z}) \tag{3.11}$$

with Plücker coordinates A, B, C and c, b, a. Then $\mu \pmod{m}$ and $\nu \pmod{n}$ defined by

$$\begin{pmatrix} mn & 0 & 0 \\ -\mu n & n & 0 \\ * & \nu & 1 \end{pmatrix} = \gamma \begin{pmatrix} a & b & c \\ 2c & a & b \\ 2b & 2c & a \end{pmatrix} \tag{3.12}$$

will satisfy $\mu^3 \equiv 2 \pmod{m}$ and $\nu^3 \equiv 2 \pmod{n}$. In addition $\gcd(m, n, 6) = 1$ and $\gcd(m, n, \mu - \nu) = 1$. Conversely, given such m, n, $\mu \pmod{m}$ and $\nu \pmod{n}$, there will correspond unique α and α' in the above way.

In the statement of this theorem, we are recalling that cosets in (3.11) are parametrized by their Plücker coordinates. Here by a coset with coordinates A, B, C, c, b, and a we mean that a representative of this coset γ satisfies

$$\gamma = \begin{pmatrix} A & B & C \\ * & * & * \\ * & * & * \end{pmatrix}, \quad \gamma^{-1} = \begin{pmatrix} * & * & c \\ * & * & b \\ * & * & a \end{pmatrix}, \tag{3.13}$$

see [BFG88] for example. We recall in addition that A, B, C, c, b, and a are eligible Plücker coordinates if and only if $\gcd(A, B, C) = \gcd(a, b, c) = 1$ and

$$cA + bB + aC = 0. \tag{3.14}$$

The coprimality condition is included in the statement of theorem 3.4, and we observe that the requirement that $\alpha\alpha' \in \mathbb{Z}$ is exactly that

$$\begin{aligned} bA + aB + 2cC = 0 \\ cA + bB + aC = 0, \end{aligned} \tag{3.15}$$

so we do indeed have admissible Plücker coordinates. However, we notice that in fact A, B, C, c, b, and a satisfy an additional constraint, and so only special cosets in (3.11) are used in parametrizing the roots of the cubic congruence. This is a major difference from the quadratic setting of chapter 2, where the fact that every coset was used in the parametrization played a large role in the proof of the equidistribution of the roots.

We will return to this later in chapter 5, for now we state a corollary of theorem 3.4 in the case of $n = 1$:

Corollary 3.5. *Let $(a, b, c) \in \mathbb{Z}^3$ be in a fundamental domain for the action of the units of $\mathbb{Z}[2^{1/3}]$ on $\mathbb{Z}^3$ identified with $\mathbb{Z}[2^{1/3}]$ via the basis $\{1, 2^{1/3}, 2^{2/3}\}$. Suppose that*

$$\gcd(a^2 - 2bc, 2c^2 - ab, b^2 - ac) = 1, \tag{3.16}$$

so that there are integers u, v, and w satisfying

$$(a^2 - bc)u + (2c^2 - ab)v + (b^2 - ac)w = 1. \tag{3.17}$$

Then the integers

$$\begin{aligned} m &= a^3 + 2b^3 + 4c^3 - 6abc \\ \mu &= 2(b^2 - ac)u + (a^2 - 2bc)v + (2c^2 - ab)w \end{aligned} \tag{3.18}$$

satisfy

$$\mu^3 \equiv 2 \pmod{m}, \tag{3.19}$$

and also the additional congruence

$$\mu^2 \equiv 2(2c^2 - ab)u + 2(b^2 - ac)v + (a^2 - 2bc)w \pmod{m}. \tag{3.20}$$

We remark that different solutions u, v, and w to (3.17) will only give different representatives of μ and/or μ^2, and that every possible modulus m and root $\mu \pmod{m}$ of $X^3 \equiv 2 \pmod{m}$ arises via (3.18).

As in chapter 2, we can use the parametrization to derive an approximation to the roots, theorem 3.7. To prove this theorem, it is necessary to control the sizes of the relevant parameters used in theorem 3.4. To this end we first prove the following proposition:

Proposition 3.6. *Let $\alpha = a + b2^{1/3} + c2^{2/3} \in \mathbb{Z}[2^{1/3}]$ and $\alpha' = A + B2^{1/3} + C2^{2/3} \in \mathbb{Z}[2^{1/3}]$ satisfy the conditions of theorem 3.4. Then there is a choice of fundamental domain, say $\mathcal{D}$, for the action of the units in $\mathbb{Z}[2^{1/3}]$ such that $\alpha \in \mathcal{D}$ implies that*

$$\begin{aligned} a, b, c &\ll N(\alpha)^{1/3} = m^{1/3}n^{2/3} \\ A, B &\ll N(\alpha')^{1/3} = m^{2/3}n^{1/3} \\ C &\asymp N(\alpha')^{1/3} = m^{2/3}n^{1/3}. \end{aligned} \tag{3.21}$$

After restricting α to be in this fundamental domain, we can apply an LU decomposition to the matrix γ of theorem 3.4, and, using proposition 3.6 to control the error terms, the result is the following theorem:

Theorem 3.7. *Let $\alpha = a + b2^{1/3} + c2^{2/3} \in \mathbb{Z}[2^{1/3}]$ and $\alpha' = A + B2^{1/3} + C2^{2/3} \in \mathbb{Z}[2^{1/3}]$ satisfy the conditions of theorem 3.4 with $\alpha \in \mathcal{D}$, where $\mathcal{D}$ is as in proposition 3.6. Further, let $\mu \pmod{m}$ be one of the roots corresponding to α and α' as in theorem 3.4. Then*

$$\frac{\mu}{m} = -\frac{W}{C} + O\left(\frac{1}{m}\right) \pmod{1}, \tag{3.22}$$

where W is defined by the congruences

$$\begin{aligned} AW &\equiv -b \pmod{C} \\ BW &\equiv c \pmod{C}. \end{aligned} \tag{3.23}$$

We remark that these congruences have exactly one solution $\pmod{C}$ because $\gcd(A, B, C) = 1$ and $Ac + Bb + cC = 0$.

We observe that in light of proposition 3.6, we have here an approximation to $\frac{\mu}{m}$ to within $O\left(\frac{1}{m}\right)$ by a fraction with denominator of size $m^{2/3}n^{1/3}$. Even in the best case in terms of the size of the denominator, $n = 1$, even this is far from optimal in the sense of Dirichlet's theorem on Diophantine approximation. There are many fractions with denominator of size $m^{2/3}$ inside an interval of length $\frac{1}{m}$, so we have to ask what is the significance of this particular fraction?

There are two distinct approaches that we will take towards answering this question. One approach, which we will take up later in chapter 5 centers around the observation that there are presumably many different approximations to $\frac{\mu}{m}$, naively one for each n, all contained within $O\left(\frac{1}{m}\right)$ of $\frac{\mu}{m}$. The other approach, which we will consider for the remainder of this chapter, centers instead on the observation that in the case $n = 1$ we also get, via corollary 3.5, an approximation to $\frac{\mu^2}{m}$. This is contained in the following theorem:

Theorem 3.8. *Let $\alpha = a + b2^{1/3} + c2^{2/3} \in \mathbb{Z}[2^{1/3}$ satisfy the conditions of corollary 3.5 with $\alpha \in \mathcal{D}$, where $\mathcal{D}$ is as in proposition 3.6. Further, let u, v, w be as in corollary 3.5 and $\mu \pmod{m}$ be the corresponding root. Then*

$$\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right) = \left(\frac{bu - cv}{b^2 - ac}, \frac{bv - au}{b^2 - ac}\right) + O\left(\frac{1}{m}\right) \quad (\text{mod } \mathbb{Z}^2), \tag{3.24}$$

where the implied constant is absolute.

A first observation is that the approximation to the point $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ is a pair of fractions with the same denominator. We will call such a point in $\mathbb{R}^2$ a torsion point because these are exactly the (representatives of) the torsion points in $\mathbb{R}^2/\mathbb{Z}^2$, the denominator being the torsion.

Statistically speaking, disc of radius $\frac{1}{m}$ in $\mathbb{R}^2$ will probably have no more than $\ll 1$ torsion points in $\mathbb{R}^2/\mathbb{Z}^2$ with torsion of size $\ll m^{2/3}$. This is because there are $\asymp m^2$ such torsion points, and we should expect them to be, roughly speaking, evenly distributed between the m^2 squares in $\mathbb{R}^2/\mathbb{Z}^2$ with side length $\frac{1}{m}$. The problem is that this spacing property does not always hold, for example a disk of radius $\frac{1}{m}$ centered on an axis is really no different than an interval in the 1-dimensional setting. Proposition

3.9 below characterizes in a sense this kind of exceptional circumstance; it states that torsion points not contained in any integral line with small coefficients are well-spaced.

Proposition 3.9. *Let Q be a positive real number and let $\left(\frac{r}{q}, \frac{s}{q}\right)$ be a q-torsion point in $\mathbb{R}^2/\mathbb{Z}^2$. Further, define the lattice Λ by*

$$\Lambda = \left\{(A, B) \in \mathbb{Z}^2 \ : \ Ar + Bs \in q\mathbb{Z}\right\}, \tag{3.25}$$

which is a projection of the lattice of integral lines containing $\left(\frac{r}{q}, \frac{s}{q}\right)$. Then the distance between $\left(\frac{r}{q}, \frac{s}{q}\right)$ and any other torsion point with torsion $\leq Q$ is at least

$$\frac{1}{qQ} \min\left\{||\boldsymbol{v}|| \ : \ \boldsymbol{v} \in \Lambda, \boldsymbol{v} \neq \boldsymbol{0}\right\}. \tag{3.26}$$

Because the way we prove theorem 3.8 in section 3.3.2, which is only superficially different from the way we prove theorem 3.7 in section 3.3.1, exhibits the approximation to $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ naturally as the intersection of two lines in $\mathbb{R}^2$, proposition 3.9 is easy to apply in our setting. The result is the following theorem:

Theorem 3.10. *For a positive real number M and any disc $D \subset \mathbb{R}^2/\mathbb{Z}^2$ of radius $\frac{1}{M}$, we have*

$$\#\left\{\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right) \in D \ : \ M < m \leq 2M, \ \mu^3 \equiv 2 \pmod{M}\right\} \ll 1, \tag{3.27}$$

where the implied constant is absolute.

We caution however, that a key step in the proof of this theorem is to show that at most $O(1)$ different roots of the congruence correspond to a given approximation. Here the issue is handled without much difficultly, however when we consider the other approach mentioned above in chapter 5, the analogous question is unfortunately still unresolved.

Nevertheless, we finish the chapter in section 3.5 by using the spacing property to derive a two dimensional large sieve inequality for the roots of the congruence:

Theorem 3.11. *Let K and L be positive real numbers, and let $\alpha_{k,l}$ be a sequence of* *complex numbers. Then*

$$\sum_{M<m\leq 2M}\sum_{\mu^3\equiv 2(m)}\left|\sum_{|k|\leq K}\sum_{|l|\leq L}\alpha_{k,l}e\left(\frac{k\mu+l\mu^2}{m}\right)\right|^2 \\ \ll (M+K)(M+L)\sum_{|k|\leq K}\sum_{|l|\leq L}|\alpha_{k,l}|^2, \tag{3.28}$$

where the implied constant is absolute.

We note that the this inequality is optimal up to the implied constant when K and L are at least M. Indeed, in this regime we can set

$$\alpha_{k,l}=e\left(\frac{-k\mu_0-l\mu_0^2}{m_0}\right) \tag{3.29}$$

for one of the roots $\mu_0 \pmod{m_0}$, then the right side of (3.28) will be $(KL)^2$, which is the size of just a single term on the left. On the other hand, we note that by Cauchy's inequality, the left side of (3.28) is trivially

$$\leq KLM\sum_{|k|\leq K}\sum_{|l|\leq L}|\alpha_{k,l}|^2, \tag{3.30}$$

so when $KL\leq M$, theorem 3.11 gives nothing nontrivial.

3.1 Correspondence between roots and ideals

The proof of theorem 3.1 will start in the same manner as the proof of theorem 2.1, working with sub-lattices, later ideals, I in the cubic ring $\mathbb{Z}[2^{1/3}]$. Fixing the $\mathbb{Z}$-basis $\{1,2^{1/3},2^{2/3}\}$ of $\mathbb{Z}[2^{1/3}]$, we pick, as we did in the quadratic setting, the unique $\mathbb{Z}$-basis $\{\beta_1,\beta_2,\beta_3\}$ of I in Hermite normal form. Here this means that

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix}=A\begin{pmatrix}1\\ 2^{1/3}\\ 2^{2/3}\end{pmatrix}, \tag{3.31}$$

with A an integer matrix of the form

$$A=\begin{pmatrix}a_{11}&0&0\\ a_{21}&a_{22}&0\\ a_{31}&a_{32}&a_{33}\end{pmatrix} \tag{3.32}$$

with $a_{11}, a_{22}, a_{33} > 0$ and $0 \leq a_{21}, a_{31} < a_{11}$, $0 \leq a_{32} < a_{22}$.

Proceeding as we did in the section 2.1, we note that since $2^{1/3}$ generates $\mathbb{Z}[2^{1/3}]$, I being an ideal is equivalent to $2^{1/3}I$ being a sublattice of I. In other words, we need $2^{1/3}$ to act by an integral matrix with respect to the basis (3.31). And since

$$2^{1/3}\begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}, \tag{3.33}$$

we see that I being an ideal is equivalent to

$$A\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{pmatrix}A^{-1} = \begin{pmatrix} -\frac{a_{21}}{a_{22}} & \frac{a_{11}}{a_{22}} & 0 \\ -\frac{a_{21}^2}{a_{11}a_{22}} + \frac{a_{21}a_{32}}{a_{11}a_{33}} - \frac{a_{22}a_{31}}{a_{11}a_{33}} & \frac{a_{21}}{a_{22}} - \frac{a_{32}}{a_{33}} & \frac{a_{22}}{a_{33}} \\ 2\frac{a_{33}}{a_{11}} - \frac{a_{21}a_{31}}{a_{11}a_{22}} + \frac{a_{21}a_{32}^2}{a_{11}a_{22}a_{33}} - \frac{a_{31}a_{32}}{a_{11}a_{33}} & \frac{a_{31}}{a_{22}} - \frac{a_{32}^2}{a_{22}a_{33}} & \frac{a_{32}}{a_{33}} \end{pmatrix} \tag{3.34}$$

being an integer matrix.

From the $(1,2)$ and $(2,1)$ entries, we see that some necessary conditions for I to be an ideal are that $a_{33} \mid a_{22}$ and $a_{22} \mid a_{11}$. Moreover, from the $(1,1)$ and $(3,3)$ entries, $a_{22} \mid a_{21}$ and $a_{33} \mid a_{32}$. And finally, from the $(3,2)$ entry, $a_{33} \mid a_{31}$. From these conditions, we see that for I to be an ideal, it is necessary that we be able to re-write A as

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} mna_{33} & 0 & 0 \\ 0 & na_{33} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -\mu & 1 & 0 \\ \lambda & \nu & 1 \end{pmatrix}. \tag{3.35}$$

With these substitutions, the matrix (3.34) becomes

$$\begin{pmatrix} \mu & m & 0 \\ -\frac{1}{m}(\mu^2 + \mu\nu + \lambda) & -(\mu + \nu) & n \\ \frac{1}{mn}(2 + \lambda(\mu - \nu) - \mu\nu^2) & \frac{1}{n}(\lambda - \nu^2) & \nu \end{pmatrix}. \tag{3.36}$$

As in the quadratic case, we lose nothing by assuming $a_{33} = 1$ as it simply factors out of I and does not appear in the integrality conditions of (3.36). And as in the quadratic case, this assumption corresponds to the ideal I not being divisible by any rational integers. Although in the quadratic setting, the role of a_{33} was n, here, working with 3×3 matrices, n plays a more significant role.

From the $(2,1)$ and $(3,2)$ entries of (3.36), we see that for I to be an ideal, λ must satisfy the two congruences

$$\begin{aligned} &\lambda \equiv -\mu^2 - \mu\nu \pmod{m} \\ &\lambda \equiv \nu^2 \pmod{n}. \end{aligned} \tag{3.37}$$

In order for these congruences to be consistent, we of course must have

$$\mu^2 + \mu\nu + \nu^2 \equiv 0 \pmod{\gcd(m,n)}. \tag{3.38}$$

If (3.38) is satisfied, then solving (3.37) shows that λ will be of the form

$$\lambda = \nu^2 \frac{\overline{m}m}{\gcd(m,n)} - (\mu^2 + \mu\nu)\frac{\overline{n}n}{\gcd(m,n)} + \kappa \frac{mn}{\gcd(m,n)} \pmod{mn}, \tag{3.39}$$

where κ is defined modulo $\gcd(m,n)$ and $\overline{m}$, $\overline{n}$ are some choice of integers so that

$$\overline{m}m + \overline{n}n = \gcd(m,n). \tag{3.40}$$

We remark that the choice of $\overline{m}$ and $\overline{n}$ are not unique, but all choices can be obtained by respectively adding and subtracting $\frac{ln}{\gcd(m,n)}$ and $\frac{lm}{\gcd(m,n)}$ to some specific choice of $\overline{m}$ and $\overline{n}$, where l is an arbitrary integer. Hence a different choice of $\overline{m}$ and $\overline{n}$ will correspond to adding $l\frac{\mu^2+\mu\nu+\nu^2}{\gcd(m,n)}$ to κ.

Substituting (3.39) into the $(3,1)$ entry of the matrix (3.36) yields the following necessary condition for I to be an ideal:

$$(\nu^3 - 2)\frac{\overline{m}m}{\gcd(m,n)} + (\mu^3 - 2)\frac{\overline{n}n}{\gcd(m,n)} - \kappa(\mu - \nu)\frac{mn}{\gcd(m,n)} \equiv 0 \pmod{mn}. \tag{3.41}$$

We can rewrite the left side of (3.41) in two ways using (3.40), to either emphasis divisibility by m or by n. First we can write it as

$$\mu^3 - 2 - (\mu - \nu)\frac{\mu^2 + \mu\nu + \nu^2}{\gcd(m,n)}\overline{m}m - \kappa(\mu - \nu)\frac{mn}{\gcd(m,n)}, \tag{3.42}$$

which shows that $\mu^3 \equiv 2 \pmod{m}$ is necessary for I to be an ideal. Second, we can write the left side of (3.41) as

$$\nu^3 - 2 + (\mu - \nu)\frac{\mu^2 + \mu\nu + \nu^2}{\gcd(m,n)}\overline{n}n - \kappa(\mu - \nu)\frac{mn}{\gcd(m,n)}, \tag{3.43}$$

which shows that $\nu^3 \equiv 2 \pmod{n}$ is also necessary.

Assuming the conditions $\mu^3 \equiv 2 \pmod{m}$ and $\nu^3 \equiv 2 \pmod{n}$, we can divide the congruence (3.41) through by $\frac{mn}{\gcd(m,n)}$ to obtain the necessary condition

$$\kappa(\mu-\nu) \equiv \frac{\mu^3-2}{m}\overline{n} + \frac{\nu^3-2}{n}\overline{m} \pmod{\gcd(m,n)}. \tag{3.44}$$

We note that using different choices of $\overline{m}$ and $\overline{n}$ and applying the corresponding change to κ discussed in the remark following (3.40) leaves the condition (3.44) unchanged.

Clearly (3.44) will have a unique solution in κ if and only if $\gcd(m,n,\mu-\nu)=1$, which we, recalling the introduction to the chapter interpret, as μ and ν being genuinely different roots of the cubic congruence since it is equivalent to μ not being congruent to ν modulo all primes that divide both m and n. On the other hand, if I is an ideal and there is a prime p dividing all of m, n, and $\mu-\nu$, then (3.38) implies that

$$3\mu^2 \equiv 0 \pmod{p}. \tag{3.45}$$

If $p \neq 3$, then (3.45) implies that $\mu \equiv 0 \pmod{p}$, which together with $\mu^3 \equiv 2 \pmod{p}$ shows that $p=2$. This is the issue addressed in the introduction to this chapter, and its resolution there followed from a consideration of the ramification type of 2 and 3 in $\mathbb{Z}[2^{1/3}]$.

Let us now quickly discuss the sufficient conditions for a sublattice I of $\mathbb{Z}[2^{1/3}]$ with basis matrix A as in (3.35) with respect to the basis $\{1, 2^{1/3}, 2^{2/3}\}$ of $\mathbb{Z}[2^{1/3}]$, thus finishing the proof of theorem 3.1. If $\mu \pmod{m}$ and $\nu \pmod{n}$ satisfy $\mu^3 \equiv 2 \pmod{m}$ and $\nu^3 \equiv 2 \pmod{n}$, then

$$\mu^3-\nu^3 = (\mu-\nu)(\mu^2+\mu\nu+\nu^2) \equiv 0 \pmod{\gcd(m,n)}. \tag{3.46}$$

If in addition we assume that $\gcd(m,n,\mu-\nu)=1$, then (3.46) implies (3.38), and in addition we can find a unique $\kappa \pmod{\gcd(m,n)}$ satisfying (3.44), for some specific choice of $\overline{m}$, $\overline{n}$. Now since (3.38) is satisfied, setting λ as in (3.39) makes the matrix (3.36) into an integral matrix. This finishes the proof of theorem 3.1.

We close this section by quickly proving corollaries 3.2 and 3.3. Corollary 3.2 follows immediately from (3.35). Indeed, (3.35) shows that if I is an ideal, then the matrix A can be brought into Smith normal form simply by multiplying on the right by a lower triangular matrix in $SL_3(\mathbb{Z})$. The invariant factors are then seen to be as claimed. Corollary 3.3 is also very easy from what we have done. Under the assumption that I is an ideal for which $n = 1$, we can take $\nu = 0$, and then (3.37) shows immediately that in this case $\lambda \equiv -\mu^2 \pmod{m}$.

3.2 Parametrization of the roots

3.2.1 Proof of theorem 3.4

We start by proving theorem 3.4, and after we will see how to obtain corollary 3.5 from it. The key observation that powers the parametrization is the same as in the quadratic case: because $\mathbb{Z}[2^{1/3}]$ has class number one, every ideal I has a generator, which can then be used to find a natural $\mathbb{Z}$-basis. The main work involved in proving theorem 3.4 then is to compute the matrix in $SL_3(\mathbb{Z})$ that relates the basis obtained in terms of the generator of the ideal to the basis in Hermite normal form that contains information about the roots of the cubic congruences via theorem 3.1.

Suppose $I \subset \mathbb{Z}[2^{1/3}]$ is an ideal that is not divisible by any rational integer. Then, as mentioned above, the fact that $\mathbb{Z}[2^{1/3}]$ has class number one implies that there is an integer $\alpha = a + b2^{1/3} + c2^{1/3} \in \mathbb{Z}[2^{1/3}]$, unique up to multiplication by units, such that I is generated by α. And since $I = (\alpha)$, we get the following natural basis of I,

$$\begin{pmatrix} \alpha \\ \alpha 2^{1/3} \\ \alpha 2^{2/3} \end{pmatrix} = \begin{pmatrix} a & b & c \\ 2c & a & b \\ 2b & 2c & a \end{pmatrix} \begin{pmatrix} 1 \\ 2^{1/3} \\ 2^{2/3} \end{pmatrix}. \tag{3.47}$$

Having the two bases, (3.47) and the one from theorem 3.1, there must be a matrix $\gamma \in GL_3(\mathbb{Z})$ such that

$$\begin{pmatrix} mn & 0 & 0 \\ -\mu n & n & 0 \\ * & \nu & 1 \end{pmatrix} = \gamma \begin{pmatrix} a & b & c \\ 2c & a & b \\ 2b & 2c & a \end{pmatrix}. \tag{3.48}$$

And actually, if we assume that the norm of α, which is of course the determinant of the matrix in (3.47), is positive, then $\gamma \in SL_3(\mathbb{Z})$. We can ensure that $N(\alpha) > 0$ by replacing α by $-\alpha$ if necessary, and in fact we will later choose α to be in a specific fundamental domain for the units that satisfies this condition.

Immediately from (3.48) we see that the top row of γ must be orthogonal to the second and third columns of the matrix in (3.47). That is if the top row of γ has coefficients A, B, and C, then they must satisfy

$$\begin{aligned} \gcd(A, B, C) &= 1 \\ cA + bB + aC &= 0 \\ bA + aB + 2cC &= 0, \end{aligned} \tag{3.49}$$

and these are in fact enough to determine A, B, and C up to an overall change in sign. We observe that the orthogonality conditions are are equivalent to the product of α and $\alpha' = A + B2^{1/3} + C2^{2/3}$ being a rational integer. Indeed,

$$\alpha\alpha' = (aA + 2cB + 2bC) + (bA + aB + 2cC)2^{1/3} + (cA + bB + aC)2^{2/3}. \tag{3.50}$$

Moreover, we see that since $mn > 0$, A, B, and C are completely determined by the requirement that $\alpha\alpha'$ be a positive, rational integer with α' not divisible by any rational integer.

From (3.48) we also see that the last column of γ^{-1} must be the same as the last column of the matrix in (3.47) – in particular we have that $\gcd(a, b, c) = 1$, which just re-affirms that α is not divisible by any rational integer. Now this data, the first row of γ and the last column of γ^{-1} is enough to determine the coset in

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} \backslash SL_3(\mathbb{Z}) \tag{3.51}$$

containing γ, indeed A, B, C and c, b, a are exactly the Plücker coordinates of the coset containing γ. Moreover, we see that since we want μ and ν to be considered as residue classes modulo m and n, respectively, the matrix from theorem 3.1, that is the matrix on the left of (3.48), is only defined up to multiplication on the left by matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix}. \tag{3.52}$$

In other words, given A, B, C and c, b, a satisfying the above requirements, any choice of representative γ for the coset in (3.51) with these as its Plücker coordinates will satisfy (3.48) for some integer representatives of $\mu \pmod m$ and $\nu \pmod n$. This finishes the proof of theorem 3.4.

3.2.2 Proof of corollary 3.5

Corollary 3.5 is a relatively easy consequence of theorem 3.4 in the special case $n = 1$, and hence $\nu = 0$. Before specializing to this case, we note that in general we see from the orthogonality conditions in (3.49) that the vector (A, B, C) must be proportional to the cross-product of (c, b, a) and $(b, a, 2c)$. That is,

$$(a^2 - 2bc, 2c^2 - ab, b^2 - ac) = l(A, B, C) \tag{3.53}$$

where

$$l = \gcd(a^2 - 2bc, 2c^2 - ab, b^2 - ac). \tag{3.54}$$

We claim that in fact $l = n$. Indeed, we can write the determinant of the matrix in (3.47) as

$$mn^2 = a^3 + 2b^3 + 4c^3 - 6abc = a(a^2 - 2bc) + 2c(2c^2 - ab) + 2b(b^2 - ac), \tag{3.55}$$

so substituting (3.54) and using the fact that

$$mn = aA + 2cB + 2bC \tag{3.56}$$

shows that $l = n$.

We now see that $n = 1$ is equivalent to $\gcd(a^2 - 2bc, 2c^2 - ab, b^2 - ac) = 1$, the condition on a, b and c in corollary 3.5. Moreover, we recall from corollary 3.3 that in the case $n = 1$ we have $\lambda = -\mu^2$, and so (3.48) in this case takes the form

$$\begin{pmatrix} m & 0 & 0 \\ -\mu & 1 & 0 \\ -\mu^2 & 0 & 1 \end{pmatrix} = \gamma \begin{pmatrix} a & b & c \\ 2c & a & b \\ 2b & 2c & a \end{pmatrix}. \tag{3.57}$$

In contrast to section 3.2.1, γ is now considered as a coset in

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & 0 & 1 \end{pmatrix} \backslash SL_3(\mathbb{Z}). \tag{3.58}$$

In fact we see that

$$\gamma^{-1} = \begin{pmatrix} u & b & c \\ v & a & b \\ w & 2c & a \end{pmatrix} \tag{3.59}$$

where u, v, and w are integers so that

$$u(a^2 - 2bc) + v(2c^2 - ab) + w(2c^2 - ab) = 1, \tag{3.60}$$

which exist by the coprimality condition. Further we note that different choices of u, v and w satisfying (3.60) give different representatives of the coset in (3.58), which in turn give different representatives of μ and μ^2 modulo m. The parametrization as stated in corollary 3.5 now follows from explicitly calculating

$$\begin{pmatrix} m \\ -\mu \\ -\mu^2 \end{pmatrix} = \gamma^{-1} \begin{pmatrix} a \\ 2c \\ 2b \end{pmatrix} = \begin{pmatrix} a^2 - 2bc & 2c^2 - ab & b^2 - ac \\ bw - av & au - cw & cv - bu \\ 2cv - aw & bw - 2cu & au - bv \end{pmatrix} \begin{pmatrix} a \\ 2c \\ 2b \end{pmatrix}. \tag{3.61}$$

3.3 Approximation of the roots

3.3.1 *LU* decomposition

We will perform a slight variation on the LU decomposition for the matrix γ as above. The reason for doing a variation instead of a standard LU decomposition will be made clear in chapter 5. Setting

$$\gamma = \begin{pmatrix} A & B & C \\ U & V & W \\ X & Y & Z \end{pmatrix}, \tag{3.62}$$

we first compute

$$\gamma = \begin{pmatrix} 1 & 0 & 0 \\ \frac{W}{C} & 1 & 0 \\ \frac{Z}{C} & -\frac{u}{c} & 1 \end{pmatrix} \begin{pmatrix} A & B & C \\ \frac{b}{C} & -\frac{c}{C} & 0 \\ \frac{1}{c} & 0 & 0 \end{pmatrix}, \tag{3.63}$$

where we recall that

$$\gamma^{-1} = \begin{pmatrix} u & x & c \\ v & y & b \\ w & z & a \end{pmatrix}. \tag{3.64}$$

The decomposition (3.63) is of course only valid if C and c are not zero 0. We will ensure the condition $C \neq 0$ later in picking the fundamental domain for the action of the units, but if $c = 0$, we can instead decompose

$$\gamma = \begin{pmatrix} 1 & 0 & 0 \\ \frac{W}{C} & 1 & 0 \\ \frac{Z}{C} & 0 & 1 \end{pmatrix} \begin{pmatrix} A & B & C \\ \frac{b}{C} & 0 & 0 \\ \frac{y}{C} & \frac{x}{C} & 0 \end{pmatrix}. \tag{3.65}$$

In the first case, $c \neq 0$, applying the decomposition to the parametrization of theorem 3.4 yields

$$\begin{pmatrix} mn & 0 & 0 \\ -\mu n & n & 0 \\ \lambda & \nu & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{W}{C} & 1 & 0 \\ \frac{Z}{C} & -\frac{u}{c} & 1 \end{pmatrix} \begin{pmatrix} mn & 0 & 0 \\ -\frac{B}{C}n & n & 0 \\ \frac{a}{c} & \frac{b}{c} & 1 \end{pmatrix}, \tag{3.66}$$

after recalling that $(nA, nB, nC) = (a^2 - 2bc, 2c^2 - ab, b^2 - ac)$. Similarly the decomposition (3.65) yields in the case $c = 0$ that

$$\begin{pmatrix} mn & 0 & 0 \\ -\mu n & n & 0 \\ \lambda & \nu & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{W}{C} & 1 & 0 \\ \frac{Z}{C} & 0 & 1 \end{pmatrix} \begin{pmatrix} mn & 0 & 0 \\ -\frac{B}{C}n & n & 0 \\ \frac{y}{C} & \frac{x}{C} & 1 \end{pmatrix}. \tag{3.67}$$

In either decomposition, we see that

$$\frac{\mu}{m} = -\frac{W}{C} + \frac{B}{Cm}. \tag{3.68}$$

Theorem 3.7 clearly follows from (3.68) once we show that $B \ll C$. And this clearly follows from proposition 3.6, which will be proved in section 3.3.3.

3.3.2 Intersection of lines

Before finishing the proof of theorem 3.7 by proving proposition 3.6, we first prove theorem 3.8, still assuming proposition 3.6. This theorem could be proved by doing an LU decomposition, and in fact a close inspection of section 3.3.1 already gives what we need. But instead, to obtain an approximation to $\frac{\mu}{m}$ in the case $n = 1$, we return to the equation (3.48) with γ as in (3.59). We have

$$\begin{pmatrix} u & b & c \\ v & a & b \\ w & 2c & a \end{pmatrix} \begin{pmatrix} m & 0 & 0 \\ -\mu & 1 & 0 \\ -\mu^2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ 2c & a & b \\ 2b & 2c & a \end{pmatrix}. \tag{3.69}$$

Examining this equality for the $(1,1)$ entry on the right-hand side, we obtain

$$u - b\frac{\mu}{m} - c\frac{\mu^2}{m} = \frac{a}{m} \tag{3.70}$$

upon dividing by m. We expect, and will ensure later by picking $\alpha = a + b2^{1/3} + c2^{2/3}$ in the fundamental domain given by proposition 3.6, that $a \ll m^{1/3}$. Accordingly, we expect the right hand side of (3.70) to be small, specifically $\ll m^{-2/3}$. We can interpret this geometrically as the point $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ lying close to the line $bX + cY = u$.

Similarly, by inspecting the $(2,1)$ and $(3,1)$ entries of the right side of (3.69), we expect that $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ will also lie close to the lines $aX + bY = v$ and $2cX + aY = w$. Now, if the triangle with sides these three lines has at least one corner, that is one of the points

$$\left(\frac{bu - cv}{b^2 - ac}, \frac{bv - au}{b^2 - ac}\right), \left(\frac{cv - au}{2c^2 - ab}, \frac{2cu - bv}{2c^2 - ab}\right), \left(\frac{au - bv}{a^2 - 2bc}, \frac{av - 2cu}{a^2 - 2bc}\right), \tag{3.71}$$

with an angle which is not too small, then the point $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ will be close to this corner. In fact proposition 3.6 gives sufficient control on the angle to show that the first of these is a good approximation.

That said, we prove theorem 3.8 explicitly by solving for $\frac{\mu}{m}$ and $\frac{\mu^2}{m}$ from one of the pairs of the three equations coming from (3.69) above. We have

$$\frac{1}{m}\begin{pmatrix} b & c \\ a & b \end{pmatrix}\begin{pmatrix} \mu \\ \mu^2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} - \frac{1}{m}\begin{pmatrix} a \\ 2c \end{pmatrix}, \tag{3.72}$$

so

$$\frac{1}{m}\begin{pmatrix} \mu \\ \mu^2 \end{pmatrix} = \frac{1}{b^2 - ac}\begin{pmatrix} bu - cv \\ bv - au \end{pmatrix} + \frac{1}{m(b^2 - ac)}\begin{pmatrix} 2c^2 - ab \\ a^2 - 2bc \end{pmatrix}. \tag{3.73}$$

Recalling that when $n = 1$, $A = a^2 - 2bc$, $B = 2c^2 - ab$, and $C = b^2 - ab$, we see that proposition 3.6 provides sufficient estimates to conclude theorem 3.8.

3.3.3 Fundamental domain

For $\beta \in \mathbb{Q}(2^{1/3})$, let $\beta^{(1)}$ be the real embedding and $\beta^{(j)}$, $j = 2, 3$, be the complex embeddings. And for $C > 0$ a constant to be determined, set

$$\mathcal{D}_1 = \left\{\beta \in K \ : \ C|N(\beta)|^{1/3} < \beta^{(1)} \leq C\varepsilon^{(1)}|N(\beta)|^{1/3}\right\}, \tag{3.74}$$

where $\varepsilon = 1 + 2^{1/3} + 2^{2/3}$ is the fundamental unit. $\mathcal{D}_1$ is clearly a fundamental domain for the action of the units in $\mathbb{Z}[2^{1/3}]$.

For $\beta \in \mathbb{Q}(2^{1/3})$, we have

$$|\beta^{(2)}|^2 = |\beta^{(2)}\beta^{(3)}| = \frac{|N(\beta)|}{|\beta^{(1)}|}, \tag{3.75}$$

so for $\beta \in \mathcal{D}_1$, we have

$$|\beta^{(2)}| \leq C^{-1/2}N(\beta)^{1/3}. \tag{3.76}$$

Hence

$$|\mathrm{Tr}(\beta))| = \left|\beta^{(1)} + \beta^{(2)} + \beta^{(3)}\right| \leq (\varepsilon^{(1)}C + 2C^{-1/2})N(\beta)^{1/3}, \tag{3.77}$$

and

$$|\mathrm{Tr}(\beta)| \geq \beta^{(1)} - 2\left|\beta^{(2)}\right| \geq (C - 2C^{-1/2})N(\beta)^{1/3}. \tag{3.78}$$

So picking $C = 2$, say, gives

$$|Tr(\beta)| \asymp N(\beta)^{1/3} \tag{3.79}$$

for $\beta \in \mathcal{D}_1$.

We now define the fundamental domain for α by

$$\mathcal{D} = \left\{ \beta \in \mathbb{Q}(2^{1/3}) \ : \ 2^{-1/3}\beta^{-1} \in \mathcal{D}_1 \right\}. \tag{3.80}$$

By how $\mathcal{D}_1$ was defined, it is not hard to see that if $\alpha \in \mathcal{D}$, then all of the embeddings satisfy $|\alpha^{(j)}| \asymp |N(\alpha)|^{1/3} = m^{1/3}n^{2/3}$. And since a, b, and c can be written as fixed linear combinations of the embeddings, we have that a, b, and c are all $\ll m^{1/3}n^{2/3}$.

Now, since $\alpha' = \frac{mn}{\alpha'}$, we see that if $\alpha \in \mathcal{D}$, then $2^{1/3}\alpha' \in \mathcal{D}_1$. So not only are the embeddings of α' all $\asymp |N(\alpha')|^{1/3} = m^{2/3}n^{1/3}$, whence A, B, and C are all $\ll m^{2/3}n^{1/3}$, but also from (3.79) we have $C = \frac{1}{6}\mathrm{Tr}(2^{1/3}\alpha') \asymp m^{2/3}n^{1/3}$.

3.4 Spacing between the roots

We begin by proving proposition 3.9. Let $\left(\frac{r}{q}, \frac{s}{q}\right)$ and $\left(\frac{r_1}{q_1}, \frac{s_1}{q_1}\right)$ be representatives of distinct torsion points in $\mathbb{R}^2/\mathbb{Z}^2$. Let $AX + BY = C$ with A, B, C integers such that $gcd(A, B, C) = 1$ be the equation of the line between the two. The coprimality condition on the coefficients implies that

$$\mathbb{Z}^3 \cap \mathrm{Null}\begin{pmatrix} q & q_1 \\ -r & -r_1 \\ -s & -s_1 \end{pmatrix} = \mathbb{Z}\begin{pmatrix} C & A & B \end{pmatrix}. \tag{3.81}$$

On the other hand the cross product of $\begin{pmatrix} q & -r & -s \end{pmatrix}$ and $\begin{pmatrix} q_1 & -r_1 & -s_1 \end{pmatrix}$ is in this null space, so we can conclude that

$$\begin{pmatrix} rs_1 - r_1 s & qs_1 - q_1 s & rq_1 - r_1 q \end{pmatrix} = k\begin{pmatrix} C & A & B \end{pmatrix}, \tag{3.82}$$

for some integer k. Since the torsion points are distinct, we know that $k \neq 0$, so in fact $|k| \geq 1$. We have

$$\left|\frac{r}{q} - \frac{r_1}{q_1}\right| = \frac{|rq_1 - r_1 q|}{qq_1} \geq \frac{|B|}{qq_1}, \tag{3.83}$$

and similarly

$$\left|\frac{s}{q} - \frac{s_1}{q_1}\right| \geq \frac{|A|}{qq_1}. \tag{3.84}$$

From (3.83) and (3.84) we see that the size of $|A|$ and $|B|$ from lines $AX + BY = C$ passing through a representative of a torsion point control the spacing from this representative to a representative of any another torsion point. So fixing a torsion point, we can lower bound the distance from it to any other torsion point by considering the lattice of lines passing through it. Moreover, we observe that if $AX + BY = C$ passes through a representative $\left(\frac{r}{q}, \frac{s}{q}\right)$, then $AX + BY = C + kA + lB$ passes through another representative $\left(\frac{r}{q} + k, \frac{s}{q} + l\right)$, where k and l are integers. Hence the set of $\begin{pmatrix} A & B \end{pmatrix}$ under consideration will not depend on the choice of representative.

Before applying this lemma to our approximations in theorem 3.8, we remark that the set of all $\begin{pmatrix} A & B \end{pmatrix}$ such that some fixed representative $\left(\frac{r}{q}, \frac{s}{q}\right)$ lies on a line $AX + BY = C$ forms a sublattice of $\mathbb{Z}^2$. As mentioned previously, this lattice is independent of the representative $\left(\frac{r}{q}, \frac{s}{q}\right)$ chosen for the torsion point $\mathbb{R}^2/\mathbb{Z}^2$. And if this lattice, properly oriented and normalized to have co-volume 1, does not lie too high in the cusp of $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$, then the shortest vector in the lattice will have norm about the square root of the co-volume.

Let's consider the point $\left(\frac{bu-cv}{b^2-ac}, \frac{bv-au}{b^2-ac}\right)$, the approximation to $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ given by theorem 3.8. From the way it was constructed in section 3.3.2, as the intersection of the lines $bX + cY = u$ and $aX + bY = v$, we can see that the lattice discussed in the previous paragraph contains

$$\mathrm{Span}_{\mathbb{Z}} \left\{ \begin{pmatrix} b & c \end{pmatrix}, \begin{pmatrix} a & b \end{pmatrix} \right\}. \tag{3.85}$$

To see that the lattice of lines containing the point is no bigger, we simply note that the row vectors $\begin{pmatrix} u & b & c \end{pmatrix}$ and $\begin{pmatrix} v & a & b \end{pmatrix}$ can be completed by a third vector to make a matrix in $SL_3(\mathbb{Z})$.

The co-volume of this lattice is $b^2 - ac$, which we have forced by proposition 3.6 to be $\asymp m^{2/3}$. Recalling that this proposition also ensures that a, b, and c are all $\ll m^{1/3}$, we also have

$$b^2 - ac \leq \sqrt{(b^2 + c^2)(a^2 + b^2)} \ll m^{1/3}\sqrt{b^2 + c^2}, \tag{3.86}$$

so

$$\sqrt{b^2 + c^2} \asymp m^{1/3}, \tag{3.87}$$

and similarly

$$\sqrt{a^2 + b^2} \asymp m^{1/3}. \tag{3.88}$$

Suppose we scale and rotate this lattice so that the vector $\begin{pmatrix} b & c \end{pmatrix}$ becomes $\begin{pmatrix} 1 & 0 \end{pmatrix}$, thereby identifying the basis of the lattice with a point in the upper half-plane $\mathbb{H}$, the image of $\begin{pmatrix} a & b \end{pmatrix}$ under this scaling and rotation. After this transformation, the co-volume of the lattice is $\gg 1$, whence the point in $\mathbb{H}$ has height$\gg 1$ above the x-axis. Moreover, since $\sqrt{a^2+b^2} \asymp \sqrt{b^2+c^2}$, the point also has distance $\ll 1$ from the origin. As such, the point lies in a fixed, compact region of $\mathbb{H}$, whence, even after quotienting out by the action of $SL_2(\mathbb{Z})$ on the basis, the lattice lies in a fixed region, bounded away from the cusp.

In accordance with the remarks above, we know that the shortest vector in the lattice will have norm $\asymp$ square root of the co-volume, so here the shortest vector will be $\asymp m^{1/3}$. Combining this with proposition 3.9 we have that the torsion point given by the approximation of theorem 3.8 is spaced by at least $\frac{1}{m^{1/3}Q}$ from any other torsion point with torsion $\leq Q$.

We can now finish the proof of theorem 3.10, the spacing result for the set of points

$$S = \left\{ \left(\frac{\mu}{m}, \frac{\mu^2}{m} \right) \ : \ \mu^3 \equiv 2 \pmod{m}, \ M < m \leq 2M \right\}. \tag{3.89}$$

First we show that we can recover the point $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$ up to $O(1)$ possibilities from the approximation in theorem 3.8. Given the approximation, we find the lattice of lines containing it, which is (3.85). Clearly if we have the basis of the lattice as given in (3.85) we can recover a, b, and c and hence m and $\mu \pmod{m}$. However, we do know that this basis, properly normalized, lies in a fixed compact subset of $SL_2(\mathbb{R})$. And since $SL_2(\mathbb{Z})$ acts discontinuously on $SL_2(\mathbb{R})$, the number of bases of our lattice contained in this compact subset will be bounded, proving the claim.

Now, from the spacing property of the approximations, we see that around each approximation there is a disc of radius $\gg \frac{1}{M}$ that contains no other approximation, and so there are at most $O(1)$ approximations in any disk of radius $O\left(\frac{1}{M}\right)$. And from the above, each approximation can arise from at most $O(1)$ points $\left(\frac{\mu}{m}, \frac{\mu^2}{m}\right)$, so theorem 3.10 follows.

3.5 Large sieve inequality

Utilizing the duality principle, we see that it is enough for theorem 3.11 to prove that for any sequence of complex numbers $b_{m,\mu}$,

$$\begin{aligned}\sum_{k\le K}\sum_{l\le L}\left|\sum_{M<m\le 2M}\sum_{\mu^3\equiv 2(m)} b_{m,\mu}e\left(\frac{k\mu+l\mu^2}{m}\right)\right|^2 \\ \ll (M+K)(M+L)\sum_{M<m\le 2M}\sum_{\mu^3\equiv 2(m)}|b_{m,\mu}|^2.\end{aligned} \tag{3.90}$$

Let $f:\mathbb{R}\to\mathbb{R}$ be a smooth function such that $f(x)\ge 0$ for all x, $f(x)\ge 1$ for $0\le x\le 1$, and $\hat{f}$, the Fourier transform of f, is compactly supported. Then the left hand side of (3.90) is

$$\le \sum_k f\left(\frac{k}{K}\right)\sum_l f\left(\frac{l}{L}\right)\left|\sum_{M<m\le 2M}\sum_{\mu^3\equiv 2(m)} b_{m,\mu}e\left(\frac{k\mu+l\mu^2}{m}\right)\right|^2. \tag{3.91}$$

Expanding out the square, (3.91) becomes

$$\sum_{\substack{M<m\le 2M\\ \mu^3\equiv 2(m)}}\sum\sum_{\substack{M<m_1\le 2M\\ \mu_1^3\equiv 2(m_1)}}\sum b_{m,\mu}\bar{b}_{m_1,\mu_1}\mathcal{B}(m,\mu,m_1,\mu_1)\mathcal{B}'(m,\mu,m_1,\mu_1), \tag{3.92}$$

where

$$\mathcal{B}(m,\mu,m_1,\mu_1)=\sum_k f\left(\frac{k}{K}\right)e\left(k\left(\frac{\mu}{m}-\frac{\mu_1}{m_1}\right)\right), \tag{3.93}$$

and

$$\mathcal{B}'(m,\mu,m_1,\mu_1)=\sum_l f\left(\frac{l}{L}\right)e\left(l\left(\frac{\mu^2}{m}-\frac{\mu_1^2}{m_1}\right)\right). \tag{3.94}$$

Applying Poisson summation to (3.93), we have

$$\mathcal{B}(m,\mu,m_1,\mu_1)=K\sum_k \hat{f}\left(K\left(k-\left(\frac{\mu}{m}-\frac{\mu_1}{m_1}\right)\right)\right). \tag{3.95}$$

Now, by the compact support of $\hat{f}$, only k for which

$$\left|k-\left(\frac{\mu}{m}-\frac{\mu_1}{m_1}\right)\right|\ll\frac{1}{K} \tag{3.96}$$

will contribute to the sum in (3.95). If $K\gg 1$, then at most one k will appear, and even then, only when

$$\left\|\frac{\mu}{m}-\frac{\mu_1}{m_1}\right\|\ll\frac{1}{K}, \tag{3.97}$$

where we use $||\cdot||$ to denote the distance to the nearest integer, which gives the metric on $\mathbb{R}/\mathbb{Z}$. Hence for $K \gg 1$ we have

$$\mathcal{B}(m,\mu,m_1,\mu_1) \ll K \mathbb{1}_{\left|\left|\frac{\mu}{m}-\frac{\mu_1}{m_1}\right|\right| \ll \frac{1}{K}}. \tag{3.98}$$

In fact, this bound clearly works for all K, perhaps by adjusting the implied constants.

By the same reasoning, we have the similar bound for $\mathcal{B}'$,

$$\mathcal{B}'(m,\mu,m_1,\mu_1) \ll L \mathbb{1}_{\left|\left|\frac{\mu^2}{m}-\frac{\mu_1^2}{m_1}\right|\right| \ll \frac{1}{L}}. \tag{3.99}$$

Denoting by $R_{m,\mu}$ the set of points (x,y) in $\mathbb{R}^2/\mathbb{Z}^2$ satisfying

$$\left|\left|\frac{\mu}{m}-x\right|\right| \ll \frac{1}{K}, \quad \left|\left|\frac{\mu^2}{m}-y\right|\right| \ll \frac{1}{L}, \tag{3.100}$$

with the same implied constants as in the indicator functions of (3.98) and (3.99), the left hand side of (3.90) is

$$\ll KL \sum\sum\sum\sum_{\substack{M<m,m_1\leq 2M \\ \mu^3\equiv 2(m),\ \mu_1^3\equiv 2(m_1) \\ \left(\frac{\mu_1}{m_1},\frac{\mu_1^2}{m_2}\right)\in R_{m,\mu}}} |b_{m,\mu}b_{m_1,\mu_1}|. \tag{3.101}$$

Applying $|b_{m,\mu}b_{m_1,\mu_1}| \leq \frac{1}{2}|b_{m,\mu}|^2 + \frac{1}{2}|b_{m_1,\mu_1}|^2$ and exploiting the symmetry between m,μ and m_1,μ_1, we see that (3.101) is

$$\leq KL \sum_{M<m\leq 2M} \sum_{\mu^3\equiv 2(m)} |b_{m,\mu}|^2 \sum\sum_{\substack{M<m_1\leq 2M,\ \mu_1^3\equiv 2(m_1) \\ \left(\frac{\mu_1}{m_1},\frac{\mu_1^2}{m_1}\right)\in R_{m,\mu}}} 1. \tag{3.102}$$

We can cover the rectangle $R_{m,\mu}$ by $\ll \left(\frac{M}{K}+1\right)\left(\frac{M}{L}+1\right)$ discs of radius $\frac{1}{M}$, and in each of these discs there are $\ll 1$ points $\left(\frac{\mu_1}{m_1},\frac{\mu_1^2}{m_1}\right)$, according to theorem 3.10. We have that (3.102) is then

$$\ll KL\left(\frac{M}{K}+1\right)\left(\frac{M}{L}+1\right) \sum_{M<m\leq 2M} \sum_{\mu^3\equiv 2(m)} |b_{m,\mu}|^2, \tag{3.103}$$

from which (3.91), and hence theorem 3.11, follows.

Chapter 4

Roots of general polynomial congruences

In this chapter we will be generalizing the results of chapter 3 to a general polynomial. To be specific, we will be studying the roots of the congruence

$$F(X) = X^d - a_1 X^{d-1} - \cdots - a_d \equiv 0 \pmod{m}, \tag{4.1}$$

where we will assume that $d > 2$. Just as the ring $\mathbb{Z}[2^{1/3}]$ was the main tool for studying roots of the congruence $X^3 \equiv 2 \pmod{m}$, we will be considering the ring $\mathbb{Z}[\alpha]$. Here α is a root of the polynomial F thought of as a vector in $\mathbb{R}^d$ with coordinates the real embeddings and the real and imaginary parts of the complex embeddings.

We will only generalize the results of chapter 3 when $n = 1$, that is the ideals I for which $\mathbb{Z}[2^{1/3}]/I$ is cyclic. We begin in this chapter by proving an analogue of corollary 3.2, which characterizes the special case that we will be working with in what follows.

Proposition 4.1. *Let I be the sublattice of $\mathbb{Z}[\alpha]$ with basis $\{\beta_1, \ldots, \beta_d\}$ given by*

$$\beta_i = \sum_{j=1}^{d} b_{ij} \alpha^{j-1}, \tag{4.2}$$

where the matrix $B = (b_{ij})_{1 \le i,j \le d}$ is in (lower triangular) Hermite normal form, so $b_{ij} = 0$ if $j > i$ and $0 \le b_{ij} < b_{jj}$. Then for I to be an ideal of $\mathbb{Z}[\alpha]$, it is necessary that b_{ii} divides b_{ij} and b_{jj} for all $1 \le j \le i$. In particular, if I is an ideal, then the b_{ii} are the invariant factors of $\mathbb{Z}[\alpha]/I$.

Having this characterization of the ideals we will be working with, we prove the following analogue of corollary 3.3:

Theorem 4.2. *Let $I \subset \mathbb{Z}[\alpha]$ be an ideal such that the quotient $\mathbb{Z}[\alpha]/I$ is additively* *cyclic. Then I has a unique basis $\{\beta_1, \ldots, \beta_d\}$ of the form*

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{pmatrix} = \begin{pmatrix} m & 0 & \cdots & 0 \\ -\mu & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\mu^{d-1} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \alpha \\ \vdots \\ \alpha^{d-1}, \end{pmatrix} \tag{4.3}$$

where $m > 0$ and μ is a residue class modulo m satisfying the polynomial congruence (4.1). Conversely, given $m > 1$ and $\mu \pmod m$ satisfying (4.1), the sublattice I of $\mathbb{Z}[\alpha]$ given by the basis $\{\beta_1, \ldots, \beta_d\}$ as in (4.3) is an ideal such that $\mathbb{Z}[\alpha]/I$ is cyclic.

Continuing as we did in chapters 2 and 3, we use this correspondence between certain ideals and roots of the congruence to derive a parametrization. The idea is largely the same, relating two bases for an ideal by a matrix in $SL_d(\mathbb{Z})$. But since we no longer have that our ring $\mathbb{Z}[\alpha]$ has class number one, the setup for the parametrization is more complicated, it is as follows:

Theorem 4.3. *Fix a complete system of representatives I_l for the narrow class group of $\mathbb{Z}[\alpha]$. For each l, fix bases $\{\beta_1, \ldots, \beta_d\}$ and $\{\overline{\beta}_1, \ldots, \overline{\beta}_d\}$ of I_l and I_l^{-1}, respectively, so that*

$$\operatorname{sign} \det(\beta_i^{(j)})_{\substack{1 \le i \le d \\ 1 \le j \le d}} = \operatorname{sign} \det(\overline{\beta}_i^{(j)})_{\substack{1 \le i \le d \\ 1 \le j \le d}} = \operatorname{sign} \det((\alpha^{(j)})^{i-1})_{\substack{1 \le i \le d \\ 1 \le j \le d}}. \tag{4.4}$$

Define the integers b_{ijk} by

$$\overline{\beta}_i \beta_j = \sum_{k=1}^{d} b_{ijk} \alpha^{k-1}, \tag{4.5}$$

and set

$$B_i = (b_{ijk})_{\substack{1 \le j \le d \\ 1 \le k \le d}}, \quad 1 \le i \le d. \tag{4.6}$$

Let c_i, $1 \le i \le d$, be integers so that, with

$$C = \sum_{i=1}^{d} c_i B_i, \tag{4.7}$$

the integers, say c'_j, $1 \le j \le d$, forming the first row of $(\det C)C^{-1}$ satisfy

$$\gcd(c'_1, \ldots, c'_d) = 1, \tag{4.8}$$

assume in addition that the c_i *are so that*

$$\gamma = \sum_{i=1}^{d} c_i \overline{\beta}_i \tag{4.9}$$

is in a specific fundamental domain for the action on $\mathbb{Q}(\alpha)$ *of the positive norm units in* $\mathbb{Z}[\alpha]$.

Let u_j, $1 \leq j \leq d$, *be integers so that*

$$\sum_{j=1}^{d} u_j c_j' = 1, \tag{4.10}$$

and denote the (j,k) *entry of the matrix* C *by* c_{jk}. *Then*

$$\begin{pmatrix} m \\ \mu \\ \vdots \\ \mu^{d-1} \end{pmatrix} = \begin{pmatrix} u_1 & c_{12} & \cdots & c_{1d} \\ u_2 & c_{22} & \cdots & c_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ u_d & c_{d2} & \cdots & c_{dd} \end{pmatrix}^{-1} \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{d1} \end{pmatrix} \tag{4.11}$$

parametrizes all the m *and* $\mu \pmod{m}$ *satisfying* $F(\mu) \equiv 0 \pmod{m}$ *and corresponding to invertible ideals via theorem 4.2. This parametrization is unique, noting that different choices of the integers* u_j *give different representatives of* $\mu \pmod{m}$.

We remark that a weakness in the above parametrization is the absence of a characterization of the roots $\mu \pmod{m}$ for which the corresponding ideal is invertible. We imagine, although we do not do so here, that the theory of conductors would show that there is fixed moduli m_0 and roots $\mu_j \pmod{m_0}$ so that the ideal corresponding to μ (mod m) is invertible if and only if $\gcd(m, m_0, \mu - \mu_j) = 1$, which is to say that μ is genuinely different from the roots $\mu_j \pmod{m_0}$.

Nevertheless, we continue as in chapter 3 by finding an approximation to the points $\left(\frac{\mu}{m}, \ldots, \frac{\mu^{d-1}}{m}\right)$, at least those that appear in the parametrization in theorem 4.3.

Theorem 4.4. *With the notation as in theorem 4.3, let* C_{i1} *denote the* $(d-1) \times (d-1)$ *sub-matrix obtained from* C *by removing the* i*th row and first column, and set* $\boldsymbol{u}_i$ *to be the vector* $(u_1, \ldots, u_d)$ *with the* i*th entry removed. Then for some* i,

$$\begin{pmatrix} \frac{\mu}{m} \\ \vdots \\ \frac{\mu^{d-1}}{m} \end{pmatrix} = C_{1i}^{-1} \boldsymbol{u}_i + O\left(\frac{1}{m}\right) \pmod{\mathbb{Z}^{d-1}}. \tag{4.12}$$

We remark that a weakness of this theorem is that we do not specify, as we did in chapters 2 and 3, what exactly the approximation is, instead just giving an option of d points, at least one (although most likely all) of which is close to the point we are trying to approximate. This defect can be remedied when $\mathbb{Z}[\alpha]$ has at least one real embedding by picking the fundamental domain for the action of the units appropriately, but we do not outline how this is done here.

This approximation, combined with the following proposition about the spacing between torsion points in $\mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$, the analogue of proposition 3.9, is used to derive a spacing property, theorem 4.6 below.

Proposition 4.5. *Let Q be a positive real number and let $\left(\frac{r_1}{q}, \dots, \frac{r_{d-1}}{q}\right)$ be a q-torsion point in $\mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$. Further, let Λ be the lattice in $\mathbb{R}^{d-1}$ consisting of the $(1, j)$ entries of the Plücker coordinates of the integral lines containing $\left(\frac{r_1}{q}, \dots, \frac{r_{d-1}}{q}\right)$. Then Λ is well-defined, and the distance between $\left(\frac{r_1}{q}, \dots, \frac{r_{d-1}}{q}\right)$ and any distinct torsion point with torsion $\leq Q$ is at least*

$$\frac{1}{qQ} \min\{||\boldsymbol{v}|| \ : \ \boldsymbol{v} \in \Lambda, \boldsymbol{v} \neq 0\}. \tag{4.13}$$

This proposition uses heavily the Plücker coordinates of lines in $\mathbb{R}^{d-1}$, or perhaps really the corresponding projective space. We recall that the Plücker coordinates of the line passing through points $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^{d-1}$ with homogeneous coordinates $(x_0, x_1, \dots, x_{d-1})$ and $(y_0, y_1, \dots y_{d-1})$ are the $\binom{d}{2}$ determinants of the 2×2 sub-matrices of

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_{d-1} \\ y_0 & y_1 & \cdots & y_{d-1} \end{pmatrix}. \tag{4.14}$$

In the statement of proposition 4.5, we refer to the $(1, j)$ Plücker coordinates, by which we simply mean the determinants of the sub-matrix of (4.14) formed by the first and jth column. We remark, and we will make use of later, the fact that we can dually describe a line in $\mathbb{R}^{d-1}$ as the intersection of $d-2$ planes. If these planes are given by equations $u_j + \sum_{j=1}^{d-1} c_{ij} X_j = 0$, $1 \leq j \leq d-2$, then the Plükcer coordinates can also be realized as the determinants of the sub-matrices of

$$\begin{pmatrix} u_1 & c_{11} & \cdots & c_{1(d-1)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{d-2} & c_{(d-2)1} & \cdots & c_{(d-2)(d-1)} \end{pmatrix} \tag{4.15}$$

obtained by removing two columns. Up to scaling and sign, the coordinates obtained from the planes by removing the ith and jth column will be equal to the (i,j) coordinate as described above from the point matrix (4.14).

Applied to the approximation in theorem 4.4, proposition 4.5 yields the following theorem:

Theorem 4.6. *Let M be a positive real number and let D be a disc in $\mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$ with radius $\frac{1}{M}$. Then the number of $\left(\frac{\mu}{m},\ldots,\frac{\mu^{d-1}}{m}\right) \in D$ with $F(\mu) \equiv 0 \pmod{m}$ and $M < m \leq 2M$ is bounded by a constant depending only on the polynomial F.*

A key step in the proof of this theorem, similar to the proofs of theorems 2.6 and theorem 3.10, is the fact that one can obtain the root $\mu \pmod{m}$ up to bounded number of possibilities, the bound depending on the congruence. The proof of this fact in this general setting is very similar to the proof in section 3.4. It is interesting to note that in general recovering the root $\mu \pmod{m}$ from the lattice of lines passing the approximation uses a linear independence argument that only works when $d > 2$, see the end of section 4.4. While this appears strange at first, this lack of linear independence is used crucially the proof of theorem 2.4, the equidistribution of the quadratic roots.

We close the introduction to this chapter by quoting the following consequence of theorem 4.6, whose proof we will omit since it is essentially no different from the proof of theorem 3.11.

Theorem 4.7. *Let K_j, $1 \leq j \leq d-1$ be positive real numbers and let $\beta_{k_1\cdots k_{d-1}}$ be a sequence of complex numbers. Then*

$$\sum_{M<m\leq 2M} \sum_{F(\mu)\equiv 0(m)} \left| \sum_{|k_1|\leq K_1} \cdots \sum_{|k_{d-1}|\leq K_{d-1}} \beta_{k_1\cdots k_{d-1}} e\left(\frac{k_1\mu + \cdots k_{d-1}\mu^{d-1}}{m}\right)\right|^2 \\ \ll (M+K_1)\cdots(M+K_{d-1}) \sum_{|k_1|\leq K_1} \cdots \sum_{|k_{d-1}|\leq K_{d-1}} |\beta_{k_1\cdots k_{d-1}}|^2. \tag{4.16}$$

4.1 Correspondence between roots and ideals

In this section we prove theorem 4.2. We start with a sublattice I of the ring $\mathbb{Z}[\alpha]$, which has a unique basis $\{\beta_1, \dots, \beta_d\}$ in Hermite normal form. This is to say that

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} = B \begin{pmatrix} 1 \\ \vdots \\ \alpha^{d-1} \end{pmatrix} \tag{4.17}$$

where $B = (b_{ij})$ is lower-triangular and $b_{ij} = 0$ if $j > i$, with the diagonal entries $b_{jj} > 0$ and $0 \le b_{ij} < b_{jj}$.

Just as in the cubic case in the previous chapters, the sublattice I being an ideal is equivalent to a certain matrix being integral. Here the matrix is BAB^{-1} where A is the matrix by which α acts, which is in rational canonical form. Explicitly we have $A = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 & \text{if } 1 \le i \le d-1, j = i+1 \\ a_{d-j+1} & \text{if } i = d \\ 0 & \text{otherwise.} \end{cases} \tag{4.18}$$

Letting $B^{-1} = (b'_{ij})$, we observe that for $1 \le i \le d-1$, the (i, j) entry of BAB^{-1} is

$$\sum_{k=j-1}^{i} b_{ik} b'_{(k+1)j}, \tag{4.19}$$

where for convenience we set $b_{i0} = 0$. In particular, the $(i, i+1)$ entry is simply $b_{ii} b'_{(i+1)(i+1)} = \frac{b_{ii}}{b_{(i+1)(i+1)}}$. From this we see that for I to be an ideal, it is necessary to have

$$b_{11} = m_1 b_{22} = m_1 m_2 b_{33} = \cdots = \prod_{i=1}^{d} m_i. \tag{4.20}$$

The diagonal entries of BAB^{-1} are not much more difficult to compute due to the fact that $b'_{i(i-1)} = -\frac{b_{i(i-1)}}{b_{ii} b_{(i-1)(i-1)}}$. With this, the (i, i) entry is

$$\frac{b_{i(i-1)}}{b_{ii}} - \frac{b_{(i+1)i}}{b_{(i+1)(i+1)}}. \tag{4.21}$$

Applied with $i = 1$, we see that for I to be an ideal, it is necessary that $b_{21} = c_{21} b_{22}$ for some integer c_{21}. Continuing inductively, we see that for all $2 \le i \le d$, $b_{i(i-1)} = c_{i(i-1)} b_{ii}$ is necessary.

Continuing this analysis by trying to obtain exactly the necessary and sufficient conditions for I to be an ideal is a bit unwieldy, so we will instead focus on the analogue of $n = 1$, that is corollary 3.3, in chapter 3. Our first step is to classify the ideals we are interested in, proving proposition 4.1.

4.1.1 Proof of proposition 4.1

Our method will be to prove that b_{ii} divides $b_{i(i-j)}$ by first inducting on i and then on j; the case $j = 1$ for arbitrary i has already been handled. Let $j > 1$ and assume the divisibility condition for all smaller j and arbitrary i. Now the base case for inducting on i is $i = j+1$ and to prove the divisibility here we consider the $(j, 1)$ entry of BAB^{-1}, which is

$$\sum_{k=1}^{j} b_{jk} b'_{(k+1)1}. \tag{4.22}$$

Since

$$b'_{(k+1)1} = (-1)^k \left(\prod_{l=1}^{k+1} b_{ll} \right)^{-1} \det(b_{rs})_{\substack{2 \le r \le k+1 \\ 1 \le s \le \min\{k,r\}}}, \tag{4.23}$$

we can apply the induction hypothesis for the b_{rs} to see that for $k < j$, $b'_{(k+1)1}$ is a fraction with denominator b_{11}. And for $k = j$, we can perform a co-factor expansion along the bottom row, noting that for $s > 1$ we can apply the induction hypothesis to see that

$$b'_{(j+1)1} = \frac{\text{integer}}{b_{11}} \pm \frac{b_{(j+1)1}}{b_{11} b_{(j+1)(j+1)}}. \tag{4.24}$$

Putting these facts into (4.22) and applying both the inductive hypothesis for $k < j$ to write $b_{jk} = c_{jk} b_{jj}$ and also the previously noted $b_{11} = b_{jj} \prod_{l=1}^{j-1} m_l$, to see that the $(j, 1)$ entry of BAB^{-1} has the form

$$\frac{\text{integer}}{\prod_{l=1}^{j-1} m_l} \pm \frac{b_{(j+1)1}}{b_{(j+1)(j+1)} \prod_{l=1}^{j-1} m_l}. \tag{4.25}$$

From this it is clearly necessary for $b_{(j+1)(j+1)}$ to divide $b_{(j+1)1}$ – proving the base case for this induction.

The general case for the induction on $i > j + 1$ follows similarly. We consider now the $(i - 1, i - j)$ entry of BAB^{-1}, which is

$$\sum_{k=i-j-1}^{i-1} b_{(i-1)k} b'_{(k+1)(i-j)}. \tag{4.26}$$

Here we have

$$b'_{(k+1)(i-j)} = (-1)^{i+j+k+1} \left(\prod_{l=i-j}^{k+1} b_{ll} \right)^{-1} \det(b_{rs})_{\substack{i-j+1 \le r \le k+1 \\ i-j \le s \le \min\{r,k\}}} , \tag{4.27}$$

except for the first term, $k = i - j - 1$, where clearly $b'_{(i-j)(i-j)} = b^{-1}_{(i-j)(i-j)}$. For all except the last term, $k = i - 1$, we can apply the j inductive hypothesis to see that each $b'_{(k+1)(i-j)}$ is a fraction with denominator $b_{(i-j)(i-j)}$. Now, applying the j hypothesis for all $i - j - 1 < k < i - 1$ and the i inductive hypothesis for $k = i - j - 1$, we see that all terms in (4.26) are fractions with denominator $\prod_{l=i-j}^{i-2} m_l$. For the last term, we perform a co-factor expansion along the bottom row, $r = k + 1 = i$ of the matrix in (4.27), applying the j inductive hypothesis to the entries with $s > i - j$ to see, as we did in the base case, that the $(i - 1, i - j)$ entry of BAB^{-1} has the form

$$\frac{\text{integer}}{\prod_{l=i-j}^{i-2} m_l} \pm \frac{b_{i(i-j)}}{b_{ii} \prod_{l=i-j}^{i-2} m_l}. \tag{4.28}$$

This shows that it is necessary that b_{ii} divides $b_{i(i-j)}$, thus completing the induction.

We finish the proof of the proposition by simply noting that because of the necessary divisibility conditions, if I is an ideal, then the basis given by B can be made diagonal by multiplying on the right by a (lower triangular) matrix in $SL_n(\mathbb{Z})$. Thus the remaining diagonal entries, the m_j, are the invariant factors of the quotient $\mathbb{Z}[\alpha]/I$.

4.1.2 Proof of theorem 4.2

To simplify the calculations, we will from now on assume that the quotient of $\mathbb{Z}[\alpha]$ by I is cyclic. If I is an ideal this assumption via proposition 4.1 implies that $m_j = 1$ except for $j = 1$. Let us set $m_1 = m$. By our assumption that B is in Hermite normal form, we also see that the cyclicity implies that all the off-diagonal entries in B are 0 outside of the first column.

With this simplifying assumption, we observe first that

$$b'_{ij} = \begin{cases} \frac{1}{m} & \text{if } i = j = 1 \\ -\frac{b_{i1}}{m} & \text{if } i > 1, j = 1 \\ 1 & \text{if } i = j > 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4.29}$$

And now, for $1 \leq i \leq d-1$, the (i,j) entry of BAB^{-1} will be

$$\begin{cases} -b_{21} & \text{if } i = j = 1 \\ -\frac{1}{m}(b_{i1}b_{21} + b_{(i+1)1}) & \text{if } i > 1, j = 1 \\ b_{i1} & \text{if } j = 2 \\ 1 & \text{if } i = j-1 > 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4.30}$$

Only the second case of the above gives an integrality condition, which, setting $b_{21} = -\mu$, is satisfied if and only if

$$b_{i1} \equiv -\mu^{i-1} \pmod{m}. \tag{4.31}$$

for all $2 \leq i \leq n$.

It remains to see the integrality conditions arising from the bottom row of BAB^{-1}. A relatively quick calculation shows that the (d,j) entry of BAB^{-1} is

$$\begin{cases} -\frac{1}{m}\left(b_{d1}b_{21} - a_d + \sum_{l=2}^{d} a_{d-l+1}b_{l1}\right) & \text{if j = 1} \\ b_{d1} + a_{d-1} & \text{if } j = 2 \\ a_{d-j+1} & \text{if } j > 2, \end{cases} \tag{4.32}$$

where we recall that $\alpha^d = a_1\alpha^{d-1} + \cdots + a_d$. Substituting (4.31) into (4.32), we see that the integrality conditions are satisfied if and only if μ is a root of the polynomial congruence

$$\mu^d \equiv a_1\mu^{d-1} + \cdots + a_d \pmod{m}. \tag{4.33}$$

This concludes the proof of theorem 4.2.

4.2 Parametrization of the roots

In this section we prove theorem 4.3. As mentioned in the statement of that theorem, we will only consider the roots $\mu \pmod{m}$ that correspond, via theorem 4.2, to invertible ideals. It would be nice to have a more concrete characterization of the $\mu \pmod{m}$ in terms of m and μ themselves, but we unfortunately have not yet been able to do this.

Let us recall the notation from the set up for theorem 4.3. Fix a complete system of representative ideals I_l, for the narrow class group of $\mathbb{Z}[\alpha]$, and that $\{\beta_{kl} : k = 1, \ldots, d\}$ is a basis for I_l. Furthermore, let us denote by $\{\overline{\beta}_{kl} : k = 1, \ldots, d\}$ a basis for the inverse, I_l^{-1}, and we assume that the order of the β_{kl} and $\overline{\beta}_{kl}$ is so that

$$\operatorname{sign}\det(\beta_i^{(j)})_{\substack{1\le i\le d\\ 1\le j\le d}} = \operatorname{sign}\det(\overline{\beta}_i^{(j)})_{\substack{1\le i\le d\\ 1\le j\le d}} = \operatorname{sign}\det((\alpha^{(j)})^{i-1})_{\substack{1\le i\le d\\ 1\le j\le d}}. \tag{4.34}$$

And since $\overline{\beta}_{il}\beta_{jl} \in \mathbb{Z}[\alpha]$, there are integers b_{ijkl} so that

$$\beta_{il}\overline{\beta}_{jl} = \sum_{k=1}^{d} b_{ijkl}\alpha^{k-1}. \tag{4.35}$$

Now every every ideal $I \subset \mathbb{Z}[\alpha]$ is equivalent to exactly one of the I_l, so there is also a $\gamma \in I_l^{-1}$ with positive norm, unique up to multiplication by the positive-norm units in $\mathbb{Z}[\alpha]$, such that $I = \gamma I_l$. As in the previous chapter, this γ gives rise to a basis, $\{\gamma\beta_{il} : i = 1, \ldots, d\}$ for I. Now, since

$$\gamma = \sum_{i=1}^{d} c_i\overline{\beta}_{il}, \tag{4.36}$$

we observe that the jth basis vector of I can be expressed as

$$\gamma\beta_{jl} = \sum_{i=1}^{d}\sum_{k=1}^{d} c_i b_{ijkl}\alpha^{k-1}. \tag{4.37}$$

This shows that

$$\begin{pmatrix} \gamma\beta_{1l} \\ \gamma\beta_{2l} \\ \vdots \\ \gamma\beta_{dl} \end{pmatrix} = C \begin{pmatrix} 1 \\ \alpha \\ \vdots \\ \alpha^{d-1} \end{pmatrix} \tag{4.38}$$

where the (j, k) entry of C is

$$c_{ij} = \sum_{i=1}^{d} c_i b_{ijkl}. \tag{4.39}$$

We remark that C is a linear combination of d fixed matrices,

$$C = \sum_{i=1}^{d} c_i B_i, \tag{4.40}$$

and each collection of the fixed matrices B_i a narrow ideal class of $\mathbb{Z}[\alpha]$.

Since they come from two different bases of the same ideal I, the matrix B, which contains the root $\mu \pmod{m}$ of the polynomial congruence, and the matrix C must be related by a matrix in $GL_d(\mathbb{Z})$, in fact a matrix in $SL_d(\mathbb{Z})$ because of the sign conditions $m > 0$ and (4.34). This is to say that $\overline{C}B = C$ for some $\overline{C} \in SL_n(\mathbb{Z})$. Here the shape of $\overline{C}$ is easy to determine since B only differs from the identity in the first column; we see that $\overline{C}$ agrees with C on all but the first column. Moreover, this first column of $\overline{C}$, the entries of which we denote by u_j, is determined up to multiplication on the right by matrices of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & 1 \end{pmatrix} \tag{4.41}$$

by the requirement that $\det \overline{C} = 1$.

We remark first that the existence of integers making this determinant condition hold is equivalent to the determinants of the d sub-matrices, which we denote by C_j, of C formed by the last $d-1$ columns being coprime. Taken with the correct sign, these determinants also form the first row of $\overline{C}^{-1}$ of $(\det C)C^{-1}$, and we denote them by c_j', $j = 1, \ldots, d$. The second remark is, by recalling the equation $\overline{C}B = C$, the ambiguity in the first column of $\overline{C}$ is actually the same ambiguity that arises by considering the μ^j in the first column of B as residue classes modulo m – different choices for representatives of these residues corresponds exactly to the different choices of the first column of $\overline{C}$. This is enough to prove theorem 4.3.

4.3 Approximations to the roots

In the statement of theorem 4.4, we make a statement that for some i, referring to a row to be removed from the matrix C, we obtain a good approximation to the point

$\left(\frac{\mu}{m}, \ldots, \frac{\mu^{d-1}}{m}\right)$. For sake of exposition, however, we will show below only the manipulations necessary if $i = 1$ was the case. Since unlike in chapters 2 and 3, the author does not know how to pick a fundamental domain so that $i = 1$, or some other fixed value, always gives the approximation, we instead at the end of this section see why some i will give an approximation, and leave it to the reader to imagine the necessary calculations for general $i \neq 1$.

We start by rearranging the first column of the matrix equation $\overline{C}B = C$ to see that

$$\begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix} \begin{pmatrix} \frac{\mu}{m} \\ \vdots \\ \frac{\mu^{d-1}}{m} \end{pmatrix} = \begin{pmatrix} u_2 \\ \vdots \\ u_d \end{pmatrix} - \frac{1}{m} \begin{pmatrix} c_{21} \\ \vdots \\ c_{d1} \end{pmatrix}. \tag{4.42}$$

Now, supposing we can pick γ in a fundamental domain in such a way that $c_{ij} \ll m^{1/d}$, and we shall do so below, we can interpret these equations as the vector $\left(\frac{\mu}{m}, \ldots, \frac{\mu^{d-1}}{m}\right) \in \mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$ being close to the $d-1$ planes $c_{j2}X_1 + \cdots + c_{jd}X_{d-1} = u_j$, $2 \leq j \leq d$. Under the condition that

$$\det(c_{ij})_{\substack{2 \leq i \leq d \\ 2 \leq j \leq d}} \gg m^{1-1/d} \tag{4.43}$$

we will see that the vector in fact lies close to the intersection of these $d-1$ planes, this is the content of theorem 4.4.

From (4.42) all that needs to be proved is that

$$\begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix}^{-1} \begin{pmatrix} c_{21} \\ \vdots \\ c_{d1} \end{pmatrix} \ll 1. \tag{4.44}$$

To see this, we rearrange the first column of the equation $CC^{-1} = I$ to obtain

$$\begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix} \begin{pmatrix} c'_{21} \\ \vdots \\ c'_{d1} \end{pmatrix} = -\frac{c'_1}{\det C} \begin{pmatrix} c_{21} \\ \vdots \\ c_{d1} \end{pmatrix}, \tag{4.45}$$

where the c'_{ij} are the entries of C^{-1}, and we recall that $c_{1j} = \frac{c'_j}{\det C}$ are the entries of the first row. Now, (4.45) shows that

$$\begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix}^{-1} \begin{pmatrix} c_{21} \\ \vdots \\ c_{d1} \end{pmatrix} = -\frac{\det C}{c'_1} \begin{pmatrix} c'_{21} \\ \vdots \\ c'_{d1} \end{pmatrix}, \tag{4.46}$$

so to prove (4.44) it will suffices to have $c'_1 \gg m^{1-1/d}$ and $\det C c'_{j1} \ll m^{1-1/d}$.

Since $\det C c'_{j1}$ is a polynomial of degree $d-1$ in the c_i defining γ, the second required bound follows from $c_i \ll m^{1/d}$, which we show below. The author however does not know how to ensure the first bound, so instead we will show that at least one of the c'_i will be $\gg m^{1-1/d}$. Then redoing the above with the ith instead of the first row removed will give the desired approximation.

We obtain the desired bounds for the c_i and the c'_i by first noticing that the c_i are fixed linear combinations of the embeddings of γ, and the c'_i are fixed linear combinations of the embeddings of $N(\gamma)\gamma^{-1}$. This last fact follows by rewriting (4.38) as

$$C^{-1} \begin{pmatrix} \beta_{1l} \\ \vdots \\ \beta_{dl} \end{pmatrix} = \begin{pmatrix} \gamma^{-1} \\ \vdots \\ \gamma^{-1}\alpha^{d-1} \end{pmatrix}, \tag{4.47}$$

so from the first row we see that

$$\gamma^{-1} = \frac{1}{\det C} \sum_{j=1}^{d} c'_j \beta_{jl}, \tag{4.48}$$

and we note that also from (4.38), $\det C = N(I_l)N(\gamma)$.

Now we can easily pick γ to be in a fundamental domain for the action of the positive norm units in such a way that the embeddings of γ are all $\asymp m^{1/d}$, and whence the embeddings of $N(\gamma)\gamma^{-1}$ are all $\asymp m^{1-1/d}$, recalling that $N(\gamma) = m/N(I_l)$, I_l being fixed. Once this is done the bounds for the c_i and one of the c'_j follows from the observations of the previous paragraph.

This fundamental domain can be seen from Dirichlet's unit theorem. In the logarithmic embedding of $\mathbb{Z}[\alpha]$ into $\mathbb{R}^d$, the units cut out a rank $d-1$ lattice on the plane orthogonal to the vector $(1, 1, \ldots, 1)$, and the positive norm units of course cut out a finite index sub-lattice. Now a fundamental domain for the action of the units can be taken to be the region that projects, parallel to $(1, 1, \ldots, 1)$, onto a fundamental parallelopiped of this lattice. Then it is clear that the γ for which this logarithmic embedding lies in this region, the logarithmic embedding differs from a constant vector by at most $O(1)$. Exponentiating this estimate gives the desired bounds for the embeddings of γ.

4.4 Spacing between the roots

Definition 4.8. *For a positive integer q, a q-*torsion point *in $\mathbb{R}^{d-1}$ is a vector $\boldsymbol{x}$ such that $q\boldsymbol{x} \in \mathbb{Z}^{d-1}$ but $q'\boldsymbol{x} \notin \mathbb{Z}^{d-1}$ for any $q' < q$.*

In light of this definition, every torsion point has the form

$$\boldsymbol{x} = \left(\frac{r_1}{q}, \ldots, \frac{r_{d-1}}{q} \right) \tag{4.49}$$

where r_j are integers such that $\gcd(q, r_1, \ldots, r_{d-1}) = 1$. Accordingly, we can identify a torsion points in $\mathbb{R}^{d-1}$ with vectors $\boldsymbol{r} = (q, r_1, \ldots, r_{d-1}) \in \mathbb{Z}^d$ having coprime coordinates.

Now, given two torsion points $\boldsymbol{x}$ and $\boldsymbol{x}'$, we consider the Plücker coordinates of the line containing both. These coordinates are the $\binom{d}{2}$ quantities formed by taking 2×2 determinants from the matrix

$$\begin{pmatrix} q & r_1 & \cdots & r_{d-1} \\ q' & r_1' & \cdots & r_{d-1}' \end{pmatrix}. \tag{4.50}$$

The $d-1$ of these that include the first column, say

$$s_{1j} = \det \begin{pmatrix} q & r_j \\ q' & r_j' \end{pmatrix}, \tag{4.51}$$

control the distance between $\boldsymbol{x}$ and $\boldsymbol{x}'$. Indeed,

$$||\boldsymbol{x} - \boldsymbol{x}'|| = \frac{1}{qq'} \left(s_{11}^2 + \cdots + s_{1d-1}^2 \right)^{1/2}. \tag{4.52}$$

Fixing a torsion point $\boldsymbol{x}$, we can lower bound the distance between $\boldsymbol{x}$ and any other torsion point by considering the lattice of lines with integral valued Plücker coordinates pass through $\boldsymbol{x}$ – or rather the projection of this lattice onto the $d-1$ dimensional space spanned by the coordinates having the form of s_{1j} in (4.51). We remark that the

Plücker coordinates of two lines cannot in general be added to obtain the coordinates of a third line because the eligible coordinates are subject to a system of quadratic equations, the Plücker relations. However the coordinates corresponding to the lines passing through a fixed point does form a linear subspace. These observations are the content of proposition 4.5.

For our purposes, the torsion point will be the point on the right side of (4.12), assuming as we did in the previous section that the i giving the approximation to $\left(\frac{\mu}{m}, \dots, \frac{\mu^{d-1}}{m}\right)$ is $i = 1$, namely

$$\begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix}^{-1} \begin{pmatrix} u_2 \\ \vdots \\ u_d \end{pmatrix}, \tag{4.53}$$

which has torsion c_1'. This gives our torsion point as the intersection of the $d-1$ planes

$$c_{j2}X_1 + \cdots + c_{jd}X_{d-1} = u_j, \tag{4.54}$$

$2 \leq j \leq d$, and so it is more natural to consider the lines passing through our point dually: as the intersection of sets of $d-2$ hyper-planes containing the point. The Plücker coordinates of the line can be determined naturally from this dual perspective as well. For example, the Plücker coordinates of one line are given, at least up to order and sign, by the determinants from the $d-2 \times d-2$ minors of the matrix

$$\begin{pmatrix} -u_3 & c_{32} & \cdots & c_{3d} \\ \vdots & \vdots & \ddots & \vdots \\ -u_d & c_{d2} & \cdots & c_{dd} \end{pmatrix}. \tag{4.55}$$

Moreover, this, together with the other $d-2$ lines obtained by removing rows other than the row corresponding to the $j = 2$ plane in (4.54), form a basis for the lattice of lines containing the torsion point.

We observe that the $d-1$ determinants of the minors formed by removing the first column along with another are exactly the s_{1j} coordinates, see (4.51), used to determine the spacing. We note that this basis does not depend on the choice of u_j, and so does not depend on the choice of representative of the torsion point modulo $\mathbb{Z}^{d-1}$. Hence a basis of the lattice used in proposition 4.5 to lower bound the spacing is given by the columns of

$$c_1' \begin{pmatrix} c_{22} & \cdots & c_{2d} \\ \vdots & \ddots & \vdots \\ c_{d2} & \cdots & c_{dd} \end{pmatrix}^{-1}, \tag{4.56}$$

which we denote by $\boldsymbol{c}_{1j}$, $1 \leq j \leq d-1$.

We prove the following lemma for the case $i = 1$ giving the approximation of theorem 4.4, and we leave it to the reader to imagine the similar proofs for general i.

Lemma 4.9. *If $c_1' \gg m^{1-1/n}$, then the matrix (4.56) normalized by $c_1'^{(d-2)/(d-1)}$ to have determinant 1 lies in a fixed compact set in $SL_{d-1}(\mathbb{R})$, and hence the smallest vector in the lattice with basis $\{\boldsymbol{c}_{11}, \ldots, \boldsymbol{c}_{1(d-2)}\}$ has size $\gg c_1'^{(d-2)/(d-1)} \gg m^{1-2/d}$.*

Proof. It is easy to see that the determinant of the matrix (4.56) is $c_1'^{d-2}$, and so by Hadamard's inequality

$$c_1'^{d-2} \leq ||\boldsymbol{c}_{11}|| \cdots ||\boldsymbol{c}_{1(d-1)}||. \tag{4.57}$$

On the other hand, since each of the $\boldsymbol{c}_{1j}$ have coordinates polynomials of degree $d-2$ in the c_i, which we recall are $\ll m^{1/d}$, we have $||\boldsymbol{c}_{1j}|| \ll m^{(d-2)/d}$. Replacing all but one of the $||\boldsymbol{c}_{1j}||$ in (4.57) by this bound, we have that, under the hypothesis $c_1' \gg m^{1-1/d}$,

$$m^{(d-1)(d-2)/d} \ll ||\boldsymbol{c}_{1j}|| m^{(d-2)^2/d} \ll m^{(d-1)(d-2)/d}, \tag{4.58}$$

so $||\boldsymbol{c}_{1j}|| \asymp m^{1-2/d} \asymp c_1'^{(d-2)/(d-1)}$. These estimates are enough to show that upon normalizing so that the determinant is 1, the resulting matrix is in a compact subset of $SL_{d-1}(\mathbb{R})$.

Clearly then the (normalized) lattice with this basis lies in a compact subset of $SL_{d-1}(\mathbb{Z}) \backslash SL_{d-1}(\mathbb{R})$, and so does not approach the cusp in any direction. This shows that the smallest vector in the lattice has size $\gg$ the $(d-1)$th root of the determinant, finishing the proof of the lemma. □

Having this lemma, the proof of theorem 4.6 is almost finished. Indeed, for each of the points $\left(\frac{\mu}{m}, \ldots, \frac{\mu^{d-1}}{m}\right)$ contained in a ball of radius $\frac{1}{M}$, all the approximations given by theorem 4.4 will be contained in a ball of radius $O(\frac{1}{M})$. However, by lemma 4.9 and proposition 4.5, each of these approximations will be spaced by at least

$$\gg \frac{1}{M^{2(d-1)/d}} M^{(d-2)/d} = \frac{1}{M} \tag{4.59}$$

from any other distinct torsion point, so there can be at most $\ll 1$ of these approximations in this ball. The theorem is then proved if we can show that at most $\ll 1$ of the points $\left(\frac{\mu}{m}, \dots, \frac{\mu^{d-1}}{m}\right)$ can correspond to a given one of the approximations.

We start by noting that a torsion point in $\mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$ determines the lattice of integral lines containing it, and because $SL_{d-1}(\mathbb{Z})$ acts discontinuously on $SL_{d-1}(\mathbb{R})$, the number of bases of the lattice lying in the compact set of lemma 4.9 will be bounded by a constant that depends only on the compact set. For each one of these bases, there are d candidates, one for each of the possible i giving the approximation of theorem 4.4. Further, we need to test each of the narrow ideal classes, but once this is done, we claim that the basis of the lattice determines the c_i and whence the m and $\mu \pmod{m}$; this would show that the number is indeed bounded by a constant depending only on the congruence.

To see this final step, we recall that the matrix C is a linear combination of matrices B_i depending on the ideal class, and the coefficients are exactly the c_i. So, for the example we have been working with, the question of recovering the c_i from the matrix (4.56) is a question about the linear independence of the corresponding $(d-1) \times (d-1)$ sub-matrices of the B_i. Denoting these sub-matrices by B_{1i}, suppose there were numbers c_i so that

$$\sum_{i=1}^{d} c_i B_{1i} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}. \tag{4.60}$$

From the definition of the B_i, see the statement of theorem 4.3, this means that the corresponding $\gamma = \sum c_i \overline{\beta}_i$ satisfies

$$\gamma\beta_j \in \mathbb{Z}, \quad 2 \leq j \leq d. \tag{4.61}$$

However, for a fixed γ, the set of points $\beta \in \mathbb{Z}[\alpha]$ for which $\gamma\beta \in \mathbb{Z}$ forms a line. On the other hand, since the β_j span a full rank lattice, at most one of them can lie on such a line. This is a clearly a contradiction when $d > 2$, the setting which we are considering here.

Chapter 5

Future directions

This chapter diverges slightly from the previous chapters in that we do not reach a desired goal, but rather present a few hopefully interesting observations and calculations that have unfortunately not lead the author to any concrete results. In section 5.1 we describe a connection between the roots of the cubic congruence $\mu^3 \equiv 2 \pmod{m}$ and the binary cubic form $X^3 - 2Y^3$. We begin with the following proposition:

Proposition 5.1. *Pairs of* $\alpha = a + b2^{1/3} + c2^{2/3}$, $\alpha' = A + B2^{1/3} + C2^{2/3} \in \mathbb{Z}[2^{1/3}]$ *such that* $\alpha\alpha' \in \mathbb{Z}$ *are naturally parametrized by integral matrices*

$$\begin{pmatrix} b & c \\ B & C \end{pmatrix}, \tag{5.1}$$

with non-zero determinant, typically not ± 1*, such that the image of* $X^3 - 2Y^3$ *under the action of the matrix is an integral binary cubic form.*

Moreover, such matrices corresponding to α *and* α' *not divisible by any rational integers are decomposed uniquely as*

$$\begin{pmatrix} b & c \\ B & C \end{pmatrix} = \begin{pmatrix} dg^2 & kg \\ 0 & g \end{pmatrix} \begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix} \tag{5.2}$$

where

$$\begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix} \in \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \backslash SL_2(\mathbb{Z}), \tag{5.3}$$

$\gcd(dg, k) = 1$ *and* $B'^3 - 2C'^3 \equiv 0 \pmod{d}$.

We recall in this proposition that the matrix (5.1) acts on a binary cubic form $F(X, Y)$ as

$$F(X, Y) \mapsto \frac{1}{bC - cB} F\left(bX + BY, cX + CY\right), \tag{5.4}$$

so the integrality condition in proposition 5.1 is non-trivial. Of course the reason why we are interested in the α and α' as in the proposition is exactly because they are used to parametrize the roots of the cubic congruence in theorem 3.4. A deficiency in that was noted in chapter three was that the cosets used in parametrizing the roots were not arbitrary because their Plücker coordinates satisfied an additional constraint, see (3.15). A consequence of this is that in the approximation $-\frac{W}{C}$ to the root $\frac{\mu}{m}$ given by theorem 3.7, W and C are tied together in a complicated way. The following proposition is an attempt to explicate this connection.

Proposition 5.2. *Suppose $\alpha = a+b2^{1/3}+c2^{2/3}$ and $\alpha' = A+B2^{1/3}+C2^{1/3}$ correspond to the roots $\mu \pmod m$ and $\nu \pmod n$ as in theorem 3.4, and that b, c, B, and C correspond to α and α' as in proposition 5.1. With the decomposition (5.2), we have*

$$\begin{aligned} m &= -\frac{1}{bC-bC}\left(B^3-2C^3\right) = -\frac{1}{d}\left(B'^3-2C'^3\right) \\ n &= \frac{1}{bC-cB}\left(b^3-2c^3\right) = \frac{1}{d}\left((dgb'+kB)^3-2(dgc'+kC')^3\right) \\ A &= \frac{1}{d}\left((b'B'^2-2c'C'^2)dg+(B'^3-2C'^3)k\right) \\ W &\equiv dgc'^2 \equiv \overline{m}B \pmod{C}, \end{aligned} \tag{5.5}$$

where $-\frac{W}{C}$ is the approximation to $\frac{\mu}{m}$ and $\overline{m}$ is the multiplicative inverse of m modulo C'.

We remark that in the last line of (5.5) we see that the approximation to $\frac{\mu}{m}$ can also be written as $-\frac{\overline{m}B}{C}$. That this approximates $\frac{\mu}{m}$ can in fact be seen directly from the above propositions in a way quite different than how we proved theorem 3.7 in section 3.3. Indeed, from the first line of (5.5) we have

$$-dm = B'^3-2C'^3 \equiv 0 \pmod{m}, \tag{5.6}$$

so because $\gcd(B',C')=1$ implies $\gcd(m,C')=1$ we have

$$B'\overline{C'} \equiv \mu \pmod{m} \tag{5.7}$$

where

$$m\overline{m}+C'\overline{C'}=1. \tag{5.8}$$

Rearranging the above, we have

$$\frac{\mu}{m} \equiv \frac{B'\overline{C'}}{m} = -\frac{\overline{m}B'}{C'} - \frac{B'}{C'm} \pmod{1}, \tag{5.9}$$

which gives the approximation when $B' \ll C'$.

In section 5.2 we count ideals in $\mathbb{Z}[2^{1/3}]$ according to the corresponding m and n in theorem 3.1. Although it is a bit of a digression we first prove the following proposition, which computes the co-type zeta function for $\mathbb{Z}[2^{1/3}]$:

Proposition 5.3. *Let*

$$\zeta_{\mathbb{Z}[2^{1/3}]}(s_1, s_2, s_3) = \sum_{0 \neq I \subset \mathbb{Z}[2^{1/3}]} d_1(I)^{-s_1} d_2(I)^{-s_2} d_3(I)^{-s_3}, \tag{5.10}$$

where $d_1(I)$, $d_2(I)$, and $d_3(I)$ are the invariant factors of $\mathbb{Z}[2^{1/3}]/I$, ordered so that $d_3(I) \mid d_2(I) \mid d_1(I)$. We have

$$\begin{aligned} &\zeta_{\mathbb{Z}[2^{1/3}]}(s_1, s_2, s_3) = \\ &\quad = (1 + 2^{-s_1} + 2^{-s_1-s_2})(1 + 3^{-s_1} + 3^{-s_1-s_2})\zeta(s_1 + s_2 + s_3) \\ &\qquad \times \prod_{p \in \mathcal{P}_1} \left(\frac{1 + 2p^{-s_1} + 2p^{-s_1-s_2} + p^{-2s_1-s_2}}{(1 - p^{-s_1})(1 - p^{-s_1-s_2})} \right) \prod_{p \in \mathcal{P}_2} \left(\frac{1 - p^{-2s_1-s_2}}{(1 - p^{-s_1})(1 - p^{-s_1-s_2})} \right), \end{aligned} \tag{5.11}$$

where $\zeta(s)$ is the Riemann zeta function.

After proving this proposition in section 5.2.1, we move on in section 5.2.2 to prove the following:

Proposition 5.4. *Let ϕ be a fixed smooth, compactly supported function on $(0, \infty)$. Then for a real number $N \geq 1$ and a root $\mu \pmod{m}$ of $\mu^3 \equiv 2 \pmod{m}$ with $\gcd(m, 6) = 1$, we have*

$$\sum_{\substack{0 \neq I \subset \mathbb{Z}[2^{1/3}] \\ d \nmid I, \forall d \in \mathbb{Z} \\ m(I) = m, \mu(I) = \mu}} \phi\left(\frac{n(I)}{N}\right) = K g(m) N \int_0^\infty \phi(x) dx + O(\tau(m) N^{1/2+\epsilon}) \tag{5.12}$$

where for and ideal I we denote by $m(I)$, $\mu(I)$, and $n(I)$ the corresponding m, μ, and n given by theorem 3.1, $g(m)$ is an arithmetic function given by

$$g(m) = \prod_{\substack{p \mid m \\ p \in \mathcal{P}_1}} \left(1 - \frac{1}{p+2}\right) \prod_{\substack{p \mid m \\ p \in \mathcal{P}_2}} \left(1 - \frac{1}{p}\right), \tag{5.13}$$

and K is a constant,

$$\begin{aligned} K &= \frac{2\pi}{9\sqrt{3}} \log(1 + 2^{1/3} + 2^{2/3}) \\ &\quad \times \prod_{p\in\mathcal{P}_1} \left(1 - \frac{3}{p^2} + \frac{2}{p^3}\right) \prod_{p\in\mathcal{P}_2} \left(1 - \frac{1}{p^2}\right) \prod_{p\in\mathcal{P}_3} \left(1 - \frac{1}{p^3}\right) \\ &\approx 0.507396. \end{aligned} \tag{5.14}$$

Roughly speaking proposition 5.4 says that a given root $\mu \pmod{m}$ shows up via theorem 3.1 in roughly N ideals with norm mn^2, with $n \asymp N$. This allows us to see the comments following theorem 3.7 in a different light. Since each approximation $-\frac{W}{C}$ is within $O\left(\frac{1}{m}\right)$, and by proposition 3.6 $C \asymp m^{2/3}N^{1/3}$, we see that there can be at most

$$\ll \frac{1}{m}\left(m^{2/3}N^{1/3}\right)^2 = m^{1/3}N^{2/3} \tag{5.15}$$

distinct approximations. On the other hand since each root $\mu \pmod{m}$ occurs in roughly N ideals, we have N approximations $-\frac{W}{C}$, not necessarily distinct. In fact, the above shows that when N is large compared to m, the multiplicity of each $-\frac{W}{C}$ is at least $N^{1/3}m^{-1/3}$, and it seems reasonable, although the author has not been able to prove, that this is also roughly an upper bound for the multiplicity when N is larger than m.

Proposition 5.2 gives some insight into this multiplicity issue. Indeed, we see that m and the approximation $-\frac{W}{C}$ are independent of the k in the decomposition (5.2). And, inspecting either the expression for n or the expression for A in (5.5), we see that if $n \asymp N$, $A \ll m^{2/3}N^{1/3}$, then k is restricted to be in an interval of length $m^{-1/3}N^{1/3}$. It is clear that this k gives multiplicity to the approximations $-\frac{W}{C}$, and it seems plausible to the author that it is essentially the only source of multiplicity. We remark that if this was proven, then a 1 dimensional spacing property and large sieve for the roots of cubic congruences would follow easily from (5.15).

We end the chapter and dissertation in section 5.3 with a miscellaneous calculation applying the results of the previous sections in this chapter to transform the Weyl sum of the $\frac{\mu}{m}$. The main ideas are first to sum over ideals instead of the roots, which, from proposition 5.4, weights each root $\mu \pmod{m}$ by about $g(m)N$. Second, the sum over ideals is replaced by a sum over the matrices as in proposition 5.1. Finally, we execute Poisson summation on the variable k in the decomposition (5.2), exploiting the independence of the approximation from k. The result from these transformations is a situation close to how the proof of equidistribution in the quadratic case began. But unfortunately the author does not know how to proceed in the cubic case because the sum is still contaminated by the restriction that d divides $B'^3 - 2C'^3$. We note that we can interpret the quadratic case as only using $d = 1$, a tremendously simplified situation. A counting argument shows that an analogous simplification cannot hold in the cubic setting simply because there are far more roots of cubic congruences than there are numbers represented by a binary cubic form.

5.1 Connection to binary cubic forms

In chapter 3 our point of view had been given α, that is the generator of the ideal I, one can find α', which is then used to construct the matrix γ need to parametrize the roots $\mu \pmod{m}$ and $\nu \pmod{n}$. We recall that $\alpha' = A + B2^{1/3} + C2^{2/3}$ was determined from $\alpha = a + b2^{1/3} + c2^{2/3}$ with $\gcd(a, b, c)$ by the requirements

$$\alpha\alpha' \in \mathbb{Z}_{>0}, \quad \gcd(A, B, C) = 1. \tag{5.16}$$

However it is clear that one could equally well determine α from α' by the exact same requirements; the map taking α to α' is an involution. This suggests looking for a more symmetric way to parametrize the roots of our cubic congruence, and indeed we note that the orthogonality requirements in (3.49) can be written as

$$\begin{pmatrix} B & b \\ C & c \end{pmatrix} \begin{pmatrix} a \\ A \end{pmatrix} = - \begin{pmatrix} 2cC \\ bB \end{pmatrix}, \tag{5.17}$$

so b, c, B, and C together can be used to determine a and A! The big question, however, is the following: how are we to describe the b, c, B, and C for which (5.17) has integer solutions a and A? It turns out that binary cubic forms will be our main tool in this regard.

We start with the following definition from Delone and Faddeev, [DF64].

Definition 5.5. *For a binary cubic form $f(X, Y) = aX^3 + bX^2Y + cXY^2 + dY^3$, the left and right roots of f are, respectively, the roots of $f(X, a)$ and $f(d, -Y)$.*

The significance of this definition for us is contained in the following lemma,

Lemma 5.6. *The α and α' related to an ideal I as in the above are the left and right roots of the binary cubic form*

$$f(X,Y) = nX^3 - 3aX^2Y + 3AXY^2 - mY^3. \tag{5.18}$$

Proof. Given α, the set of points $\beta\mathbb{Z}[2^{1/3}]$ such that $\alpha\beta \in \mathbb{Z}$ forms a line in $\mathbb{Z}[2^{1/3}]$, so picking α' as above, in particular having $\gcd(A,B,C) = 1$, generates this line. Now, suppose

$$\alpha^3 - 3a\alpha^2 + a_1\alpha - mn^2 = 0 \tag{5.19}$$

is the minimal polynomial of α. Note that we are using the fact that $3a$ is the trace of α. Then, from the remark above about the definition of α', we have

$$\alpha^2 - 3a\alpha + a_1 = n\alpha'. \tag{5.20}$$

And similarly, if

$$\alpha'^3 - 3A\alpha'^2 + A_1\alpha' - m^2n = 0 \tag{5.21}$$

is the minimal polynomial of α', then

$$\alpha'^2 - 3A\alpha' + A_1 = m\alpha. \tag{5.22}$$

Note that we have used the fact that the norm of α' is m^2n; this follows from $\alpha\alpha' = mn$ by taking norms.

Putting this together, we have the multiplication table

$$\begin{aligned} \alpha\alpha' &= mn \\ \alpha^2 &= -a_1 + 3a\alpha + n\alpha' \\ \alpha'^2 &= -A_1 + m\alpha + 3A\alpha', \end{aligned} \tag{5.23}$$

which shows that together with 1, α and α' generate a cubic ring. According to the Delone-Faddeev correspondence, [DF64], α and α' are the left and right roots of the binary cubic form (5.18). □

In terms of the binary cubic form (5.18), determining a and A from b, c, B, and C is related to expressing f by its Lagrange resolvent. We have the following lemma from [DF64]:

Lemma 5.7 (Delone-Faddeev, 1964). *With $f(X,Y)$ as above, we have*

$$f(X,Y) = \frac{1}{3\Delta}\left((\xi_1 X + \xi_2 Y)^3 - (\eta_1 X + \eta_2 Y)^3\right), \tag{5.24}$$

where

$$\begin{aligned}
\xi_1 &= \alpha^{(1)} + \omega\alpha^{(2)} + \omega^2\alpha^{(3)} \\
\xi_2 &= \alpha'^{(1)} + \omega\alpha'^{(2)} + \omega^2\alpha'^{(3)} \\
\eta_1 &= \alpha^{(1)} + \omega^2\alpha^{(2)} + \omega\alpha^{(3)} \\
\eta_2 &= \alpha'^{(1)} + \omega^2\alpha'^{(2)} + \omega\alpha'^{(3)} \\
\Delta &= \det\begin{pmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{pmatrix},
\end{aligned} \tag{5.25}$$

$\alpha^{(j)}$, $\alpha'^{(j)}$ are the embeddings of α and α', and $\omega \neq 1$ is a cube root of unity.

Since a quick calculation shows that a choice of order for the embeddings and a choice of cube root of unity gives

$$\begin{aligned}
\xi_1 &= 3\sqrt[3]{2}b \\
\xi_2 &= 3\sqrt[3]{2}B \\
\eta_1 &= 3\sqrt[3]{2}c \\
\eta_2 &= 3\sqrt[3]{2}C,
\end{aligned} \tag{5.26}$$

we see from (5.24) we see that given a fixed binary cubic form, for us

$$f_0(X,Y) = X^3 - 2Y^3 \tag{5.27}$$

will be particularly useful, the action of a matrix

$$\beta = \begin{pmatrix} b & c \\ B & C \end{pmatrix} \tag{5.28}$$

is given by

$$\begin{aligned}
f_\beta(X,Y) &= \frac{1}{bC - cB} f_0(bX + BY, cX + CY) \\
&= \frac{1}{bC - cB}((b^3 - 2c^3)X^3 + 3(b^2B - 2c^2C)X^2Y \\
&\quad + 3(bB^2 - 2cC^2)XY^2 + (B^3 - 2C^3)Y^3).
\end{aligned} \tag{5.29}$$

Comparing (5.18) with (5.24) and applying (5.29), we now see that a pair α and α' used in the parametrization of the roots of our cubic congruence give rise to a matrix (5.28) that maps the binary cubic form f_0 into an integral binary cubic form. This is summarized in proposition 5.1.

To prove proposition 5.2, we make a few more observations regarding the matrices (5.28). To start, we let

$$\gamma = \begin{pmatrix} A & B & C \\ U & V & W \\ X & Y & Z \end{pmatrix} \tag{5.30}$$

and

$$\gamma^{-1} = \begin{pmatrix} u & x & c \\ v & y & b \\ w & z & a \end{pmatrix}, \tag{5.31}$$

where $\gamma \in SL_3(\mathbb{Z})$ is the matrix used in the parametrization of the roots of the congruence, theorem 3.4.

Setting $g = \gcd(B, C)$, we observe by taking determinants of minors that $g \mid c$ and $g \mid x$. Moreover, by multiplying by γ^{-1} in (4.9) and examining the $(1, 2)$ entry, we see that $nx \equiv b \pmod{c}$ – this shows that $g \mid b$ and hence $g \mid \gcd(b, c)$. Now again by taking minors we see that $\gcd(b, c) \mid C$ and also $\gcd(b, c) \mid W$. And rearranging (4.9) by taking inverses as

$$\begin{pmatrix} n & 0 & 0 \\ \mu n & mn & 0 \\ * & -\nu mn & mn^2 \end{pmatrix} \gamma = \begin{pmatrix} a^2 - 2bc & 2c^2 - ab & b^2 - ac \\ 2(b^2 - ac) & a^2 - 2bc & 2c^2 - ab \\ 2(2c^2 - ab) & 2(b^2 - ac) & a^2 - 2bc \end{pmatrix}, \tag{5.32}$$

we use (3.53) with the fact that there $l = n$ to see that $mW \equiv B \pmod{C}$, from which $\gcd(b, c) \mid B$ and hence $g = \gcd(b, c)$ as well.

Equating coefficients in (5.18) and (5.29), we have

$$g^3 \mid bB^2 - 2cC^2 = A(bC - cB). \tag{5.33}$$

Since $\gcd(A, B, C) = \gcd(A, g) = 1$ we must have $bC - cB = dg^3$ for some integer d.

Accordingly, the matrix (5.28) decomposes as

$$\begin{pmatrix} b & c \\ B & C \end{pmatrix} = \begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix} \begin{pmatrix} dg & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix}, \tag{5.34}$$

where k is an integer, $gB' = B$, $gC' = C$, and b', c' are some choice of integers so that $b'C' - c'B' = 1$, different choices corresponding to translates of the variable k by a multiple of dg. We note that having $\gcd(b, c) = 1$ in this notation is equivalent to $\gcd(dg, k) = 1$. Moreover such a decomposition is unique for a specific choice of b' and c'.

With

$$\beta' = \begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix}, \tag{5.35}$$

we have

$$f_\beta(X, Y) = \frac{1}{d} f_{\beta'}(dgX, kX + Y), \tag{5.36}$$

where

$$f_{\beta'}(X, Y) = (b'X + B'Y)^3 - 2(c'X + C'Y)^3 = n'X^3 - 3a'X^2Y + 3A'XY^2 - m'Y^3, \tag{5.37}$$

say. Explicitly, we have

$$\begin{aligned} n &= \frac{1}{d}(n'd^3g^3 - 3a'd^2g^2 + 3A'dgk^2 - m'k^3) \\ a &= \frac{1}{d}(a'd^2g^2 - 2A'dgk + m'k^2) \\ A &= \frac{1}{d}(A'dg - m'k) \\ m &= \frac{1}{d}m'. \end{aligned} \tag{5.38}$$

From these equations, we see that the only divisibility condition is that $d \mid m'$, or

$$B'^3 - 2C'^3 \equiv 0 \pmod{d}. \tag{5.39}$$

Proposition 5.2 apart from the last line of (5.5) follows from what we have done. To see this last line we recall that $W \pmod{C}$ is defined by the following

$$\begin{aligned} BW &\equiv c \pmod{C} \\ AW &\equiv -b \pmod{C}. \end{aligned} \tag{5.40}$$

The first equation is equivalent to

$$W \equiv -dgc'^2 \pmod{C'}, \tag{5.41}$$

so then putting (5.38) into the second line of (5.40) gives

$$A'g(-dgc'^2) - mk(-dgc'^2 + lC') \equiv -dg^2b' - kgB' \pmod{C}. \tag{5.42}$$

Now

$$mdc'^2 \equiv -B'^3c'^2 \equiv -B' \pmod{C'}, \tag{5.43}$$

so the above simplifies to

$$-A'dg^2c'^2 - mklC' \equiv -dg^2b' \pmod{C}. \tag{5.44}$$

Since also

$$A'c'^2 \equiv b'B'^2c'^2 \equiv b' \pmod{C'}, \tag{5.45}$$

(5.44) becomes

$$-dg^2b' - mklC' \equiv -dg^2b' \pmod{C}, \tag{5.46}$$

so $l \equiv 0$, whence

$$W \equiv -dgc'^2 \pmod{C}. \tag{5.47}$$

Finally, to complete the last line of (5.5), we note that since

$$d \equiv -\overline{m}B'^3 \pmod{C'}, \tag{5.48}$$

where $\overline{m}$ is the inverse of m modulo C', we have

$$W \equiv \overline{m}B \pmod{C}. \tag{5.49}$$

5.2 Ideal and root counts

5.2.1 Co-type zeta function

For an ideal $I \subset \mathbb{Z}[2^{1/3}]$, we let $d_1(I)$, $d_2(I)$, and $d_3(I)$ denote the invariant factors. That is

$$\mathbb{Z}[2^{1/3}]/I \cong \mathbb{Z}/d_1(I)\mathbb{Z} \oplus \mathbb{Z}/d_2(I)\mathbb{Z} \oplus \mathbb{Z}/d_3(I)\mathbb{Z} \tag{5.50}$$

with $d_3(I) \mid d_2(I) \mid d_2(I)$. In the study of subgroup growth, a lot of attention has been paid to certain zeta functions associated to the invariant factors, see [Pet07] and [CKK17]. In our situation, counting ideals in the ring $\mathbb{Z}[2^{1/3}]$, we define the co-type zeta function as

$$\zeta_{\mathbb{Z}[2^{1/3}]}(s_1, s_2, s_3) = \sum_{0\neq I\subset\mathbb{Z}[2^{1/3}]} d_1(I)^{-s_1} d_2(I)^{-s_2} d_3(I)^{-s_3}, \tag{5.51}$$

where the sum is of course over ideals I.

In the language of theorem 3.1 and corollary 3.2, we see that $d_3(I)$ is the largest integer divisor of I, and, applying the theorem to $I/d_3(I)$, we have $m = d_1(I)/d_2(I)$, $n = d_2(I)/d_3(I)$. We have

$$\begin{aligned}\zeta_{\mathbb{Z}[2^{1/3}]}(s_1, s_2, s_3) &= \sum_{0\neq I\subset\mathbb{Z}[2^{1/3}]} m(I)^{-s_1} n(I)^{-s_1-s_2} d_3(I)^{-s_1-s_2-s_3} \\ &= \zeta(s_1+s_2+s_3) \sum_{\substack{0\neq I\subset\mathbb{Z}[2^{1/3}] \\ d\nmid I, \forall d\in\mathbb{Z}}} m(I)^{-s_1} n(I)^{-s_1-s_2},\end{aligned} \tag{5.52}$$

where $\zeta(s)$ is the Riemann zeta function. We apply theorem 3.1 to arrange this sum as

$$\begin{aligned}\zeta_{\mathbb{Z}[2^{1/3}]}(s_1, s_2, s_3) = &\sum_{\gcd(m,6)=1} \sum_{\mu^3\equiv 2(m)} m^{-s_1} \sum_{n\geq 1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} n^{-s_1-s_2} \\ &+ \sum_{\gcd(m,6)=2} \sum_{\mu^3\equiv 2(m)} m^{-s_1} \sum_{\gcd(n,2)=1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} n^{-s_1-s_2} \\ &+ \sum_{\gcd(m,6)=3} \sum_{\mu^3\equiv 2(m)} m^{-s_1} \sum_{\gcd(n,3)=1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} n^{-s_1-s_2} \\ &+ \sum_{\gcd(m,6)=6} \sum_{\mu^3\equiv 2(m)} m^{-s_1} \sum_{\gcd(n,6)=1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} n^{-s_1-s_2} \\ = &\ S_1 + S_2 + S_3 + S_4,\end{aligned} \tag{5.53}$$

say. Starting with S_1, we have

$$\begin{aligned}S_1 = &\sum_{\gcd(m,6)=1} \sum_{\mu^3\equiv 2(m)} m^{-s_1} (1+2^{-s_1-s_2})(1+3^{-s_1-s_2}) \\ &\times \prod_{\substack{p\mid m \\ p\in\mathcal{P}_1}} \left(1 + \frac{2p^{-s_1-s_2}}{1-p^{-s_1-s_2}}\right) \prod_{\substack{p\nmid m \\ p\in\mathcal{P}_1}} \left(1 + \frac{3p^{-s_1-s_2}}{1-p^{-s_1-s_2}}\right) \prod_{\substack{p\nmid m \\ p\in\mathcal{P}_2}} \left(1 + \frac{p^{-s_1-s_2}}{1-p^{-s_1-s_2}}\right),\end{aligned} \tag{5.54}$$

where $\mathcal{P}_1$ is the set of primes in $\mathbb{Z}$ that split completely in $\mathbb{Z}[2^{1/3}]$ and $\mathcal{P}_2$ is the set of those that factor into a degree 1 times a degree 2 prime; neither $\mathcal{P}_1$ nor $\mathcal{P}_2$ contain 2 or 3. Explicitly, we have $\mathcal{P}_2$ is the set of all primes other than 2 that are $\equiv 2 \pmod 3$, and $\mathcal{P}_1$ is the set of all primes that can be represented by the binary quadratic form $X^2 + 27Y^2$.

We arrange this as

$$\begin{aligned} S_1 &= (1+2^{-s_1-s_2})(1+3^{-s_1-s_2}) \prod_{p\in\mathcal{P}_1} \left(1 + \frac{3p^{-s_1-s_2}}{1-p^{-s_1-s_2}}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1}{1-p^{-s_1-s_2}}\right) \\ &\quad \times \sum_{\gcd(m,6)=1} \sum_{\mu^3\equiv 2(m)} m^{-s_1} \prod_{\substack{p\mid m\\ p\in\mathcal{P}_1}} \left(\frac{1+p^{-s_1-s_2}}{1+2p^{-s_1-s_2}}\right) \prod_{\substack{p\mid m\\ p\in\mathcal{P}_2}} (1-p^{-s_1-s_2}) \\ &= (1+2^{-s_1-s_2})(1+3^{-s_1-s_2}) \\ &\quad \times \prod_{p\in\mathcal{P}_1} \left(\frac{1+2p^{-s_1}+2p^{-s_1-s_2}+p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1-p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right). \end{aligned} \tag{5.55}$$

Similar calculations show that

$$\begin{aligned} S_2 &= 2^{-s_1}(1+3^{-s_1-s_2}) \\ &\quad \times \prod_{p\in\mathcal{P}_1} \left(\frac{1+2p^{-s_1}+2p^{-s_1-s_2}+p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1-p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \\ S_3 &= 3^{-s_1}(1+2^{-s_1-s_2}) \\ &\quad \times \prod_{p\in\mathcal{P}_1} \left(\frac{1+2p^{-s_1}+2p^{-s_1-s_2}+p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1-p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \\ S_4 &= 6^{-s_1} \prod_{p\in\mathcal{P}_1} \left(\frac{1+2p^{-s_1}+2p^{-s_1-s_2}+p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1-p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right). \end{aligned} \tag{5.56}$$

Putting these into (5.53), we obtain

$$\begin{aligned} &\zeta_{\mathbb{Z}[2^{1/3}]}(s_1,s_2,s_3) = \\ &\quad = (1+2^{-s_1}+2^{-s_1-s_2})(1+3^{-s_1}+3^{-s_1-s_2})\zeta(s_1+s_2+s_3) \\ &\qquad \times \prod_{p\in\mathcal{P}_1} \left(\frac{1+2p^{-s_1}+2p^{-s_1-s_2}+p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1-p^{-2s_1-s_2}}{(1-p^{-s_1})(1-p^{-s_1-s_2})}\right). \end{aligned} \tag{5.57}$$

5.2.2 Number of n for a given $\mu \pmod{m}$

For a fixed m and $\mu \pmod{m}$, we will want to count the number of n and $\nu \pmod{n}$ to which, together with m and $\mu \pmod{m}$, correspond to an ideal as in theorem 3.1. Let ϕ be a smooth function, compactly supported in the interval $(0,\infty)$, which has Mellin transform

$$\hat{\phi}(s) = \int_0^\infty \phi(x)x^{s-1}dx, \tag{5.58}$$

which by integration by parts is $\ll (1+|s|)^{-j}$ for any $j>0$. The quantity we would like to estimate is the following sum over ideals I,

$$G(m,\mu,N) = \sum_{\substack{0\neq I\subset\mathbb{Z}[2^{1/3}] \\ d\nmid I\forall d\in\mathbb{Z} \\ m(I)=m,\mu(I)=\mu}} \phi\left(\frac{n(I)}{N}\right) \tag{5.59}$$

where N is a positive real number. If $\gcd(m,6)=1$, which we will assume from here on, we have by theorem 3.1

$$G(m,\mu,N) = \sum_{n\geq 1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} \phi\left(\frac{n}{N}\right). \tag{5.60}$$

Applying Mellin inversion, we have

$$\begin{aligned}
G(m,N) &= \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} \left(\sum_{n\geq 1} \sum_{\substack{\nu^3\equiv 2(n) \\ \gcd(m,n,\mu-\nu)=1}} n^{-s}\right) N^s\hat{\phi}(s)ds \\
&= \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} \Bigg((1+2^{-s})(1+3^{-s}) \prod_{\substack{p\mid m \\ p\in\mathcal{P}_1}} \left(1+\frac{2p^{-s}}{1-p^{-s}}\right) \prod_{\substack{p\nmid m \\ p\in\mathcal{P}_1}} \left(1+\frac{3p^{-s}}{1-p^{-s}}\right) \\
&\qquad \times \prod_{\substack{p\nmid m \\ p\in\mathcal{P}_2}} \left(\frac{1}{1-p^{-s}}\right)\Bigg) N^s\hat{\phi}(s)ds. \\
&= \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} \left(\prod_{\substack{p\mid m \\ p\in\mathcal{P}_1}} \left(\frac{1+p^{-s}}{1+2p^{-s}}\right) \prod_{\substack{p\mid m \\ p\in\mathcal{P}_2}} (1-p^{-s})\right) \\
&\qquad \times (1+2^{-s})(1+3^{-s}) \prod_{p\in\mathcal{P}_1} \left(1+\frac{3p^{-s}}{1-p^{-s}}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1}{1-p^{-s}}\right) N^s\hat{\phi}(s)ds.
\end{aligned} \tag{5.61}$$

Now in order to be explicit as possible we will relate the Euler product in this last line of (5.61), which not surprisingly the Dirichlet series for counting the roots of $\nu^3 \equiv 2$

(mod n), to the Dedekind zeta function of $\mathbb{Z}[2^{1/3}]$. We have the following

$$\begin{aligned} D(s) &= \sum_{n\geq} \frac{1}{n^s} \#\{\nu \pmod n \ : \ \nu^3 \equiv 2 \pmod n\} \\ &= (1+2^{-s})(1+3^{-s}) \prod_{p\in\mathcal{P}_1} \left(1 + \frac{3p^{-s}}{1-p^{-s}}\right) \prod_{p\in\mathcal{P}_2} \left(\frac{1}{1-p^{-s}}\right) \\ &= (1-2^{-2s})(1-3^{-2s}) \\ &\quad \left(\prod_{p\in\mathcal{P}_1} \left(1-3p^{-2s}+2p^{-3s}\right) \prod_{p\in\mathcal{P}_2} \left(1-p^{-2s}\right) \prod_{p\in\mathcal{P}_3} \left(1-p^{-3s}\right)\right) \zeta_{\mathbb{Z}[2^{1/3}]}(s). \end{aligned} \tag{5.62}$$

The Euler product in the last line of (5.62) clearly converges absolutely and uniformly in $\mathrm{Re}(s) \geq \frac{1}{2} + \epsilon$, and so $D(s)$ inherits the properties of the Dedekind zeta function $\zeta_{\mathbb{Z}[2^{1/3}]}(s)$ in this region: $D(s)$ is meromorphic in $\mathrm{Re}(s) > \frac{1}{2}$ with a simple pole at $s = 1$.

Because of the rapid decay of $\hat{\phi}$ we can shift the contour in (5.61) to $\mathrm{Re}(s) = \frac{1}{2} + \epsilon$, picking up the residue at $s = 1$, to obtain

$$G(m, \mu, N) = g(m) \left(\mathrm{Res}_{s=1} D(s)\right) \hat{\phi}(1) N + O_\epsilon\left(\tau(m) N^{1/2+\epsilon}\right), \tag{5.63}$$

where

$$g(m) = \prod_{\substack{p|m \\ p\in\mathcal{P}_1}} \left(1 - \frac{1}{p+2}\right) \prod_{\substack{p|m \\ p\in\mathcal{P}_2}} \left(1 - \frac{1}{p}\right), \tag{5.64}$$

We note that $g(m)$ satisfies $(\log\log m)^{-1} \ll g(m) \ll 1$ and, for explicitness,

$$\begin{aligned} \mathrm{Res}_{s=1} D(s) &= \frac{2}{3} \prod_{p\in\mathcal{P}_1} \left(1 - \frac{3}{p^2} + \frac{2}{p^3}\right) \prod_{p\in\mathcal{P}_2} \left(1 - \frac{1}{p^2}\right) \prod_{p\in\mathcal{P}_3} \left(1 - \frac{1}{p^3}\right) \\ &\quad \times \frac{\pi}{3\sqrt{3}} \log(1 + 2^{1/3} + 2^{2/3}) \\ &\approx 0.507396 \end{aligned} \tag{5.65}$$

by the class number formula.

5.3 A transformation for a variant of the Weyl sum

With ϕ a smooth, compactly supported function on $(0, \infty)$ and M, N positive real numbers, we begin with the following sum over ideals,

$$S(M, N) = \sum_{\substack{0 \neq I \subset \mathbb{Z}[2^{1/3}] \\ d \nmid I \forall d \in \mathbb{Z} \\ \gcd(m(I), 6) = 1}} \phi\left(\frac{m(I)}{M}\right) \phi\left(\frac{n(I)}{N}\right) e\left(\frac{h\mu(I)}{m(I)}\right), \tag{5.66}$$

where h is a nonzero integer. We first arrange $S(M,N)$ as

$$S(M,N) = \sum_{\gcd(m,6)=1} \phi\left(\frac{m}{M}\right) \sum_{\mu^3 \equiv 2 (m)} e\left(\frac{h\mu}{m}\right) \sum_{\substack{0 \neq I \subset \mathbb{Z}[2^{1/3}] \\ d \nmid I \forall d \in \mathbb{Z} \\ m(I)=m, \mu(I)=\mu}} \phi\left(\frac{n(I)}{N}\right). \tag{5.67}$$

By proposition 5.4, we can evaluate the inner sum to obtain

$$S(M,N) = \text{constant}\cdot\hat{\phi}(1)N \sum_{\gcd(m,6)=1} g(m)\phi\left(\frac{m}{M}\right) \sum_{\mu^3 \equiv 2(m)} e\left(\frac{h\mu}{m}\right) + O\left(MN^{1/2}(MN)^{\epsilon}\right), \tag{5.68}$$

which yields

$$\sum_{\gcd(m,6)=1} g(m)\phi\left(\frac{m}{M}\right) \sum_{\mu^3 \equiv 2(m)} e\left(\frac{h\mu}{m}\right) = \frac{\text{constant}}{\hat{\phi}(1)N} S(M,N) + O\left(MN^{-1/2}(MN)^{\epsilon}\right). \tag{5.69}$$

The left side of (5.69) is clearly a slight variant of the Weyl sum for the roots $\mu \pmod{m}$ since the factor $g(m)$ is of little significance. We will attempt to estimate this variant of the Weyl sum by transforming $S(M,N)$ using the connection to binary cubic forms outlined in section 5.1.

To this end, we replace the sum over ideals I in $S(M,N)$ by a sum over $\alpha' = A + B2^{1/3} + C2^{2/3} \in \mathbb{Z}[2^{1/3}]$, as in the parametrization of theorem 3.4. Of course we need to handle the redundancy in α caused by the units, and instead of sharply cutting α to be in a fundamental domain we introduce a smooth function $\Psi(A,B,C)$, which we will take to be[1]

$$\Psi(A,B,C) = \psi\left(\frac{A + B2^{1/3} + C2^{2/3}}{|A^3 + 2B^3 + 4C^3 - 6ABC|^{1/3}}\right) \tag{5.70}$$

with ψ given by the following lemma.

Lemma 5.8. *There is a fixed smooth function $\psi(x)$ compactly supported in the interval $\left(\frac{3}{2}, 10\right)$, say, such that for all $x \in (0,\infty)$,*

$$\sum_{j=-\infty}^{\infty} \psi(\varepsilon^j x) = 1, \tag{5.71}$$

where $\varepsilon = 1 + 2^{1/3} + 2^{2/3}$ is the fundamental unit in $\mathbb{Z}[2^{1/3}]$. Moreover, we have $\psi^{(j)}(x) \ll 1$ for all j.

[1] In the argument of ψ we are confusing for the moment $2^{1/3}$ with its real embedding.

Proof. We let ψ_1 be a smooth function supported in $\left(\frac{3}{2}, 10\right)$ that is strictly positive in the interval $(2, 2\varepsilon))$, which comes from the fundamental domain used in proposition 3.6. The shows that for any x,

$$\sum_{j=-\infty}^{\infty} \psi_1(\varepsilon^j x) > 0. \tag{5.72}$$

Now, setting

$$\psi(x) = \left(\sum_{j=-\infty}^{\infty} \psi_1(\varepsilon^j x)\right)^{-1} \psi_1(x), \tag{5.73}$$

we see that ψ satisfies (5.71), and also the estimates for the derivatives since only a bounded number of j will contribute to the sum in the denominator of (5.73). □

With this Ψ, applying the calculations in section 5.1 yields

$$\begin{aligned} S(M,N) = \sum\sum\sum_{\substack{\gcd(A,B,C)=1 \\ \gcd\left(\frac{B^3-2C^3}{bC-cB},6\right)=1}} &\phi\left(\frac{B^3-2C^3}{(bC-cB)M}\right)\phi\left(-\frac{b^3-2c^3}{(bC-cB)N}\right) \\ &\times e\left(-\frac{hW}{C}\right)\Psi(A,B,C) + O(|h|N), \end{aligned} \tag{5.74}$$

where we have used the approximation $\frac{\mu}{m} = -\frac{W}{C} + O\left(\frac{1}{M}\right)$ of theorem 3.7. We recall that $\alpha = a + b2^{1/3} + c2^{2/3}$ is determined from α' by the requirement that $\gcd(a,b,c) = 1$ and $\alpha\alpha' \in \mathbb{Z}_{>0}$, and we also recall that $U \pmod{C}$ is determined by

$$\begin{aligned} AW &\equiv -b \pmod{C} \\ BW &\equiv c \pmod{c}. \end{aligned} \tag{5.75}$$

Again applying the calculations in section 5.1, we replace the sum over A, B, and C by a sum over b, c, B and C. Recalling that the eligible of these are all written uniquely as

$$\begin{pmatrix} b & c \\ B & C \end{pmatrix} = \begin{pmatrix} dg^2 & kg \\ 0 & g \end{pmatrix} \begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix} \tag{5.76}$$

where

$$\begin{pmatrix} b' & c' \\ B' & C' \end{pmatrix} \in \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \backslash SL_2(\mathbb{Z}) \tag{5.77}$$

is a fixed representative, $d \mid B'^3 - 2C'^3$, $\gcd(k, dg) = 1$, and

$$\gcd\left(\frac{B'^3 - 2C'^3}{d}, 6\right) = 1. \tag{5.78}$$

We remark that different choices of the representative (5.77) shift k by a multiple of dg. With this parametrization, we have

$$\begin{aligned} A &= g\left(b'B'^2 - 2c'C'^2\right) - k\frac{B'^3 - 2C'^3}{d} \\ W &\equiv -dgc'^2 \pmod{gC'}., \end{aligned} \tag{5.79}$$

and so

$$\begin{aligned} S(M,N) = &\sum_{\substack{B',C' \ll M^{2/3}N^{1/3} \\ \gcd(B',C')=1}}\sum \sum_{\substack{d \mid B'^3 - 2C'^3 \\ \gcd\left(\frac{B'^3-2C'^3}{d},6\right)=1}} e\left(h\frac{dc'^2}{C'}\right) \\ &\times \sum_{gB',gC' \asymp M^{2/3}N^{1/3}} \sum_{\gcd(k,dg)=1} \phi\left(\frac{B'^3-2C'^3}{dM}\right) \\ &\quad \times \phi\left(\frac{(dgb'+kB')^3 - 2(dgc'+kC')^3}{dN}\right) \\ &\quad\quad \times \Psi\left(g\left(b'B'^2 - 2c'C'^2\right) - k\frac{B'^3-2C'^3}{d}, gB', gC'\right) \\ &\quad\quad + O\left(|h|N\right). \end{aligned} \tag{5.80}$$

(5.81)

We break the sum over k into arithmetic progressions $k \equiv \kappa \pmod{dg}$ and apply Poisson summation to each to obtain

$$\begin{aligned} S(M,N) = &\sum_{\substack{B',C' \ll M^{2/3}N^{1/3} \\ \gcd(B',C')=1}}\sum \sum_{\substack{d \mid B'^3 - 2C'^3 \\ \gcd\left(\frac{B'^3-2C'^3}{d},6\right)=1}} e\left(h\frac{dc'^2}{C'}\right)\phi\left(\frac{B'^3-2C'^3}{dM}\right) \\ &\times \sum_{\substack{gB' \asymp M^{2/3}N^{1/3} \\ gC' \asymp M^{2/3}N^{1/3}}} \frac{1}{dg} \sum_{\kappa (dg)^*} \sum_{k} e\left(\frac{k\kappa}{dg}\right) \\ &\quad \int_{-\infty}^{\infty} \phi\left(\frac{(dgb'+xB')^3 - 2(dgc'+xC')^3}{dN}\right) \\ &\quad\quad \times \Psi\left(g\left(b'B'^2 - 2c'C'^2\right) - x\frac{B'^3-2C'^3}{d}, gB', gC'\right) e\left(-\frac{kx}{dg}\right) dx \\ &\quad\quad + O\left(|h|N\right). \end{aligned} \tag{5.82}$$

We can bring the sum over κ to the inside and evaluate the Ramanujan sum, obtaining

$$\begin{aligned} S(M,N) = & \sum_{\substack{B',C' \ll M^{2/3}N^{1/3} \\ \gcd(B',C')=1}}\sum \sum_{\substack{d \mid B'^3-2C'^3 \\ \gcd\left(\frac{B'^3-2C'^3}{d},6\right)=1}} e\left(h\frac{dc'^2}{C'}\right)\phi\left(\frac{B'^3-2C'^3}{dM}\right) \\ & \times \sum_{\substack{gB' \asymp M^{2/3}N^{1/3} \\ gC' \asymp M^{2/3}N^{1/3}}} \frac{1}{dg}\sum_{f \mid dg}\mu\left(\frac{dg}{f}\right)f \\ & \sum_{k \equiv 0(f)} \int_{-\infty}^{\infty} \phi\left(\frac{(dgb'+xB')^3-2(dgc'+xC')^3}{dN}\right) \\ & \times \Psi\left(g\left(b'B'^2-2c'C'^2\right)-x\frac{B'^3-2C'^3}{d}, gB', gC'\right)e\left(-\frac{kx}{dg}\right)dx \\ & + O\left(|h|N\right). \end{aligned} \tag{5.83}$$

We note that from proposition 3.6, Ψ restricts A to be in an interval of length $M^{2/3}N^{1/3}$, which corresponds to x in the integral of (5.83) being restricted to an interval of length $M^{-1/3}N^{1/3}$. Hence $M^{-1/3}N^{1/3}$ serves as an upper bound for this integral, and integration by parts an arbitrary number of times shows that, after replacing $k \leftarrow df$, only k for which

$$k \ll \frac{dg}{f}\frac{M^{1/3}}{N^{1/3}}(MN)^{\epsilon} \tag{5.84}$$

will contribute. It follows that $M^{-1/3}N^{1/3}$ times right side of (5.84) serves as an upper bound for the sum over all non-zero k, and the total contribution is then seen to be no more than $M^{4/3}N^{2/3}(MN)^{\epsilon}$, which we note is smaller than the trivial bound when N is larger than M. We have

$$\begin{aligned} S(M,N) = & \sum_{\substack{B',C' \ll M^{2/3}N^{1/3} \\ \gcd(B',C')=1}}\sum \sum_{\substack{d \mid B'^3-2C'^3 \\ \gcd\left(\frac{B'^3-2C'^3}{d},6\right)=1}} e\left(h\frac{dc'^2}{C'}\right)\phi\left(\frac{B'^3-2C'^3}{dM}\right) \\ & \times \sum_{\substack{gB' \asymp M^{2/3}N^{1/3} \\ gC' \asymp M^{2/3}N^{1/3}}} \frac{\varphi(dg)}{dg}\int_{-\infty}^{\infty} \phi\left(\frac{(dgb'+xB')^3-2(dgc'+xC')^3}{dN}\right) \\ & \times \Psi\left(g\left(b'B'^2-2c'C'^2\right)-x\frac{B'^3-2C'^3}{d}, gB', gC'\right)dx \\ & + O\left(|h|N + M^{4/3}N^{2/3}(MN)^{\epsilon}\right). \end{aligned} \tag{5.85}$$

References

[BBM17] Valentin Blomer, Jack Buttcane, and Péter Maga, *Applications of the Kuznetsov formula on* GL(3*) II: the level aspect*, Math. Ann. **369** (2017), no. 1-2, 723–759. MR 3694659

[BFG88] Daniel Bump, Solomon Friedberg, and Dorian Goldfeld, *Poincar series and kloosterman sums for sl(3, z)*, Acta Arithmetica **50** (1988), no. 1, 31–89 (eng).

[Blo13] Valentin Blomer, *Applications of the Kuznetsov formula on $GL(3)$*, Invent. Math. **194** (2013), no. 3, 673–729. MR 3127065

[But12] Jack Buttcane, *Sums of SL(3,Z) Kloosterman Sums*, ProQuest LLC, Ann Arbor, MI, 2012, Thesis (Ph.D.)–University of California, Los Angeles. MR 3022586

[Byk87] V. A. Bykovskii, *Spectral decompositions of certain automorphic functions and their number-theoretic applications*, Journal of Soviet Mathematics **36** (1987), no. 1, 8–21.

[CKK17] Gautam Chinta, Nathan Kaplan, and Shaked Koplewitz, *The cotype zeta function of $\mathbb{Z}^d$*, arXiv e-prints (2017), arXiv:1708.08547.

[DF64] B. N. Delone and D. K. Faddeev, *The theory of irrationalities of the third degree*, Translations of Mathematical Monographs, Vol. 10, American Mathematical Society, Providence, R.I., 1964. MR 0160744

[DFI95] W. Duke, J. B. Friedlander, and H. Iwaniec, *Equidistribution of roots of a quadratic congruence to prime moduli*, Annals of Mathematics **141** (1995), no. 2, 423–441.

[FI97] Etienne Fouvry and Henryk Iwaniec, *Gaussian primes*, Acta Arith. **79** (1997), no. 3, 249–287. MR 1438827

[FI98] John Friedlander and Henryk Iwaniec, *The polynomial $X^2 + Y^4$ captures its primes*, Ann. of Math. (2) **148** (1998), no. 3, 945–1040. MR 1670065

[Hej86] Dennis A. Hejhal, *Roots of quadratic congruences and eigenvalues of the non-Euclidean Laplacian*, The Selberg trace formula and related topics (Brunswick, Maine, 1984), Contemp. Math., vol. 53, Amer. Math. Soc., Providence, RI, 1986, pp. 277–339. MR 853563

[Hoo63] Christopher Hooley, *On the number of divisors of quadratic polynomials.*, Acta Mathematica **110** (1963), no. 1, 97.

[Hoo64] C. Hooley, *On the distribution of the roots of polynomial congruences*, Mathematika **11** (1964), 39–49. MR 0163874

[Hoo78] Christopher Hooley, *On the greatest prime factor of a cubic polynomial*, J. Reine Angew. Math. **303/304** (1978), 21–50. MR 514671

[IK04] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR 2061214

[Iwa78] Henryk Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. Math. **47** (1978), no. 2, 171–188. MR 0485740

[Li10] Xiaoqing Li, *A spectral mean value theorem for* GL(3), J. Number Theory **130** (2010), no. 11, 2395–2403. MR 2678854

[Pet07] V. M. Petrogradsky, *Multiple zeta functions and asymptotic structure of free abelian groups of finite rank*, J. Pure Appl. Algebra **208** (2007), no. 3, 1137–1158. MR 2283452

[T0́0] Árpád Tóth, *Roots of quadratic congruences.*, IMRN: International Mathematics Research Notices **2000** (2000), no. 14, 719.

[Ter88] Audrey Terras, *Harmonic analysis on symmetric spaces and applications. II*, Springer-Verlag, Berlin, 1988. MR 955271

[Wel18] M. C. Welsh, *Spacing and A Large Sieve Type Inequality for Roots of a Cubic Congruence*, ArXiv e-prints (2018).