Generalizations of the Gap Principle and the Thue-Siegel Principle, with Applications to Diophantine Equations

by

Anton Mosunov

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Pure Mathematics

Waterloo, Ontario, Canada, 2019

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis we develop generalizations of two well-known principles from the theory of Diophantine approximation, namely the gap principle and the Thue-Siegel principle. Our results find their applications in the theory of Diophantine equations. Let α be an algebraic number over $\mathbb{Q}$ and let $F(X, Y)$ be the homogenization of the minimal polynomial of α. In the special case when $\mathbb{Q}(\alpha)/\mathbb{Q}$ is a Galois extension of degree at least seven, we establish absolute bounds on the number of solutions of certain equations of Thue and Thue-Mahler type, which involve $F(X, Y)$. Consequently, we give theoretical evidence in support of Stewart's conjecture (1991). More generally, if every conjugate β of α is such that the degree of β over $\mathbb{Q}(\alpha)$ is small relative to the degree of α over $\mathbb{Q}$, we establish bounds of the form $C\gamma$, where C is an absolute constant and γ is a natural parameter associated with α that does not exceed the degree of α over $\mathbb{Q}$. We expect this parameter to be small, perhaps even bounded by an absolute constant.

Acknowledgements

I am grateful to Prof. Cameron L. Stewart for his elegant lectures, through which I learned about the fascinating area of Diophantine approximation; additionally, I value his masterful supervision and the insightful comments on my research and thesis. I thank my dear wife Asia for her love and the courage to overcome an almost seven thousand kilometre distance for the sole purpose of being nearby. I feel beholden to my mother, who has always supported me during my long lasting academic journey. I am indebted to Prof. David McKinnon and Prof. Jason P. Bell for their numerous consultations. Finally, I would like to thank Prof. Kevin Hare, Prof. Jeffrey Shallit and Prof. Jeffrey Thunder for reviewing this manuscript, as well as Patrick Naylor and Dr. Stanley Yao Xiao for many productive discussions.

Посвящение

Моим бабушке и дедушке, Тащиан Галине Александровне (урожд. Анохиной) и Генриху Львовичу, искренним и честным людям удивительной судьбы.

To my grandmother and grandfather, Tashchian Galina Aleksandrovna (born Anokhina) and Genrikh Lvovich, sincere and honest people who lived remarkable lives.

Тащиан Галина Александровна (1930 – 2015)
Тащиан Генрих Львович (1931 – 2015)

Table of Contents

Chapter 1

Introduction

1.1 Diophantine Approximation and its Applications to the Theory of Diophantine Equations

The theory of Diophantine approximation is a fascinating area of mathematics, which studies approximation properties of numbers. Over the past two hundred years its development led to numerous breakthroughs in number theory, including the discovery of the first transcendental number and the development of methods for solving various Diophantine equations.

In Diophantine approximation we are primarily interested in the question of how well real numbers can be approximated by rationals, and its variations and generalizations. If α is a real number and x/y is a rational number with $y \geq 1$, then the quality of approximation of α by x/y can be measured by means of a quantity μ such that the inequality

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{y^{\mu}} \tag{1.1.1}$$

is satisfied. The larger μ is, the better the approximation of x/y with respect to α is. It was observed by Dirichlet that for $\mu = 2$ the inequality above can be achieved for infinitely many integers x and y, as long as α is real and irrational. On the other hand, Liouville pointed out that, if α is an irrational algebraic number of degree d and $\mu > d$, then (1.1.1) has only finitely many solutions in integers x and y with $y \geq 1$. In other words, algebraic numbers cannot be approximated by rationals too well. This observation

enabled Liouville to construct the first transcendental number $\sum_{n=1}^{+\infty} 10^{-n!}$, known today as a *Liouville number*, that was well-approximated by rationals [23].

In this thesis, we generalize certain principles from the theory of Diophantine approximation and apply them to analyze Diophantine equations of Thue and Thue-Mahler type. A *Thue equation* is an equation of the form

$$F(x, y) = m, \tag{1.1.2}$$

where $F(X, Y) \in \mathbb{Z}[X, Y]$ is a homogeneous polynomial of degree $d \geq 3$ with non-zero discriminant $D(F)$, m is a fixed positive integer, and x, y are integer variables. In 1909 Thue [33] established that there is a finite upper bound on the number of solutions of (1.1.2), provided that F is irreducible. He observed that large solutions to (1.1.2), — that is, solutions (x, y) satisfying $|y| \geq C_1$ for some number $C_1 = C_1(m, \mu, F)$, — must satisfy the inequality (1.1.1), with α being one of the roots of $F(X, 1)$. The problem essentially reduced to counting distinct solutions x/y to the inequality (1.1.1).

Perhaps surprisingly, it is not difficult to count distinct x/y satisfying (1.1.1) with y varying in a fixed range. Indeed, if it so happens that $C_1 \leq y_1 < y_2 \leq C_2$, then the fact that $x_1/y_1 \neq x_2/y_2$ yields

$$\frac{1}{y_1 y_2} \leq \left|\frac{x_1}{y_1} - \frac{x_2}{y_2}\right| \leq \left|\alpha - \frac{x_1}{y_1}\right| + \left|\alpha - \frac{x_2}{y_2}\right| < \frac{1}{y_1^{\mu}} + \frac{1}{y_2^{\mu}} < \frac{2}{y_1^{\mu}},$$

resulting in the inequality

$$2y_2 > y_1^{\mu-1},$$

which is known as the *gap principle*. For $\mu > 2$ this inequality states that, if two distinct rationals satisfy (1.1.1), then their denominators must be far apart from each other. Unfortunately, as the quantity C_2 can be arbitrarily large, the gap principle itself does not provide any insight on the number of distinct solutions to (1.1.1). However, it was established by Thue that with C_1 sufficiently large one may take $C_2 = y_1^{\eta}$ for some $\eta > 1$, where x_1/y_1 is a solution to (1.1.1) with the smallest denominator $y_1 \geq C_1$. This phenomenon is known as the *Thue-Siegel principle*. Upon combining the gap principle with the Thue-Siegel principle it is possible to count distinct solutions x/y to the inequality (1.1.1), and hence to the Thue equation (1.1.2).

Since Thue's time, the estimates on the number of solutions of (1.1.2) have been improved significantly. In 1933, assuming that F is irreducible, Mahler established the existence of a number C, dependent only on F, such that the number of primitive solutions to (1.1.2), — that is, solutions with x and y coprime, — does not exceed $C^{1+\omega(m)}$, where

$\omega(m)$ denotes the number of distinct prime divisors of m [24]. In fact, his result was even stronger: if instead of (1.1.2) we consider the equation

$$F(x,y) = p_1^{k_1} \cdots p_t^{k_t}, \tag{1.1.3}$$

where $p_1, p_2, \ldots, p_t$ are distinct fixed prime numbers, then it follows from Mahler's argument that the number of integer solutions $(x, y, k_1, k_2, \ldots, k_t)$ to (1.1.3), with x, y coprime and k_i non-negative, does not exceed C^{1+t}. The equation (1.1.3) is called a *Thue-Mahler equation*. Further improvements to this estimate have been made by Erdős and Mahler [12], and Lewis and Mahler [21].

It was conjectured by Siegel that the number of primitive solutions to (1.1.2) should not depend on the coefficients of F. Siegel's conjecture was established in 1984 by Evertse [15], who proved that the number of primitive solutions to (1.1.3) does not exceed

$$2 \cdot 7^{d^3(2t+3)}, \tag{1.1.4}$$

where a binary form F of degree d was assumed to be divisible by at least three pairwise linearly independent linear forms in some algebraic number field. An estimate on the number of solutions to (1.1.2) thus follows by replacing the number t in (1.1.4) with $\omega(m)$.

When integers x and y are arbitrary, the number of solutions to (1.1.2) can be large. For example, in 2008 Stewart [32] proved that when F is of degree 3 and $D(F) \neq 0$ then there is a positive number $c = c(F)$ such that the number of solutions to (1.1.2) is at least $c(\log m)^{1/2}$. However, if we restrict our attention only to primitive solutions, then their number does not seem to increase with the growth of m. In 1987 it was conjectured by Erdős, Stewart and Tijdeman [13] that the number of primitive solutions to (1.1.2) does not exceed some constant, which depends only on d. In the same year Bombieri and Schmidt [4] proved that the number of primitive solutions to (1.1.2) does not exceed

$$Cd^{1+\omega(m)},$$

where the constant C is absolute. In 1991 Stewart [31] replaced $\omega(m)$ in the above estimate with $\omega(g)$, where g is a divisor of m satisfying $g \gg_F m^{(4+d)/3d}$ (this is the statement of [31, Theorem 1] with $\varepsilon = 1/2$). In the same paper, Stewart conjectured the following.

Conjecture 1.1. (Stewart, [31, Section 6]) *There exists an absolute constant c_0 such that for any binary form $F \in \mathbb{Z}[X, Y]$ with nonzero discriminant and degree at least three there exists a number $C = C(F)$, such that if m is an integer larger than C, then the Thue equation (1.1.2) has at most c_0 solutions in coprime integers x and y.*

There has not been much recent work towards establishing Stewart's conjecture. The most notable step forward can be found in the work of Thunder [34]. Based on [31] he gives a heuristic that supports the conjecture of Stewart when the degree of the form F is at least five. By generalizing the gap principle and the Thue-Siegel principle outlined above, we develop new methods for estimating the number of primitive solutions of (1.1.2) and (1.1.3) in the case $t = 1$, thus providing theoretical evidence in support of Stewart's conjecture.

1.2 Organization of the Thesis and Contributions

The thesis is organized as follows. In Chapter 2 we develop a generalized gap principle for all absolute values of $\mathbb{Q}$. We apply our generalized gap principle and the Thue-Siegel principle [3] to prove Theorems 2.2 and 2.3, where we establish absolute bounds on the number of solutions of certain equations of Thue and Thue-Mahler type. The bounds established in Theorems 2.2, 2.3 depend on the size of $\mathrm{Aut}'\,|F|$, a finite group of size at most 24 associated with a binary form F. In Chapter 3 we compute $\mathrm{Aut}'\,|F|$ explicitly for binary forms associated with $2\cos(2\pi/n)$. The main result of this chapter is stated in Theorem 3.1.

We use methods developed in Chapter 2 to produce a generalized gap principle for rationals $x_1/y_1, x_2/y_2$ in lowest terms that approximate algebraic numbers α, β, respectively, provided that $\beta \in \mathbb{Q}(\alpha)$. It is an interesting and challenging problem to remove this restriction. In Chapters 4 and 5 we generalize the Thue-Siegel principle (Theorem 4.1) and the gap principle (Lemmas 5.13, 5.14) to the case when the degree of β over $\mathbb{Q}(\alpha)$ is small relative to the degree of α over $\mathbb{Q}$. These principles are combined to prove Theorems 5.5 and 5.6, the first of which is stated in terms of explicitly computable numbers only. In both theorems our bounds can be written in the form $C\gamma$, where C is an absolute constant and γ is some natural parameter which does not exceed the degree of a binary form F and can be significantly smaller.

Chapter 2

Absolute Bounds for the Number of Solutions of Certain Equations of Thue and Thue-Mahler Type

Let

$$F(X,Y) = c_d X^d + c_{d-1} X^{d-1} Y + \cdots + c_1 X Y^{d-1} + c_0 Y^d$$

denote a homogeneous polynomial of degree d with integer coefficients. For an arbitrary polynomial $f \in \mathbb{Z}[X]$, we define the *content* of f to be the greatest common divisor of its coefficients. Thus, in the case of F defined above, the content is equal to $\gcd(c_0, c_1, \ldots, c_d)$. We say that F is *irreducible* if the equality $F = GH$, where $G, H \in \mathbb{Q}[X,Y]$ are homogeneous polynomials, implies that either G or H has degree zero.

In this chapter we study equations of the form $|F(x,y)| = tp^z$, with p^z a prime power and t a small integer variable. We demonstrate that it is possible to provide an absolute bound on the number of their solutions, provided that F is irreducible of degree $d \geq 7$ and the size of the Galois group of $F(X,1)$ over $\mathbb{Q}$ is equal to d.

In order to state the main results given in Theorems 2.2 and 2.3, we need to introduce the notion of an automorphism of a binary form. For a 2×2 matrix $M = \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right)$, with complex entries, define the binary form $F_M(X,Y)$ by

$$F_M(X,Y) = F(sX + uY, tX + vY).$$

Definition 2.1. We say that $M = \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbb{C})$ is an *automorphism of* F (resp., $|F|$) if $F_M = F$ (resp., $F_M = F$ or $F_M = -F$). If K is a field containing $\mathbb{Q}$, the set of all

automorphisms of F (resp., $|F|$) with $s,t,u,v \in K$ is denoted by $\mathrm{Aut}_K F$ (resp., $\mathrm{Aut}_K |F|$). We also define

$$\mathrm{Aut}' F = \left\{ \frac{1}{\sqrt{|sv-tu|}} \begin{pmatrix} s & u \\ t & v \end{pmatrix} : s,t,u,v \in \mathbb{Z} \right\} \cap \mathrm{Aut}_{\overline{\mathbb{Q}}} F. \tag{2.0.1}$$

Analogously, we define $\mathrm{Aut}' |F|$.

One can easily verify that $\mathrm{Aut}' |F|$ is a group. In Lemmas 2.4, 2.5 we will show that for $d \geq 3$ this group is finite and contains at most 24 elements.

For an arbitrary finite set X, let $\#X$ denote its cardinality. We prove the following.

Theorem 2.2. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d \geq 7$ and content one. Let α be a root of $F(X,1)$ and assume that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois. Let λ be taken from Table 2.1 for $7 \leq d \leq 16$ and $\lambda = 1 - 16.2/d$ for $d \geq 17$. Let p be prime, k a positive integer, and consider the Diophantine equation*

$$|F(x,y)| = tp^k. \tag{2.0.2}$$

Provided that p^k is sufficiently large, the number of solutions to (2.0.2) in integers (x,y,t) such that

$$\gcd(x,y) = 1, \;\; 1 \leq t \leq (p^k)^\lambda$$

is at most $\#\,\mathrm{Aut}' |F|$. In particular, it does not exceed 24. More precisely, for any two solutions $(x_1,y_1,t_1),(x_2,y_2,t_2)$ there exists a matrix $M = |sv-tu|^{-1/2} \cdot \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right)$ in $\mathrm{Aut}' |F|$ such that

$$\frac{x_2}{y_2} = \frac{sx_1 + uy_1}{tx_1 + vy_1}.$$

d	7	8	9	10	11	12	13	14	15	16
λ	0.004	0.068	0.124	0.172	0.216	0.254	0.289	0.321	0.349	0.375

Table 2.1: Values of λ corresponding to d in the range $7 \leq d \leq 16$.

Theorem 2.3. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d \geq 7$ and content one. Let α be a root of $F(X,1)$ and assume that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois. Let λ be such that*

$$0 \leq \lambda < 1 - 8.1/(d+2).$$

Let p be prime, and consider the Diophantine equation

$$|F(x,y)| = tp^z. \tag{2.0.3}$$

Provided that p is sufficiently large, the number of solutions to (2.0.3) in integers (x,y,z,t) such that

$$\gcd(x,y) = 1, \quad z \geq 1, \quad 1 \leq t \leq (p^z)^\lambda$$

is at most

$$\#\operatorname{Aut}'|F| \cdot \left\lfloor 1 + \frac{11.51 + 1.5\log d + \log\left((d-2.05)/(1+\lambda)\right)}{\log((d-2.05)/(1+\lambda) - 0.5d)} \right\rfloor.$$

If we let $\lambda = 0.5 - 4.05/(d+2)$, then it is a consequence of Theorem 2.3 and the fact that $\operatorname{Aut}'|F| \leq 24$ (see Lemma 2.5) that the number of solutions in integers (x,y,z,t) to (2.0.3) does not exceed 1992 for all $d \geq 7$ and it does not exceed 72 for all $d \geq 10^{15}$.

2.1 Automorphisms of Binary Forms

In this section we establish several results about automorphisms of a binary form $F(X,Y)$. At the end we prove Proposition 2.7, where we explain the relation between automorphisms of F and the roots of $F(X,1)$.

Lemma 2.4. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d \geq 3$. Then $\operatorname{Aut}_{\mathbb{Q}}|F|$ is $\mathrm{GL}_2(\mathbb{Q})$-conjugate to one of the groups from Table 2.2.*

Proof. Let

$$M = \frac{1}{N}\begin{pmatrix} s_1 & u_1 \\ t_1 & v_1 \end{pmatrix},$$

where N, s_1, t_1, u_1, v_1 are integers such that $\gcd(s_1,t_1,u_1,v_1) = 1$ and $N \in \mathbb{N}$. Since $F_M = \pm F$, we see that $|D(F_M)| = |D(F)|$, where $D(F)$ is the discriminant of F. Then it follows from the formula $D(F_M) = (\det M)^{d(d-1)}D(F)$ that $|\det(M)| = 1$. Therefore $N = \sqrt{|s_1v_1 - t_1u_1|}$.

Now suppose that $M \in \operatorname{Aut}_{\mathbb{Q}}|F|$ has infinite order. Define $S = NM \in \mathrm{M}_2(\mathbb{Z})$ and for $k \in \mathbb{N}$ let s_k, t_k, u_k, v_k be such that $S^k = \left(\begin{smallmatrix} s_k & u_k \\ t_k & v_k \end{smallmatrix}\right)$. Fix some coprime pair $(x_0,y_0) \in \mathbb{Z}^2$ and let $h = |F(x_0,y_0)|$. Then for any $k \in \mathbb{N}$ we have

$$h = |F(x_0,y_0)| = |F_{M^k}(x_0,y_0)| = \frac{1}{N^{dk}}|F_{S^k}(x_0,y_0)| = \frac{1}{N^{dk}}|F(s_kx_0 + u_ky_0, t_kx_0 + v_ky_0)|.$$

Let $g_k = \gcd(s_k x_0 + u_k y_0, t_k x_0 + v_k y_0)$. Then

$$\left| F\left(\frac{s_k x_0 + u_k y_0}{g_k}, \frac{t_k x_0 + v_k y_0}{g_k} \right) \right| = \frac{N^{kd} h}{g_k^d},$$

where the quantities on either side of the above equality are positive integers. If $N = 1$ then g_k^d divides h for all k. In particular, there exists some divisor h' of h such that $h' = h/g_k^d$ for infinitely many k. But then we obtain infinitely many solutions to the Thue equation $|F(x, y)| = h'$. Since F is irreducible and of degree at least three, this contradicts Thue's Theorem [33]. If $N > 1$, let $p_1, \ldots, p_n$ denote the distinct primes that divide N. Since $N^{kd} h/g_k^d$ is an integer for all k, there exists some divisor h' of h such that $N^{kd} h/g_k^d$ is of the form $p_1^{z_1} \cdots p_n^{z_n} h'$ for infinitely many k. Therefore the Thue-Mahler equation

$$|F(x, y)| = p_1^{z_1} \cdots p_n^{z_n} h'$$

has infinitely many solutions $(x, y, z_1, \ldots, z_n)$, with $\gcd(x, y) = 1$. Since F is irreducible and of degree at least three, this contradicts Mahler's Theorem [24], which means that $M \in \mathrm{Aut}_{\mathbb{Q}} |F|$ always has finite order.

Now suppose that $\mathrm{Aut}_{\mathbb{Q}} |F|$ contains at least 13 distinct elements $M_1, \ldots, M_{13}$, each of which has finite order. By Schur's Theorem [8], any finitely generated torsion subgroup of $\mathrm{GL}_n(\mathbb{C})$ is finite.[1] Hence $\langle M_1, \ldots, M_{13} \rangle$ is a finite subgroup of $\mathrm{GL}_2(\mathbb{Q})$. However, it is known that every finite subgroup of $\mathrm{GL}_2(\mathbb{Q})$ has to be $\mathrm{GL}_2(\mathbb{Q})$-conjugate to one of the groups listed in Table 2.2 [25]. Since all these subgroups have at most 12 elements, we reach a contradiction. Therefore $\mathrm{Aut}_{\mathbb{Q}} |F|$ is a finite subgroup of $\mathrm{GL}_2(\mathbb{Q})$, and so it is $\mathrm{GL}_2(\mathbb{Q})$-conjugate to one of the groups listed in Table 2.2. □

Lemma 2.5. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a binary form of degree $d \geq 3$. Let $\mathrm{Aut}' |F|$ be as in (2.0.1). Then $\mathrm{Aut}' |F| \cong \boldsymbol{C}_n$ or $\mathrm{Aut}' |F| \cong \boldsymbol{D}_n$, where $n \in \{1, 2, 3, 4, 6, 8, 12\}$.*

Proof. Note that $\mathrm{Aut}_{\mathbb{Q}} |F|$ is a subgroup of $\mathrm{Aut}' |F|$. Furthermore, for any $M \in \mathrm{Aut}' |F|$ we have $M^2 \in \mathrm{Aut}_{\mathbb{Q}} |F|$. By Lemma 2.4, $\mathrm{Aut}_{\mathbb{Q}} |F|$ is finite, and so any $M \in \mathrm{Aut}' |F|$ has finite order. In fact, since the orders of elements in $\mathrm{Aut}_{\mathbb{Q}} |F|$ are $\{1, 2, 3, 4, 6\}$, the only possible orders of elements in $\mathrm{Aut}' |F|$ are $\{1, 2, 3, 4, 6, 8, 12\}$.

Next, recall a classical result that any finite subgroup of $\mathrm{GL}_2(\mathbb{R})$ is $\mathrm{GL}_2(\mathbb{R})$-conjugate to a finite subgroup of the orthogonal group $O_2(\mathbb{R})$. Since finite subgroups of $O_2(\mathbb{R})$ correspond to rotations and reflections on a plane, we conclude that each finite subgroup of $\mathrm{GL}_2(\mathbb{R})$,

[1] I am grateful to Patrick Naylor for his help on this part of the argument.

Group	Generators	Group	Generators
$\boldsymbol{C}_1$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\boldsymbol{D}_1$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$\boldsymbol{C}_2$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\boldsymbol{D}_2$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
$\boldsymbol{C}_3$	$\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$	$\boldsymbol{D}_3$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$
$\boldsymbol{C}_4$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\boldsymbol{D}_4$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
$\boldsymbol{C}_6$	$\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$	$\boldsymbol{D}_6$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$

Table 2.2: Representatives of equivalence classes of finite subgroups of $\mathrm{GL}_2(\mathbb{Q})$ under $\mathrm{GL}_2(\mathbb{Q})$-conjugation.

including $\mathrm{Aut}'\,|F|$, is isomorphic to either a cyclic group $\boldsymbol{C}_n$ of order n or a dihedral group $\boldsymbol{D}_n$ of order $2n$.

Now suppose that G contains at least 25 distinct elements $M_1, \ldots, M_{25}$. By Schur's Theorem [8], any finitely generated torsion subgroup of $\mathrm{GL}_n(\mathbb{C})$ is finite. Hence $\langle M_1, \ldots, M_{25} \rangle$ is a finite subgroup of $\mathrm{GL}_2(\mathbb{R})$, so it is isomorphic to either $\boldsymbol{C}_n$ or $\boldsymbol{D}_n$ for some n. In the former case we see that $n \geq 25$, while in the latter case $n \geq 13$. In both cases we obtain a contradiction, since the largest order that an element of $\mathrm{Aut}'\,|F|$ can have is 12. Therefore $\mathrm{Aut}'\,|F|$ contains at most 24 elements. □

Example 2.6. Let us give an example of a group of the form (2.0.1) that is not a subgroup of $\mathrm{GL}_2(\mathbb{Q})$. Consider

$$G = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{3} & 2/\sqrt{3} \end{pmatrix} \right\rangle.$$

Then $G \cong \boldsymbol{D}_{12}$. If we choose coprime integers a, b so that $a \equiv 3b \pmod{10}$, then any

(reciprocal) binary form

$$\begin{aligned}F(X,Y) &= a(X^{12}+Y^{12}) - 6aXY(X^{10}+Y^{10}) \\ &+ \frac{231a+2b}{5}X^2Y^2(X^8+Y^8) - (176a+2b)X^3Y^3(X^6+Y^6) \\ &+ \frac{495a+5b}{2}X^4Y^4(X^4+Y^4) + 2bX^5Y^5(X^2+Y^2) \\ &- \frac{1122a+29b}{5}X^6Y^6\end{aligned}$$

will have integer coefficients and satisfy the equation $F_M = F$ for any $M \in G$. Consequently, if (x,y) is a solution of $F(x,y) = m$, then so are $(y,-x+y)$, $(-x+y,-x)$, $(-x,-y)$, $(-y,x-y)$, $(x-y,x)$, (y,x), $(-x+y,y)$, $(-x,-x+y)$, $(-y,-x)$, $(x-y,-y)$, $(x,x-y)$. This phenomenon was observed by Stewart in [31, Section 6] with respect to binary forms invariant under $\boldsymbol{D}_6$, which is a subgroup of G. In addition to these 12 solutions, we have $F(x',y') = 729m$ for any $(x',y') \in \{(x+y,-x+2y), (-x+2y,-2x+y), (-2x+y,-x-y), (-x-y,x-2y), (x-2y,2x-y), (2x-y,x+y), (-x+2y,x+y), (-2x+y,-x+2y), (-x-y,-2x+y), (x-2y,-x-y), (2x-y,x-2y), (x+y,2x-y)\}$.

Proposition 2.7. *Let $F(X,Y) = c_dX^d + c_{d-1}X^{d-1} + \cdots + c_0Y^d \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d \geq 3$. Let $\alpha_1, \ldots, \alpha_d$ be the roots of $F(X,1)$. Then there exists j such that*

$$\alpha_j = \frac{v\alpha_1 - u}{-t\alpha_1 + s}$$

for some integers s,t,u,v if and only if the matrix

$$\frac{1}{\sqrt{|sv-tu|}}\begin{pmatrix} s & u \\ t & v\end{pmatrix}, \quad |sv-tu| = \left|\frac{F(s,t)}{c_d}\right|^{2/d}$$

is an automorphism of $|F|$.

Proof. Since $F(X,1)$ is irreducible, its Galois group acts transitively on the roots $\alpha_1, \alpha_2, \ldots, \alpha_d$. Therefore

$$\frac{v\alpha_1 - u}{-t\alpha_1 + s}, \frac{v\alpha_2 - u}{-t\alpha_2 + s}, \ldots, \frac{v\alpha_d - u}{-t\alpha_d + s}$$

is a permutation of $\alpha_1, \ldots, \alpha_d$. Thus

$$\begin{aligned} F(X,Y) &= c_d \prod_{i=1}^{d} \left(X - \frac{v\alpha_i - u}{-t\alpha_i + s} Y \right) \\ &= \frac{c_d}{\prod_{i=1}^{d}(s - t\alpha_i)} \prod_{i=1}^{d} ((-t\alpha_i + s)X - (v\alpha_i - u)Y) \\ &= \frac{c_d}{F(s,t)} F(sX + uY, tX + vY) \\ &= \pm \eta^d F(sX + uY, tX + vY), \end{aligned}$$

where $\eta = |c_d / F(s,t)|^{1/d} \in \mathbb{R}$. Since F is homogeneous, we see that the matrix $M = \eta \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix} \right)$ is an automorphism of $|F|$. Since

$$D(F_M) = (\det M)^{d(d-1)} D(F)$$

and $F_M = \pm F$, we see that $|D(F_M)| = |D(F)|$, and so $|\det M| = 1$. Therefore $|\eta|^2 \cdot |sv - tu| = 1$, leading us to a conclusion that $\eta = |\eta| = |sv - tu|^{-1/2}$.

Conversely, suppose that $M = |sv - tu|^{-1/2} \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix} \right) \in \mathrm{Aut}'\, |F|$. Then

$$F_M(X,Y) = \frac{c_d}{|sv - tu|^{d/2}} \prod_{i=1}^{d} (sX + uY - \alpha_i(tX + vY)).$$

We see that the polynomial $F_M(X,1)$ vanishes at $(v\alpha_i - u)/(-t\alpha_i + s)$ for $i = 1, \ldots, d$. Since $F_M = \pm F$, the polynomials $F_M(X,1), F(X,1)$ have the same roots, so there exists some index j such that $\alpha_j = (v\alpha_1 - u)/(-t\alpha_1 + s)$. □

2.2 Preliminary Results

This section contains five lemmas, which we utilize in Section 2.3 to explore the properties of minimal polynomials, as well as in Section 2.5, where we establish Archimedean and non-Archimedean gap principles. Before we proceed, let us introduce some definitions and notation.

Definition 2.8. For an arbitrary polynomial $R \in \mathbb{Z}[X_1, X_2, \ldots, X_n]$, we let $H(R)$ denote the maximum of Archimedean absolute values of its coefficients, and refer to this quantity

as a *naive height*, or simply a *height*, of R. For an algebraic number α with the minimal polynomial f, we write $H(\alpha) = H(f)$. For a point $(x_1, x_2, \ldots, x_n) \in \mathbb{Z}^n$, we define

$$H(x_1, x_2, \ldots, x_n) = \max_{i=1,2,\ldots,n} \{|x_i|\},$$

and refer to this quantity as the *height* of $(x_1, x_2, \ldots, x_n)$.

Definition 2.9. Let α be an algebraic number of degree d over $\mathbb{Q}$ and let $\alpha = \alpha_1, \ldots, \alpha_d$ be the conjugates of α. The *house* of α, denoted $\overline{|\alpha|}$, is defined to be

$$\overline{|\alpha|} = \max\{|\alpha_1|, \ldots, |\alpha_d|\},$$

where c_α is the leading coefficient of the minimal polynomial of α.

Definition 2.10. Let α be an algebraic number of degree d over $\mathbb{Q}$ and let $\alpha = \alpha_1, \ldots, \alpha_d$ be the conjugates of α. The *Mahler measure* of α, denoted $M(\alpha)$, is defined to be

$$M(\alpha) = c_\alpha \prod_{i=1}^{d} \max\{1, |\alpha_i|\}.$$

In this section, as well as all the subsequent ones, we will write

$$D_{i,j} = \frac{1}{i!j!}\frac{\partial^{i+j}}{\partial X^i \partial Y^j}, \quad D_i = \frac{1}{i!}\frac{d^i}{dX^i}.$$

Lemma 2.11. (Liouville's Theorem, [30, Theorem 1E]) *Let $\alpha \in \mathbb{C}$ be an algebraic number of degree d over $\mathbb{Q}$. Then for all integers x, y, $y \neq 0$ such that $x \neq y\alpha$ the inequality*

$$\left|\alpha - \frac{x}{y}\right| > \frac{1}{2^{d+1} H(\alpha) \max\{1, |\alpha|\}^{d-1} H(x,y)^d} \tag{2.2.1}$$

holds.

Proof. Let

$$f(X) = c_d X^d + \cdots + c_1 X + c_0$$

be the minimal polynomial of α. Since $f(\alpha) = 0$, it follows from Taylor's Theorem that

$$f\left(\frac{x}{y}\right) = \sum_{i=1}^{d} D_i f(\alpha) \left(\frac{x}{y} - \alpha\right)^i.$$

By the triangle inequality and the identity

$$\sum_{i=1}^{n} \binom{i}{r} = \binom{n+1}{r+1}, \tag{2.2.2}$$

also known under the name of the hockey-stick identity, we have

$$\begin{aligned} |D_i f(\alpha)| &\leq H(\alpha) \sum_{j=i}^{d} \binom{j}{i} \max\{1, |\alpha|\}^{j-i} \\ &\leq H(\alpha) \max\{1, |\alpha|\}^{d-i} \sum_{j=i}^{d} \binom{j}{i} \\ &= H(\alpha) \binom{d+1}{i+1} \max\{1, |\alpha|\}^{d-i}. \end{aligned} \tag{2.2.3}$$

We may assume that $|\alpha - x/y| \leq 1$, for otherwise we are done. Since $x \neq y\alpha$, it must be the case that $f(x/y) \neq 0$. Therefore

$$\begin{aligned} \frac{1}{H(x,y)^d} &\leq \left| f\left(\frac{x}{y}\right) \right| \\ &\leq \left| \alpha - \frac{x}{y} \right| \sum_{i=1}^{d} |D_i f(\alpha)| \\ &\leq \left| \alpha - \frac{x}{y} \right| H(\alpha) \sum_{i=1}^{d} \binom{d+1}{i+1} \max\{1, |\alpha|\}^{d-i} \\ &\leq \left| \alpha - \frac{x}{y} \right| H(\alpha) \max\{1, |\alpha|\}^{d-1} \sum_{i=2}^{d+1} \binom{d+1}{i} \\ &< \left| \alpha - \frac{x}{y} \right| 2^{d+1} H(\alpha) \max\{1, |\alpha|\}^{d-1}. \end{aligned}$$

□

Lemma 2.12. (p-adic Liouville theorem) *Let p be a rational prime and $\alpha \in \mathbb{Q}_p$ a p-adic algebraic number of degree d over $\mathbb{Q}$. Let $|\ |_p$ denote the p-adic absolute value on $\mathbb{Q}_p$, normalized so that $|p|_p = p^{-1}$. Then for all integers x, y such that $x \neq y\alpha$, the inequality*

$$|y\alpha - x|_p \geq \frac{|c_\alpha|_p^{d-1}}{(d+1)H(\alpha)H(x,y)^d} \tag{2.2.4}$$

holds, where c_α is the leading coefficient of the minimal polynomial of α.

Proof. Let

$$f(X) = c_d X^d + \cdots + c_1 X + c_0$$

be the minimal polynomial of α and $F(X,Y) = Y^d f(X/Y)$ its homogenization. Note that $c_d = c_\alpha$. Since $f(\alpha) = 0$, it follows from Taylor's Theorem that

$$\begin{aligned} F(x,y) &= (x - \alpha y) \sum_{i=1}^{d} D_i f(\alpha)(x - \alpha y)^{i-1} y^{d-i} \\ &= (x - \alpha y) \sum_{i=1}^{d} \frac{D_i f(\alpha)}{c_d^{i-1}} (c_d x - c_d \alpha y)^{i-1} y^{d-i}. \end{aligned}$$

Since $c_d\alpha$ and $c_d^{d-i} D_i f(\alpha)$ are algebraic integers, their p-adic absolute values do not exceed one, so

$$\begin{aligned} |F(x,y)|_p &\leq |y\alpha - x|_p \cdot \max_{i=1,\ldots,d} \left\{ \left| \frac{D_i f(\alpha)}{c_d^{i-1}} \right|_p \right\} \\ &= |y\alpha - x|_p \cdot \max_{i=1,\ldots,d} \left\{ \left| \frac{c_d^{d-i} D_i f(\alpha)}{c_d^{d-1}} \right|_p \right\} \\ &\leq |y\alpha - x|_p \cdot |c_d|_p^{-d-1}. \end{aligned}$$

Since $x \neq y\alpha$, it must be the case that $F(x,y) \neq 0$. By the product formula, the following trivial lower bound holds:

$$|F(x,y)|_p \geq \frac{1}{|F(x,y)|} \geq \frac{1}{(d+1)H(\alpha)H(x,y)^d}.$$

The result follows once we combine the upper and lower bounds on $|F(x,y)|_p$. □

Lemma 2.13. (Siegel's lemma, [5]) *Let N and M be positive integers with $N > M$. Let $a_{i,j}$ be integers of absolute value at most $A \geq 1$ for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. Then there exist integers $t_1, \ldots, t_N$, not all zero, such that*

$$|t_i| \leq (NA)^{\frac{M}{N-M}}, \quad \sum_{i=1}^{N} a_{i,j} t_i = 0, \quad j = 1, \ldots, M.$$

Proof. See, for example, [36, Lemma 2.7]. □

Lemma 2.14. *Let α be an algebraic number of degree d over $\mathbb{Q}$. Then for every non-negative integer r there exist rational numbers $a_{r,i}$ such that*

$$\alpha^r = a_{r,d-1}\alpha^{d-1} + \cdots + a_{r,1}\alpha + a_{r,0}.$$

Furthermore, if we denote the leading coefficient of the minimal polynomial of α by c_α, and put

$$A = 1 + \max_{0\leq i\leq d-1} \{|a_{d,i}|\},$$

then $c_\alpha^{\max\{0,r-d+1\}} a_{r,i} \in \mathbb{Z}$ and $|a_{r,i}| \leq A^{\max\{0,r-d+1\}}$ for all i such that $0 \leq i \leq d-1$.

Proof. As in [36, Proposition 2.6], our proof will proceed by induction. The result trivially holds for all r such that $0 \leq r \leq d$. Now, suppose that the statement is true for some $r \geq d$. Then

$$\begin{aligned}
\alpha^{r+1} &= \alpha \cdot (a_{r,d-1}\alpha^{d-1} + a_{r,d-2}\alpha^{d-2} + \cdots + a_{r,1}\alpha + a_{r,0}) \\
&= a_{r,d-1}\alpha^d + a_{r,d-2}\alpha^{d-1} + \cdots + a_{r,1}\alpha^2 + a_{r,0}\alpha \\
&= a_{r,d-1}\left(a_{d,d-1}\alpha^{d-1} + \cdots + a_{d,1}\alpha + a_{d,0}\right) + a_{r,d-2}\alpha^{d-1} + \cdots + a_{r,1}\alpha^2 + a_{r,0}\alpha \\
&= \left(a_{r,d-1}a_{d,d-1} + a_{r,d-2}\right)\alpha^{d-1} + \cdots + (a_{r,d-1}a_{d,1} + a_{r,0})\alpha + a_{r,d-1}a_{d,0}.
\end{aligned}$$

We conclude that

$$a_{r+1,0} = a_{r,d-1}a_{d,0}, \quad a_{r+1,i} = a_{r,d-1}a_{d,i} + a_{r,i-1}, \quad 1 \leq i \leq d-1.$$

Multiplying both sides of the above equalities by c_α^{r-d+2}, we see that $c_\alpha^{r-d+2}a_{r+1,i} \in \mathbb{Z}$ for all i such that $0 \leq i \leq d-1$. Also, for $1 \leq i \leq d-1$,

$$|a_{r+1,i}| \leq |a_{r,d-1}| \cdot |a_{d,i}| + |a_{r,i-1}| \leq A^{r-d+1} \cdot \max\{1, |a_{d,i}|\} + A^{r-d+1} = A^{r-d+2}.$$

An analogous estimate holds for $|a_{r+1,0}|$, and so we conclude that the inequality $|a_{r,i}| \leq A^{\max\{0,r-d+1\}}$ is true for all r and i. □

Let α be an algebraic number and $\mathcal{O}$ the ring of integers of $\mathbb{Q}(\alpha)$. Let c_α denote the leading coefficient of the minimal polynomial of α. We define

$$\theta_\alpha = \#\mathcal{O}/\mathbb{Z}[c_\alpha\alpha]. \tag{2.2.5}$$

That is, θ_α is equal to the cardinality of the quotient ring $\mathcal{O}/\mathbb{Z}[c_\alpha\alpha]$. By definition, this quantity is finite. Recall the notion of a house of algebraic number from Definition 2.9.

Lemma 2.15. *Let α be an algebraic number of degree d over $\mathbb{Q}$ and*

$$\beta = b_{d-1}\alpha^{d-1} + \cdots + b_1\alpha + b_0,$$

where $b_0, b_1, \ldots, b_{d-1} \in \mathbb{Q}$. Let c_α, c_β denote the leading coefficients of the minimal polynomials of α, β, respectively. Then

$$\max_{1\le i\le d}\{|b_i|\} \le d \cdot \overline{|\beta|} \cdot \max_{1\le j\le d} \prod_{\substack{1\le i\le d\\ i\ne j}} \frac{1+|\alpha_i|}{|\alpha_i-\alpha_j|}.$$

Furthermore,

$$\theta_\alpha c_\beta \beta \in \mathbb{Z}[c_\alpha \alpha],$$

where θ_α is defined in (2.2.5). In particular, $\theta_\alpha c_\beta b_i \in \mathbb{Z}$ for all $i = 0, 1, \ldots, d-1$.

Proof. Let $\alpha = \alpha_1, \ldots, \alpha_d$ denote the conjugates of α. For $j = 1, \ldots, d$, let

$$\beta_j = \sum_{i=0}^{d-1} b_i \alpha_j^i.$$

Then each β_j is a conjugate of $\beta = \beta_1$. Further,

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{d-1} \\ 1 & \alpha_2 & \alpha_2^2 & \ldots & \alpha_2^{d-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_d & \alpha_d^2 & \ldots & \alpha_d^{d-1} \end{pmatrix} \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{d-1} \end{pmatrix}. \tag{2.2.6}$$

Let us denote the Vandermonde matrix on the right-hand side of the above expression by V. Then it follows from the inequality (4.1) in [16] that

$$\|V^{-1}\|_\infty \le \max_{1\le j\le d} \prod_{\substack{1\le i\le d\\ i\ne j}} \frac{1+|\alpha_i|}{|\alpha_i-\alpha_j|},$$

where $\|\cdot\|_\infty$ denotes the matrix infinity norm.

Now, let $V^{-1} = (v_{ij})$. Then it follows from (2.2.6) that $b_i = \sum_{j=1}^d v_{ij}\beta_j$, so

$$|b_i| \le \sum_{j=1}^d |v_{ij}| \cdot |\beta_j| \le d \cdot \overline{|\beta|} \cdot \max_{1\le j\le d} \prod_{\substack{1\le i\le d\\ i\ne j}} \frac{1+|\alpha_i|}{|\alpha_i-\alpha_j|}.$$

Next, note that $\theta_\alpha c_\beta \beta = (\#\mathcal{O}/\mathbb{Z}[c_\alpha \alpha]) c_\beta \beta \in \mathbb{Z}[c_\alpha \alpha]$ due to the fact that $c_\beta \beta \in \mathcal{O}$. Finally, observe that

$$\theta_\alpha c_\beta \beta = \sum_{i=0}^{d-1} \theta_\alpha c_\beta b_i \alpha^i \in \mathbb{Z}[c_\alpha \alpha].$$

Since $\mathbb{Z}[c_\alpha \alpha] \subseteq \mathbb{Z}[\alpha]$, it must be the case that each coefficient $\theta_\alpha c_\beta b_i$ is an integer. □

The following lemma is a consequence of [34, Lemma 2].

Lemma 2.16. *Let p be a rational prime and let $\overline{\mathbb{Q}_p}$ denote the algebraic closure of the field of p-aidc numbers $\mathbb{Q}_p$. Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible homogeneous polynomial of degree $d \geq 2$ and content one, and denote the roots of $F(X,1)$ by $\alpha_1, \ldots, \alpha_d \in \overline{\mathbb{Q}_p}$. Let x and y be coprime integers. If i_0 is an index with*

$$\frac{|x - \alpha_{i_0} y|_p}{\max\{1, |\alpha_{i_0}|_p\}} = \min_{1 \leq i \leq d} \left\{ \frac{|x - \alpha_i y|_p}{\max\{1, |\alpha_i|_p\}} \right\},$$

then

$$\frac{|x - \alpha_{i_0} y|_p}{\max\{1, |\alpha_{i_0}|_p\}} \leq \frac{|F(x,y)|_p}{|D(F)|_p^{1/2}}.$$

Further, if $|F(x,y)|_p < |D(F)|_p^{1/2}$, then the index i_0 above is unique and $\alpha_{i_0} \in \mathbb{Q}_p$.

2.3 Minimal Pairs

Let α be an algebraic number of degree d over $\mathbb{Q}$ and $\beta \in \mathbb{Q}(\alpha)$. With a pair (α, β) we associate two polynomials $P, Q \in \mathbb{Z}[X]$, which possess certain minimal properties listed in Definition 2.17. The properties of minimal pairs summarized in Proposition 2.19 will play a crucial role in the development of a generalized gap principle introduced in Section 2.5.

Definition 2.17. Let α be an algebraic number of degree d and $\beta \in \mathbb{Q}(\alpha)$. We say that two polynomials P, Q, not both identically equal to zero, form a *minimal pair* for (α, β), if they satisfy the following four properties:

(1) $P, Q \in \mathbb{Z}[X]$.

(2) $P(\alpha) + \beta Q(\alpha) = 0$.

(3) The quantity $\max\{\deg P, \deg Q\}$ is minimal among all polynomials satisfying the properties (1), (2).

(4) The quantity $\max\{H(P), H(Q)\}$ is minimal among all polynomials satisfying properties (1), (2), (3).

If P, Q is a minimal pair for (α, β) we write

$$r(\alpha, \beta) = \max\{\deg P, \deg Q\}.$$

Example 2.18. If P, Q is a minimal pair for (α, β) then $-P, -Q$ is also a minimal pair for (α, β). This already demonstrates that minimal pairs are not unique. Furthermore, the uniqueness is not guaranteed even if we impose an additional condition that the leading coefficient of Q is equal to one. Indeed, let

$$\alpha = 2\cos\left(\frac{2\pi}{15}\right), \quad \beta = 2\cos\left(\frac{4\pi}{15}\right).$$

Then both

$$P_1(X) = -X^2 + 2, \quad Q_1(X) = 1$$

and

$$P_2(X) = -X^2 + 2X - 1, \quad Q_2(X) = X^2 - X - 1$$

are minimal pairs for (α, β).

Let P, Q be a minimal pair for (α, β) and define a polynomial

$$R(X, Y) = P(X) + YQ(X).$$

Polynomials of such form were used by Thue [33] for the purpose of establishing the first instance of the Thue-Siegel principle. More precisely, they were constructed as to achieve high vanishing at the point (α, α), i.e., $D_i R(\alpha, \alpha) = 0$ for $i = 0, 1, \ldots, \ell$ for some large ℓ (see the exposition of Thue's method in [36, Chapter 2]). In turn, we construct $R(X, Y)$ so to achieve $R(\alpha, \beta) = 0$ for arbitrary $\beta \in \mathbb{Q}(\alpha)$, for the purpose of obtaining a generalized gap principle.

Proposition 2.19. *Let α be an algebraic number of degree d over $\mathbb{Q}$ and $\beta \in \mathbb{Q}(\alpha)$. Let P, Q be a minimal pair for (α, β) and put $r = r(\alpha, \beta)$. Then the polynomials P, Q and their Wronskian $W = PQ' - QP'$ possess the following properties.*

1.

$$0 \leq r \leq \lfloor d/2 \rfloor. \tag{2.3.1}$$

2. *P and Q are coprime.*

3. *If $\hat{P}, \hat{Q} \in \mathbb{Z}[X]$ satisfy $\hat{P}(\alpha) + \beta\hat{Q}(\alpha) = 0$ and $\max\{\deg \hat{P}, \deg \hat{Q}\} \leq d - 1 - r$, then $\hat{P} = GP$, $\hat{Q} = GQ$ for some $G \in \mathbb{Z}[X]$.*

4. *Let $\alpha = \alpha_1, \ldots, \alpha_d$ be the conjugates of α. Let c_α, c_β be the leading coefficients of the minimal polynomials of α, β, respectively. Then*

$$\max\{H(P), H(Q)\} \leq$$
$$2^{-d/2} \left(d(d+2)^2 (c_\alpha + H(\alpha))^{d/2} \cdot \theta_\alpha c_\beta |\overline{\beta}| \cdot \max_{1 \leq j \leq d} \prod_{\substack{1 \leq i \leq d \\ i \neq j}} \frac{1+|\alpha_i|}{|\alpha_i - \alpha_j|} \right)^d, \tag{2.3.2}$$

where θ_α is defined in (2.2.5).

5. *If $\alpha \in \mathbb{C}$ then*

$$|W(\alpha)| \geq \max\{1, |\alpha|\}^{2r-1} (c_\alpha^{d-1} M(\alpha))^{-(2r-1)} (4r^3 \max\{H(P), H(Q)\}^2)^{-(d-1)}, \tag{2.3.3}$$

where $M(\alpha)$ denotes the Mahler measure of α (see Definition 2.10).

If $\alpha \in \mathbb{Q}_p$ then

$$|W(\alpha)|_p \geq (d+1)^{-(2r-1)/2} (2r)^{-3d/2} H(\alpha)^{-(2r+d-2)} (r \max\{H(P), H(Q)\}^2)^{-d}. \tag{2.3.4}$$

Proof. Let us prove each of the above statements.

1. Let $s = \lfloor d/2 \rfloor$ and

$$\hat{P}(X) = \sum_{i=0}^{s} a_i X^i, \quad \hat{Q}(X) = \sum_{i=0}^{s} a_{s+1+i} X^i. \tag{2.3.5}$$

We view the $2s+2$ integer coefficients $a_0, \ldots, a_{2s+1}$ as variables. Since α is algebraic of degree d over $\mathbb{Q}$ and $\beta \in \mathbb{Q}(\alpha)$, the equation $\hat{P}(\alpha) + \beta\hat{Q}(\alpha) = 0$ defines d linear equations over $\mathbb{Q}$. Since $2s + 2 > d$, the existence of a non-trivial solution to the system of d linear equations over $\mathbb{Q}$ in $2s+2$ variables is guaranteed by Lemma 2.13. Therefore there exist non-zero polynomials $\hat{P}, \hat{Q}$ such that $\max\{\deg \hat{P}, \deg \hat{Q}\} \leq s$. Consequently, the polynomials P, Q with $\max\{\deg P, \deg Q\}$ minimal satisfy

$$\max\{\deg P, \deg Q\} \leq \max\{\deg \hat{P}, \deg \hat{Q}\} \leq s.$$

2. Let $G = \gcd(P, Q)$ and suppose that $\deg G \geq 1$. Then certainly $G(\alpha) \neq 0$, because α has degree d and $\deg G \leq \deg P < d$. Put $\hat{P} = P/G$ and $\hat{Q} = Q/G$. Then

$$\hat{P}(\alpha) + \beta\hat{Q}(\alpha) = 0$$

and

$$\max\{\deg \hat{P}, \deg \hat{Q}\} < \max\{\deg P, \deg Q\},$$

in contradiction to our assumption that $\max\{\deg P, \deg Q\}$ is minimal. Therefore $\deg G = 0$ and so P, Q are coprime.

3. Since

$$P(\alpha) + \beta Q(\alpha) = \hat{P}(\alpha) + \beta\hat{Q}(\alpha) = 0,$$

we have

$$P(\alpha)\hat{Q}(\alpha) - Q(\alpha)\hat{P}(\alpha) = 0.$$

Since α has degree d and

$$\begin{aligned}\deg\left(P\hat{Q} - Q\hat{P}\right) &\leq \max\{\deg P, \deg Q\} + \max\{\deg \hat{P}, \deg \hat{Q}\} \\ &\leq r + (d - 1 - r) \\ &\leq d - 1,\end{aligned}$$

we conclude that $P\hat{Q} - Q\hat{P}$ is identically equal to zero. If $\hat{Q} = 0$, then $\hat{P} = 0$, and so $G = 0$. Otherwise $P/Q = \hat{P}/\hat{Q}$. If we put $G = \gcd(\hat{P}, \hat{Q})$, then it becomes clear that $\hat{P} = GP$, $\hat{Q} = GQ$.

4. Define $b_i, c_{k,i} \in \mathbb{Q}$ as follows:

$$\begin{aligned}\alpha^k &= c_{k,d-1}\alpha^{d-1} + \cdots + c_{k,1}\alpha + c_{k,0}, \\ \beta &= b_{d-1}\alpha^{d-1} + \cdots + b_1\alpha + b_0.\end{aligned}$$

Let $\hat{P}, \hat{Q}$ be as in (2.3.5). Then

$$\begin{aligned}\hat{P}(\alpha) + \beta\hat{Q}(\alpha) &= \sum_{i=0}^{s} a_i\alpha^i + \left(\sum_{i=0}^{d-1} b_i\alpha^i\right) \cdot \left(\sum_{j=0}^{s} a_{s+1+j}\alpha^j\right) \\ &= \sum_{i=0}^{s} a_i\alpha^i + \sum_{j=0}^{s} a_{s+1+j} \sum_{i=0}^{d-1} b_i\alpha^{i+j} \\ &= \sum_{i=0}^{s} a_i\alpha^i + \sum_{j=0}^{s} a_{s+1+j} \left(\sum_{i=j}^{d-1} b_{i-j}\alpha^i + \sum_{k=d}^{d-1+j} b_{k-j}\alpha^k\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{s} a_i \alpha^i + \sum_{j=0}^{s} a_{s+1+j} \sum_{i=j}^{d-1} b_{i-j} \alpha^i + \sum_{j=0}^{s} a_{s+1+j} \sum_{k=d}^{d-1+j} b_{k-j} \sum_{i=0}^{d-1} c_{k,i} \alpha^i \\
&= \sum_{i=0}^{s} \left(a_i + \sum_{j=0}^{i} b_{i-j} a_{s+1+j} + \sum_{j=0}^{s} \sum_{k=d}^{d-1+j} b_{k-j} c_{k,i} a_{s+1+j} \right) \alpha^i + \\
&+ \sum_{i=s+1}^{d-1} \sum_{j=0}^{s} \left(b_{i-j} + \sum_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right) a_{s+1+j} \alpha^i \\
&= \sum_{i=0}^{d-1} L_i(\vec{a}) \alpha^i,
\end{aligned}
$$

where $\vec{a} = (a_0, a_1, \ldots, a_{2s+1})$ and

$$
L_i(\vec{a}) = \begin{cases} a_i + \sum\limits_{j=0}^{i} \left(b_{i-j} + \sum\limits_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right) a_{s+1+j} + \sum\limits_{j=i+1}^{s} \left(\sum\limits_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right) a_{s+1+j}, & \text{if } i \le s; \\ \sum\limits_{j=0}^{s} \left(b_{i-j} + \sum\limits_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right) a_{s+1+j} & \text{if } i \ge s+1. \end{cases}
$$

We conclude that the equation $\hat{P}(\alpha) + \beta \hat{Q}(\alpha) = 0$ is equivalent to the system of $d-1$ linear equations $L_0(\vec{a}) = \ldots = L_{d-1}(\vec{a}) = 0$ over $\mathbb{Q}$.

Put

$$
B = \max_{0 \le i \le d-1} \{|b_i|\}, \quad C = \max_{\substack{0 \le k \le d-1+s \\ 0 \le i \le d-1}} \{|c_{k,i}|\}.
$$

Then we can bound the (rational) coefficients of $L_i(\vec{a})$ from above by $B(1+sC)$:

$$
\left| b_{i-j} + \sum_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right| \le B + jBC \le B(1+sC),
$$

$$
\left| \sum_{k=d}^{d-1+j} b_{k-j} c_{k,i} \right| \le jBC \le sBC < B(1+sC).
$$

By Lemma 2.15 we have $\theta_\alpha c_\beta b_i \in \mathbb{Z}$ for all i and

$$
B \le d \cdot \overline{|\beta|} \cdot \max_{1 \le j \le d} \prod_{\substack{1 \le i \le d \\ i \ne j}} \frac{1 + |\alpha_i|}{|\alpha_i - \alpha_j|}.
$$

Further, by Lemma 2.14 we have $c_\alpha^{\max\{0,k-d+1\}} c_{k,i} \in \mathbb{Z}$ for all i, k and

$$C \le \left(1 + c_\alpha^{-1} H(\alpha)\right)^s.$$

Hence the linear forms

$$\hat{L}_i(\vec{a}) = \theta_\alpha c_\beta c_\alpha^s L_i(\vec{a})$$

have integer coefficients and the size of these coefficients is at most

$$\theta_\alpha c_\beta c_\alpha^s B(1 + sC).$$

In turn, this quantity does not exceed

$$A = (c_\alpha^s + s(c_\alpha + H(\alpha))^s) \cdot d \cdot \theta_\alpha c_\beta \overline{|\beta|} \cdot \max_{1 \le j \le d} \prod_{\substack{1 \le i \le d \\ i \ne j}} \frac{1 + |\alpha_i|}{|\alpha_i - \alpha_j|}.$$

By Lemma 2.13 and the inequality $c_\alpha^s + s(c_\alpha + H(\alpha))^s \le (d/2+1)(c_\alpha + H(\alpha))^s$, we have

$$\begin{aligned} \max\{H(\hat{P}), H(\hat{Q})\} &= \max_{0 \le i \le 2s+1} \{|a_i|\} \\ &\le ((2s+2)A)^{d/(2s+2-d)} \\ &\le \left(\frac{d(d+2)^2}{2} (c_\alpha + H(\alpha))^{d/2} \theta_\alpha c_\beta \overline{|\beta|} \cdot \max_{1 \le j \le d} \prod_{\substack{1 \le i \le d \\ i \ne j}} \frac{1 + |\alpha_i|}{|\alpha_i - \alpha_j|} \right)^d. \end{aligned}$$

Now that we know an upper bound on $\max\{H(\hat{P}), H(\hat{Q})\}$, we can determine the upper bound on $\max\{H(P), H(Q)\}$ by considering the following two cases.

Case 1. Suppose that $\max\{\deg \hat{P}, \deg \hat{Q}\} > d - 1 - r$. Then it follows from Part 1 and the inequality $\max\{\deg \hat{P}, \deg \hat{Q}\} \le \lfloor d/2 \rfloor$ that

$$d \le r + \max\{\deg \hat{P}, \deg \hat{Q}\} \le 2\lfloor d/2 \rfloor.$$

Thus d is even and $\max\{\deg \hat{P}, \deg \hat{Q}\} = r = d/2$. Therefore the pair $\hat{P}, \hat{Q}$ satisfies Properties (1), (2), (3) in Definition 2.17. By Property (4), the polynomials P, Q satisfy

$$\max\{H(P), H(Q)\} \le \max\{H(\hat{P}), H(\hat{Q})\},$$

and so the result follows.

Case 2. Suppose that $\max\{\deg \hat{P}, \deg \hat{Q}\} \leq d-1-r$. Then we can use Part 3 to conclude that $\hat{P} = GP$, $\hat{Q} = GQ$ for some $G \in \mathbb{Z}[X]$. Since either $\hat{P}$ or $\hat{Q}$ is non-zero, we have $H(G) \geq 1$. By Gelfond's lemma [2, Lemma 1.6.11],

$$H(P) \leq H(G)H(P) \leq 2^{\deg(GP)} H(GP) \leq 2^{d/2} H(\hat{P}).$$

An analogous estimate for $H(Q)$ yields the result.

5. Since P, Q are coprime and $r \geq 1$, they are linearly independent over $\mathbb{Q}$, so the Wronskian $W = PQ' - QP'$ is not identically equal to zero. Since α has degree d and

$$\begin{aligned} \deg W &= \deg(PQ' - QP') \\ &\leq \max\{\deg P, \deg Q\} + \max\{\deg P', \deg Q'\} \\ &\leq d/2 + (d/2 - 1) \\ &\leq d - 1, \end{aligned}$$

we conclude that $W(\alpha) \neq 0$.

With the basic properties of heights listed in [36, Section 2.4.1] we find the upper bound on $H(W)$:

$$\begin{aligned} H(W) &\leq H(PQ') + H(QP') \\ &\leq r(H(P)H(Q') + H(Q)H(P')) \\ &\leq 2r^2 H(P)H(Q) \\ &\leq 2r^2 \max\{H(P), H(Q)\}^2. \end{aligned}$$

Suppose that $\alpha \in \mathbb{C}$. Then $c_\alpha^{\deg W} W(\alpha)$ is a non-zero algebraic integer, so

$$N_{\mathbb{Q}(\alpha)/\mathbb{Q}} \left(c_\alpha^{\deg W} W(\alpha) \right) = c_\alpha^{d \deg W} \prod_{i=1}^{d} W(\alpha_i)$$

is a non-zero rational integer. Thus

$$\begin{aligned} |W(\alpha)|^{-1} &\leq c_\alpha^{d \deg W} \prod_{i=2}^{d} |W(\alpha_i)| \\ &\leq c_\alpha^{d \deg W} \prod_{i=2}^{d} (\deg W + 1) H(W) \max\{1, |\alpha_i|\}^{\deg W} \\ &\leq (2rH(W))^{d-1} \left(\frac{c_\alpha^{d-1} M(\alpha)}{\max\{1, |\alpha|\}} \right)^{2r-1}. \end{aligned}$$

Suppose that $\alpha \in \mathbb{Q}_p$. By [27, Theorem 1.3.2] there exist polynomials $\varphi(X), \psi(X) \in \mathbb{Z}[X]$ such that $\deg \varphi < \deg W$, $\deg \psi < d$, and

$$\varphi(X)f(X) + \psi(X)W(X) = \operatorname{Res}(f, W).$$

Since $\operatorname{Res}(f, W) \neq 0$ and α is a root of $f(X)$ we see that $\psi(\alpha)W(\alpha) = \operatorname{Res}(f, W)$. Since $c_\alpha^{d-1}\psi(\alpha)$ is an algebraic integer, its p-adic absolute value does not exceed one, so

$$|W(\alpha)|_p \geq |c_\alpha^{d-1}\psi(\alpha)W(\alpha)|_p = |c_\alpha^{d-1} \operatorname{Res}(f, W)|_p.$$

Further, it follows from Hadamard's inequality, as well as the upper bound on $H(W)$ established previously, that

$$\begin{aligned} |\operatorname{Res}(f, W)| &\leq (\deg f + 1)^{\deg W/2}(\deg W + 1)^{\deg f/2} H(\alpha)^{\deg W} H(W)^{\deg f} \\ &\leq (d+1)^{(2r-1)/2}(2r)^{d/2} H(\alpha)^{2r-1}(2r^2 \max\{H(P), H(Q)\}^2)^d. \end{aligned}$$

Combining the lower bound on $|W(\alpha)|_p$ with the upper bound on $|\operatorname{Res}(f, W)|$ yields the result:

$$\begin{aligned} |W(\alpha)|_p &\geq |c_\alpha^{d-1} \operatorname{Res}(f, W)|_p \\ &\geq |c_\alpha^{d-1} \operatorname{Res}(f, W)|^{-1} \\ &\geq H(\alpha)^{-(d-1)} |\operatorname{Res}(f, W)|^{-1} \\ &\geq (d+1)^{-(2r-1)/2}(2r)^{-d/2} H(\alpha)^{-(2r+d-2)}(2r^2 \max\{H(P), H(Q)\}^2)^{-d}. \end{aligned}$$

□

We conclude this section by addressing the question of computation of a minimal pair. For a pair of algebraic numbers (α, β) such that $\beta \in \mathbb{Q}(\alpha)$, define

$$\mathcal{P}(\alpha, \beta) = \left\{ P(X) + YQ(X) \colon P, Q \in \mathbb{Q}[X], \deg P \leq \frac{d}{2}, \deg Q \leq \frac{d}{2}, P(\alpha) + \beta Q(\alpha) = 0 \right\}. \tag{2.3.6}$$

It is straightforward to verify that $\mathcal{P}(\alpha, \beta)$ is a finite-dimensional vector space over $\mathbb{Q}$. Furthermore, in view of Proposition 2.19 part 1, it contains a polynomial $P(X) + YQ(X)$, where P, Q is a minimal pair for (α, β). We outline the procedure to determine P, Q in Algorithm 1.

Algorithm 1 Computation of a minimal pair

Require: A pair of algebraic numbers (α, β) such that $\beta \in \mathbb{Q}(\alpha)$
Ensure: P, Q is a minimal pair for (α, β)

1: Put $r = \lfloor \deg \alpha/2 \rfloor$.
2: Determine a basis $\{R_1, \ldots, R_k\}$ of the vector space $\mathcal{P}(\alpha, \beta)$ defined in (2.3.6).
3: For each $j = 1, \ldots, k$, define $c_0, c_1, \ldots, c_{2r+1}$ so that

$$R_j(X, Y) = \sum_{i=0}^{r} c_{2i,j} X^i + Y \sum_{i=0}^{r} c_{2i+1,j} X^i.$$

4: Compute the reduced row echelon form $M = (m_{ij})$ of a matrix

$$\begin{pmatrix} c_{2r+1,1} & c_{2r,1} & \cdots & c_{1,1} & c_{0,1} \\ c_{2r+1,2} & c_{2r,2} & \cdots & c_{1,2} & c_{0,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{2r+1,k} & c_{2r,k} & \cdots & c_{1,k} & c_{0,k} \end{pmatrix}.$$

5: Define

$$\hat{R}_{k+1-j}(X, Y) = \sum_{i=0}^{r} m_{2i,j} X^i + Y \sum_{i=0}^{r} m_{2i+1,j} X^i,$$

so that $\{\hat{R}_1, \ldots, \hat{R}_k\}$ is a new basis of $\mathcal{P}(\alpha, \beta)$ satisfying $\deg_X \hat{R}_1 \leq \ldots \leq \deg_X \hat{R}_k$. Note that, by construction, $\deg_X \hat{R}_1 < \deg_X \hat{R}_3$.
6: If $\deg_X \hat{R}_2 > \deg_X \hat{R}_1$ return NP, NQ, where N is equal to the least common multiple of the denominators of coefficients of $\hat{R}_1(X, Y) = P(X) + YQ(X)$. Otherwise proceed to Step 7.
7: Determine a basis B of the finest $\mathbb{Z}$-lattice $\mathcal{L}$ contained inside the vector space spanned by $\hat{R}_1, \hat{R}_2$.
8: Apply LLL lattice basis reduction algorithm to B, with naive height as a norm function, to compute the Minkowski-reduced basis B' of $\mathcal{L}$ [20].
9: Return P, Q, where $P(X) + YQ(X)$ is a polynomial in B' that has the smallest height.

2.4 A Gap Principle in the Presence of Vanishing

Let α, β be algebraic numbers over $\mathbb{Q}$ and $R(X,Y) \in \mathbb{Z}[X,Y]$ a non-zero irreducible polynomial such that $\deg_X R \geq 1$, $\deg_Y R \geq 1$, and $R(\alpha,\beta) = 0$. If $x_1/y_1, x_2/y_2$ are good rational approximations to α, β, respectively, such that $H(x_2,y_2) \geq H(x_1,y_1)$, then a gap principle $H(x_2,y_2) \geq C^{-1}H(x_1,y_1)^r$ holds for some $C > 0, r > 1$, provided that $R(x_1/y_1, x_2/y_2) \neq 0$ and $\max\{\deg_X R, \deg_Y R\}$ is small. This phenomenon will be studied in detail in Chapter 5. However, one cannot deduce such a conclusion when $R(x_1/y_1, x_2/y_2) = 0$. In this section, we prove Proposition 2.20, where we show that, *despite the vanishing* of R at a rational point $(x_1/y_1, x_2/y_2)$, it is still possible to produce a gap principle, provided that R is irreducible, $\deg_X R \geq 2$ and $\deg_Y R \leq 1$.

Proposition 2.20. *Let $P, Q \in \mathbb{Z}[X]$ be coprime and such that $r = \max\{\deg P, \deg Q\} \geq 1$. Let $x_1/y_1, x_2/y_2$ be rational numbers in lowest terms such that $H(x_2,y_2) \geq H(x_1,y_1)$ and*

$$P\left(\frac{x_1}{y_1}\right) + \frac{x_2}{y_2} Q\left(\frac{x_1}{y_1}\right) = 0. \tag{2.4.1}$$

Then

$$H(x_2,y_2) \geq C^{-1} H(x_1,y_1)^r,$$

where

$$C = C(P,Q) = \left(2^r (r+1)^{3(r+1)/2} \max\{H(P), H(Q)\}^{2(r+2)}\right)^r. \tag{2.4.2}$$

The proof of Proposition 2.20 is given at the end of the section. It follows directly from Lemma 2.21 and Corollary 2.23, which we will now establish.

Lemma 2.21. *Let $P, Q \in \mathbb{Z}[X]$ be coprime polynomials of degrees r and s, respectively, such that $r \geq \max\{1, s\}$. Let c_P be the leading coefficient of P and*

$$P(X,Y) = Y^r P(X/Y), \quad Q(X,Y) = Y^r (X/Y).$$

Then for all coprime integers x and y the number $g = \gcd\left(P(x,y), Q(x,y)\right)$ divides

$$\varrho = \varrho(P,Q) = \left| c_P^{r-s} \operatorname{Res}(P,Q) \right|,$$

where $\operatorname{Res}(P,Q)$ denotes the resultant of P and Q. Furthermore, if P and Q do not have a linear factor in common then

$$1 \leq \varrho \leq |c_P|^{r-s} (r+1)^{s/2} (s+1)^{r/2} \max\{H(P), H(Q)\}^{r+s}.$$

Proof. Suppose that a prime power p^n exactly divides g. Since x and y are coprime, either x or y is not divisible by p. Suppose that p does not divide y. By [27, Theorem 1.3.2] there exist polynomials $\varphi(X), \psi(X) \in \mathbb{Z}[X]$ such that

$$\varphi(X)P(X) + \psi(X)Q(X) = \operatorname{Res}(P, Q).$$

Let $t = \max\{\deg \varphi, \deg \psi\}$. We evaluate the polynomial on the left-hand side at $X = x/y$ and multiply both sides of the above equality by y^{r+t}:

$$y^t\varphi(x/y)P(x, y) + y^t\psi(x/y)Q(x, y) = \operatorname{Res}(P, Q)y^{r+t}.$$

By definition of t, the numbers $y^t\varphi(x/y)$ and $y^t\psi(x/y)$ are integers. Since p does not divide y and p^n divides both $P(x, y), Q(x, y)$, we conclude that p^n divides $\operatorname{Res}(P, Q)$.

Suppose that p divides y. Then p does not divide x, and so by analogy with the previous case we see that p^n divides $\operatorname{Res}(P(1, X), Q(1, X))$. Let $\mathcal{R}(f) = X^{\deg f} f(1/X)$ denote the reciprocal of a polynomial f. Then

$$P(1, X) = \mathcal{R}(P), \;\; Q(1, X) = X^{r-s}\mathcal{R}(Q),$$

and so

$$\begin{aligned}\operatorname{Res}(P(1, X), Q(1, X)) &= \operatorname{Res}\big(\mathcal{R}(P), X^{r-s}\mathcal{R}(Q)\big)\\ &= \operatorname{Res}(\mathcal{R}(P), X)^{r-s} \operatorname{Res}(\mathcal{R}(P), \mathcal{R}(Q))\\ &= c_P^{r-s}(-1)^{rs} \operatorname{Res}(P, Q).\end{aligned}$$

Therefore p^n divides $\big|c_P^{r-s} \operatorname{Res}(P, Q)\big|$, and the result follows.

Finally, since P, Q are coprime and $r \geq 1$, we have $\operatorname{Res}(P, Q) \neq 0$, so $\varrho \geq 1$. Applying Hadamard's inequality and $r \geq s$, we obtain

$$\begin{aligned}|\operatorname{Res}(P, Q)| &\leq (r + 1)^{s/2}(s + 1)^{r/2} H(P)^s H(Q)^r\\ &\leq (r + 1)^{s/2}(s + 1)^{r/2} \max\{H(P), H(Q)\}^{r+s}.\end{aligned}$$

□

Lemma 2.22. *Let*

$$P(X, Y) = \prod_{i=1}^{r} (\alpha_i X + \beta_i Y), \;\; Q(X, Y) = \prod_{j=1}^{r} (\gamma_j X + \delta_j Y)$$

be binary forms of degree $r \geq 1$, with complex coefficients. Let

$$C = C(P,Q) = \frac{\min_{i,j}\{|\alpha_i\delta_j - \beta_i\gamma_j|\}}{\max_{i,j}\left\{\max\{|\alpha_i| + |\gamma_j|, |\beta_i| + |\delta_j|\}\right\}}. \tag{2.4.3}$$

Suppose that $P(X,Y), Q(X,Y)$ do not have a linear factor in common, so that $C > 0$. Then for all pairs $(x,y) \in \mathbb{C}^2$ we have

$$\max\{|P(x,y)|, |Q(x,y)|\} \geq C^r H(x,y)^r.$$

Proof. We claim that either

$$\min_{i=1,\dots,r}\{|\alpha_i x + \beta_i y|\} \geq C|y| \text{ or } \min_{j=1,\dots,r}\{|\gamma_j x + \delta_j y|\} \geq C|y|.$$

For suppose not. Then for all i, j we have

$$\begin{aligned}
|(\alpha_i\delta_j - \beta_i\gamma_j)y| &= |\alpha_i(\gamma_j x + \delta_j y) - \gamma_j(\alpha_i x + \beta_i y)| \\
&\leq (|\alpha_i| + |\gamma_j|)\max\{|\alpha_i x + \beta_i y|, |\gamma_j x + \delta_j y|\} \\
&< (|\alpha_i| + |\gamma_j|)C|y| \\
&\leq \min\{|\alpha_i\delta_j - \beta_i\gamma_j|\}|y|,
\end{aligned}$$

so we reach a contradiction. Without loss of generality suppose that $\min\{|\alpha_i x + \beta_i y|\} \geq C|y|$. Then

$$|P(x,y)| = \prod_{i=1}^{r} |\alpha_i x + \beta_i y| \geq \min\{|\alpha_i x + \beta_i y|\}^r \geq C^r |y|^r.$$

Analogously, either

$$\min_{i=1,\dots,r}\{|\alpha_i x + \beta_i y|\} \geq C|x| \text{ or } \min_{j=1,\dots,r}\{|\gamma_j x + \delta_j y|\} \geq C|x|.$$

In the first case we can immediately conclude that $|P(x,y)| \geq C^r H(x,y)^r$, and the result follows. Otherwise we have $|Q(x,y)| \geq C^r|x|^r$. Combining this inequality with $|P(x,y)| \geq C^r|y|^r$ yields the result. □

Corollary 2.23. *Let $P(X), Q(X) \in \mathbb{Z}[X]$ be coprime polynomials of degrees r and s, respectively, such that $r \geq \max\{1, s\}$. Let c_P, c_Q be the leading coefficients of P, Q, respectively, and*

$$P(X,Y) = Y^r P(X/Y), \;\; Q(X,Y) = Y^r Q(X/Y).$$

Then for all pairs $(x,y) \in \mathbb{C}^2$ we have

$$\max\{|P(x,y)|, |Q(x,y)|\} \geq \frac{|c_P c_Q| H(x,y)^r}{\left(2^{r-1}(r+1)^{(3r-1)/2}\left(|c_P|^{1/r}\cdot\overline{|P|} + |c_Q|^{1/r}\cdot\overline{|Q|}\right)\max\{H(P), H(Q)\}^{2r}\right)^r}.$$

Proof. Let $\mu_1, \dots, \mu_r$ be the roots of $P(X)$ and write

$$P(X,Y) = c_P \prod_{i=1}^{r} (X - \mu_i Y) = \prod_{i=1}^{r} (\alpha_i X + \beta_i Y),$$

where $\alpha_i = c_P^{1/r}$, $\beta_i = -c_P^{1/r}\mu_i$. We consider the following two cases.

Case 1. Suppose that $s = 0$, i.e., $Q(X) = c_Q$. Then

$$Q(X,Y) = c_Q Y^r = \prod_{j=1}^{r} (\gamma_j X + \delta_j Y),$$

where $\gamma_j = 0$, $\delta_j = c_Q^{1/r}$. The constant C in (2.4.3) can be estimated from below as follows:

$$C = \frac{|c_P c_Q|^{1/r}}{\max_{i,j} \{\max\{|c_P|^{1/r}, |c_P|^{1/r}|\mu_i| + |c_Q|^{1/r}\}\}} \geq \frac{|c_P c_Q|^{1/r}}{|c_P|^{1/r}\overline{|P|} + |c_Q|^{1/r} \cdot \overline{|Q|}},$$

where $\overline{|Q|} = 1$ by definition. The result follows from Lemma 2.22.

Case 2. Suppose that $s \geq 1$. Let $\nu_1, \dots, \nu_s$ be the roots of $Q(X)$ and write

$$Q(X,Y) = c_Q Y^{r-s} \prod_{j=1}^{s} (X - \nu_j Y) = \prod_{j=1}^{r} (\gamma_j X + \delta_j Y),$$

where

$$\gamma_j = \begin{cases} c_Q^{1/r}, & \text{if } 1 \leq i \leq s; \\ 0, & \text{if } s+1 \leq i \leq r, \end{cases} \qquad \delta_j = \begin{cases} -c_Q^{1/r}\nu_i, & \text{if } 1 \leq i \leq s; \\ c_Q^{1/r}, & \text{if } s+1 \leq i \leq r. \end{cases}$$

The constant C in (2.4.3) can be estimated from below as follows:

$$C = \frac{\min_{i,j}\{|\alpha_i\delta_j - \beta_i\gamma_j|\}}{\max_{i,j} \{\max\{|\alpha_i| + |\gamma_j|, |\beta_i| + |\delta_j|\}\}} \geq \frac{|c_P c_Q|^{1/r} \min\{1, \min_{i,j}\{|\mu_i - \nu_j|\}\}}{|c_P|^{1/r} \cdot \overline{|P|} + |c_Q|^{1/r} \cdot \overline{|Q|}}. \quad (2.4.4)$$

By [6, Theorem A],

$$\min_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} \{|\mu_i - \nu_j|\} \geq 2^{1-r}(r+1)^{(1-3r)/2} \max\{H(P), H(Q)\}^{-2r}.$$

Since $P, Q \in \mathbb{Z}[X]$ and $r \geq 1$, we have $\max\{H(P), H(Q)\} \geq 1$, so the quantity on the right-hand side of the above inequality does not exceed one. Combining this lower bound on $\min_{i,j}\{|\mu_i - \nu_j|\}$ with (2.4.4), we obtain

$$C \geq \frac{|c_P c_Q|^{1/r}}{2^{r-1}(r+1)^{(3r-1)/2} \left(|c_P|^{1/r} \cdot \overline{|P|} + |c_Q|^{1/r} \cdot \overline{|Q|}\right) \max\{H(P), H(Q)\}^{2r}}.$$

The result follows from Lemma 2.22. □

Proof of Proposition 2.20. From equation (2.4.1) it follows that $Q(x_1/y_1) \neq 0$, for otherwise $P(x_1/y_1) = 0$, which means that P, Q are not coprime. Let

$$P(X,Y) = Y^r P(X/Y), \;\; Q(X,Y) = Y^r Q(X/Y).$$

Since $|y_1| \geq 1$, it must be the case that $Q(x_1, y_1) = y_1^r Q(x_1/y_1) \neq 0$, so

$$\frac{x_2}{y_2} = -\frac{P(x_1/y_1)}{Q(x_1/y_1)} = -\frac{P(x_1,y_1)}{Q(x_1,y_1)}.$$

Since x_2 and y_2 are coprime and $P(x_1,y_1), Q(x_1,y_1)$ are integers, we see that

$$|x_2| = \frac{|P(x_1,y_1)|}{g}, \;\; |y_2| = \frac{|Q(x_1,y_1)|}{g},$$

where $g = \gcd\left(P(x_1,y_1), Q(x_1,y_1)\right)$. Let c_P, c_Q be the leading coefficients of P, Q, respectively. By Lemma 2.21,

$$\begin{aligned} g &\leq \max\{|c_P|, |c_Q|\}^r (r+1)^r \max\{H(P), H(Q)\}^{2r} \\ &\leq (r+1)^r \max\{H(P), H(Q)\}^{3r}. \end{aligned}$$

Thus

$$H(x_2,y_2) = \frac{\max\{|P(x_1,y_1)|, |Q(x_1,y_1)|\}}{g} \geq \frac{\max\{|P(x_1,y_1)|, |Q(x_1,y_1)|\}}{(r+1)^r \max\{H(P), H(Q)\}^{3r}}.$$

Now, note that, for any polynomial $h \in \mathbb{Z}[X]$ of degree at least one with the leading coefficient c_h,

$$\overline{|h|} \leq \frac{M(h)}{c_h} \leq \frac{(\deg h + 1) H(h)}{c_h},$$

where $M(h)$ denotes the Mahler measure of h. Consequently,

$$|c_P|^{1/r} \cdot \overline{|P|} + |c_Q|^{1/r} \cdot \overline{|Q|} \leq 2(r+1) \max\{H(P), H(Q)\}. \tag{2.4.5}$$

Since P, Q are coprime, Corollary 2.23 applies. We utilize it along with (2.4.5) to conclude that

$$\begin{aligned} H(x_2,y_2) &\geq \frac{\max\{|P(x_1,y_1)|, |Q(x_1,y_1)|\}}{(r+1)^r \max\{H(P), H(Q)\}^{3r}} \\ &\geq \frac{|c_P c_Q| H(x,y)^r}{\left(2^{r-1}(r+1)^{(3r+1)/2} \left(|c_P|^{1/r} \cdot \overline{|P|} + |c_Q|^{1/r} \cdot \overline{|Q|}\right) \max\{H(P), H(Q)\}^{2r+3}\right)^r} \\ &\geq C^{-1} H(x_1,y_1)^r. \end{aligned}$$

□

2.5 A Generalized Gap Principle

In this section we establish a generalized gap principle for both Archimedean and non-Archimedean absolute values on $\mathbb{Q}$. Note that the numbers C_1, C_2, C_3 appearing in the statements of Lemmas 2.24, 2.25 depend on $C_0, \mu, \alpha, \beta, r$ and $\max\{H(P), H(Q)\}$. Applying the estimates (2.3.1) and (2.3.2) given in Proposition 2.19, it is possible to remove the dependency on r and $\max\{H(P), H(Q)\}$.

Lemma 2.24. (A generalized Archimedean gap principle) *Let α be a complex algebraic number of degree $d \geq 3$ over $\mathbb{Q}$ and let β be irrational and in $\mathbb{Q}(\alpha)$. Let P, Q be a minimal pair for α, β and $r = \max\{\deg P, \deg Q\}$. Let μ be a real number for which*

$$\begin{aligned} 2 < \mu \leq d, &\quad \text{if } r = 1, \\ \max\{d/r, r+1\} < \mu \leq d, &\quad \text{if } r \geq 2. \end{aligned}$$

Let C_0 be a positive real number and put

$$C_1 = \left(C_0 2^{2d+r+1}(r+1) r^{3(d-1)} (c_\alpha^{d-1} M(\alpha))^{2r-1} \max\{H(P), H(Q)\}^{2d}\right)^{1/\mu}, \tag{2.5.1}$$

$$C_2 = C_0 2^{r+1} \max\{1, |\alpha|\}^r (2 + |\beta|) \max\{H(P), H(Q)\}. \tag{2.5.2}$$

If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms, $H(x_2, y_2) \geq H(x_1, y_1) \geq C_1$ and

$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{C_0}{H(x_1, y_1)^\mu}, \quad \left|\beta - \frac{x_2}{y_2}\right| < \frac{C_0}{H(x_2, y_2)^\mu},$$

then one of the following holds.

1. $H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu - r}$.
2. $r = 1$ *and there exist integers s, t, u, v with $sv - tu \neq 0$ such that*

$$\beta = \frac{s\alpha + t}{u\alpha + v} \quad \text{and} \quad \frac{x_2}{y_2} = \frac{sx_1 + ty_1}{ux_1 + vy_1}.$$

3. $r \geq 2$ *and $H(x_1, y_1) < C_3$, where*

$$C_3 = \left(2^d H(\alpha) \max\{1, |\alpha|\}^{d+1} (C_1 C_4)^\mu\right)^{\frac{1}{r\mu - d}} \tag{2.5.3}$$

and $C_4 = C(P, Q)$ is defined in (2.4.2).

Proof. Recall that for any $h \in \mathbb{Z}[X]$ and$i$ such that $0 \le i \le \deg h$ the inequality

$$|D_i h(\alpha)| \le H(h)\binom{\deg h + 1}{i+1} \max\{1, |\alpha|\}^{\deg h - i} \tag{2.5.4}$$

holds.[2]

Define $R(X,Y) = P(X) + YQ(X)$, so that $R(\alpha,\beta) = 0$. If $R(x_1/y_1, x_2/y_2) \neq 0$, then it follows from the triangle inequality, $|\alpha - x_1/y_1| < 1$, $|\beta - x_2/y_2| < 1$, and (2.5.4), that

$$\begin{aligned}
\frac{1}{H(x_1,y_1)^r H(x_2,y_2)} &\le \left| R\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right| \\
&\le \sum_{i=0}^{r}\sum_{j=0}^{1} |D_{i,j}R(\alpha,\beta)| \left|\alpha - \frac{x_1}{y_1}\right|^i \left|\beta - \frac{x_2}{y_2}\right|^j \\
&< \frac{C_0}{H(x_1,y_1)^\mu} \sum_{i=0}^{r} \left(|D_{i,0}R(\alpha,\beta)| + |D_{i,1}R(\alpha,\beta)|\right) \\
&\le \frac{C_0}{H(x_1,y_1)^\mu} \sum_{i=0}^{r} \left(|D_i P(\alpha)| + (1+|\beta|)\cdot|D_i Q(\alpha)|\right) \\
&\le \frac{C_0}{H(x_1,y_1)^\mu} \sum_{i=0}^{r} \left(\binom{r+1}{i+1} H(P) + (1+|\beta|)\binom{r+1}{i+1} H(Q)\right) \max\{1,|\alpha|\}^{r-i} \\
&< \frac{C_2}{H(x_1,y_1)^\mu},
\end{aligned}$$

where C_2 is defined in (2.5.2). Since $\mu > r+1$ we obtain the gap principle $H(x_2,y_2) > C_2 H(x_1,y_1)^{\mu - r}$, so case 1 holds.

Suppose that $R(x_1/y_1, x_2/y_2) = 0$. If $r = 1$, then by definition $R(X,Y) = (sX+t) - Y(uX+v)$ for some integers s,t,u,v. Note that $sv - tu \neq 0$, for otherwise the number β would have to be rational. Since $R(\alpha,\beta) = R(x_1/y_1, x_2/y_2) = 0$, case 2 holds.

It remains to establish case 3. We will prove that if $r \ge 2$ then $H(x_1,y_1)$ is bounded above by C_3 given in (2.5.3). We begin by showing that

$$\left|\beta - \frac{x_2}{y_2}\right| > \frac{2C_0}{C_1^\mu \max\{1,|\alpha|\}^2} \left|\alpha - \frac{x_1}{y_1}\right|. \tag{2.5.5}$$

[2]For the derivation of this inequality see (2.2.3).

Note that

$$\left|\beta - \frac{x_2}{y_2}\right| = \left|\frac{P(\alpha)}{Q(\alpha)} - \frac{P(x_1/y_1)}{Q(x_1/y_1)}\right| = \frac{|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|}{|Q(\alpha)Q(x_1/y_1)|}. \tag{2.5.6}$$

Thus, in order to establish (2.5.5), we need to estimate $|Q(\alpha)|$, $|Q(x_1/y_1)|$ from above and $|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|$ from below. The first two are easy and essentially follow from the triangle inequality and (2.5.4):

$$|Q(\alpha)| \le (r+1)\max\{H(P), H(Q)\}\max\{1, |\alpha|\}^r, \tag{2.5.7}$$

$$\begin{aligned}\left|Q\left(\frac{x_1}{y_1}\right)\right| &\le \sum_{i=0}^{r} |D_i Q(\alpha)| \cdot \left|\alpha - \frac{x_1}{y_1}\right|^i \\ &< \sum_{i=0}^{r} |D_i Q(\alpha)| \\ &\le H(Q) \sum_{i=0}^{r} \binom{r+1}{i+1} \max\{1, |\alpha|\}^{r-i} \\ &< 2^{r+1} \max\{H(P), H(Q)\}\max\{1, |\alpha|\}^r.\end{aligned} \tag{2.5.8}$$

It remains to estimate $|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|$ from below. Let $W = PQ' - QP'$ denote the Wronskian of P and Q. By Taylor's Theorem, (2.5.4) and (2.3.3),

$$\begin{aligned}&\left|P(\alpha)Q\left(\tfrac{x_1}{y_1}\right) - Q(\alpha)P\left(\tfrac{x_1}{y_1}\right)\right| \\ &= \left|P(\alpha)\sum_{i=0}^{r} D_i Q(\alpha)\left(\tfrac{x_1}{y_1} - \alpha\right)^i - Q(\alpha)\sum_{i=0}^{r} D_i P(\alpha)\left(\tfrac{x_1}{y_1} - \alpha\right)^i\right| \\ &= \left|\alpha - \tfrac{x_1}{y_1}\right| \cdot \left|\sum_{i=0}^{r-1} \left(P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)\right)\left(\tfrac{x_1}{y_1} - \alpha\right)^i\right| \\ &= \left|\alpha - \tfrac{x_1}{y_1}\right| \cdot \left|W(\alpha) + \left(\tfrac{x_1}{y_1} - \alpha\right)\sum_{i=1}^{r-1} \left(P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)\right)\left(\tfrac{x_1}{y_1} - \alpha\right)^{i-1}\right| \\ &> \left|\alpha - \tfrac{x_1}{y_1}\right| \left(|W(\alpha)| - \tfrac{C_0}{H(x_1,y_1)^\mu}\sum_{i=1}^{r-1} |P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)|\right) \\ &\ge \left|\alpha - \tfrac{x_1}{y_1}\right| \left(|W(\alpha)| - \tfrac{C_0}{H(x_1,y_1)^\mu} 2(r+1)\max\{H(P), H(Q)\}^2 \sum_{i=1}^{r-1} \tbinom{r+1}{i+2} \max\{1, |\alpha|\}^{2r-i-1}\right) \\ &> \left|\alpha - \tfrac{x_1}{y_1}\right| \left((4r^3\max\{H(P), H(Q)\}^2)^{-(d-1)} \left(\tfrac{c_\alpha^{d-1} M(\alpha)}{\max\{1,|\alpha|\}}\right)^{-(2r-1)} - \right. \\ &\left. - \tfrac{C_0}{H(x_1,y_1)^\mu} 2^{r+2}(r+1)\max\{H(P), H(Q)\}^2 \max\{1, |\alpha|\}^{2r-2}\right) \\ &\ge \left|\alpha - \tfrac{x_1}{y_1}\right| \cdot C_0 C_1^{-\mu} 2^{r+2}(r+1)\max\{H(P), H(Q)\}^2 \max\{1, |\alpha|\}^{2r-2},\end{aligned}$$

where the last inequality follows from $H(x_1, y_1) \geq C_1$. Combining the above result with (2.5.6), (2.5.7) and (2.5.8) yields

$$\left|\beta - \frac{x_2}{y_2}\right| = \frac{|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|}{|Q(\alpha)Q(x_1/y_1)|} > \frac{2C_0}{C_1^{\mu}\max\{1, |\alpha|\}^2}\left|\alpha - \frac{x_1}{y_1}\right|.$$

By Proposition 2.20, $H(x_2, y_2) \geq C_4^{-1}H(x_1, y_1)^r$. Combining this inequality with (2.5.5) yields

$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{C_1^{\mu}\max\{1, |\alpha|\}^2}{2C_0}\left|\beta - \frac{x_2}{y_2}\right| < \frac{C_1^{\mu}\max\{1, |\alpha|\}^2}{2H(x_2, y_2)^{\mu}} \leq \frac{(C_1C_4)^{\mu}\max\{1, |\alpha|\}^2}{2H(x_1, y_1)^{r\mu}}.$$

Thus we obtain an upper bound on $|\alpha - x_1/y_1|$. On the other hand, by Lemma 2.11 we have the lower bound (2.2.1). Combining upper and lower bounds,

$$\frac{1}{2^{d+1}H(\alpha)\max\{1, |\alpha|\}^{d-1}H(x_1, y_1)^d} \leq \left|\alpha - \frac{x_1}{y_1}\right| < \frac{(C_1C_4)^{\mu}\max\{1, |\alpha|\}^2}{2H(x_1, y_1)^{r\mu}}.$$

Since $\mu > d/r$,

$$H(x_1, y_1) < \left(2^d H(\alpha)\max\{1, |\alpha|\}^{d+1}(C_1C_4)^{\mu}\right)^{1/(r\mu - d)} = C_3,$$

so case 3 holds. □

Lemma 2.25. (A generalized non-Archimedean gap principle) *Let p be a rational prime. Let $\alpha \in \mathbb{Q}_p$ be a p-adic algebraic number of degree $d \geq 3$ over $\mathbb{Q}$ and let β be irrational and in $\mathbb{Q}(\alpha)$. Let P, Q be a minimal pair for α, β and $r = \max\{\deg P, \deg Q\}$. Denote the leading coefficients of the minimal polynomials of α, β by c_α, c_β, respectively. Let μ be such that*

$$\begin{aligned} 2 < \mu \leq d, &\quad \text{if } r = 1, \\ \max\{d/r, r+1\} < \mu \leq d, &\quad \text{if } r \geq 2, \end{aligned}$$

Let C_0 be a positive real number,

$$C_1 = \left(C_0 2^{3d/2+1}(d+1)^{(2r-1)/2}r^{5d/2}c_\alpha^{2r-1}H(\alpha)^{2r+d-2}\max\{H(P), H(Q)\}^{2d}\right)^{1/\mu}, \tag{2.5.9}$$

$$C_2 = 2C_0(r+1)c_\alpha^r c_\beta \max\{H(P), H(Q)\}. \tag{2.5.10}$$

If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms, $H(x_2, y_2) \geq H(x_1, y_1) \geq C_1$ and

$$|y_1\alpha - x_1|_p < \frac{C_0}{H(x_1, y_1)^{\mu}}, \quad |y_2\beta - x_2|_p < \frac{C_0}{H(x_2, y_2)^{\mu}},$$

then one of the following holds.

1. $H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu - r}$.

2. $r = 1$ *and there exist integers* s, t, u, v *with* $sv - tu \neq 0$ *such that*

$$\beta = \frac{s\alpha + t}{u\alpha + v} \quad \textit{and} \quad \frac{x_2}{y_2} = \frac{sx_1 + ty_1}{ux_1 + vy_1}.$$

3. $r \geq 2$ *and* $H(x_1, y_1) < C_3$, *where*

$$C_3 = \left((d+1)H(\alpha)c_\alpha^{d-r}(C_1C_4)^\mu\right)^{1/(r\mu - d)} \tag{2.5.11}$$

and $C_4 = C(P, Q)$ *is defined in (2.4.2).*

Proof. Define $R(X, Y) = P(X) + YQ(X)$, so that $R(\alpha, \beta) = 0$. Suppose that $R(x_1/y_1, x_2/y_2) \neq 0$. Then the following trivial lower bound holds:

$$\begin{aligned}\left| y_1^r y_2 R\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right|_p &\geq \frac{1}{|y_1^r y_2 R(x_1/y_1, x_2/y_2)|} \\ &\geq \frac{1}{2(r+1)\max\{H(P), H(Q)\}H(x_1, y_1)^r H(x_2, y_2)}.\end{aligned}$$

Note that for each $(i, j) \in \{0, \ldots, r\} \times \{0, 1\}$ the p-adic number $c_\alpha^{r-i} c_\beta^{1-j} D_{ij} R(\alpha, \beta)$ is an algebraic integer. Thus its p-adic absolute value does not exceed one. Via the application of Taylor's Theorem we obtain the following upper bound:

$$\begin{aligned}\left| y_1^r y_2 R\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right|_p &\leq \max_{(i,j)\neq(0,0)} \left\{ |D_{ij}R(\alpha, \beta)|_p \cdot |y_1\alpha - x_1|_p^i \cdot |y_2\beta - x_2|_p^j \right\} \\ &= \max_{(i,j)\neq(0,0)} \left\{ \left| \frac{c_\alpha^{r-i} c_\beta^{1-j} D_{ij} R(\alpha, \beta)}{c_\alpha^{r-i} c_\beta^{1-j}} \right|_p \cdot |y_1\alpha - x_1|_p^i \cdot |y_2\beta - x_2|_p^j \right\} \\ &\leq |c_\alpha^r c_\beta|_p^{-1} \max_{(i,j)\neq(0,0)} \left\{ |y_1 c_\alpha \alpha - c_\alpha x_1|_p^i \cdot |y_2 c_\beta \beta - c_\beta x_2|_p^j \right\} \\ &\leq c_\alpha^r c_\beta \max\{|y_1 c_\alpha \alpha - c_\alpha x_1|_p, |y_2 c_\beta \beta - c_\beta x_2|_p\} \\ &\leq c_\alpha^r c_\beta \max\{|c_\alpha|_p, |c_\beta|_p\} \max\{|y_1\alpha - x_1|_p, |y_2\beta - x_2|_p\} \\ &< \frac{C_0 c_\alpha^r c_\beta}{H(x_1, y_1)^\mu},\end{aligned}$$

Upon combining the upper and lower bounds we obtain the inequality

$$\frac{1}{2(r+1)\max\{H(P), H(Q)\}H(x_1, y_1)^r H(x_2, y_2)} < \frac{C_0 c_\alpha^r c_\beta}{H(x_1, y_1)^\mu},$$

which is equivalent to $H(x_2, y_2) > C_2 H(x_1, y_1)^{\mu - r}$, where C_2 is defined in (2.5.10). Therefore case 1 holds.

Suppose that $R(x_1/y_1, x_2/y_2) = 0$. If $r = 1$, then by definition $R(X, Y) = (sX + t) - Y(uX + v)$ for some integers s, t, u, v. Note that $sv - tu \neq 0$, for otherwise the number β would have to be rational. Since $R(\alpha, \beta) = R(x_1/y_1, x_2/y_2) = 0$, case 2 holds.

It remains to establish case 3. We will prove that if $r \geq 2$ then $H(x_1, y_1)$ is bounded above by a number specified in (2.5.11). We begin by showing that

$$|y_2\beta - x_2|_p > \frac{C_0 c_\alpha^{r-1}}{C_1^\mu} |y_1\alpha - x_1|_p . \tag{2.5.12}$$

The equation $R(x_1/y_1, x_2/y_2) = 0$ implies that

$$x_2 = \operatorname{sign}(Q(x_1, y_1)) \frac{P(x_1, y_1)}{g}, \quad y_2 = \frac{|Q(x_1, y_1)|}{g},$$

where $g = \gcd(P(x_1, y_1), Q(x_1, y_1))$. Consequently,

$$|y_2\beta - x_2|_p = \frac{|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p}{|gQ(\alpha)|_p} \tag{2.5.13}$$

Thus, in order to establish (2.5.12), we need to estimate $|g|_p$, $|Q(\alpha)|_p$ from above and $|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p$ from below. The first two are easy: since g is an integer, we have

$$|g|_p \leq 1. \tag{2.5.14}$$

Since $c_\alpha^r Q(\alpha)$ is an algebraic integer,

$$|Q(\alpha)|_p \leq |c_\alpha|_p^{-r} \leq c_\alpha^r. \tag{2.5.15}$$

It remains to estimate $|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p$ from below. Before we proceed, note that for any i the number $c_\alpha^{2r-i-1}(P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha))$ is an algebraic integer, so its p-adic absolute value does not exceed one. Consequently,

$$\begin{aligned} \left| \frac{P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)}{c_\alpha^i} \right|_p &= \left| \frac{c_\alpha^{2r-i-1}(P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha))}{c_\alpha^{2r-1}} \right|_p \\ &\leq |c_\alpha|_p^{-(2r-1)} \\ &\leq c_\alpha^{2r-1}, \end{aligned}$$

Now let $W = PQ' - QP'$ denote the Wronskian of P and Q. By Taylor's Theorem and (2.3.4),

$$\begin{aligned}
&|P(\alpha)Q(x_1,y_1) - Q(\alpha)P(x_1,y_1)|_p \\
&= \left| P(\alpha) \sum_{i=0}^{r} D_i Q(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} - Q(\alpha) \sum_{i=0}^{r} D_i P(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} \right|_p \\
&= |y_1\alpha - x_1|_p \left| \sum_{i=0}^{r-1} (P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)) (x_1 - \alpha y_1)^i y_1^{r-1-i} \right|_p \\
&\geq |y_1\alpha - x_1|_p \left(|W(\alpha)|_p - |y_1\alpha - x_1|_p \max_{i=0,\dots,r-1} \left\{ \left| \tfrac{P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)}{c_\alpha^i} \right|_p \right\} \right) \\
&> |y_1\alpha - x_1|_p \left(\left| c_\alpha^{d-1} \operatorname{Res}(f,W) \right|_p - \tfrac{C_0}{H(x_1,y_1)^\mu} c_\alpha^{2r-1} \right) \\
&\geq |y_1\alpha - x_1|_p \cdot \left((d+1)^{-(2r-1)/2} (2r)^{-3d/2} H(\alpha)^{-(2r+d-2)} (r \max\{H(P), H(Q)\}^2)^{-d} - \tfrac{C_0}{H(x_1,y_1)^\mu} c_\alpha^{2r-1} \right) \\
&\geq |y_1\alpha - x_1|_p \cdot C_0 C_1^{-\mu} c_\alpha^{2r-1},
\end{aligned}$$

where the last inequality follows from $H(x_1, y_1) \geq C_1$. Combining the above result with (2.5.13), (2.5.14) and (2.5.15), we obtain

$$|y_2\beta - x_2|_p = \frac{|P(\alpha)Q(x_1,y_1) - Q(\alpha)P(x_1,y_1)|_p}{|gQ(\alpha)|_p} > \frac{C_0 c_\alpha^{r-1}}{C_1^\mu} |y_1\alpha - x_1|_p.$$

By Proposition 2.20, $H(x_2, y_2) \geq C_4^{-1} H(x_1, y_1)^r$. Combining this inequality with (2.5.12) yields

$$|y_1\alpha - x_1|_p < \frac{C_1^\mu}{C_0 c_\alpha^{r-1}} |y_2\beta - x_2|_p < \frac{C_1^\mu}{c_\alpha^{r-1} H(x_2,y_2)^\mu} \leq \frac{(C_1C_4)^\mu}{c_\alpha^{r-1} H(x_1,y_1)^{r\mu}}.$$

Thus we obtain an upper bound on $|y_1\alpha - x_1|_p$. On the other hand, by Lemma 2.12 we have the lower bound (2.2.4). Combining upper and lower bounds,

$$\frac{1}{(d+1)H(\alpha)c_\alpha^{d-1} H(x_1,y_1)^d} \leq |y_1\alpha - x_1|_p < \frac{(C_1C_4)^\mu}{c_\alpha^{r-1} H(x_1,y_1)^{r\mu}}.$$

Since $r \geq 2$, we have

$$H(x_1, y_1) < \left((d+1)H(\alpha) c_\alpha^{d-r} (C_1C_4)^\mu \right)^{1/(r\mu - d)} = C_3,$$

so case 3 holds. □

2.6 Counting Approximations of Large Height

In this section we prove Theorem 2.31, where we count approximations x/y, with large height, to distinct algebraic numbers $\alpha_1, \dots, \alpha_n$ such that $\mathbb{Q}(\alpha_i) = \mathbb{Q}(\alpha_1)$ for all $i = 1, 2, \dots, n$.

Before we proceed, we need to prove a variant of the well-known result of Lewis and Mahler [21], and recall the statement of the Thue-Siegel principle. In each of the results stated below, we fix a field $K = \mathbb{C}$ or $K = \mathbb{Q}_p$ and let $|\ |$ denote the standard absolute value on K; that is, $|\ |$ is the Archimedean absolute value if $K = \mathbb{C}$ and $|\ | = |\ |_p$ is the p-adic absolute value if $K = \mathbb{Q}_p$, normalized so that $|p| = p^{-1}$.

Lemma 2.26. *Let*

$$F(X, Y) = c_d X^d + c_{d-1} X^{d-1} Y + \cdots + c_0 Y^d$$

be a binary form of degree $d \geq 2$ with integer coefficients such that $c_0 c_d \neq 0$. Let x, y be non-zero integers. There exists a root α of $F(X, 1)$ such that

$$\min\left\{\left|\alpha - \frac{x}{y}\right|, \left|\alpha^{-1} - \frac{y}{x}\right|\right\} \leq \frac{C|F(x, y)|}{H(x, y)^d},$$

where

$$C = \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2}}{|D(F)|^{1/2}}.$$

Proof. Let α be a root of $F(X, 1)$ that minimizes $|\alpha - x/y|$. By [31, Lemma 3],

$$\left|\alpha - \frac{x}{y}\right| \leq \frac{C|F(x, y)|}{|y|^d}.$$

If $|y| \geq |x|$ then $H(x, y) = |y|$, and so the result holds. Otherwise, since $c_0 c_d \neq 0$, we see that the roots of $F(X, 1)$ and $F(1, X)$ are non-zero, meaning that all roots of $F(1, X)$ are of the form α^{-1}, where α is a root of $F(X, 1)$. If we let β^{-1} be a root of $F(1, X)$ that minimizes $|\beta^{-1} - y/x|$, then it follows from [31, Lemma 3] that

$$\left|\beta^{-1} - \frac{y}{x}\right| \leq \frac{C|F(x, y)|}{|x|^d}.$$

Since $|x| > |y|$, the result follows. □

Lemma 2.27. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let α, β be distinct numbers, each algebraic over $\mathbb{Q}$. Let C_0, μ be positive real numbers.*

If x/y is a rational number such that $H(x, y) \geq (2C_0/|\alpha - \beta|)^{1/\mu}$ and

$$\left|\alpha - \frac{x}{y}\right| < \frac{C_0}{H(x, y)^\mu}$$

then

$$\left|\beta - \frac{x}{y}\right| \geq \frac{C_0}{H(x,y)^\mu}.$$

Proof. Suppose that the statement is false. Then it follows from the triangle inequality that

$$|\alpha - \beta| \leq \left|\alpha - \frac{x}{y}\right| + \left|\beta - \frac{x}{y}\right| < \frac{2C_0}{H(x,y)^\mu},$$

and so $H(x,y) < (2C_0/|\alpha-\beta|)^{1/\mu}$, leading us to a contradiction. □

Corollary 2.28. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and let $\overline{K}$ denote the algebraic closure of K. Denote the standard absolute value on K by $|\ |$. Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree $d \geq 2$ with roots $\alpha_1, \ldots, \alpha_d \in \overline{K}$. Let C_0, μ be positive real numbers and define*

$$C = C(C_0, \mu, f) = \left(C_0(d+1)^{3d/2} H(f)^{d-1}\right)^{1/\mu}.$$

If x/y is a rational number such that $H(x,y) \geq C$ and

$$\left|\alpha_i - \frac{x}{y}\right| < \frac{C_0}{H(x,y)^\mu}$$

for some $i \in \{1, \ldots, d\}$ then

$$\left|\alpha_j - \frac{x}{y}\right| \geq \frac{C_0}{H(x,y)^\mu}$$

for all $j \neq i$.

Proof. By Lemma 2.27, it is sufficient to verify that the inequality

$$C \geq (2C_0)^{1/\mu} \cdot \min_{1 \leq i < j \leq d} \{|\alpha_i - \alpha_j|\}^{-1/\mu}$$

holds.

If $K = \mathbb{C}$ then it follows from Mahler's lemma [26, Lemma 1.1] that

$$\min_{1 \leq i < j \leq d} \{|\alpha_i - \alpha_j|\} \geq \frac{\sqrt{3}}{(d+1)^{(2d+1)/2} H(f)^{d-1}}.$$

If $K = \mathbb{Q}_p$ then it follows from [26, Lemma 2.3] that

$$\min_{1 \leq i < j \leq d} \{|\alpha_i - \alpha_j|\} \geq \frac{1}{d^{3d/2} H(f)^{d-1}}.$$

Thus

$$\begin{aligned}\frac{2C_0}{\min\limits_{1\le i<j\le d}\{|\alpha_i-\alpha_j|\}} &\le 2C_0\max\left\{3^{-1/2}(d+1)^{(2d+1)/2}, d^{3d/2}\right\}H(f)^{d-1}\\ &< C_0(d+1)^{3d/2}H(f)^{d-1}\\ &= C^{\mu},\end{aligned}$$

as claimed. □

Lemma 2.29. (Thue-Siegel Principle) *Let $K=\mathbb{C}$ or $\mathbb{Q}_p$, and denote the standard absolute value on K by $|\ |$. Let $\alpha\in K$ be an algebraic number of degree $d\ge 3$ over $\mathbb{Q}$ and $\beta\in\mathbb{Q}(\alpha)$. Let t,τ be such that*

$$\frac{2+\sqrt{2d^3+2d^2-4d}}{d(d+1)} < t < \sqrt{\frac{2}{d}},\quad \sqrt{2-dt^2} < \tau < t-\frac{2}{d}, \tag{2.6.1}$$

and put $\lambda = 2/(t-\tau)$, so that $\lambda < d$. Define

$$A_1 = \frac{t^2}{2-dt^2}\left(\log M(\alpha)+\frac{d}{2}\right),\quad A_2 = \frac{t^2}{2-dt^2}\left(\log M(\beta)+\frac{d}{2}\right).$$

Let $x_1/y_1, x_2/y_2$ be rational numbers that satisfy the inequalities

$$\left|\alpha-\frac{x_1}{y_1}\right| < \frac{1}{(4e^{A_1}H(x_1,y_1))^{\lambda}},\quad \left|\beta-\frac{x_2}{y_2}\right| < \frac{1}{(4e^{A_2}H(x_2,y_2))^{\lambda}}.$$

If $K=\mathbb{Q}_p$ we also impose the condition $|x_1|=|x_2|=1$. Then

$$\log(4e^{A_2})+\log H(x_2,y_2) \le \delta^{-1}\left(\log(4e^{A_1})+\log H(x_1,y_1)\right).$$

Proof. Note that since $d\ge 3$ the intervals in (2.6.1) are guaranteed to be non-empty, so the statement is not vacuous. When $|\alpha|\le 1$ the proof is as in [3] with $\alpha_1\ne\alpha_2$ and comments from [4]. If $|\alpha|>1$ we have

$$\left|\alpha^{-1}-\frac{y_1}{x_1}\right| \le |\alpha|^{-1}|y_1/x_1|\left(4e^{A_1}H(x_1,y_1)\right)^{-\lambda} < \left(3e^{A_1}H(x_1,y_1)\right)^{-\lambda}$$

for both $K=\mathbb{R}$ and $K=\mathbb{Q}_p$. Indeed, the first case was considered in [4]. The second case is true due to our additional assumption that $|x_1|=1$. Analogous observations apply to $\beta, x_2/y_2$ in place of $\alpha, x_1/y_1$. □

Definition 2.30. Let α be an irrational number. The *orbit* of α is the collection

$$\operatorname{orb}(\alpha) = \left\{ \frac{v\alpha - u}{-t\alpha + s} \colon s,t,u,v \in \mathbb{Z}, sv - tu \neq 0 \right\}.$$

Theorem 2.31. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let $\alpha_1 \in K$ be an algebraic number of degree $d \geq 3$ over $\mathbb{Q}$ and $\alpha_2, \alpha_3, \ldots, \alpha_n$ be distinct elements of $\mathbb{Q}(\alpha_1)$, different from α_1, each of degree d. Let μ be such that*

$$0.5d + 1 < \mu \leq d.$$

Let C_0 be a real number such that $C_0 > (4e^A)^{-1}$, where

$$A = 500^2 \left(\log \max_{i=1,\ldots,n} \{M(\alpha_i)\} + \frac{d}{2} \right). \tag{2.6.2}$$

There exists an explicitly computable positive number C_1, which depends on $C_0, \mu, \alpha_1, \alpha_2, \ldots, \alpha_n$, but not on p in the case $K = \mathbb{Q}_p$, with the following property.

The total number of rationals x/y in lowest terms, which satisfy $H(x,y) \geq C_1$,

$$\left| \alpha_j - \frac{x}{y} \right| < \frac{C_0}{H(x,y)^\mu} \tag{2.6.3}$$

for some $j \in \{1, 2, \ldots, n\}$, and $|x| = 1$ if $K = \mathbb{Q}_p$, is less than

$$\gamma \left\lfloor 1 + \frac{11.51 + 1.5 \log d + \log \mu}{\log(\mu - 0.5d)} \right\rfloor,$$

where

$$\gamma = \max\{\gamma_1, \ldots, \gamma_n\}, \quad \gamma_i = \#\{j \colon \alpha_j \in \operatorname{orb}(\alpha_i)\}. \tag{2.6.4}$$

Proof. Throughout the proof we will be adjusting our choice of C_1 four times. More precisely, the value of C_1 is chosen so to satisfy (2.6.5), (2.6.8), (2.6.11), and (2.6.13).

Let C_1 be such that

$$C_1 \geq (2C_0 / \min\{|\alpha_i - \alpha_j|\})^{1/\mu}. \tag{2.6.5}$$

Then it follows from Lemma 2.27 that for each x/y satisfying (2.6.3) the index $j \in \{1, 2, \ldots, n\}$ is unique.

Let $x_1/y_1, x_2/y_2, \ldots, x_\ell/y_\ell$ be the list of rational numbers that satisfy the following conditions.

1. $C_1 \le H(x_1, y_1) \le H(x_2, y_2) \le \ldots \le H(x_\ell, y_\ell)$.
2. $\gcd(x_j, y_j) = 1$ for all $j = 1, 2, \ldots, \ell$.
3. For each $j \in \{1, 2, \ldots, \ell\}$, there exists the index $i_j \in \{1, 2, \ldots, n\}$ such that
$$\left| \alpha_{i_j} - \frac{x_j}{y_j} \right| < \frac{C_0}{H(x_j, y_j)^\mu}.$$
By the discussion above, this index is unique.
4. For every $j, k \in \{1, 2, \ldots, \ell\}$, if $\alpha_{i_k} \in \mathrm{orb}(\alpha_{i_j})$, i.e., $\alpha_{i_k} = (s\alpha_{i_j} + t)/(u\alpha_{i_j} + v)$ for some integers s, t, u, v, then $x_k/y_k \ne (sx_j + ty_j)/(ux_j + vy_j)$.

Due to the fourth condition this list need not be uniquely defined. This fact, however, does not affect our estimates. The fourth property requires additional clarification: to each approximation in the list
$$x_1/y_1, x_2/y_2, \ldots, x_\ell/y_\ell$$
correspond several approximations, which we call *derived*. To be more precise, from x_j/y_j one can naturally construct a (possibly bad) rational approximation to arbitrary $\alpha \in \mathrm{orb}(\alpha_{i_j})$ as follows. Let
$$\alpha = \frac{s\alpha_{i_j} + t}{u\alpha_{i_j} + v} \quad \text{and} \quad \frac{x'_j}{y'_j} = \frac{sx_j + ty_j}{ux_j + vy_j}$$
for some integers s, t, u, v. Then
$$\alpha - \frac{x'_j}{y'_j} = \frac{tu - sv}{(u\alpha + v)(u(x_j/y_j) + v)} \left(\alpha_{i_j} - \frac{x_j}{y_j} \right),$$
so rational approximations to α and α_{i_j} are connected. Thus, by imposing condition (4), we insist that x'_j/y'_j does not appear in the list $x_1/y_1, x_2/y_2, \ldots, x_\ell/y_\ell$.

In order to account for the presence of derived rational approximations, we introduce the value γ_i defined in (2.6.4). Note that the value γ_{i_j} is equal to the number of rational approximations derived from x_j/y_j, including x_j/y_j itself. Consequently, if we let N denote the total number of rationals satisfying the conditions specified in the hypothesis, then N does not exceed $\sum_{j=1}^{\ell} \gamma_{i_j}$. Therefore
$$N \le \sum_{i=1}^{\ell} \gamma_{i_j} \le \gamma\ell,$$

where γ is defined in (2.6.4). Thus it remains to estimate ℓ.

To derive an upper bound on ℓ, we begin by applying a generalized gap principle to the ordered pair $(\alpha_{i_k}, \alpha_{i_{k+1}})$. Let

$$C_2 = \max_{j,k}\{C_2(C_0, \mu, \alpha_j, \alpha_k)\},$$

where $C_2(C_0, \mu, \alpha_j, \alpha_k)$ is defined in (2.5.2) if $K = \mathbb{R}$ or in (2.5.10) if $K = \mathbb{Q}_p$, with $(\alpha, \beta) = (\alpha_j, \alpha_k)$. Let $r(j,k)$ be the number r from Proposition 2.19 with $(\alpha, \beta) = (\alpha_j, \alpha_k)$. Define

$$E = \mu - \max_{j,k}\{r_{jk}\}. \tag{2.6.6}$$

It follows from (2.3.1) that

$$E \geq \mu - 0.5d. \tag{2.6.7}$$

Note that if $K = \mathbb{Q}_p$ then $|y_1| \leq 1$ and so

$$|y_1\alpha - x_1| = |y_1| \cdot \left|\alpha - \frac{x_1}{y_1}\right| < \frac{C_0}{H(x_1, y_1)^\mu}.$$

Analogously,

$$|y_2\beta - x_2| < \frac{C_0}{H(x_2, y_2)^\mu}.$$

Let $C_1(j,k), C_3(j,k)$ denote the constants (2.5.1), (2.5.3) if $K = \mathbb{C}$ or (2.5.9), (2.5.11) if $K = \mathbb{Q}_p$. We adjust C_1 by choosing it so that

$$C_1 \geq \max\left\{\max_{j,k}\{C_1(j,k)\}, \max_{\substack{j,k \\ r(j,k)\geq 2}}\{C_3(j,k)\}\right\}. \tag{2.6.8}$$

With such a choice of C_1, case 3 in Lemmas 2.24 or 2.25 cannot hold. With our choice of x_i/y_i we have also excluded case 2. Therefore only case 1 remains, i.e., it is possible to apply our generalized gap principle to the ordered pair $(\alpha, \beta) = (\alpha_{i_k}, \alpha_{i_{k+1}})$:

$$\log H(x_{k+1}, y_{k+1}) > (\mu - r(i_k, i_{k+1}))\log H(x_k, y_k) - \log C_2(i_k, i_{k+1}) \geq E\log H(x_k, y_k) - \log C_2$$

for any $k \in \{1, 2, \ldots, \ell - 1\}$, where E is defined in (2.6.6). Consequently,

$$\begin{aligned}
\log H(x_\ell, y_\ell) &> E\log H(x_{\ell-1}, y_{\ell-1}) - \log C_2 \\
&> E^2\log H(x_{\ell-2}, y_{\ell-2}) - (1+E)\log C_2 \\
&> \cdots \\
&> E^{\ell-1}\log H(x_1, y_1) - (1 + E + \cdots + E^{\ell-2})\log C_2.
\end{aligned}$$

Thus we obtain the following lower bound on $\log H(x_\ell, y_\ell)$:

$$\log H(x_\ell, y_\ell) > E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1}-1}{E-1} \log C_2. \tag{2.6.9}$$

Next, we apply the Thue-Siegel principle from Lemma 2.29 to the pair $(\alpha, \beta) = (\alpha_{i_1}, \alpha_{i_\ell})$. Observe that, since all α_i have degree d, we have $\mathbb{Q}(\alpha_{i_1}) = \mathbb{Q}(\alpha_{i_\ell})$, so $\alpha_{i_\ell} \in \mathbb{Q}(\alpha_{i_1})$. For $a = 1/500$, set

$$t = \sqrt{\frac{2}{d+a^2}}, \quad \tau = 2at.$$

Then

$$\lambda = \frac{2}{t-\tau} = \frac{2}{(1-2a)t} < 1.42\sqrt{d}.$$

Further,

$$\frac{t^2}{2-dt^2} = \frac{1}{a^2} = 500^2,$$

$$A_1 = 500^2 \left(\log M(\alpha_{i_1}) + \frac{d}{2} \right), \quad A_\ell = 500^2 \left(\log M(\alpha_{i_\ell}) + \frac{d}{2} \right),$$

$$\delta = \frac{dt^2 + \tau^2 - 2}{d-1} = \frac{6a^2}{(d+a^2)(d-1)}.$$

Note that

$$\delta^{-1} < 41667 d^2. \tag{2.6.10}$$

We further adjust our definition of C_1 by choosing it so that

$$C_1 \geq C_0^{\frac{1}{\mu - 1.42\sqrt{d}}} \left(4e^A\right)^{\frac{1.42\sqrt{d}}{\mu - 1.42\sqrt{d}}}, \tag{2.6.11}$$

where A is defined in (2.6.2). Now with the help of inequalities $\lambda < 1.42\sqrt{d}$ and $H(x_j, y_j) \geq C_1$ we obtain

$$\left| \alpha_{i_j} - \frac{x_j}{y_j} \right| < \frac{C_0}{H(x_j, y_j)^\mu} \leq \frac{1}{\left(4e^A H(x_j, y_j)\right)^{1.42\sqrt{d}}} < \frac{1}{\left(4e^A H(x_j, y_j)\right)^\lambda},$$

so that the hypothesis of Lemma 2.29 is satisfied. Thus we arrive at the conclusion that

$$\begin{aligned} \log H(x_\ell, y_\ell) &\leq \delta^{-1} \left(\log(4e^{A_1}) + \log H(x_1, y_1) \right) - \log(4e^{A_\ell}) \\ &< 41667 d^2 \left(\log(4e^{A_1}) + \log H(x_1, y_1) \right), \end{aligned}$$

where the last inequality follows from (2.6.10). Thus

$$\log H(x_\ell, y_\ell) < 41667d^2 \left(\log\left(4e^{A_1}\right) + \log H(x_1, y_1)\right).$$

We combine the above upper bound on $\log H(x_\ell, y_\ell)$ with the lower bound given in (2.6.9):

$$E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C_2 < 41667d^2 \left(\log\left(4e^{A_1}\right) + \log H(x_1, y_1)\right).$$

Reordering the terms yields

$$\left(E^{\ell-1} - 41667d^2\right) \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C_2 < 41667d^2 \log\left(4e^{A_1}\right). \tag{2.6.12}$$

Let us assume that

$$\ell \geq 1 + \frac{\log(41667d^2)}{\log(\mu - 0.5d)},$$

for otherwise the statement of our theorem holds. Then $E^{\ell-1} \geq 41667d^2$, so we may use the inequality $H(x_1, y_1) \geq C_1$ to replace $H(x_1, y_1)$ with C_1 in (2.6.12):

$$\left(E^{\ell-1} - 41667d^2\right) \log C_1 - \frac{E^{\ell-1} - 1}{E - 1} \log C_2 < 41667d^2 \log\left(4e^{A_1}\right).$$

From (2.6.7) we obtain

$$(\mu - 0.5d)^{\ell-1} \left(\log C_1 - \frac{\log C_2}{E - 1}\right) < 41667d^2 \log B_0 + 41667d^2 \log\left(4e^{A_1}\right) + \frac{\log C_2}{E - 1}.$$

We make a final adjustment to C_1 by choosing it so that

$$C_1 \geq C_2^{2/(E-1)}. \tag{2.6.13}$$

Then

$$(\mu - 0.5d)^{\ell-1} \frac{\log C_1}{2} < \left(41667d^2 + \frac{1}{2}\right) \log C_1 + 41667d^2 \log\left(4e^{A_1}\right),$$

leading us to a conclusion

$$(\mu - 0.5d)^{\ell-1} < 1 + 83334d^2 \left(1 + \frac{\log\left(4e^{A_1}\right)}{\log C_1}\right). \tag{2.6.14}$$

By our choice of C_1,

$$\log C_1 \geq \frac{1}{\mu - 1.42\sqrt{d}} \log C_0 + \frac{1.42\sqrt{d}}{\mu - 1.42\sqrt{d}} \log(4e^A),$$

which means that

$$\frac{\log(4e^A)}{\log C_1} \leq \frac{\mu - 1.42\sqrt{d}}{1.42\sqrt{d} + \log C_0/\log(4e^A)} < \frac{\mu - 1.42\sqrt{d}}{1.42\sqrt{d} - 1},$$

where the last inequality follows from the fact that $C_0 > (4e^A)^{-1}$. Plugging the above inequality into (2.6.14), we obtain

$$(\mu - 0.5d)^{\ell - 1} < 1 + 83334d^2 \left(1 + \frac{\mu - 1.42\sqrt{d}}{1.42\sqrt{d} - 1}\right) = 1 + 83334d^2 \frac{\mu - 1}{1.42\sqrt{d} - 1} \leq 1 + 98896d^{3/2}\mu,$$

where the last inequality follows from $d \geq 3$. We conclude that

$$\ell < 1 + \frac{\log(98897d^{3/2}\mu)}{\log(\mu - 0.5d)} < 1 + \frac{11.51 + 1.5\log d + \log \mu}{\log(\mu - 0.5d)}.$$

The result follows once we multiply the right-hand side by the constant γ defined in (2.6.4). □

It is worth noting that the idea of considering distinct algebraic numbers $\alpha_1, \ldots, \alpha_n$ that generate the same field extension of $\mathbb{Q}$ is not new. In the case $n = 2$ it is present in the Thue-Siegel principle of Bombieri [1] and Bombieri and Mueller [3], which we utilized in the proof of Theorem 2.31. It can also be found, for example, in the fundamental monograph of Schmidt [30, Theorem 6D]. However, he Schmidt was mostly interested in applications to the case $\alpha = \alpha_1 = \cdots = \alpha_n$. By using a result of Esnault and Viehweg [14], he provided an estimate for the number of rationals x/y, in lowest terms, satisfying

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{y^\mu},$$

where $\mu > 2$ is fixed [30, Theorem 9B].

Our results are two-fold in their influence. Consider Theorem 2.31 with $C_0 = 1$ and $\mu = (3d + 6)/4$. First, we established the existence of an explicitly computable constant $C_1 = C_1(\alpha_1, \ldots, \alpha_n)$ such that the number of rationals x/y, in lowest terms, satisfying both

$H(x,y) \geq C_1$ and (2.6.3) does not exceed 18γ. A combination of Corollary 2.32 and Lemma 2.5 shows that this quantity does not exceed 216, provided that $\alpha_1, \dots, \alpha_n$ are conjugates. Compare this to the estimate of Schmidt, which is, though proportional to $\log^+(\log H(\alpha))$, applies to all x/y, not just those of large height.[3] It is remarkable that Schmidt's estimate holds for all $\mu > 2$, and it is an interesting problem to generalize Theorem 2.31 to the case when $2 < \mu \leq d/2 + 1$.

Second, not only that we determined $C_1 = C_1(\alpha_1, \dots, \alpha_n)$ as above, but we also provided an insight on the form of solutions whose heights are of comparable size. This is precisely the alternative given in Lemmas 2.24 and 2.25: either the heights of approximations have to be exponentially far apart from each other or the algebraic numbers and approximations themselves have to be connected by means of some linear fractional transformation. As a consequence, we were able to establish the concluding statement in Theorem 2.2.

The next result follows directly from Theorem 2.31.

Corollary 2.32. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree $d \geq 3$. Let $\alpha = \alpha_1, \dots, \alpha_d$ be distinct roots of $f(X)$ and suppose that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois. Let μ be such that*

$$0.5d + 1 < \mu \leq d.$$

Let C_0 be a real number greater than $(4e^A)^{-1}$, with

$$A = 500^2 \left(\log M(f) + \frac{d}{2} \right),$$

$$B = \theta_\alpha \cdot \max_{1 \leq j \leq d} \prod_{\substack{1 \leq i \leq d \\ i \neq j}} \frac{1 + |\alpha_i|}{|\alpha_i - \alpha_j|},$$

where θ_α is defined in (2.2.5), and $C = C(C_0, \mu, f)$ the maximum of

$$C_0^{\frac{1}{\mu - 1.42\sqrt{d}}} (4e^A)^{\frac{1.42\sqrt{d}}{\mu - 1.42\sqrt{d}}}$$

and

$$\left(C_0 2^{d^5} (d+2)^{5d^4} H(f)^{2d^5} \left(\boxed{f} + 2 \right)^{2d^4} B^{2d^4} \right)^{4/(2\mu - d - 2)},$$

where $\boxed{f} = \max\{|\alpha_1|, \dots, |\alpha_d|\}$.

[3] Here $\log^+(x) = \log \max\{1, |x|\}$.

The total number of rationals x/y in lowest terms, which satisfy $H(x,y) \geq C$,

$$\left|\alpha_j - \frac{x}{y}\right| < \frac{C_0}{H(x,y)^\mu}$$

for some $j \in \{1,2,\dots,d\}$, and $|x| = 1$ if $K = \mathbb{Q}_p$, is less than

$$\frac{\#\operatorname{Aut}'|F|}{2} \cdot \left\lfloor 1 + \frac{11.51 + 1.5\log d + \log\mu}{\log(\mu - 0.5d)} \right\rfloor,$$

where $F(X,Y) = Y^d f(X/Y)$ is the homogenization of f.

Proof. Since $\alpha_1,\dots,\alpha_d$ are conjugates, it follows from Proposition 2.7 that the number γ defined in (2.6.4) is equal to $\operatorname{Aut}'|F|/2$. The division by 2 appears due to the automorphism $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \operatorname{Aut}'|F|$, which maps (x,y) to $(-x,-y)$. Since $x/y = (-x)/(-y)$, half of the automorphisms can be disregarded. Thus it remains to verify that the number C exceeds the constant C_1 from Theorem 2.31. That is, we would like to ensure that the inequalities (2.6.5), (2.6.8), (2.6.11), (2.6.13) are satisfied, with C in place of C_1.

Before proceeding, we need to introduce some notation. Let P_i, Q_i be a minimal pair for (α_1,α_i), put $r_i = \max\{\deg P_i, \deg Q_i\}$ and $C_4(i) = C(P_i,Q_i)$, where $C(P_i,Q_i)$ is defined in (2.4.2). Let $\mathcal{O}$ denote the ring of integers of $\mathbb{Q}(\alpha_1)$. Put $C_1(i) = C_1(1,i)$, $C_2(i) = C_2(1,i)$, and $C_3(i) = C_3(1,i)$. In view of the fact that $\alpha_1,\dots,\alpha_d$ are conjugates it is sufficient to consider only $C_1(i), C_2(i), C_3(i)$ instead of $C_1(j,k), C_2(j,k), C_3(j,k)$ for all possible j,k.

Now we are ready to obtain our estimates. First, note that the inequalities $0.5d + 1 < \mu \leq d$ imply

$$\frac{1}{\mu} \leq \frac{1}{2\mu - d} \leq \frac{4}{2\mu - d - 2}. \tag{2.6.15}$$

With (2.6.15) it is straightforward to verify that

$$C \geq \left(C_0(d+2)^{3d/2} H(f)^{d-1}\right)^{1/\mu},$$

so it follows from Corollary 2.28 that the inequality (2.6.5) is satisfied. By definition of C, the inequality (2.6.11) holds as well.

It follows from (2.3.2) that

$$\max_{1\leq i\leq d}\{\max\{H(P_i),H(Q_i)\}\} < 2^{d(d-1)/2}\left((d+2)^3 H(f)^{(d+2)/2}\left(\lceil f\rceil + 2\right)B\right)^d. \tag{2.6.16}$$

It follows from (2.4.2), $r_i \le d/2$ and (2.6.16) that

$$\max_{1\le i\le d}\{C_4(i)\} \le \left(2^{-(d+6)/4}(d+2)^{3(d+2)/4}\max_{1\le i\le d}\{\max\{H(P_i),H(Q_i)\}\}^{d+4}\right)^{d/2} \tag{2.6.17}$$
$$\le \left(2^{(2d^3+6d^2-9d-6)/4}(d+2)^{(12d^2+51d+6)/4}H(f)^{(d^3+6d^2+8d)/2}\left(\lceil f\rceil+2\right)^{(d+4)d}B^{(d+4)d}\right)^{d/2}.$$

It follows from (2.5.1), (2.5.9), $r_i \le d/2$, $c_d \le H(f)$, $M(f) \le (d+1)H(f)$ and (2.6.16) that

$$\max_{1\le i\le d}\{C_1(i)\}^{\mu} \le C_0 2^{-d/2+3}(d+2)^{4d-3}c_d^{(d-1)^2}H(f)^{2(d-1)}\max_{1\le i\le d}\{\max\{H(P_i),H(Q_i)\}\}^{2d} \tag{2.6.18}$$
$$\le C_0 2^{(2d^3-2d^2-d+6)/2}(d+2)^{6d^2+4d-3}H(f)^{d^3+3d^2-1}\left(\lceil f\rceil+2\right)^{2d^2}B^{2d^2}.$$

It follows from (2.5.2), (2.5.10), $r_i \le d/2$, $c_d \le H(f)$ and (2.6.16) that

$$\begin{aligned}\max_{1\le i\le d}\{C_2(i)\} &\le C_0(d+2)\left(2c_d\left(\lceil f\rceil+2\right)\right)^{(d+2)/2}\max_{1\le i\le d}\{\max\{H(P_i),H(Q_i)\}\}\\ &\le C_0 2^{(d^2+2)/2}(d+2)^{3d+1}H(f)^{(d+1)(d+2)/2}\left(\lceil f\rceil+2\right)^{(3d+2)/2}B^d\\ &\le C^{(2\mu-d-2)/4}.\end{aligned}$$

Combining this inequality with (2.6.7) implies that (2.6.13) is satisfied.

Since $\mu/(2\mu-d) \ge 2$, we see that

$$\max_{1\le i\le d}\{C_1(i)\} \le \max_{\substack{1\le i\le d\\ r_i\ge 2}}\{C_3(i)\}.$$

It follows from (2.5.3), (2.5.11), $\mu \le d$, $c_d \le H(f)$, (2.6.15), (2.6.17) and (2.6.18) that

$$\begin{aligned}\max_{\substack{1\le i\le d\\ r_i\ge 2}}\{C_3(i)\}^{2\mu-d} &\le 2^d c_d^{d-2}H(f)(\lceil f\rceil+2)^{d+1}\max_{1\le i\le d}\{C_1(i)\}^{\mu}\cdot\max_{1\le i\le d}\{C_4(i)\}^d\\ &\le C_0 2^{(2d^5+6d^4-d^3-14d^2+4d+24)/8}(d+2)^{(12d^4+51d^3+54d^2+32d-24)/8}\times\\ &\times H(f)^{(d^5+6d^4+12d^3+12d^2+4d-8)/4}\left(\lceil f\rceil+2\right)^{(d^4+4d^3+4d^2+2d+2)/2}B^{d^2(d^2+4d+4)/2}\\ &\le C^{2\mu-d},\end{aligned}$$

so (2.6.8) is satisfied. □

The following result is an improvement to [17, Theorem 1] in the case when $F(X,Y)$ is an irreducible binary form of degree $d \geq 3$ such that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$, where α is a root of $F(X,1)$, is a Galois extension. More precisely, let m be a positive integer. Then [17, Theorem 1] states that there exists an explicitly computable constant $C' = C'(m,F)$ such that the number of primitive solutions to the Thue inequality

$$0 < |F(x,y)| \leq m \tag{2.6.19}$$

satisfying $H(x,y) \geq C'$ does not exceed $5d$. In turn, if we set $\mu = (3d+2)/4$ and use the upper bound $\#\operatorname{Aut}'|F| \leq 24$ established in Lemma 2.5, then it follows from Corollary 2.33 that there exists an explicitly computable constant $C = C(m,F)$ such that the number of primitive solutions to (2.6.19) does not exceed 432.

Corollary 2.33. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d \geq 3$. Let $\alpha_1, \dots, \alpha_d$ be distinct roots of $F(X,1)$ and assume that the field extension $\mathbb{Q}(\alpha_1)/\mathbb{Q}$ is Galois. For a positive integer m consider the Thue inequality (2.6.19). Let μ be a real number such that $0.5d + 1 < \mu \leq d$. Define*

$$A = 500^2 \left(\log M(F) + \frac{d}{2} \right), \quad B = \theta_\alpha \cdot \prod_{\substack{1 \leq i \leq d \\ i \neq j}} \frac{1 + |\alpha_i|}{|\alpha_i - \alpha_j|},$$

where θ_α is defined in (2.2.5), and let C the maximum of

$$(4e^A)^{\frac{1.42\sqrt{d}}{\mu - 1.42\sqrt{d}}},$$

$$\left(2^{d^5} (d+2)^{5d^4} H(f)^{2d^5} \left(\boxed{f} + 2 \right)^{2d^4} B^{2d^4} \right)^{4/(2\mu - d - 2)},$$

and

$$\left(\frac{2^d d^{(d-1)/2} M(F)^{d-2} m}{|D(F)|^{1/2}} \right)^{\frac{1}{d-\mu}}.$$

Then the number of solutions (x,y) to (2.6.19) such that

$$\gcd(x,y) = 1, \quad H(x,y) \geq C$$

does not exceed

$$\#\operatorname{Aut}'|F| \cdot \left\lfloor 1 + \frac{11.51 + 1.5 \log d + \log \mu}{\log(\mu - 0.5d)} \right\rfloor.$$

Proof. Let $\alpha_1, \alpha_2, \ldots, \alpha_d$ be the roots of $F(X,1)$. By Lemma 2.26, there exists an index $j \in \{1,2,\ldots,d\}$ such that

$$\min\left\{\left|\alpha_j - \frac{x}{y}\right|, \left|\alpha_j^{-1} - \frac{y}{x}\right|\right\} \leq \frac{2^{d-1}d^{(d-1)/2}M(F)^{d-2}m}{|D(F)|^{1/2}H(x,y)^d} < \frac{1}{H(x,y)^{\mu}},$$

where the last inequality follows from $H(x,y) \geq C$. To count the number of solutions to this inequality, we apply Corollary 2.32 twice, first to $C_0 = 1, \mu, f(X) = F(X,1)$, and then to $C_0 = 1, \mu, f(X) = F(1,X)$. □

2.7 Proof of Theorem 2.2

By Roth's Theorem [29], for every root α of $F(X,1)$ there exist only finitely many non-zero integers x, y such that $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x,y)^{-2.05}$. Since

$$|F(x,y)| \leq (d+1)H(F)H(x,y)^d$$

and $|F(x,y)| = tp^k$, we have

$$\frac{tp^k}{(d+1)H(F)} \leq H(x,y)^d.$$

Hence by choosing a large enough p^k we can increase $H(x,y)$ and make it so large that the inequality $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x,y)^{-2.05}$ is no longer satisfied for every root α of $F(X,1)$.

Define

$$C_0 = \frac{2^{d-1}d^{(d-1)/2}M(F)^{d-2}}{|D(F)|^{1/2}}.$$

Let (x,y,t) be a solution to (2.0.2). By Lemma 2.26,

$$\min\left\{\left|\alpha - \frac{x}{y}\right|, \left|\alpha^{-1} - \frac{y}{x}\right|\right\} \leq \frac{C_0 tp^k}{H(x,y)^d}.$$

From our choice of p^k and the above inequality it follows that

$$\frac{1}{H(x,y)^{2.05}} < \min\left\{\left|\alpha - \frac{x}{y}\right|, \left|\alpha^{-1} - \frac{y}{x}\right|\right\} \leq \frac{C_0 tp^k}{H(x,y)^d},$$

which is equivalent to

$$H(x,y) < (C_0 t p^k)^{1/(d-2.05)}. \tag{2.7.1}$$

Since $t \leq (p^k)^\lambda$,

$$tp^k \leq (p^k)^{1+\lambda} \leq |F(x,y)|_p^{-(1+\lambda)}.$$

Combining this inequality with (2.7.1), we get

$$\begin{aligned} H(x,y)^{d-2.05} &< C_0 t p^k \\ &\leq C_0 |F(x,y)|_p^{-(1+\lambda)}. \end{aligned}$$

We conclude that

$$|F(x,y)|_p < \frac{C_0^{1/(1+\lambda)}}{H(x,y)^\mu}, \tag{2.7.2}$$

where

$$\mu = \frac{d-2.05}{1+\lambda}.$$

We take p^k sufficiently large that

$$p^k > |D(F)|.$$

Then

$$|F(x,y)|_p \leq p^{-k} < |D(F)|^{-1} \leq |D(F)|_p.$$

By Lemma 2.16 there exists a unique p-adic root $\alpha \in \mathbb{Q}_p$ of $F(X,1)$ such that

$$\frac{|y\alpha - x|_p}{\max\{1, |\alpha|_p\}} \leq \frac{|F(x,y)|_p}{|D(F)|_p^{1/2}}.$$

Since $c_d\alpha$ is an algebraic integer, we see that $|c_d\alpha|_p \leq 1$, so $\max\{1, |\alpha|_p\} \leq |c_d|_p^{-1}$. Combining this inequality with (2.7.2), we obtain

$$\begin{aligned} |y\alpha - x|_p &< \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x,y)|_p \\ &\leq \frac{C_1}{H(x,y)^\mu}, \end{aligned}$$

where

$$C_1 = C_0^{1/(1+\lambda)} c_d |D(F)|^{1/2}.$$

Now, let (x_1, y_1, t_1), (x_2, y_2, t_2) be two solutions to (2.0.2) ordered so that $H(x_2, y_2) \geq H(x_1, y_1)$. Then it follows from the discussion above that there exist p-adic roots $\alpha, \beta \in \mathbb{Q}_p$ of $F(X, 1)$ such that

$$|y_1\alpha - x_1|_p < \frac{C_1}{H(x_1, y_1)^\mu}, \quad |y_2\beta - x_2|_p < \frac{C_1}{H(x_2, y_2)^\mu}.$$

By Lemma 2.25, there exists a number C_2, which depends on C_1, μ, F, but not on p, such that if $H(x_2, y_2) \geq H(x_1, y_1) \geq C_2$, then either $H(x_2, y_2) > C_3 H(x_1, y_1)^{\mu - d/2}$ for some positive number C_3, or α, β and $x_1/y_1, x_2/y_2$ are connected by means of a linear fractional transformation. By choosing p^k sufficiently large we can always ensure that $H(x_1, y_1) \geq C_2$. We obtain an upper bound on $H(x_2, y_2)$ by combining (2.7.1) with the inequality $|F(x_1, y_1)| \leq (d+1)H(F)H(x_1, y_1)^d$:

$$\begin{aligned} H(x_2, y_2) &< (C_0 t_2 p^k)^{1/(d-2.05)} \\ &\leq \left(C_0 (p^k)^{1+\lambda}\right)^{1/(d-2.05)} \\ &\leq \left(C_0 (t_1 p^k)^{1+\lambda}\right)^{1/(d-2.05)} \\ &= \left(C_0 |F(x_1, y_1)|^{1+\lambda}\right)^{1/(d-2.05)} \\ &\leq \left(C_0 \left((d+1)H(F)H(x_1, y_1)^d\right)^{1+\lambda}\right)^{1/(d-2.05)}. \end{aligned}$$

Merging the above upper bound with the lower bound $H(x_2, y_2) > C_3 H(x_1, y_1)^{\mu - d/2}$ results in the inequality

$$C_3 H(x_1, y_1)^{\mu - d/2 - d(1+\lambda)/(d-2.05)} < \left(C_0 \left((d+1)H(F)\right)^{1+\lambda}\right)^{1/(d-2.05)}.$$

From our choice of λ it follows that the exponent of $H(x_1, y_1)$ is positive, and so $H(x_1, y_1)$ is bounded. Therefore by making p^k (and therefore $H(x_1, y_1)$) sufficiently large we can always ensure that the inequality $H(x_2, y_2) > C_4 H(x_1, y_1)^{\mu - d/2}$ does not hold. Then α, β and $x_1/y_1, x_2/y_2$ are connected by means of a linear fractional transformation:

$$\beta = \frac{v\alpha - u}{-t\alpha + s}, \quad \frac{x_2}{y_2} = \frac{vx_1 - uy_1}{-tx_1 + sy_1},$$

where $s, t, u, v \in \mathbb{Z}$ and $sv - tu \neq 0$. By Lemma 2.7, the matrix

$$M = \frac{1}{\sqrt{|sv - tu|}} \begin{pmatrix} s & u \\ t & v \end{pmatrix}$$

is an automorphism of $|F|$, so it is an element of $\mathrm{Aut}'\,|F|$. Hence the number of solutions (x, y, t) to (2.0.2) is at most $\#\,\mathrm{Aut}'\,|F|$.

2.8 Proof of Theorem 2.3

The beginning of the proof is similar to the proof of Theorem 2.2. By Roth's Theorem [29], for every root α of $F(X,1)$ there exist only finitely many non-zero integers x, y such that $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x,y)^{-2.05}$. Now let (x,y,z,t) be a solution of (2.0.3). Since $|F(x,y)| \leq (d+1)H(F)H(x,y)^d$ and $|F(x,y)| = tp^z$, we have

$$\frac{p}{(d+1)H(F)} \leq \frac{tp^z}{(d+1)H(F)} = \frac{|F(x,y)|}{(d+1)H(F)} \leq H(x,y)^d.$$

Hence by choosing a large enough p we can increase $H(x,y)$ and make it so large that the inequality $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x,y)^{-2.05}$ is no longer satisfied for every root α of $F(X,1)$.

Let (x,y,z,t) be a solution of (2.0.3). As in the proof of Theorem 2.2, for our choice of p the inequality

$$H(x,y) < (C_0 t p^z)^{1/(d-2.05)} \tag{2.8.1}$$

holds, where

$$C_0 = \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2}}{|D(F)|^{1/2}}.$$

Since

$$tp^z \leq (p^z)^{1+\lambda} \leq |F(x,y)|_p^{-(1+\lambda)},$$

it follows from (2.8.1) that

$$|F(x,y)|_p < \frac{C_0^{1/(1+\lambda)}}{H(x,y)^\mu},$$

where

$$\mu = \frac{d - 2.05}{1+\lambda}.$$

We take p sufficiently large that

$$p > |D(F)|.$$

Then

$$|F(x,y)|_p \leq p^{-1} < |D(F)|^{-1} \leq |D(F)|_p.$$

By Lemma 2.16 there exists a unique p-adic root $\alpha \in \mathbb{Q}_p$ of $F(X,1)$ such that

$$|y\alpha - x|_p \leq \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x,y)|_p < \frac{C_1}{H(x,y)^\mu},$$

where

$$C_1 = C_0^{1/(1+\lambda)} c_d |D(F)|^{1/2}.$$

Note that C_1 is independent of p. Further, in order to ensure that $p \nmid x$ and $p \nmid y$, we adjust our choice of p as follows:

$$p > |c_0 c_d|.$$

Indeed, if $p \mid y$, then p does not divide x, because x and y are coprime. Since $z \geq 1$, it is evident from equation that

$$c_d x^d + y(c_{d-1}x^{d-1} + \cdots + c_0 y^{d-1}) = \pm t p^z$$

that p divides c_d, in contradiction to our choice of p. Then $|y|_p = 1$, and so for any $\alpha \in \mathbb{Q}_p$ we have

$$\left|\alpha - \frac{x}{y}\right|_p = |y\alpha - x|_p.$$

Analogously, we can show that $|x|_p = 1$. Therefore

$$\left|\alpha - \frac{x}{y}\right|_p < \frac{C_1}{H(x,y)^\mu}.$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_d$ be the roots of $F(X, 1)$. Applying Corollary 2.32 to $C_1, \mu, f(X) = F(X, 1)$, we conclude that there exists a positive number C_2, which depends on C_1, μ, F, but not on p, such that the number of rationals x/y in lowest terms satisfying $H(x, y) \geq C_2$, $|x|_p = 1$, and

$$\left|\alpha_j - \frac{x}{y}\right|_p < \frac{C_1}{H(x,y)^\mu} \tag{2.8.2}$$

for some $j \in \{1, 2, \ldots, d\}$ is less than

$$\#\operatorname{Aut}'|F| \cdot \left\lfloor 1 + \frac{11.51 + 1.5\log d + \log\mu}{\log(\mu - 0.5d)} \right\rfloor.$$

If we choose p so that $p \geq (d+1)H(F)C_2^d$, then

$$C_2^d \leq \frac{p}{(d+1)H(F)} \leq \frac{hp^z}{(d+1)H(F)} = \frac{|F(x,y)|}{(d+1)H(F)} \leq H(x,y)^d,$$

so the inequality $H(x, y) \geq C_2$ is satisfied. Since all solutions (x, y, z, t) to (2.0.3), including those that satisfy $H(x, y) \geq C_2$, also satisfy (2.8.2), the result follows.

Chapter 3

Automorphisms of Binary Forms Associated with $2\cos(2\pi/n)$

Let n be a positive integer such that $\varphi(n) \geq 6$, where $\varphi(n)$ is the Euler's totient function. Let

$$f(X) = X^d + c_{d-1}X^{d-1} + \cdots + c_1X + c_0$$

denote the minimal polynomial of the algebraic integer $2\cos(2\pi/n)$, whose degree d is $\varphi(n)/2$. Let $F(X,Y) = Y^d f(X/Y)$ denote the homogenization of $f(X)$.

Recall Definition 2.1 from Chapter 2, where we introduced the notion of an automorphism of a binary form. In this chapter, we determine $\mathrm{Aut}_{\mathbb{Q}} F$ and $\mathrm{Aut}' F$ for all binary forms $F(X,Y)$ of degree $d \geq 3$ defined above. For $n \geq 3$, we also compute $\mathrm{Aut}_{\mathbb{Q}} T_n$ for a binary form $T_n(X,Y)$, which is the homogenization of the n-th Chebyshev polynomial of the first kind. We apply Theorems 2.2, 2.3 to $F(X,Y)$ associated with $2\cos(2\pi/n)$ to obtain absolute bounds on the number of solutions of equations (2.0.2) and (2.0.3).

Recall the definitions of $\boldsymbol{C}_1, \boldsymbol{C}_2, \boldsymbol{C}_3, \boldsymbol{C}_4$ and $\boldsymbol{D}_2, \boldsymbol{D}_4$ given in Table 2.2. Our main results are stated in Theorems 3.1 and 3.2. Since the only positive integers n that satisfy $3 \leq \varphi(n)/2 \leq 4$ are $7, 9, 14, 15, 16, 18, 20, 24, 30$, we see that Theorem 3.1 covers all possible cases.

Theorem 3.1. *For a positive integer n such that $\varphi(n) \geq 6$, let $F(X,Y)$ denote the homogenization of the minimal polynomial of $2\cos(2\pi/n)$. Let $d = \varphi(n)/2$ denote its degree.*

1. If $d \geq 5$ is odd, then $\mathrm{Aut}' F = \{I\} \cong \boldsymbol{C}_1$, where I denotes the 2×2 identity matrix.

2. *If $d \geq 6$ is even and $n \not\equiv 0 \pmod 4$, then* $\operatorname{Aut}' F = \{\pm I\} \cong \boldsymbol{C}_2$.

3. *If $n = 7$ or 18, the binary forms corresponding to $2\cos(2\pi/7)$ and $2\cos(\pi/9)$ are $X^3 + X^2Y - 2XY^2 - Y^3$ and $X^3 - 3XY^2 - Y^3$, respectively. In either case,*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle \cong \boldsymbol{C}_3.$$

4. *If $n = 9$ or 14, the binary forms corresponding to $2\cos(2\pi/9)$ and $2\cos(\pi/7)$ are $X^3 - 3XY^2 + Y^3$ and $X^3 - X^2Y - 2XY^2 + Y^3$, respectively. In either case,*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle \cong \boldsymbol{C}_3.$$

5. *If $n = 15$, the binary form corresponding to $2\cos(2\pi/15)$ is $X^4 - X^3Y - 4X^2Y^2 + 4XY^3 + Y^4$. In this case,*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \right\rangle \cong \boldsymbol{C}_4.$$

6. *If $n = 30$, the binary form corresponding to $2\cos(\pi/15)$ is $X^4 + X^3Y - 4X^2Y^2 - 4XY^3 + Y^4$. In this case,*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \right\rangle \cong \boldsymbol{C}_4.$$

7. *If $n \equiv 0 \pmod 4$ and $n \neq 24$, then*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle \cong \boldsymbol{D}_2.$$

8. *If $n = 24$, the binary form corresponding to $2\cos(\pi/12)$ is $X^4 - 4X^2Y^2 + Y^4$. In this case,*

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle \cong \boldsymbol{D}_4.$$

Theorem 3.2. *For an integer $n \geq 3$, let $T_n(X, Y) \in \mathbb{Z}[X, Y]$ denote the homogenization of the n-th Chebyshev polynomial of the first kind.*

1. *If $n \geq 3$ is odd, then*

$$\operatorname{Aut}_{\mathbb{Q}} T_n = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle \cong \boldsymbol{C}_2.$$

2. If $n \geq 4$ is even, then

$$\mathrm{Aut}_{\mathbb{Q}}\, T_n = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle \cong \boldsymbol{D}_2.$$

Combining Theorem 3.1 with Theorems 2.2, 2.3 established in Chapter 2, we obtain the following.

Corollary 3.3. *Let n be a positive integer such that $\varphi(n) \geq 14$ and let $F(X,Y)$ be the homogenization of the minimal polynomial of $2\cos(2\pi/n)$. Put $d = \varphi(n)/2$. Let λ be taken from Table 2.1 for $7 \leq d \leq 16$ and $\lambda = 1 - 16.2/d$ for $d \geq 17$. Let p be prime, k a positive integer, and consider the Diophantine equation*

$$|F(x,y)| = tp^k. \tag{3.0.1}$$

Provided that p^k is sufficiently large, the number of solutions to (3.0.1) in integers (x,y,t) such that

$$\gcd(x,y) = 1, \;\; 1 \leq t \leq (p^k)^\lambda$$

is either 0, or 2 when $4 \nmid n$, or 4 when $4 \mid n$. If solutions exist, they are of the form $(x,y),(-x,-y)$ if $4 \nmid n$ and of the form $(x,y),(-x,-y),(x,-y),(-x,y)$ if $4 \mid n$.

Corollary 3.4. *Let n be a positive integer such that $\varphi(n) \geq 14$ and let $F(X,Y)$ be the homogenization of the minimal polynomial of $2\cos(2\pi/n)$. Put $d = \varphi(n)/2$. Let λ be such that*

$$0 \leq \lambda < 1 - 8.1/(d+2).$$

Let p be prime, and consider the Diophantine equation

$$|F(x,y)| = tp^z. \tag{3.0.2}$$

Provided that p is sufficiently large, the number of solutions to (3.0.2) in integers (x,y,z,t) such that

$$\gcd(x,y) = 1, \;\; z \geq 1, \;\; 1 \leq t \leq (p^z)^\lambda$$

is at most

$$N \cdot \left\lfloor 1 + \frac{11.51 + 1.5\log d + \log(d - 2.05 - d\lambda/(1+\lambda))}{\log(d/2 - 2.05 - d\lambda/(1+\lambda))} \right\rfloor,$$

where $N = 2$ if $4 \nmid n$ and $N = 4$ if $4 \mid n$.

If we let $\lambda = 0.5 - 4.05/(d+2)$, then it is a consequence of Corollary 3.4 that the number of solutions in integers (x,y,z,t) to (3.0.2) does not exceed 332 for all $d \geq 7$ and it does not exceed 12 for all $d \geq 10^{15}$.

3.1 Preliminary Results

Let n be a positive integer such that $\varphi(n) \geq 4$. In this section, we summarize the properties of algebraic numbers of the form $2\cos(2\pi/n)$, their minimal polynomials $f(X)$, and the binary forms $F(X,Y)$ associated with them.

Let i be the imaginary unit and let

$$\zeta_n = \exp\left(\frac{2\pi i}{n}\right) = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)$$

denote a primitive n-th root of unity. Then

$$\Phi_n(X) = \prod_{\substack{1\leq \ell < n \\ \gcd(\ell,n)=1}} \left(X - \zeta_n^\ell\right)$$

is the minimal polynomial of ζ_n. Note that $\deg \Phi_n = \varphi(n)$. Further,

$$\zeta_n^{-1} = \exp\left(-\frac{2\pi i}{n}\right) = \cos\left(-\frac{2\pi}{n}\right) + i\sin\left(-\frac{2\pi}{n}\right) = \cos\left(\frac{2\pi}{n}\right) - i\sin\left(\frac{2\pi}{n}\right),$$

which means that $\zeta_n + \zeta_n^{-1} = 2\cos(2\pi/n) \in \mathbb{R}$. Since both ζ_n and ζ_n^{-1} are algebraic integers, so is $\zeta_n + \zeta_n^{-1}$.

Let $f(X)$ denote the minimal polynomial of $\zeta_n + \zeta_n^{-1}$. The action of the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ on $\zeta_n + \zeta_n^{-1}$ allows us to determine the conjugates of $\zeta_n + \zeta_n^{-1}$, which are $\zeta_n^\ell + \zeta_n^{-\ell} = 2\cos(2\pi\ell/n)$ for ℓ such that $1 \leq \ell < n/2$ and $\gcd(\ell, n) = 1$. We conclude that the minimal polynomial $f(X)$ of $2\cos(2\pi/n)$ is of the form

$$f(X) = \prod_{\substack{1\leq \ell < n/2 \\ \gcd(\ell,n)=1}} \left(X - 2\cos\left(\frac{2\pi\ell}{n}\right)\right).$$

Since for $n \geq 3$ the interval $[1, n/2)$ contains exactly one half of all the integers less than n that are coprime to n, we see that $f(X)$ has degree $d = \varphi(n)/2$.

Lemma 3.5. *Let $(\mathbb{Z}/n\mathbb{Z})^*$ denote the group of units of $\mathbb{Z}/n\mathbb{Z}$. Define the group*

$$G_n = (\mathbb{Z}/n\mathbb{Z})^*/\{\pm 1\}.$$

Then $\mathrm{Gal}(\mathbb{Q}(2\cos(2\pi/n))/\mathbb{Q}) \cong G_n$.

Proof. We index the roots of $f(X)$ through the elements $\{\pm\ell\}$ of G_n as follows:

$$\alpha_{\{\pm\ell\}} = \zeta_n^{\ell} + \zeta_n^{-\ell}.$$

We write α_ℓ in place of $\alpha_{\{\pm\ell\}}$, as well as α in place of α_1, for brevity. Define $\alpha_0 = \beta_0 = 1$, and for a positive integer j let $\beta_{\ell,j} = (\alpha_\ell)^j$. Then the binomial formula implies that, for any non-negative integer k,

$$\beta_{\ell,2k} = \sum_{j=0}^{k} \binom{2k}{k-j} \alpha_{2j\ell} \text{ and } \beta_{\ell,2k+1} = \sum_{j=0}^{k} \binom{2k+1}{k-j} \alpha_{(2j+1)\ell}.$$

Note that the first of the two identities can be rewritten as follows:

$$\begin{pmatrix} \beta_0 \\ \beta_{\ell,2} \\ \beta_{\ell,4} \\ \vdots \\ \beta_{\ell,2k} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 2 & 1 & 0 & \dots & 0 \\ C_4^2 & 4 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{2k}^k & C_{2k}^{k-1} & C_{2k}^{k-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_{2\ell} \\ \alpha_{4\ell} \\ \vdots \\ \alpha_{2k\ell} \end{pmatrix},$$

where C_a^b denotes the binomial coefficient $\binom{a}{b}$. The second identity can be written in an analogous form. The $(k+1)\times(k+1)$ matrix C above is independent of the choice of ℓ, has integer coefficients and non-zero determinant, hence it is invertible. Moreover, its inverse C^{-1} also has integer coefficients, because its determinant is equal to one. By multiplying both sides of the above equation by C^{-1}, we see that each $\alpha_{2k\ell}$ can be represented as an integer linear combination of $\beta_{\ell,2j}$'s. A similar observation applies to $\alpha_{(2k+1)\ell}$ and $\beta_{\ell,2j+1}$'s. Therefore

$$\alpha_{2k\ell} = \sum_{j=0}^{k} c_{2k,j} \alpha_\ell^{2j} \text{ and } \alpha_{(2k+1)\ell} = \sum_{j=0}^{k} c_{2k+1,j} \alpha_\ell^{2j+1}$$

for some integers $c_{2k,j}$ and $c_{2k+1,j}$. From above formulas it follows that any root α_ℓ of $f(X)$ is contained in $\mathbb{Q}(\alpha)$, and as a consequence any field automorphism in $\mathrm{Gal}\,(\mathbb{Q}(\alpha)/\mathbb{Q})$ is uniquely determined by its action on α. For ℓ coprime to n, let $\sigma_\ell \in \mathrm{Gal}\,(\mathbb{Q}(\alpha)/\mathbb{Q})$ denote the field automorphism such that $\sigma_\ell\colon \alpha \mapsto \alpha_\ell$. Then for any positive even integer $2k$ we have

$$\sigma_\ell(\alpha_{2k}) = \sigma_\ell\left(\sum_{j=0}^{k} c_{2k,j}\alpha^{2j}\right) = \sum_{j=0}^{k} c_{2k,j}\sigma_\ell(\alpha)^{2j} = \sum_{j=0}^{k} c_{2k,j}\alpha_\ell^{2j} = \alpha_{2k\ell}.$$

Analogously, we can establish the formula for any odd positive integer $2k+1$ and conclude that $\sigma_\ell\colon \alpha_m \mapsto \alpha_{\ell m}$ for any m coprime to n. Finally, note that the field automorphisms of

$\mathbb{Q}(\alpha)$ that fix $\mathbb{Q}$ form a group under composition, as they satisfy the relation $\sigma_\ell \circ \sigma_j = \sigma_{\ell j}$. Consequently, the map $\mathrm{Gal}\,(\mathbb{Q}(\alpha)/\mathbb{Q}) \to G_n$, $\sigma_\ell \mapsto \{\pm\ell\}$ defines a group isomorphism. □

We conclude this section by proving Lemmas 3.6 and 3.7, which we will find useful in the proof of Theorem 3.1.

Lemma 3.6. *Let n be a positive integer and let $f(X)$ be the minimal polynomial of $2\cos(2\pi/n)$.*

1. *If $n \equiv 0 \pmod 4$ then $f(X) = g(X^2)$, where $g(X)$ is the minimal polynomial of $2 + 2\cos(4\pi/n)$.*
2. *If n is odd then $f(X) = (-1)^d g(-X)$, where $g(X)$ is the minimal polynomial of $2\cos(\pi/n)$.*

Proof. 1. Suppose that $n \equiv 0 \pmod 4$. Recall that for any $x \in \mathbb{R}$ it is the case that $2\cos^2(x) = 1 + \cos(2x)$. Therefore

$$4\cos^2\left(\frac{2\pi}{n}\right) = 2\left(1 + \cos\left(\frac{4\pi}{n}\right)\right) = 2 + 2\cos\left(\frac{2\pi}{n/2}\right).$$

Let $g(X)$ denote the minimal polynomial of $2 + 2\cos\left(2\pi/(n/2)\right)$. Note that $\deg g(X) = \varphi(n/2)/2$ and

$$g\left(4\cos^2\left(\frac{2\pi}{n}\right)\right) = 0.$$

Since for any positive integer n divisible by 4 it is the case that $\varphi(n)/2 = \varphi(n/2)$, we have

$$\deg f(X) = \frac{\varphi(n)}{2} = 2 \cdot \frac{\varphi(n/2)}{2} = 2\deg g(X) = \deg g(X^2).$$

Since the polynomials $g(X^2)$ and $f(X)$ have equal degrees, both vanish at $2\cos(2\pi/n)$, and the leading coefficient of $g(X^2)$ is positive, we conclude that $f(X) = g(X^2)$.

2. Suppose that n is odd. Note that

$$-2\cos\left(\frac{2\pi}{n}\right) = 2\cos\left(\pi + \frac{2\pi}{n}\right) = 2\cos\left(\frac{2\pi(n+2)}{2n}\right).$$

Since $\gcd(2n, n+2) = 1$, we see that $-2\cos(2\pi/n)$ is a conjugate of $2\cos(\pi/n)$. Thus, if $g(X)$ is the minimal polynomial of $2\cos(\pi/n)$, then $g(-2\cos(2\pi/n)) = 0$. But then $2\cos(2\pi/n)$ is a root of $(-1)^d g(-X)$, and since the leading coefficient of this polynomial is positive, it must be the minimal polynomial of $2\cos(2\pi/n)$.

□

Lemma 3.7. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a binary form of degree d, and let $G(X,Y) = cF(-X,Y)$, where c is a non-zero integer. Let*

$$M = \begin{pmatrix} s & u \\ t & v \end{pmatrix} \text{ and } \tilde{M} = \begin{pmatrix} s & -u \\ -t & v \end{pmatrix}.$$

Then $M \in \mathrm{Aut}_{\mathbb{Q}} F$ if and only if $\tilde{M} \in \mathrm{Aut}_{\mathbb{Q}} G$.

Proof. Let

$$T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $G(X,Y) = cF(-X,Y) = cF_T(X,Y)$. Now suppose that $M \in \mathrm{Aut}_{\mathbb{Q}} F$, i.e., $F_M = F$. Then for any 2×2 matrix A it is the case that $(F_M)_A = F_A$. Since $MT = T\tilde{M}$, we see that

$$G = cF_T = c(F_M)_T = cF_{MT} = cF_{T\tilde{M}} = c(F_T)_{\tilde{M}} = (cF_T)_{\tilde{M}} = G_{\tilde{M}}.$$

Therefore $\tilde{M} \in \mathrm{Aut}_{\mathbb{Q}} G$. The converse statement can be established analogously. □

In view of Lemma 3.6 part 2, as well as Lemma 3.7, it becomes evident that, in order to understand the automorphisms of the binary form associated with $2\cos(\pi/n)$ for odd n, it is sufficient to study the automorphisms of the binary form associated with $2\cos(2\pi/n)$.

3.2 Case $d \geq 4$ and $n \equiv 0 \pmod 4$

We begin by proving Theorem 3.1 in the case when $d \geq 4$ and $n \equiv 0 \pmod 4$. First, we will determine $\mathrm{Aut}_{\mathbb{Q}} F$ and then derive $\mathrm{Aut}' F$ from it.

According to Lemma 3.6 part 1, the minimal polynomial $f(X)$ of $2\cos(2\pi/n)$ is equal to $g(X^2)$ for some $g(X) \in \mathbb{Z}[X]$. Let $F(X,Y)$ be the homogenization of $f(X)$, and let

$$D_2 = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle. \tag{3.2.1}$$

Then $D_2 \subseteq \mathrm{Aut}_{\mathbb{Q}} F$ due to the fact that $f(X) = g(X^2)$. Further, $\mathrm{Aut}_{\mathbb{Q}} F$ is either equal to D_2, or isomorphic to $\boldsymbol{D}_4$, or isomorphic to $\boldsymbol{D}_6$, where the groups $\boldsymbol{D}_4$ and $\boldsymbol{D}_6$ are defined in Table 2.2. The three lemmas established in this section allow us to conclude that $\mathrm{Aut}_{\mathbb{Q}} F \neq D_2$ if and only if $n = 24$.

Lemma 3.8. *Let D_2 be as in (3.2.1). Every subgroup of $\mathrm{GL}_2(\mathbb{Q})$ that properly contains D_2 is either of the form*

$$\left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & t \\ -1/t & 0 \end{pmatrix} \right\rangle$$

or of the form

$$\left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1/2 & t/2 \\ -3/2t & 1/2 \end{pmatrix} \right\rangle$$

for some non-zero $t \in \mathbb{Q}$.

Proof. Let $G = \mathrm{GL}_2(\mathbb{Q})$ and let H be a finite subgroup of G that properly contains D_2. According to the classification of finite subgroups of G given in Table 2.2, every finite subgroup of G that contains a group isomorphic to $\boldsymbol{D}_2$ and has more than 4 elements is either G-conjugate to $\boldsymbol{D}_4$ or $\boldsymbol{D}_6$. We consider these two cases separately.

1. Suppose that H is G-conjugate to $\boldsymbol{D}_4$. That is, there exists some matrix $A \in G$ such that $H = A\boldsymbol{D}_4A^{-1}$. Since $D_2 \subsetneq H$, we also have

   $$D_2 = ANA^{-1}$$

   for some subgroup N of $\boldsymbol{D}_4$ that is isomorphic to $\boldsymbol{D}_2$. Note that $\boldsymbol{D}_4$ contains exactly two subgroups isomorphic to $\boldsymbol{D}_2$, namely $\boldsymbol{D}_2$ itself and D_2. Thus we consider two separate cases, i.e., $N = \boldsymbol{D}_2$ and $N = D_2$.

   (a) Suppose that $D_2 = A\boldsymbol{D}_2A^{-1}$. A straightforward calculation shows that every matrix $A \in G$ such that $D_2 = A\boldsymbol{D}_2A^{-1}$ must be of the form

   $$\begin{pmatrix} a & -a \\ b & b \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} a & a \\ b & -b \end{pmatrix}$$

   for some non-zero $a, b \in \mathbb{Q}$. Independently of the form of A, we have

   $$\begin{aligned} H &= \left\langle A \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} A^{-1}, A \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A^{-1} \right\rangle \\ &= \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & a/b \\ -b/a & 0 \end{pmatrix} \right\rangle \end{aligned}$$

   fo some non-zero $a, b \in \mathbb{Q}$. Upon setting $t = a/b$ the result follows.

(b) Suppose that $D_2 = AD_2A^{-1}$. A straightforward calculation shows that every matrix $A \in G$ such that $D_2 = AD_2A^{-1}$ must be of the form

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ or } \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

for some non-zero $a, b \in \mathbb{Q}$. Consequently,

$$\begin{aligned} H &= \left\langle \begin{pmatrix} 0 & a/b \\ b/a & 0 \end{pmatrix}, \begin{pmatrix} 0 & a/b \\ -b/a & 0 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & a/b \\ -b/a & 0 \end{pmatrix} \right\rangle \end{aligned}$$

for some non-zero $a, b \in \mathbb{Q}$. Upon setting $t = a/b$ the result follows.

2. Suppose that H is G-conjugate to $\boldsymbol{D}_6$. That is, there exists some matrix $A \in G$ such that $H = A\boldsymbol{D}_6A^{-1}$. Since $D_2 \subsetneq D_6$, we also have

$$D_2 = ANA^{-1}$$

for some subgroup N of $\boldsymbol{D}_6$ that is isomorphic to $\boldsymbol{D}_2$. Note that $\boldsymbol{D}_6$ contains exactly three subgroups isomorphic to $\boldsymbol{D}_2$, namely $\boldsymbol{D}_2$ itself,

$$D_2^{(1)} = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \right\rangle,$$

and

$$D_2^{(2)} = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \right\rangle.$$

Thus we consider three separate cases, i.e., $N = \boldsymbol{D}_2$, $N = D_2^{(1)}$, and $N = D_2^{(2)}$.

(a) Suppose that $D_2 = A\boldsymbol{D}_2A^{-1}$ for some $A \in G$. As it was explained previously, every matrix A which satisfies $D_2 = A\boldsymbol{D}_2A^{-1}$ must be of the form

$$\begin{pmatrix} a & -a \\ b & b \end{pmatrix} \text{ or } \begin{pmatrix} a & a \\ b & -b \end{pmatrix}$$

for some non-zero $a, b \in \mathbb{Q}$. Therefore

$$\begin{aligned} H &= \left\langle A \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} A^{-1}, A \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} A^{-1} \right\rangle \\ &= \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1/2 & a/2b \\ -3b/2a & 1/2 \end{pmatrix} \right\rangle \end{aligned}$$

for some non-zero $a, b \in \mathbb{Q}$. Upon setting $t = a/b$ the result follows.

(b) Suppose that $D_2 = AD_2^{(1)}A^{-1}$ for some $A \in G$. A straightforward calculation shows that A must be of the form

$$\begin{pmatrix} a & -2a \\ b & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} a & 0 \\ b & -2b \end{pmatrix}$$

for some non-zero $a, b \in \mathbb{Q}$. Therefore

$$H = \left\langle \begin{pmatrix} -1/2 & -3a/2b \\ -b/2a & 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 & 3a/2b \\ -b/2a & 1/2 \end{pmatrix} \right\rangle$$

for some non-zero $a, b \in \mathbb{Q}$. Upon setting $t = 3a/b$ the result follows.

(c) Suppose that $D_2 = A\boldsymbol{D}_2^{(2)}A^{-1}$ for some $A \in G$. A straightforward calculation shows that A must be of the form

$$\begin{pmatrix} -2a & a \\ 0 & b \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & a \\ -2b & b \end{pmatrix}$$

for some non-zero $a, b \in \mathbb{Q}$. Therefore

$$H = \left\langle \begin{pmatrix} -1/2 & -3a/2b \\ -b/2a & 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 & -3a/2b \\ b/2a & 1/2 \end{pmatrix} \right\rangle$$

for some non-zero $a, b \in \mathbb{Q}$. Upon setting $t = 3a/b$ the result follows.

□

Lemma 3.9. *Let n be a positive integer such that $n \equiv 0 \pmod 4$, and let $f(X)$ be the minimal polynomial of $2\cos(2\pi/n)$. Let $F(X, Y)$ be the homogenization of $f(X)$, and suppose that D_2 as defined in (3.2.1) is a proper subgroup of $\mathrm{Aut}_{\mathbb{Q}} F$. Then $f(X)$ is a reciprocal polynomial.*

Proof. According to Lemma 3.8, there are two options for how $\mathrm{Aut}_{\mathbb{Q}} F$ can look, so we will consider two cases separately. In each case, we will make use of the formula

$$|f_m(0)| = \begin{cases} 0, & \text{if } m = 4, \\ 2, & \text{if } m = 2^k \text{ for } k \geq 3, \\ p, & \text{if } m = 4p^k \text{ for } k \geq 1, \text{ where } p \text{ is an odd prime}, \\ 1, & \text{otherwise}, \end{cases} \tag{3.2.2}$$

where $f_m(X)$ denotes the minimal polynomial of $2\cos(2\pi/m)$. The proof of the formula (3.2.2) can be found in [9].

1. Suppose that there exist integers $a \neq 0$ and $b \geq 1$ such that $\gcd(a,b) = 1$ and $M \in \mathrm{Aut}_{\mathbb{Q}} F$, where
$$M = \begin{pmatrix} 0 & a/b \\ -b/a & 0 \end{pmatrix}.$$
Then
$$\begin{aligned} F(X,Y) &= F\left(\frac{a}{b}Y, -\frac{b}{a}X\right) \\ &= (ab)^{-d} F\left(a^2 Y, -b^2 X\right). \end{aligned}$$
Thus
$$(ab)^d F(X,Y) = F(a^2 Y, -b^2 X).$$
By plugging $X = 1$ and $Y = 0$ into the above equation, we see that $c_0 = (-a/b)^d = (a/b)^d$, where c_0 denotes the constant coefficient of $f(X)$. Since $c_0 \in \mathbb{Z}$, we see that $t = a/b$ is an integer such that $t^d = c_0$. By (3.2.2) the value of c_0 is square-free, and since $d \geq 2$ is even and $t^d = c_0$, we conclude that $c_0 = 1$. Therefore $t = a/b = \pm 1$, which means that
$$\mathrm{Aut}_{\mathbb{Q}} F = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle.$$
In particular, we see that $F(X,Y) = F(Y,X)$, so $f(X) = F(X,1)$ is a reciprocal polynomial.

2. Suppose that there exist integers $a \neq 0$ and $b \geq 1$ such that $\gcd(a,b) = 1$ and $M \in \mathrm{Aut}_{\mathbb{Q}} F$, where
$$M = \begin{pmatrix} 1/2 & a/2b \\ -3b/2a & 1/2 \end{pmatrix}.$$
We will show that this is impossible.

Since $M \in \mathrm{Aut}_{\mathbb{Q}} F$,
$$\begin{aligned} F(X,Y) &= F_M(X,Y) \\ &= F\left(\frac{1}{2}X + \frac{a}{2b}Y, -\frac{3b}{2a}X + \frac{1}{2}Y\right) \\ &= (2ab)^{-d} F\left(abX + a^2 Y, -3b^2 X + abY\right). \end{aligned}$$
Thus
$$(2ab)^d F(X,Y) = F\left(abX + a^2 Y, -3b^2 X + abY\right). \tag{3.2.3}$$

By plugging $X = 0$ and $Y = 1$ into the above equation, we obtain $c_0 2^d b^d = F(a, b)$. Thus $F(a, b)$ is divisible by b. Since the leading coefficient of $F(X, Y)$ is equal to one, we see that

$$a^d \equiv F(a, 0) \equiv F(a, b) \equiv 0 \pmod{b}.$$

Then $b \mid a^d$, and since a and $b \geq 1$ are coprime, we conclude that $b = 1$ and $c_0 2^d = F(a, 1)$. By plugging $X = 1$ and $Y = 0$ into (3.2.3), we obtain $(2a)^d = F(a, -3)$. Since $f(X) = g(X^2)$, we see that

$$F(a, -3) \equiv c_0(-3)^d \equiv 0 \pmod{a^2},$$

which means that $a^2 \mid c_0 3^d$. By (3.2.2) the value of c_0 is square-free, so $a = \pm 3^r$ for some non-negative integer r. Since $\mathrm{Aut}_{\mathbb{Q}} F$ is a group, we may replace M with M^{-1}, and so without loss of generality we may assume that $a = 3^r$.

Suppose that $r \geq 3$. After plugging $a = 3^r$ and $b = 1$ into (3.2.3) we obtain

$$2^d 3^{(r-1)d} F(X, Y) = F\left(3^{r-1}X + 3^{2r-1}Y, -X + 3^{r-1}Y\right). \tag{3.2.4}$$

Thus

$$F\left(3^{r-1}X + 3^{2r-1}Y, -X + 3^{r-1}Y\right) \equiv F(0, -X) \equiv c_0(-X)^d \equiv 0 \pmod{9}.$$

Since this congruence must hold for all X, it holds for those X that are not divisible by 3, which means that 9 divides c_0. However, this result contradicts (3.2.2), which states that the value of c_0 is square-free. We conclude that the only possible values of r are $0, 1, 2$, and so the only possible values of $a = 3^r$ are $1, 3, 9$.

For $r = 0$, $X = 0$, $Y = 1$ the equation (3.2.4) gives us $F(1, 1) = 2^d$. For $r = 1$, $X = 1$, $Y = 1$ it gives us $F(1, 1) = 2^d$ once again. Finally, for $r = 2$, $X = 1$, $Y = 1$ the identity (3.2.4) is equivalent to $F(1, 1) = 3^{-d} F(15, 1)$. Since $15 - 2\cos(x) \geq 13$ for any $x \in \mathbb{R}$, in the last case we have

$$|F(1, 1)| = 3^{-d}|F(15, 1)| = 3^{-d} \prod_{\substack{1 \leq j < n/2 \\ \gcd(j,n)=1}} \left|15 - 2\cos\left(\frac{2\pi j}{n}\right)\right| \geq 3^{-d} 13^d > 2^d.$$

Thus, regardless of the value of r, we must have $|F(1, 1)| \geq 2^d$. Since $d \geq 4$, we will obtain a contradiction by proving that $|F(1, 1)| \leq 8$.

To see that this is the case, note that for any integer j we have

$$\begin{aligned}
1-2\cos\left(\frac{2\pi j}{n}\right) &= -2\cos\left(\frac{4\pi}{3}\right)-2\cos\left(\frac{2\pi j}{n}\right)\\
&= -4\cos\left(\frac{2\pi}{3}+\frac{\pi j}{n}\right)\cos\left(\frac{2\pi}{3}-\frac{\pi j}{n}\right)\\
&= 4\cos\left(\frac{\pi(2n+3j)}{3n}\right)\cos\left(\pi+\frac{\pi(2n-3j)}{3n}\right)\\
&= 4\cos\left(\frac{\pi(2n+3j)}{3n}\right)\cos\left(\frac{\pi(5n-3j)}{3n}\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
F(1,1) &= \prod_{\substack{1\le j<n/2\\ \gcd(j,n)=1}}\left(1-2\cos\left(\frac{2\pi j}{n}\right)\right)\\
&= \prod_{\substack{1\le j<n/2\\ \gcd(j,n)=1}} 2\cos\left(\frac{\pi(2n+3j)}{3n}\right)\times\prod_{\substack{1\le j<n/2\\ \gcd(j,n)=1}} 2\cos\left(\frac{\pi(5n-3j)}{3n}\right)\\
&= \prod_{\substack{1\le j<n\\ \gcd(j,n)=1}} 2\cos\left(\frac{\pi(2n+3j)}{3n}\right).
\end{aligned}$$

The result follows once we consider the following three cases:

(a) Suppose that n is divisible by 9. Then $\gcd(2n/3+j,n)=1$ if and only if $\gcd(j,n)=1$. Since n is even we have $\gcd(j,n)=\gcd(j,2n)=1$, so

$$F(1,1)=\prod_{\substack{1\le j<n\\ \gcd(j,2n)=1}} 2\cos\left(\frac{\pi(2n/3+j)}{n}\right). \tag{3.2.5}$$

By definition,

$$f_{2n}(0)=\prod_{\substack{1\le j<n\\ \gcd(j,2n)=1}} 2\cos\left(\frac{\pi j}{n}\right).$$

If we put

$$N=\prod_{\substack{1\le j<n\\ \gcd(j,2n)=1}} 2\cos\left(\frac{\pi(5n/3+j)}{n}\right),$$

then it follows from (3.2.5) and the above expression for $f_{2n}(0)$ that

$$f_{2n}(0)^2 = \prod_{\substack{1\le j<2n\\ \gcd(j,2n)=1}} 2\cos\left(\frac{\pi j}{n}\right) = F(1,1)\cdot N.$$

Since both $f_{2n}(0)$ and $F(1,1)$ are non-zero rational integers, N is a rational number. Since it is defined as a product of algebraic integers and the ring of algebraic integers is closed under multiplication, it must be that N is a rational integer. We conclude that $F(1,1)$ divides $f_{2n}(0)^2$. Since $2n$ is divisible by 72, it follows from (3.2.2) that $|f_{2n}(0)| = 1$. Therefore $|F(1,1)| = 1$.

(b) Suppose that n is divisible by 3, but it is not divisible by 9. Then the equality (3.2.5) still holds, but the numbers $2n/3 + j$ and n are not necessarily coprime. More precisely, $\gcd(2n/3 + j, n) = 3$ if $j \equiv n/3 \pmod 3$ and $\gcd(2n/3 + j, n) = 1$ if $j \equiv -n/3 \pmod 3$. Note that the case $j \equiv 0 \pmod 3$ is impossible because j and n are coprime. We write $F(1,1) = MN$, where

$$M = \prod_{\substack{1\le j<n\\ \gcd(j,n)=1\\ j\equiv n/3 \pmod 3}} 2\cos\left(\frac{\pi(2n+3j)/9}{n/3}\right), \quad N = \prod_{\substack{1\le j<n\\ \gcd(j,n)=1\\ j\equiv -n/3 \pmod 3}} 2\cos\left(\frac{\pi(2n/3+j)}{n}\right).$$

Note that $\gcd((2n + 3j)/9, n/3) = 1$ for all j coprime to n. Further,

$$f_{2n/3}(0)^3 = \prod_{\substack{1\le j<n\\ \gcd(j,2n/3)=1}} 2\cos\left(\frac{\pi j}{n/3}\right), \quad f_{2n}(0)^2 = \prod_{\substack{1\le j<2n\\ \gcd(j,2n)=1}} 2\cos\left(\frac{\pi j}{n}\right).$$

By an argument analogous to the one used in Part (a) we conclude that M divides $f_{2n/3}(0)^3$ and N divides $f_{2n}(0)^2$. Thus $|MN| \le |f_{2n/3}(0)|^3 f_{2n}(0)^2$. Since $8 \mid 2n/3$ and $24 \mid 2n$, it follows from (3.2.2) that $|f_{2n/3}(0)| \le 2$ and $|f_{2n}(0)| = 1$. We conclude that

$$|F(1,1)| \le |f_{2n/3}(0)|^3 f_{2n}(0)^2 \le 2^3 \cdot 1^2 = 8.$$

(c) Suppose that n is not divisible by 3. Then $\gcd(2n + 3j, 3n) = 1$ if and only if $\gcd(j, n) = 1$. We have

$$F(1,1) = \prod_{\substack{1\le j<n\\ \gcd(j,n)=1}} 2\cos\left(\frac{\pi(2n+3j)}{3n}\right). \tag{3.2.6}$$

Note that

$$f_{6n}(0) = \prod_{\substack{1 \le j < 3n \\ \gcd(j,6n)=1}} 2\cos\left(\frac{\pi j}{3n}\right). \tag{3.2.7}$$

Since all the factors in the product (3.2.6) appear in the product (3.2.7) exactly once, the number $f_{6n}(0)/F(1,1)$ is rational, and since it is equal to a product of algebraic integers it has to be a rational integer. We conclude that $F(1,1)$ divides $f_{6n}(0)$. Since $24 \mid 6n$, it follows from (3.2.2) that $|f_{6n}(0)| = 1$. Thus $|F(1,1)| = 1$.

□

Lemma 3.9 states that $\mathrm{Aut}_{\mathbb{Q}} F$ contains more than 4 elements whenever the minimal polynomial $f(X)$ of $2\cos(2\pi/n)$ is reciprocal. Now the statement of Theorem 3.1 in the case $n \equiv 0 \pmod 4$ follows from the next lemma, which classifies all the situations when $f(X)$ is reciprocal.

Lemma 3.10. *The minimal polynomial $f(X)$ of $2\cos(2\pi/n)$ is reciprocal if and only if $n = 3$ or $n = 24$.*

Proof. Via a straightforward calculation we can verify that for $1 \le n \le 745$ reciprocal polynomials appear only for $n = 3, 24$, and they are $X + 1$ and $X^4 - 4X^2 + 1$, respectively. It remains to prove that there are no reciprocal polynomials with $n \ge 746$.

For a positive integer n, let $g(n)$ denote Jacobsthal's function; that is, $g(n)$ is equal to the smallest positive integer m such that every sequence of m consecutive integers contains an integer coprime to n. It was proven by Kanold [19] that

$$g(n) \le 2^{\omega(n)},$$

where $\omega(n)$ denotes the number of distinct prime factors of n.[1] Combining the above upper bound with the inequality [28]

$$\omega(n) \le 1.3841 \frac{\log n}{\log\log n},$$

which holds for all $n \ge 3$, we get

$$g(n) < n^{\frac{0.96}{\log\log n}}.$$

[1] The author is grateful to Prof. Jeffrey Shallit for pointing out that better bounds exist, e.g., [18, 35]. However, Kanold's bound is sufficient for our purposes.

Now consider the interval $\left[\frac{1}{2\pi}\arccos\left(\frac{1}{4}\right), \frac{1}{4}\right)$. We claim that this interval contains a rational number j/n with j coprime to n. In other words, we would like to locate an integer j coprime to n such that

$$\frac{1}{2\pi}\arccos\left(\frac{1}{4}\right) n \le j < \frac{1}{4}n.$$

We see that such an integer j has to belong to the interval $\left[\frac{1}{2\pi}\arccos\left(\frac{1}{4}\right) n, \frac{1}{4}n\right)$, whose length exceeds $n/25$. Since our interval is half-closed, it contains at least $\lfloor n/25 \rfloor$ consecutive integers. However, for all $n \ge 746$ we have

$$n^{\frac{0.96}{\log\log n}} < \frac{n}{25} - 1,$$

and this inequality implies that

$$g(n) < n^{\frac{0.96}{\log\log n}} < \frac{n}{25} - 1 < \left\lfloor \frac{n}{25} \right\rfloor.$$

This means that the interval $\left[\frac{1}{2\pi}\arccos\left(\frac{1}{4}\right) n, \frac{1}{4}n\right)$ contains an integer j that is coprime to n. But then

$$\arccos\left(\frac{1}{4}\right) \le \frac{2\pi j}{n} < \frac{\pi}{2},$$

and consequently

$$0 < \alpha_j \le \frac{1}{2}.$$

Since $f(X)$ is reciprocal, the number α_j^{-1} is a conjugate of α_j, so there exists some ℓ such that $\alpha_\ell = \alpha_j^{-1}$. Thus $\alpha_\ell \ge 2$. On the other hand, $\alpha_\ell \le 2$, which means that $\ell = 0$. Since $\gcd(\ell, n) = 1$, we conclude that $n = 1$, and this contradicts our assumption that $n \ge 746$. □

It remains to compute $\mathrm{Aut}' F$. Suppose that $n = 24$, so that $F(X, Y) = X^4 - 4X^2Y^2 + Y^4$ and

$$\mathrm{Aut}_{\mathbb{Q}} F = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle.$$

Let $M = |sv - tu|^{-1/2} \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right) \in \mathrm{Aut}' F \setminus \mathrm{Aut}_{\mathbb{Q}} F$. Then $M^2 \in \mathrm{Aut}_{\mathbb{Q}} F$. Since $\mathrm{Aut}_{\mathbb{Q}} F$ is a proper subgroup of $\mathrm{Aut}' F$, it follows from Lemma 2.5 that $\mathrm{Aut}' F \cong \boldsymbol{D}_8$. Therefore M^2 has order 4, i.e., $M^2 = \pm \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Solving this equation in integers s, t, u, v such that

$\gcd(s,t,u,v) = 1$, we conclude that M has to be one of the following two matrices of order 8:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

However, neither of the above matrices are automorphisms of $F(X,Y) = X^4 - 4X^2Y^2 + Y^4$, so the matrix M described above does not exist. We conclude that $\operatorname{Aut}' F = \operatorname{Aut}_{\mathbb{Q}} F$.

Finally, suppose that $4 \mid n$ and $n \neq 24$, so that

$$\operatorname{Aut}_{\mathbb{Q}} F = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle.$$

Let $M = |sv - tu|^{-1/2} \left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right) \in \operatorname{Aut}' F \setminus \operatorname{Aut}_{\mathbb{Q}} F$. Then $M^2 \in \operatorname{Aut}_{\mathbb{Q}} F$. Since $\operatorname{Aut}_{\mathbb{Q}} F$ is a proper subgroup of $\operatorname{Aut}' F$, it follows from Lemma 2.5 that $\operatorname{Aut}' F \cong \boldsymbol{D}_4$. Therefore M^2 has order 4, i.e., $M^2 = \pm \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ or $M^2 = \pm \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. However, neither of these equations have solutions in integers s,t,u,v such that $\gcd(s,t,u,v) = 1$.[2] Therefore the matrix M described above does not exist. Once again, we conclude that $\operatorname{Aut}' F = \operatorname{Aut}_{\mathbb{Q}} F$.

3.3 Case $d \geq 5$ and $n \not\equiv 0 \pmod 4$

Next, we consider the case $d \geq 5$ and $n \not\equiv 0 \pmod 4$. In view of Lemma 3.6 part 2 and Lemma 3.7, we may assume that n is odd.

Let M be an element of $\operatorname{Aut}' F$, where

$$M = \frac{1}{m}\begin{pmatrix} s & u \\ t & v \end{pmatrix}, \quad m = |sv - tu|^{1/2}.$$

Then

$$m^d F(X,Y) = F(sX + uY, tX + vY),$$

which means that the polynomials $m^d f(X)$ and $F(sX+u, tX+v)$ are equal. For an integer

[2] They do have solutions in Gaussian integers.

ℓ, let $\alpha_\ell = 2\cos(2\pi\ell/n)$, and let $\alpha = \alpha_1$. Then

$$\begin{aligned} F(sX+u, tX+v) &= \prod_{\substack{1\le \ell < n/2 \\ \gcd(\ell,n)=1}} ((sX+u) - \alpha_\ell(tX+v)) \\ &= \prod_{\substack{1\le \ell < n/2 \\ \gcd(\ell,n)=1}} ((-t\alpha_\ell + s)X - (v\alpha_\ell - u)) \\ &= F(s,t) \prod_{\substack{1\le \ell < n/2 \\ \gcd(\ell,n)=1}} \left(X - \frac{v\alpha_\ell - u}{-t\alpha_\ell + s}\right). \end{aligned}$$

Since $m^d f(X)$ and $F(sX+u, tX+v)$ have the same roots, we conclude that there exists some j coprime to n such that

$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s}. \tag{3.3.1}$$

We will show that $s = v = \pm 1$ and $t = u = 0$, implying that $M \in \{\pm I\}$. Since $-I \in \mathrm{Aut}' F$ if and only if d is even, this result would allow us to conclude that $\mathrm{Aut}' F = \{I\}$ when d is odd and $\mathrm{Aut}' F = \{\pm I\}$ when d is even.

Since n is odd, it follows from the discusion in Section 3.1 that there exists $\sigma_2 \in \mathrm{Gal}\,(\mathbb{Q}(\alpha)/\mathbb{Q})$ such that $\sigma_2 \colon \alpha_\ell \mapsto \alpha_{2\ell}$ for each ℓ coprime to n. Therefore

$$\alpha_{2j} = \sigma_2(\alpha_j) = \sigma_2\left(\frac{v\alpha - u}{-t\alpha + s}\right) = \frac{v\sigma_2(\alpha) - u}{-t\sigma_2(\alpha) + s} = \frac{v\alpha_2 - u}{-t\alpha_2 + s}.$$

Since for any $x \in \mathbb{R}$ it is the case that $2\cos(2x) = (2\cos(x))^2 - 2$, we conclude that $\alpha_{2i} = \alpha_i^2 - 2$ for all i. Therefore

$$\left(\frac{v\alpha - u}{-t\alpha + s}\right)^2 - 2 = \alpha_j^2 - 2 = \alpha_{2j} = \frac{v\alpha_2 - u}{-t\alpha_2 + s} = \frac{v(\alpha^2-2) - u}{-t(\alpha^2-2) + s}.$$

From the above equality we obtain

$$\left(-t(\alpha^2-2) + s\right)\left((v\alpha - u)^2 - 2(-t\alpha + s)^2\right) = (-t\alpha + s)^2\left(v(\alpha^2-2) - u\right).$$

We conclude that the polynomial

$$\begin{aligned} &(2t^3 - t^2v - tv^2)X^4 \\ &+(-4st^2 + 2stv + 2tuv)X^3 \\ &+(2s^2t - s^2v - 2st^2 + sv^2 - 4t^3 + t^2u + 2t^2v - tu^2 + 2tv^2)X^2 \\ &+(4s^2t + 8st^2 - 2stu - 4stv - 2suv - 4tuv)X \\ &+(-2s^3 - 4s^2t + s^2u + 2s^2v + su^2 + 2tu^2) \end{aligned}$$

vanishes at α. Since the degree of α is at least 5 and the above polynomial has degree at most 4, it must be the case that this polynomial is identically equal to zero. That is,

$$\begin{cases} t(t-v)(2t+v) = 0, \\ t(-2st+sv+uv) = 0, \\ 2s^2t - s^2v - 2st^2 + sv^2 - 4t^3 + t^2u + 2t^2v - tu^2 + 2tv^2 = 0, \\ 2s^2t + 4st^2 - stu - 2stv - suv - 2tuv = 0, \\ -2s^3 - 4s^2t + s^2u + 2s^2v + su^2 + 2tu^2 = 0. \end{cases} \tag{3.3.2}$$

Depending on the value of t, we consider the following three cases.

1. Suppose that $t = 0$. Then the first two equations in (3.3.2) vanish, while the third and the fourth equations simplify to $sv(v-s) = 0$ and $suv = 0$, respectively. Note that $s \neq 0$, for otherwise the denominator of (3.3.1) vanishes. Thus the last two equations further reduce to $v(v-s) = 0$ and $uv = 0$. If $v = 0$, then the number $\alpha_j = -u/s$ is rational, in contradiction to the fact that $\deg \alpha_j \geq 5$. On the other hand, if $v = s$, then $u = 0$. Since the determinant of every automorphism is equal to ± 1, we have $sv - tu = s^2 = \pm 1$. Since s in an integer, we conclude that $s^2 = 1$, and consequently $s = v = \pm 1$ and $t = u = 0$, as claimed.

2. Suppose that $t = v$ and $t \neq 0$. Then the second equation in (3.3.2) simplifies to $v(u-s) = 0$. But then $s = u$, and
$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s} = \frac{v\alpha - u}{-v\alpha + u} = -1,$$
in contradiction to the fact that $\deg \alpha_j \geq 5$.

3. Suppose that $v = -2t$ and $t \neq 0$. Then the second equation simplifies to $v(2s+u) = 0$. But then $u = -2s$,
$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s} = \frac{-2t\alpha + 2s}{-t\alpha + s} = 2,$$
in contradiction to the fact that $\deg \alpha_j \geq 5$.

3.4 Case $d = 3, 4$ and $n \not\equiv 0 \pmod 4$

It remains to consider the case $d = 3, 4$ and $n \not\equiv 0 \pmod 4$, which corresponds to $n = 7, 9, 14, 15, 18$, and 30. In view of Lemma 3.6 part 2 and Lemma 3.7, we may restrict our

attention to odd n. We will demonstrate the case $n = 7$ in detail, and the cases $n = 9$ and $n = 15$ can be established analogously. Let

$$M = \frac{1}{m}\begin{pmatrix} s & u \\ t & v \end{pmatrix} \in \operatorname{Aut}' F,$$

where $m = |sv - tu|^{1/2}$.

1. If $n = 7$, the binary form corresponding to $2\cos(2\pi/7)$ is

$$F(X, Y) = X^3 + X^2Y - 2XY^2 - Y^3.$$

We will show that

$$\operatorname{Aut}' F = \left\langle \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle.$$

Let $\alpha_\ell = 2\cos(2\pi\ell/7)$, and note that $\alpha = \alpha_1, \alpha_2$ and α_3 are the roots of $F(X, 1)$. Furthermore, we have $\alpha_2 = \alpha^2 - 2$ and $\alpha_3 = -\alpha^2 - \alpha + 1$.

Since $M \in \operatorname{Aut}' F$, by Proposition 2.7 there exists some $j \in \{2, 3\}$ such that

$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s}.$$

Suppose that $j = 2$. Then

$$\begin{aligned} v\alpha - u &= (-t\alpha + s)\alpha_j \\ &= (-t\alpha + s)(\alpha^2 - 2) \\ &= -t\alpha^3 + s\alpha^2 + 2t\alpha - 2s \\ &= -t(-\alpha^2 + 2\alpha + 1) + s\alpha^2 + 2t\alpha - 2s \\ &= (s + t)\alpha^2 - 2s - t. \end{aligned}$$

Thus the polynomial $(s+t)X^2 - vX - (2s+t-u)$ vanishes at α, and since α has degree 3 it must be identically equal to zero. Therefore

$$\begin{cases} s + t = 0, \\ v = 0, \\ 2s + t - u = 0, \end{cases}$$

and we conclude that $t = -s$, $u = s$ and $v = 0$ for any non-zero integer s. Independently of the value of s we obtain the relation

$$\alpha_2 = \frac{v\alpha - u}{-\alpha t + s} = -\frac{1}{\alpha + 1}.$$

Further, since $\gcd(s, t, u, v) = 1$, it must be the case that $s = \pm 1$. If $s = 1$, then

$$\begin{pmatrix} s & u \\ t & v \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}.$$

However, one can easily verify that the above matrix is not an automorphism of F. If $s = -1$, then

$$M = \begin{pmatrix} s & u \\ t & v \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

One can easily verify that, in this case, $M \in \mathrm{Aut}' F$.

Suppose that $j = 3$. Then

$$v\alpha - u = (-t\alpha + s)\alpha_3 = -s\alpha^2 - (s - t)\alpha + s + t.$$

Thus the polynomial $sX^2 + (s - t + v)X - (s + t + u)$ vanishes at α, and since α has degree 3 it must be identically equal to zero. We conclude that $s = 0$, $u = -t$ and $v = t$ for any non-zero integer t. Independently of the value of t we obtain the relation $\alpha_3 = -(1 + \alpha)/\alpha$. Further, since $\gcd(s, t, u, v) = 1$, it must be the case that $t = \pm 1$. It is straightforward to check that the matrix corresponding to $t = 1$ is not an automorphism, while the matrix corresponding to $t = -1$ is equal to M^2, so it is an automorphism of F.

Since we considered all possible relations of the form $\alpha_j = (v\alpha - u)/(-t\alpha + s)$, we conclude that the only automorphisms of $F(X, Y)$ that belong $\mathrm{Aut}' F$ are I, M and M^2.

2. If $n = 9$, the binary form corresponding to $2\cos(2\pi/9)$ is

$$F(X, Y) = X^3 - 3XY^2 + Y^3.$$

We will show that

$$\mathrm{Aut}_{\mathbb{Q}} F = \left\langle \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle.$$

Let $\alpha_\ell = 2\cos(2\pi\ell/9)$, and note that $\alpha = \alpha_1, \alpha_2$ and α_4 are the roots of $F(X,1)$. Since $M \in \mathrm{Aut}' F$, by Proposition 2.7 there exists some $j \in \{2,4\}$ such that

$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s}.$$

If $j = 4$, we obtain $\alpha_4 = 1/(-\alpha+1)$ and

$$M = \begin{pmatrix} s & u \\ t & v \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Aut}' F,$$

while if $j = 2$ we obtain $\alpha_2 = (\alpha-1)/\alpha$ and $M^2 \in \mathrm{Aut}' F$.

3. If $n = 15$, the binary form corresponding to $2\cos(2\pi/15)$ is

$$F(X,Y) = X^4 - X^3Y - 4X^2Y^2 + 4XY^3 + Y^4.$$

We will show that

$$\mathrm{Aut}' F = \left\langle \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \right\rangle.$$

Let $\alpha_\ell = 2\cos(2\pi\ell/15)$, and note that $\alpha = \alpha_1, \alpha_2, \alpha_4$ and α_7 are the roots of $F(X,1)$. Since $M \in \mathrm{Aut}' F$, by Proposition 2.7 there exists some $j \in \{2,4,7\}$ such that

$$\alpha_j = \frac{v\alpha - u}{-t\alpha + s}.$$

If $j = 2$ or $j = 7$, then no relation of the above form could be obtained. If $j = 4$, then $\alpha_4 = (\alpha-2)/(\alpha-1)$ and there are two automorphisms corresponding to this relation, namely M and M^3, where

$$M = \begin{pmatrix} s & u \\ t & v \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix}.$$

3.5 Automorphisms of Binary Forms Associated with Chebyshev Polynomials

In this section, we prove Theorem 3.2. Let n be a positive integer and let $T_n(X,Y) \in \mathbb{Z}[X,Y]$ denote the homogenization of the n-th Chebyshev polynomial of the first kind. It

is well-known that the roots of $T_n(X,1)$ are given by $\cos((2j-1)\pi/2n)$ for $j=1,2,\ldots,n$. Consequently, if we let $f_m(X)$ denote the minimal polynomial of $2\cos(2\pi/m)$, then

$$T_n(X,1) = \prod_{\substack{d\mid n \\ d \text{ is odd}}} f_{4n/d}\left((2-\delta_{dn})X\right),$$

where δ_{ij} denotes the Kronecker delta function. If we let $F_m(X,Y)$ denote the homogenization of $f_m(X)$, then $T_n(X,Y)$ factors as

$$T_n(X,Y) = \prod_{\substack{d\mid n \\ d \text{ is odd}}} F_{4n/d}\left((2-\delta_{dn})X,Y\right). \tag{3.5.1}$$

Note that every binary form in the above factorization of $T_n(X,Y)$ is irreducible. The main result will follow from the formula (3.5.1), Theorem 3.1, as well as the two lemmas established below.

Lemma 3.11. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a binary form and let $A \in \mathrm{GL}_2(\mathbb{Q})$. Then $\mathrm{Aut}_{\mathbb{Q}} F_A = A^{-1}\left(\mathrm{Aut}_{\mathbb{Q}} F\right)A$.*

Proof. Suppose that $M \in \mathrm{Aut}_{\mathbb{Q}} F$. Since $F = F_M$, we have $F_A = (F_M)_A = F_{MA} = (F_A)_{A^{-1}MA}$, so $A^{-1}MA \in \mathrm{Aut}_{\mathbb{Q}} F_A$. □

Lemma 3.12. *For a positive integer n, let $T_n(X,Y)$ be the homogenization of the n-th Chebyshev polynomial of the first kind. Suppose that $G(X,Y) \in \mathbb{Z}[X,Y]$ is an irreducible binary form of degree at least 1 that divides $T_n(X,Y)$. Then $\mathrm{Aut}_{\mathbb{Q}} T_n \subseteq \mathrm{Aut}_{\mathbb{Q}} G$.*

Proof. Since $\mathrm{Aut}_{\mathbb{Q}} G = \mathrm{Aut}_{\mathbb{Q}}(rG)$ for any non-zero $r \in \mathbb{C}$, without loss of generality we may assume that G has a positive leading coefficient and content equal to one. Further, the result is trivially true for $n=1$, for if $G \mid T_1$, then $T_1 = X$ and $G = \pm X$. Therefore we may suppose that $n \geq 2$.

Let

$$M = \begin{pmatrix} s & u \\ t & v \end{pmatrix}$$

be an element of $\mathrm{Aut}_{\mathbb{Q}} T_n$ and denote the discriminant of a binary form F by $D(F)$. As is well-known,

$$D((T_n)_M) = (\det M)^{n(n-1)} D(T_n).$$

Since $T_n = (T_n)_M$ and $D(T_n) \neq 0$, we have $(\det M)^{n(n-1)} = 1$. Since $n \geq 2$, it must be the case that $\det M \neq 0$.

The proof is by contradiction, so suppose that $M \notin \operatorname{Aut}_{\mathbb{Q}} G$, which means that the binary forms G and G_M are distinct. Since G is irreducible and $\det M \neq 0$, it must be the case that G_M is also irreducible. Since $G \mid T_n$ and T_n has content one, there exists some $H(X,Y) \in \mathbb{Z}[X,Y]$ such that $T_n = GH$. Since $M \in \operatorname{Aut}_{\mathbb{Q}} F$, we have

$$T_n = (T_n)_M = (GH)_M = G_M H_M,$$

which means that $GG_M \mid T_n$ in $\mathbb{Q}[X,Y]$. Since T_n factors into irreducible forms as in (3.5.1), and both G and G_M are irreducible, the result will follow after we consider the next two cases.

1. Suppose that $G(X,Y) = X$ and $G_M(X,Y) = rF_{4n/d}(2X,Y)$ for some non-zero $r \in \mathbb{Q}$ and odd $d \mid n$ such that $d < n$. Then $G_M(X,Y) = sX + uY$, which means that $sX + uY = rF_{4n/d}(2X,Y)$ and consequently $\deg F_{4n/d} = 1$. Since $\deg F_m = 1$ if and only if $m \in \{3,4,6\}$, we conclude that $d = n$, which contradicts the assumption that $d < n$. The case when $G(X,Y) = F_{4n/d}(2X,Y)$ and $G_M(X,Y) = rX$ for some non-zero $r \in \mathbb{Q}$ and odd $d \mid n$ such that $d < n$ can be established analogously.

2. Suppose that $G(X,Y) = F_{4n/d_1}(2X,Y)$ and $G_M(X,Y) = rF_{4n/d_2}(2X,Y)$ for some non-zero $r \in \mathbb{Q}$ and two distinct odd divisors d_1, d_2 of n that are less than n. Set $k = 4n/d_1$, $\ell = 4n/d_2$,

   $$S = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},$$

   so that $G = (F_k)_S$ and $G_M = (rF_\ell)_S$. Then $G_M = (F_k)_{SM}$ and consequently $(F_k)_{SM} = (rF_\ell)_S$. Since S is invertible, we conclude that $rF_\ell = (F_k)_{SMS^{-1}}$. Consequently, there exist rational numbers a, b, c, d such that the polynomials $rF_\ell(X,1)$ and $F_k(aX+c, bX+d)$ are identical. In particular, their roots are the same, which means that

   $$2\cos(2\pi j/\ell) = \frac{-c + 2\cos(2\pi/k)d}{a - 2\cos(2\pi/k)b}$$

   for some j coprime to ℓ. Therefore $2\cos(2\pi j/\ell)$ is an element of the field $\mathbb{Q}(2\cos(2\pi/k))$. Since the Galois group of $f_\ell(X)$ is cyclic, all the conjugates of $2\cos(2\pi j/\ell)$ belong to this field as well, which means that the field $\mathbb{Q}(2\cos(2\pi/\ell))$ is entirely contained in $\mathbb{Q}(2\cos(2\pi/k))$. We repeat the argument with G and G_M interchanged, and conclude that the fields $\mathbb{Q}(2\cos(2\pi/k))$ and $\mathbb{Q}(2\cos(2\pi/\ell))$ are identical. In particular, their discriminants D_k and D_ℓ are the same.

   Before we obtain a contradiction by proving that $D_k \neq D_\ell$, we need to make two additional remarks. The first one is that, since the degrees of G and G_M are the same,

we may conclude that the degrees of F_k and F_ℓ are the same. Consequently, $\varphi(k)/2 = \deg F_k = \deg F_\ell = \varphi(\ell)/2$, which means that $\varphi(k) = \varphi(\ell)$.

The second remark is that the conditions $\varphi(k) = \varphi(\ell)$ and $k \neq \ell$ imply the existence of some odd prime which divides k but not ℓ, or vice versa. For suppose that this is not the case and

$$k = \prod_{i=1}^{t} p_i^{e_i}, \quad \ell = \prod_{j=1}^{t} p_i^{f_i}$$

for some positive integers $t, e_1, \ldots, e_t, f_1, \ldots, f_t$. Then

$$\prod_{i=1}^{t} p_i^{e_i-1}(p_i - 1) = \varphi(k) = \varphi(\ell) = \prod_{i=1}^{t} p_i^{f_i-1}(p_i - 1).$$

After dividing both sides by $\prod_{i=1}^{t}(p_i - 1)$, we obtain

$$\prod_{i=1}^{t} p_i^{e_i-1} = \prod_{i=1}^{t} p_i^{f_i-1},$$

which means that $e_i = f_i$ for all $i = 1, 2, \ldots, t$. But then $k = \ell$, in contradiction to our assumption that k and ℓ are distinct. Therefore there exists some prime that divides k but not ℓ, or the other way around. Since k and ℓ are both divisible by 4, this prime must be odd.

Now let D_m denote the discriminant of the field $\mathbb{Q}(2\cos(2\pi/m))$. It was demonstrated in [22] that

$$D_m = \begin{cases} 2^{(j-1)2^{j-2}-1}, & \text{if } m = 2^j,\ j > 2; \\ p^{(jp^j-(j+1)p^{j-1}-1)/2}, & \text{if } m = p^j \text{ or } 2p^j,\ p > 2 \text{ prime}; \\ \left(\prod_{i=1}^{\omega(m)} p_i^{e_i-1/(p_i-1)}\right)^{\frac{\varphi(m)}{2}}, & \text{if } m = \prod_{i=1}^{\omega(m)} p_i^{e_i}, \omega(m) > 1, m \neq 2p^j, \end{cases}$$

where $\omega(m)$ denotes the number of distinct prime divisors of m. Let

$$k = \prod_{i=1}^{\omega(k)} p_i^{e_i} \text{ and } \ell = \prod_{j=1}^{\omega(\ell)} q_j^{f_j}$$

be the prime factorizations of k and ℓ, respectively. Since $D_k = D_\ell$ and $\varphi(k) = \varphi(\ell)$, the above formula for the discriminant of $\mathbb{Q}\left(2\cos(2\pi/m)\right)$ yields the equality

$$\prod_{i=1}^{\omega(k)} p_i^{e_i-1/(p_i-1)} = \prod_{j=1}^{\omega(\ell)} q_j^{f_j-1/(q_j-1)}.$$

The exponents of primes in the equality above are all positive, and since we established previously that there exists some i such that $p_i \neq q_j$ for all j, or there exists some j such that $q_j \neq p_i$ for all i, this equality cannot possibly be true.

□

Now we can prove Theorem 3.2. Suppose that $n \geq 3$ is odd and $M \in \operatorname{Aut}_{\mathbb{Q}} T_n$. According to (3.5.1), every irreducible binary form different from X that divides T_n is of the form $(F_{4n/d})_S$ for some odd $d < n$ such that $d \mid n$, where

$$S = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that $(F_{24})_S$ does not divide T_n, for otherwise the equality $4n/d = 24$ would imply that n is even. It follows from Theorem 3.1 item 7, as well as Lemma 3.11, that

$$\begin{aligned} \operatorname{Aut}_{\mathbb{Q}} T_n &\subseteq \operatorname{Aut}_{\mathbb{Q}} (F_{4n/d})_S \\ &= S^{-1} \left(\operatorname{Aut}_{\mathbb{Q}} F_{4n/d} \right) S \\ &= \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle. \end{aligned}$$

Therefore there are only four options for M. Further, since X divides $T_n(X, Y)$, by Lemma 3.12 it must also be the case that $M \in \operatorname{Aut}_{\mathbb{Q}} X$. Therefore M must be of the form

$$\begin{pmatrix} 1 & 0 \\ t & v \end{pmatrix}$$

for some $t, v \in \mathbb{Q}$, which means that only two out of four options remain, namely $(t, v) = (0, -1), (0, 1)$, so the result follows.

Suppose that $n \geq 4$ is even. Then X does not divide T_n, so it follows from (3.5.1) and Theorem 3.1 items 7, 8 that

$$\left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle \subseteq \operatorname{Aut}_{\mathbb{Q}} T_n. \tag{3.5.2}$$

Suppose that the containment in (3.5.2) is strict. Then it follows from Theorem 3.1 item 8 that the binary form $(F_{24})_S$, which has degree 4, divides T_n. In this case $\operatorname{Aut}_{\mathbb{Q}} T_n = \operatorname{Aut}_{\mathbb{Q}} (F_{24})_S$. Since $T_4 \neq (F_{24})_S$, it follows from (3.5.1) that there exists some binary form

$(F_m)_S$ that divides T_n such that $m \neq 24$. If $\deg F_m \geq 4$, then $(F_m)_S$ has the automorphism group as in Theorem 3.1 item 7, so the contradiction follows from Lemma 3.12. Therefore $\deg F_m < 4$ and since the degree of F_m is even it must be the case that $\deg F_m = 2$. Because 4 divides m there are only two possible options for $(F_m)_S$, namely $(F_8)_S = 2X^2 - Y^2$ and $(F_{12})_S = 4X^2 - 3Y^2$. It is straightforward to check that $\mathrm{Aut}_{\mathbb{Q}}(F_{24})_S \not\subseteq \mathrm{Aut}_{\mathbb{Q}}(F_8)_S$ and $\mathrm{Aut}_{\mathbb{Q}}(F_8)_S \not\subseteq \mathrm{Aut}_{\mathbb{Q}}(F_{12})_S$, so once again we obtain a contradiction. We conclude that the groups in (3.5.2) are equal, as claimed.

Chapter 4

Generalization of the Thue-Siegel Principle

In order to extend the results from Chapter 2, we need to generalize both the gap principle and the Thue-Siegel principle. In this chapter, we focus on the latter.

Recall the formulation of the Thue-Siegel principle from Chapter 1: if α is an irrational algebraic number and $x_1/y_1, x_2/y_2$ are distinct rational numbers that satisfy the inequalities (1.1.1) and $y_2 > y_1 \geq C_1$ for some large number C_1, then $y_2 < y_1^{\eta}$ for some $\eta > 1$. The Thue-Siegel principle was first established by Thue [33], and later refined by Bombieri [1] and Bombieri and Mueller [3]. By improving on the work of Dyson [11], they achieved the following:

1. Instead of taking approximations $x_1/y_1, x_2/y_2$ to the same algebraic number α, the Thue-Siegel principle was extended to the case when x_1/y_1 approximates α, while x_2/y_2 approximates some $\beta \in \mathbb{Q}(\alpha)$, with $\deg \beta = \deg \alpha$.

2. A function $C(t)$ was discovered such that if $\mu > \sqrt{2d}$, $y_2 > y_1 \geq 1$,
$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{C(\alpha)}{y_1^{\mu}}, \quad \left|\beta - \frac{x_2}{y_2}\right| < \frac{C(\beta)}{y_1^{\mu}}$$
hold simultaneously, then $y_2 < y_1^{\eta}$ for some $\eta > 1$. In other words, the lower bound C_1 on y_1, y_2 has been eliminated.

We generalize the first of the two results above by considering β whose degree over $\mathbb{Q}(\alpha)$ is small relative to the degree of α over $\mathbb{Q}$. The main result of this chapter is stated in Theorem 4.1.

Theorem 4.1. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let $\alpha, \beta \in K$ be irrational algebraic numbers over $\mathbb{Q}$ and put $d = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}]$. Let $\mu > \sqrt{2d}$. There exists an explicitly computable number $C > 0$, which depends only on α, β and μ, such that if $x_1/y_1, x_2/y_2$ are rational numbers satisfying $H(x_2, y_2) \geq H(x_1, y_1) \geq C$,*

$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{1}{H(x_1, y_1)^\mu}, \quad \left|\beta - \frac{x_2}{y_2}\right| < \frac{1}{H(x_2, y_2)^\mu},$$

then

$$H(x_2, y_2) < H(x_1, y_1)^\eta,$$

where

$$\eta = \frac{256(\mu + 3)^4}{3((\mu/\sqrt{2d}) - 1)^2}.$$

Remark 4.2. Let $d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}]$ and $d_2 = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}(\alpha)]$. Without loss of generality, suppose that $d_1 \geq d_2$ (if not, then we can switch α and β). At first sight it may seem that Theorem 4.1 applies to all algebraic numbers α, β. However, when $H(x_2, y_2)$ is sufficiently large it must be the case that $\mathbb{Q}(\beta)$ is a subfield of F, where F is the splitting field of the minimal polynomial of α. Furthermore, the inequality $d_2 < d_1/2$ holds, indicating that the degree of β over $\mathbb{Q}(\alpha)$ is small relative to the degree of α over $\mathbb{Q}$.

Indeed, if $\sqrt{2d} = \sqrt{2d_1d_2} \geq d_1$, then $\mu > d_1$, and so Theorem 4.1 does not yield any improvement over Liouville's theorem. Thus it is only useful when $\sqrt{2d_1d_2} < \mu < d_1$, or equivalently $d_2 < d_1/2$. However, this condition yields a certain restriction on β: if $\mathbb{Q}(\beta) \not\subseteq F$, then

$$[\mathbb{Q}(\beta)\colon \mathbb{Q}] = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}(\alpha)] = d_2 < \mu,$$

and so once again our theorem does not yield any improvement over Liouville's theorem applied to β. Thus, if $H(x_2, y_2)$ is sufficiently large, it must be the case that $\mathbb{Q}(\beta)$ is a subfield of F.

4.1 Preliminary Results

Lemma 4.3. (See [10, Section 13.1, Theorem 4]) *Let α, β be algebraic numbers,*

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}(\alpha)]\,.$$

Then the set

$$\left\{\alpha^i\beta^j\colon 0 \leq i \leq d_1 - 1, \ \ 0 \leq j \leq d_2 - 1\right\}$$

forms a basis of $\mathbb{Q}(\alpha, \beta)$ when viewed as a vector space over $\mathbb{Q}$.

Lemma 4.4. *Let α, β be algebraic numbers over $\mathbb{Q}$ and put*

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}(\alpha)]\,.$$

Then for all non-negative integers r and s there exist rational numbers $c_{r,s,i,j}$ such that

$$\alpha^r \beta^s = \sum_{i=0}^{d_1-1} \sum_{j=0}^{d_2-1} c_{r,s,i,j} \alpha^i \beta^j.$$

Furthermore, there exist positive integers c_1 and C_1, which depend only on α and β, such that $c_1^{r+s+d_1-1} c_{r,s,i,j} \in \mathbb{Z}$ and $|c_{r,s,i,j}| \leq C_1^{r+s}$.

Proof. The existence of rational numbers $c_{r,s,i,j}$ satisfying the conditions of Lemma 4.4 follows from Lemma 4.3. The proof proceeds in three steps.

Step 1. For a non-negative integer r, write

$$\alpha^r = a_{r,d_1-1}\alpha^{d_1-1} + \cdots + a_{r,1}\alpha + a_{r,0}.$$

If we denote the leading coefficient of the minimal polynomial of α by a, and put

$$A = 1 + \max_{0 \leq i \leq d_1-1} \{|a_{d_1,i}|\}\,,$$

then it follows from Lemma 2.14 that $a^{\max\{0,r-d_1+1\}} a_{r,i} \in \mathbb{Z}$ and $|a_{r,i}| \leq A^{\max\{0,r-d_1+1\}}$ for all i such that $0 \leq i \leq d_1 - 1$.

Step 2. For a non-negative integer s, write

$$\beta^s = \sum_{j=0}^{d_2-1} \sum_{i=0}^{d_1-1} b_{s,i,j} \alpha^i \beta^j.$$

Let b be the least common multiple of the denominators of $b_{d_2,i,j}$'s, and put

$$B_0 = \max_{\substack{0 \leq i \leq d_1-1 \\ 0 \leq j \leq d_2-1}} \{|b_{d_2,i,j}|\}\,, \quad B = 1 + d_1 B_0 + d_1(d_1-1)A^{d_1-1}B_0.$$

Note that $A \leq B$. In the second step, we prove that $a^{d_1-1} b^{\max\{0,s-d_2+1\}} b_{s,i,j} \in \mathbb{Z}$ and $|b_{s,i,j}| \leq B^{\max\{0,s-d_2+1\}}$ for all i and j such that $0 \leq i \leq d_1 - 1$ and $0 \leq j \leq d_2 - 1$. We

prove these facts by induction. Our statements are trivially true for $0 \leq s \leq d_2$. If we define $b_{s,i,-1} = 0$, then for all $s \geq d_2$ we have

$$\begin{aligned}
\beta^{s+1} &= \beta \cdot \beta^s \\
&= \beta \sum_{j=0}^{d_2-1} \sum_{i=0}^{d_1-1} b_{s,i,j} \alpha^i \beta^j \\
&= \beta^{d_2} \sum_{i=0}^{d_1-1} b_{s,i,d_2-1} \alpha^i + \sum_{j=1}^{d_2-1} \sum_{i=0}^{d_1-1} b_{s,i,j-1} \alpha^i \beta^j \\
&= \left(\sum_{j=0}^{d_2-1} \sum_{i=0}^{d_1-1} b_{d_2,i,j} \alpha^i \beta^j \right) \cdot \left(\sum_{i=0}^{d_1-1} b_{s,i,d_2-1} \alpha^i \right) + \sum_{j=1}^{d_2-1} \sum_{i=0}^{d_1-1} b_{s,i,j-1} \alpha^i \beta^j \\
&= \sum_{j=0}^{d_2-1} \left(\sum_{i=0}^{d_1-1} b_{d_2,i,j} \alpha^i \right) \left(\sum_{i=0}^{d_1-1} b_{s,i,d_2-1} \alpha^i \right) \beta^j + \sum_{j=1}^{d_2-1} \sum_{i=0}^{d_1-1} b_{s,i,j-1} \alpha^i \beta^j \\
&= \sum_{j=0}^{d_2-1} \left[\sum_{k=0}^{2d_1-2} \left(\sum_{\ell=0}^{k} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1} \right) \alpha^k + \sum_{i=0}^{d_1-1} b_{s,i,j-1} \alpha^i \right] \beta^j \\
&= \sum_{j=0}^{d_2-1} \left[\sum_{k=0}^{2d_1-2} \left(\sum_{\ell=0}^{k} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1} \right) \left(\sum_{i=0}^{d_1-1} a_{k,i} \alpha^i \right) + \sum_{i=0}^{d_1-1} b_{s,i,j-1} \alpha^i \right] \beta^j
\end{aligned}$$

$$= \sum_{j=0}^{d_2-1} \sum_{i=0}^{d_1-1} \left[b_{s,i,j-1} + \sum_{k=0}^{2d_1-2} a_{k,i} \sum_{\ell=0}^{k} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1} \right] \alpha^i \beta^j.$$

We conclude that

$$\begin{aligned}
b_{s+1,i,j} &= b_{s,i,j-1} + \sum_{k=0}^{2d_1-2} a_{k,i} \sum_{\ell=0}^{k} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1} \\
&= b_{s,i,j-1} + \sum_{\ell=0}^{i} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1} + \sum_{k=d_1}^{2d_1-2} a_{k,i} \sum_{\ell=0}^{k} b_{d_2,\ell,j} b_{s,k-\ell,d_2-1}.
\end{aligned}$$

Since $a^{\max\{0,k-d_1+1\}} a_{k,i} \in \mathbb{Z}$, we conclude that $a^{d_1-1} b^{s-d_2+2} b_{s+1,i,j} \in \mathbb{Z}$. Further, since $|a_{k,i}| \leq$

$A^{\max\{0,k-d_1+1\}}$, $|b_{d_2,i,j}| \le B_0$, and $|b_{s,i,j}| \le B^{s-d_2+1}$ by the induction hypothesis, we see that

$$\begin{aligned}
|b_{s+1,i,j}| &\le |b_{s,i,j-1}| + \sum_{\ell=0}^{i} |b_{d_2,\ell,j} b_{s,k-\ell,d_2-1}| + \sum_{k=d_1}^{2d_1-2} |a_{k,i}| \sum_{\ell=0}^{k} |b_{d_2,\ell,j} b_{s,k-\ell,d_2-1}| \\
&\le B^{s-d_2+1} + (i+1) B_0 B^{s-d_2+1} + B_0 B^{s-d_2+1} \sum_{k=1}^{d_1-1} (k+1) A^k \\
&\le B^{s-d_2+1} + d_1 B_0 B^{s-d_2+1} + d_1 B_0 B^{s-d_2+1} \sum_{k=1}^{d_1-1} A^k \\
&\le B^{s-d_2+1} (1 + d_1 B_0 + d_1 (d_1-1) A^{d_1-1} B_0) \\
&= B^{s-d_2+2}.
\end{aligned}$$

Step 3. In the third step, we prove that $a^{r+d_1-1} b^{\max\{0,s-d_2+1\}} c_{r,s,i,j} \in \mathbb{Z}$ and $|c_{r,s,i,j}| \le d_1 A^r B^{\max\{0,s-d_2+1\}}$ for all i and j such that $0 \le i \le d_1 - 1$ and $0 \le j \le d_2 - 1$. We consider the expansion of $\alpha^r \beta^s$:

$$\begin{aligned}
\alpha^r \beta^s &= \sum_{j=0}^{d_2-1} \sum_{m=0}^{d_1-1} b_{s,m,j} \alpha^{r+m} \beta^j \\
&= \sum_{j=0}^{d_2-1} \sum_{m=0}^{d_1-1} b_{s,m,j} \left(\sum_{i=0}^{d_1-1} a_{r+m,i} \alpha^i \right) \beta^j \\
&= \sum_{j=0}^{d_2-1} \sum_{i=0}^{d_1-1} \left(\sum_{m=0}^{d_1-1} b_{s,m,j} a_{r+m,i} \right) \alpha^i \beta^j.
\end{aligned}$$

Therefore

$$c_{r,s,i,j} = \sum_{m=0}^{d_1-1} b_{s,m,j} a_{r+m,i},$$

Since $a^{d_1-1} b^{\max\{0,s-d_2+1\}} b_{s,m,j} \in \mathbb{Z}$ and $a^r a_{r+m,i} \in \mathbb{Z}$ for all m such that $0 \le m \le d_1 - 1$, we

conclude that $a^{r+d_1-1}b^{\max\{0,s-d_2+1\}}c_{r,s,i,j} \in \mathbb{Z}$. Further,

$$\begin{aligned}
|c_{r,s,i,j}| &\leq \sum_{m=0}^{d_1-1} |b_{s,m,j}a_{r+m,i}| \\
&\leq B^{\max\{0,s-d_2+1\}} \sum_{m=0}^{d_1-1} A^{\max\{0,r+m-d_1+1\}} \\
&\leq d_1 A^r B^{\max\{0,s-d_2+1\}} \\
&\leq d_1 B^{r+\max\{0,s-d_2+1\}},
\end{aligned}$$

where the last inequality follows from the fact that $A \leq B$. From the observations made above we conclude that the values $c_1 = \operatorname{lcm}(a,b)$ and $C_1 = d_1 B$ would satisfy the hypothesis of the lemma. □

Lemma 4.5. *Let r_1, r_2 be positive real numbers with $r_2 \leq r_1$ and let t be a real number such that $0 \leq t \leq 1$. Then the number of solutions to the inequality*

$$\frac{i}{r_1} + \frac{j}{r_2} < t \tag{4.1.1}$$

in integers i, j such that $0 \leq i \leq r_1$, $0 \leq j \leq r_2$ is less than

$$\frac{r_1 r_2}{2} t^2 + \left(r_1 + \frac{3r_2}{2}\right) t + 1.$$

Proof. For each fixed $i \in \{0, 1, \ldots, \lfloor r_1 t \rfloor\}$ only

$$j = 0, 1, \ldots, \left\lfloor r_2 t - \frac{r_2 i}{r_1} \right\rfloor$$

could satisfy the inequality (4.1.1). So the total number of solutions does not exceed

$$\begin{aligned}
\sum_{i=0}^{\lfloor r_1 t \rfloor} \left(\left\lfloor r_2 t - \frac{r_2 i}{r_1} \right\rfloor + 1 \right) &\leq (r_1 t + 1)(r_2 t + 1) - \frac{r_2}{r_1} \cdot \frac{\lfloor r_1 t \rfloor \left(\lfloor r_1 t \rfloor + 1 \right)}{2} \\
&< (r_1 t + 1)(r_2 t + 1) - \frac{r_2 t (r_1 t - 1)}{2} \\
&= \frac{r_1 r_2}{2} t^2 + \left(r_1 + \frac{3r_2}{2}\right) t + 1.
\end{aligned}$$

□

Let $(r_1, r_2) \in \mathbb{N}^2$, and let $P(X,Y) \in \mathbb{Z}[X,Y]$ be a polynomial such that $\deg_X P \leq r_1$ and $\deg_Y P \leq r_2$. For $(\alpha, \beta) \in \overline{\mathbb{Q}}^2$, write $P(X,Y)$ as

$$P(X,Y) = \sum_{i=0}^{r_1} \sum_{j=0}^{r_2} p_{i,j}(X-\alpha)^i (Y-\beta)^j.$$

Definition 4.6. We define the *index of* $P(X,Y)$ *at* (α,β) *relative to* (r_1,r_2) as

$$\operatorname{ind}(P;\alpha,\beta;r_1,r_2) = \min_{p_{i,j}\neq 0} \left(\frac{i}{r_1} + \frac{j}{r_2} \right).$$

Lemma 4.7. *Let* α, β *be algebraic numbers such that*

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha,\beta)\colon \mathbb{Q}(\alpha)],$$

$d_1 d_2 \geq 2$. *Let* t, δ *be such that*

$$0 < t < \sqrt{\frac{2}{d_1 d_2}}, \quad \frac{d_1 d_2}{2} t^2 < \delta < 1.$$

Let c_1, C_1 *be the constants from Lemma 4.4 and* r_1, r_2 *be positive integers such that the inequalities*

$$r_1 + r_2 \geq (d_1 - 1)\frac{\log c_1}{\log 2}, \tag{4.1.2}$$

and

$$d_1 d_2 \left(\frac{r_1 r_2}{2} t^2 + \left(r_1 + \frac{3r_2}{2} \right) t + 1 \right) \leq \delta r_1 r_2 \tag{4.1.3}$$

are satisfied. There exists a non-zero polynomial $P(X,Y) \in \mathbb{Z}[X,Y]$ *such that*

$$\deg_X P \leq r_1, \quad \deg_Y P \leq r_2,$$
$$H(P) < (16 c_1 C_1)^{(r_1+r_2)\delta/(1-\delta)}$$

and

$$\operatorname{ind}(P;\alpha,\beta;r_1,r_2) \geq t.$$

Proof. Note that our choice of t, δ implies $d_1 d_2 t^2/2 < \delta$, so the inequalities (4.1.2) and (4.1.3) most certainly can be satisfied if r_1, r_2 are chosen sufficiently large. We will show that there exists a non-zero polynomial

$$P(X,Y) = \sum_{i=0}^{r_1} \sum_{j=0}^{r_2} p_{i,j} X^i Y^j$$

with the properties mentioned above.

Let m be the number of pairs $(h_1, h_2) \in \{0, 1, \ldots, r_1\} \times \{0, 1, \ldots, r_2\}$ such that $(h_1/r_1 + h_2/r_2) < t$. By Lemma 4.5,

$$m < \frac{r_1 r_2}{2} t^2 + \left(r_1 + \frac{3r_2}{2} \right) t + 1. \tag{4.1.4}$$

For each such (h_1, h_2) we impose the condition $D_{h_1,h_2} P(\alpha, \beta) = 0$. Since

$$\begin{aligned} D_{h_1,h_2} P(\alpha, \beta) &= \sum_{i=h_1}^{r_1} \sum_{j=h_2}^{r_2} \binom{i}{h_1} \binom{j}{h_2} p_{i,j} \alpha^{i-h_1} \beta^{j-h_2} \\ &= \sum_{i=h_1}^{r_1} \sum_{j=h_2}^{r_2} \binom{i}{h_1} \binom{j}{h_2} p_{i,j} \left(\sum_{k=0}^{d_1-1} \sum_{\ell=0}^{d_2-1} c_{i-h_1,j-h_2,k,\ell} \alpha^k \beta^\ell \right) \\ &= \sum_{k=0}^{d_1-1} \sum_{\ell=0}^{d_2-1} \left(\sum_{i=h_1}^{r_1} \sum_{j=h_2}^{r_2} \binom{i}{h_1} \binom{j}{h_2} c_{i-h_1,j-h_2,k,\ell} p_{i,j} \right) \alpha^k \beta^\ell, \end{aligned}$$

each condition $D_{h_1,h_2} P(\alpha, \beta) = 0$ corresponds to $d_1 d_2$ linear equations indexed by $(k, \ell) \in \{0, 1, \ldots, d_1 - 1\} \times \{0, 1, \ldots, d_2 - 1\}$ of the form

$$\sum_{i=h_1}^{r_1} \sum_{j=h_2}^{r_2} \binom{i}{h_1} \binom{j}{h_2} c_{i-h_1,j-h_2,k,\ell} p_{i,j} = 0.$$

Thus there are $M = d_1 d_2 m$ linear equations over $\mathbb{Q}$ in total. To convert the above equations from $\mathbb{Q}$ to $\mathbb{Z}$ we multiply each equation by $c_1^{r_1+r_2+d_1-1}$. The coefficients of these equations are integers such that

$$\begin{aligned} \left| c_1^{r_1+r_2+d_1-1} \sum_{i=h_1}^{r_1} \sum_{j=h_2}^{r_2} \binom{i}{h_1} \binom{j}{h_2} c_{i-h_1,j-h_2,k,\ell} \right| &\leq c_1^{r_1+r_2+d_1-1} (r_1 - h_1 + 1)(r_2 - h_2 + 1) 2^{r_1+r_2} C_1^{r_1+r_2-h_1-h_2} \\ &\leq (r_1 + 1)(r_2 + 1) c_1^{d_1-1} (2c_1 C_1)^{r_1+r_2} \\ &\leq c_1^{d_1-1} (4c_1 C_1)^{r_1+r_2} \\ &\leq (8c_1 C_1)^{r_1+r_2}, \end{aligned}$$

where the last inequality follows from (4.1.2). Let $A = (8c_1 C_1)^{r_1+r_2}$ and $N = (r_1+1)(r_2+1)$

be the number of coefficients of P. Then it follows from (4.1.3) and (4.1.4) that

$$\begin{aligned} M &= d_1 d_2 m \\ &< d_1 d_2 \left(\frac{r_1 r_2}{2} t^2 + \left(r_1 + \frac{3r_2}{2} \right) t + 1 \right) \\ &\leq \delta r_1 r_2 \\ &< \delta N. \end{aligned}$$

Therefore $M/(N-M) < \delta/(1-\delta)$, so by Lemma 2.13 there exist integers $p_{i,j}$, not all zero, such that

$$H(P) = \max\{|p_{i,j}|\} \leq (NA)^{\frac{M}{N-M}} < (16 c_1 C_1)^{(r_1+r_2)\delta/(1-\delta)}.$$

□

Lemma 4.8. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$ and let $|\ |$ denote the standard absolute value on K. Let α, β be algebraic numbers over $\mathbb{Q}$, and denote the leading coefficients of the minimal polynomials of α, β by c_α, c_β, respectively. Let $P \in \mathbb{Z}[X,Y]$ be a non-zero polynomial of bi-degree (r_1, r_2) and $(i,j) \in \{0, \dots, r_1\} \times \{0, \dots, r_2\}$.*

If $K = \mathbb{C}$ then

$$|D_{i,j} P(\alpha, \beta)| \leq H(P) \binom{r_1+1}{i+1} \binom{r_2+1}{j+1} \max\{1, |\alpha|\}^{r_1-i} \max\{1, |\beta|\}^{r_2-j} \tag{4.1.5}$$

If $K = \mathbb{Q}_p$ then

$$|D_{i,j} P(\alpha, \beta)| \leq |c_\alpha|^{-(r_1-i)} |c_\beta|^{-(r_2-j)}. \tag{4.1.6}$$

Proof. Let

$$P(X,Y) = \sum_{k=0}^{r_1} \sum_{\ell=0}^{r_2} r_{k\ell} X^k Y^\ell.$$

If $K = \mathbb{C}$, then it follows from the triangle inequality and the identity (2.2.2) that

$$\begin{aligned} |D_{i,j} P(\alpha, \beta)| &\leq \sum_{k=i}^{r_1} \sum_{\ell=j}^{r_2} \binom{k}{i} \binom{\ell}{j} |r_{k\ell}| \max\{1, |\alpha|\}^{k-i} \max\{1, |\beta|\}^{\ell-j} \\ &\leq H(P) \max\{1, |\alpha|\}^{r_1-i} \max\{1, |\beta|\}^{r_2-j} \left(\sum_{k=i}^{r_1} \binom{k}{i} \right) \cdot \left(\sum_{\ell=j}^{r_2} \binom{\ell}{j} \right) \\ &= H(P) \binom{r_1+1}{i+1} \binom{r_2+1}{j+1} \max\{1, |\alpha|\}^{r_1-i} \max\{1, |\beta|\}^{r_2-j}. \end{aligned}$$

If $K = \mathbb{Q}_p$, then due to the fact that $c_\alpha\alpha, c_\beta\beta$ are algebraic integers, the number $c_\alpha^{r_1-i}c_\beta^{r_2-j}D_{i,j}P(\alpha,\beta)$ is also an algebraic integer. Therefore its p-adic absolute value does not exceed one. Thus we conclude that

$$|D_{i,j}P(\alpha,\beta)| \le |c_\alpha|^{-(r_1-i)}|c_\beta|^{-(r_2-j)}.$$

□

Lemma 4.9. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let $\alpha, \beta \in K$ be algebraic numbers over $\mathbb{Q}$, and denote the leading coefficients of the minimal polynomials of α, β by c_α, c_β, respectively. Put*

$$d_1 = [\mathbb{Q}(\alpha,\beta)\colon \mathbb{Q}]\,.$$

Let $\mu > \sqrt{2d_1d_2}$, and choose δ, t and ε, with $\varepsilon > 0$, so that

$$\frac{2}{\mu} < t < \sqrt{\frac{2}{d_1d_2}},\quad \frac{dt^2}{2} < \delta < 1,$$

$$(1+\varepsilon)(2+\varepsilon) < t\mu.$$

Suppose that there exist rational numbers x_1/y_1, x_2/y_2,

$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{1}{H(x_1,y_1)^\mu},\quad \left|\beta - \frac{x_2}{y_2}\right| < \frac{1}{H(x_2,y_2)^\mu},$$

and

$$H(x_2,y_2) \ge H(x_1,y_1) \ge \left(4C_2(16c_1C_1)^{\delta/(1-\delta)}\right)^{1/\varepsilon}, \tag{4.1.7}$$

where c_1, C_1 are the constants from Lemma 4.4 and

$$C_2 = \begin{cases} \max\{1,|\alpha|,|\beta|\}, & \text{if } K = \mathbb{C};\\ \max\{c_\alpha, c_\beta\}, & \text{if } K = \mathbb{Q}_p. \end{cases}$$

Let r_1, r_2, P be as in Lemma 4.7, where, in addition, the inequalities $r_1 + r_2 \ge 2$ and

$$r_1 \log H(x_1,y_1) \le r_2 \log H(x_2,y_2) \le (1+\varepsilon) r_1 \log H(x_1,y_1)$$

are satisfied. Then

$$\operatorname{ind}(P; x_1/y_1, x_2/y_2; r_1, r_2) \ge t - \frac{(1+\varepsilon)(2+\varepsilon)}{\mu}.$$

Proof. By Lemma 4.7, the inequality $\operatorname{ind}(P;\alpha,\beta;r_1,r_2)\geq t$ holds. Let us fix some particular h_1 and h_2 such that

$$\frac{h_1}{r_1}+\frac{h_2}{r_2}<t-\frac{(1+\varepsilon)(2+\varepsilon)}{\mu},$$

and consider the polynomial $D_{h_1,h_2}P(X,Y)$. We would like to show that it vanishes at $(x_1/y_1,x_2/y_2)$. By Taylor's theorem,

$$D_{h_1,h_2}P(X,Y)=\sum_{i=0}^{r_1-h_1}\sum_{j=0}^{r_2-h_2}D_{i+h_1,j+h_2}P(\alpha,\beta)\,(X-\alpha)^i\,(Y-\beta)^j\,.$$

At this point, we recall that $D_{i+h_1,j+h_2}P(\alpha,\beta)=0$, unless

$$\frac{i+h_1}{r_1}+\frac{j+h_2}{r_2}\geq t.$$

If we let $\Delta_\alpha=x_1/y_1-\alpha$ and $\Delta_\beta=x_2/y_2-\beta$, then for such i and j we have

$$\begin{aligned}
-\log|\Delta_\alpha^i\Delta_\beta^j| &> \mu\left(\frac{i}{r_1}r_1\log H(x_1,y_1)+\frac{j}{r_2}r_2\log H(x_2,y_2)\right)\\
&\geq \mu r_1\log H(x_1,y_1)\left(\frac{i}{r_1}+\frac{j}{r_2}\right)\\
&\geq \mu r_1\log H(x_1,y_1)\left(t-\frac{h_1}{r_1}-\frac{h_2}{r_2}\right)\\
&\geq \frac{\mu}{2+\varepsilon}\left(t-\left(t-\frac{(1+\varepsilon)(2+\varepsilon)}{\mu}\right)\right)(r_1\log H(x_1,y_1)+r_2\log H(x_2,y_2))\\
&= (1+\varepsilon)\,(r_1\log H(x_1,y_1)+r_2\log H(x_2,y_2))\,.
\end{aligned}$$

Thus for the aforementioned i and j the following inequality holds:

$$\left|\Delta_\alpha^i\Delta_\beta^j\right|<\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-1-\varepsilon}. \tag{4.1.8}$$

Now, suppose that $K=\mathbb{C}$. Then it follows from the triangle inequality, Taylor's theorem, Lemma 4.8 and (4.1.8) that

$$\begin{aligned}
\left|y_1^{r_1}y_2^{r_2}D_{h_1,h_2}P\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right)\right| &\leq H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\sum_{j=0}^{r_2-h_2}\sum_{i=0}^{r_1-h_1}|D_{i+h_1,j+h_2}P(\alpha,\beta)|\cdot\left|\Delta_\alpha^i\Delta_\beta^j\right|\\
&< 2^{r_1+r_2+2}H(P)\max\{1,|\alpha|,|\beta|\}^{r_1+r_2}\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-\varepsilon}\\
&< \left(4(16c_1C_1)^{\delta/(1-\delta)}\max\{1,|\alpha|,|\beta|\}\right)^{r_1+r_2}\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-\varepsilon}\\
&\leq 1,
\end{aligned}$$

where the last inequality follows from (4.1.7). Since $y_1^{r_1}y_2^{r_2}D_{h_1,h_2}P\left(x_1/y_1, x_2/y_2\right)$ is an integer which is less than one in absolute value, it must be equal to zero. Therefore $D_{h_1,h_2}P(X,Y)$ vanishes at $(x_1/y_1, x_2/y_2)$, as claimed.

Suppose that $K = \mathbb{Q}_p$. It follows from the non-Archimedean triangle inequality, Taylor's theorem, Lemma 4.8 and (4.1.8) that

$$\begin{aligned}\left|y_1^{r_1}y_2^{r_2}D_{h_1,h_2}P\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right)\right| &\leq \max_{\substack{0\leq i\leq r_1-h_1\\ 0\leq j\leq r_2-h_2}}\left\{|D_{i+h_1,j+h_2}P(\alpha,\beta)|\cdot|\Delta_\alpha^i\Delta_\beta^j|\right\}\\ &< \max\{c_\alpha,c_\beta\}^{r_1+r_2}\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-1-\varepsilon}.\end{aligned}$$

Suppose that $D_{h_1,h_2}P(x_1/y_1, x_2/y_2) \neq 0$. Then it follows form the product formula that

$$\begin{aligned}\left|y_1^{r_1}y_2^{r_2}D_{h_1,h_2}P\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right)\right| &\geq \left|y_1^{r_1}y_2^{r_2}D_{h_1,h_2}P\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right)\right|_\infty^{-1}\\ &\geq \left((r_1+1)(r_2+1)H(D_{h_1,h_2}P)H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-1},\\ &\geq \left(4^{r_1+r_2}H(P)H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-1}\\ &> \left(4^{r_1+r_2}(16c_1C_1)^{(r_1+r_2)\delta/(1-\delta)}H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{-1}.\end{aligned}$$

where $|\ |_\infty$ denotes the Archimedean absolute value. Combining upper and lower bounds on $|y_1^{r_1}y_2^{r_2}P(x_1/y_1, x_2/y_2)|$, we obtain

$$\frac{1}{4^{r_1+r_2}(16c_1C_1)^{(r_1+r_2)\delta/(1-\delta)}H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}} < \frac{\max\{c_\alpha,c_\beta\}^{r_1+r_2}}{\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{1+\varepsilon}},$$

or equivalently

$$\left(H(x_1,y_1)^{r_1}H(x_2,y_2)^{r_2}\right)^{\varepsilon} < \left(4(16c_1C_1)^{\delta/(1-\delta)}\max\{c_\alpha,c_\beta\}\right)^{r_1+r_2}.$$

Since $H(x_2,y_2) \geq H(x_1,y_1)$, we conclude that

$$H(x_1,y_1)^{\varepsilon} < 4(16c_1C_1)^{\delta/(1-\delta)}\max\{c_\alpha,c_\beta\},$$

which contradicts (4.1.7). Therefore $D_{h_1,h_2}P(X,Y)$ vanishes at $(x_1/y_1, x_2/y_2)$, as claimed. □

Lemma 4.10. *Let $0 < \varepsilon < 1/12$, and put $w = \varepsilon^2/24$. Let r_1, r_2 be positive integers for which $wr_1 \geq r_2$, and let $x_1/y_1, x_2/y_2$ be rational numbers such that $y_2^{r_2} \geq y_1^{r_1}$, $y_1^w \geq 64$, and $y_2^w \geq 64$. Suppose that $P(X,Y) \in \mathbb{Z}[X,Y]$ is a non-zero polynomial such that $\deg_X P \leq r_1$, $\deg_Y P \leq r_2$, and $H(P) \leq y_1^{wr_1}$. Then $\mathrm{ind}(P; x_1/y_1, x_2/y_2; r_1, r_2) \leq \varepsilon$.*

Proof. See [7, Chapter VI, Theorem IV]. □

4.2 Proof of Theorem 4.1

We choose t, δ and ε as follows:

$$t = \frac{1}{\mu} + \frac{1}{\sqrt{2d}}, \quad \delta = \frac{1}{2} + \frac{dt^2}{4}, \quad \varepsilon = \frac{3\left(\mu - \sqrt{2d}\right)}{4\sqrt{2d}(\mu+3)^2}.$$

One can easily verify that with such choice of parameters the following inequalities are satisfied:

$$\frac{2}{\mu} < t < \sqrt{\frac{2}{d}}, \quad \frac{dt^2}{2} < \delta < 1,$$
$$0 < \varepsilon < \min\left\{\frac{\sqrt{\mu^2 + 4t\mu + 6\mu + 1} - \mu - 3}{2}, \frac{1}{12}\right\}. \tag{4.2.1}$$

Let a be the leading coefficient of the minimal polynomial of α and $b > 0, b_{0,0}, \ldots, b_{d_1-1,d_2-1}$ be integers, with no factor in common, such that

$$b\beta^{d_2} = \sum_{i=0}^{d_1-1}\sum_{j=0}^{d_2-1} b_{i,j}\alpha^i\beta^j,$$

where

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha,\beta)\colon \mathbb{Q}(\alpha)].$$

Let c_α, c_β denote the leading coefficients of the minimal polynomials of α, β, respectively. Define

$$w = \varepsilon^2/24, \quad \eta = 2/w,$$
$$c_1 = \operatorname{lcm}(a, b),$$
$$C_1 = d_1 + d_1^2\left(1 + (d_1-1)(1+H(\alpha))^{d_1-1}\right) \max_{\substack{0\le i\le d_1-1\\ 0\le j\le d_2-1}}\{|b_{i,j}|/b\}.$$

$$C_2 = \begin{cases} \max\{1, |\alpha|, |\beta|\}, & \text{if } K = \mathbb{C};\\ \max\{c_\alpha, c_\beta\}, & \text{if } K = \mathbb{Q}_p, \end{cases}$$

$$C = \max\left\{(16c_1C_1)^{\frac{w^{-1}+1}{\delta^{-1}-1}}, \left(4C_2(16c_1C_1)^{\delta/(1-\delta)}\right)^{1/\varepsilon}, 64^{1/w}\right\}. \tag{4.2.2}$$

Now, suppose that x_1/y_1, x_2/y_2 are rational numbers such that

$$H(x_2, y_2) \ge H(x_1, y_1)^\eta \ge C^\eta,$$

$$\left|\alpha - \frac{x_1}{y_1}\right| < \frac{1}{H(x_1,y_1)^\mu}, \quad \left|\beta - \frac{x_2}{y_2}\right| < \frac{1}{H(x_2,y_2)^\mu}.$$

We will show that all of these conditions cannot be satisfied.

Let r_1 be an integer such that

$$r_1 \geq \max\left\{2, (d_1-1)\frac{\log c_1}{\log 2}, \frac{\log H(x_2,y_2)}{\varepsilon \log H(x_1,y_1)}\right\}.$$

Put

$$r_2 = \left\lfloor \frac{r_1 \log H(x_1,y_1)}{\log H(x_2,y_2)} \right\rfloor + 1.$$

Then

$$r_1 + r_2 > \max\left\{2, (d_1-1)\frac{\log c_1}{\log 2}\right\},$$

$$\begin{aligned} r_1 \log H(x_1,y_1) &< r_2 \log H(x_2,y_2) \\ &\leq r_1 \log H(x_1,y_1) + \log H(x_2,y_2) \\ &\leq (1+\varepsilon) r_1 \log H(x_1,y_1), \end{aligned}$$

so the conditions of Lemma 4.9 are satisfied. That is, there exists a non-zero polynomial $P(X,Y) \in \mathbb{Z}[X,Y]$ such that $H(P) < (16c_1C_1)^{(r_1+r_2)\delta/(1-\delta)}$ and

$$\operatorname{ind}(P; x_1/y_1, x_2/y_2; r_1, r_2) \geq t - \frac{(1+\varepsilon)(2+\varepsilon)}{\mu}. \tag{4.2.3}$$

Further, since

$$r_2 \log H(x_2,y_2) \leq (1+\varepsilon) r_1 \log H(x_1,y_1) \leq 2 r_1 \log H(x_1,y_1),$$

it follows form $H(x_2,y_2) \geq H(x_1,y_1)^\eta$ that

$$r_2 \leq r_1 \frac{2\log H(x_1,y_1)}{\log H(x_2,y_2)} \leq r_1 w.$$

Combining this inequality with the lower bound $H(x_1,y_1) \geq (16c_1C_1)^{\frac{w^{-1}+1}{\delta^{-1}-1}}$, we obtain

$$H(P) < (16c_1C_1)^{(r_1+r_2)\delta/(1-\delta)} \leq (16c_1C_1)^{r_1 \frac{w+1}{\delta^{-1}-1}} \leq H(x_1,y_1)^{w r_1}.$$

Consequently, the polynomial $P(X, Y)$ satisfies the conditions of Lemma 4.10, so we conclude that

$$\operatorname{ind}(P; x_1/y_1, x_2/y_2; r_1, r_2) \leq \varepsilon. \tag{4.2.4}$$

Upon combining (4.2.3) with (4.2.4) we obtain

$$t - \frac{(1+\varepsilon)(2+\varepsilon)}{\mu} \leq \varepsilon.$$

One can easily verify that this inequality contradicts our choice of ε in (4.2.1).

Chapter 5

Bounds on the Number of Solutions to a Wider Class of Equations of Thue and Thue-Mahler Type

In Chapter 2, we produced absolute bounds on the number of solutions of certain equations of Thue and Thue-Mahler type by exploring properties of *minimal pairs* associated with a pair of algebraic numbers (α, β), and then applying these properties to produce generalized Archimedean and non-Archimedean gap principles. One of the limitations of our method was that the number β would have to be taken from the field $\mathbb{Q}(\alpha)$. In this chapter, we push our theory to its (currently visible) limits and introduce the notion of a *minimal polynomial* of a pair of algebraic numbers (α, β). If the degree of β over $\mathbb{Q}(\alpha)$ is small relative to the degree of α over $\mathbb{Q}$, we are, once again, able to produce generalizations of Archimedean and non-Archimedean gap principles, and then combine them with the generalized Thue-Siegel principle established in Chapter 4 to produce bounds on the number of solutions of equations of Thue and Thue-Mahler type. Though our bounds are not absolute, they still yield improvements over what is presently available in the literature.

In order to state our main result, we need to introduce the notion of a *minimal polynomial* of a pair of algebraic numbers (α, β) and explain which minimal polynomials are considered to be *μ-special* and *μ-exceptional*.

Definition 5.1. Let α, β be algebraic numbers over $\mathbb{Q}$. A *minimal polynomial* $R(X_1, X_2)$ of (α, β) is a non-zero polynomial that satisfies the following properties.

(1) $R(X_1, X_2) \in \mathbb{Z}[X_1, X_2]$.

(2) $R(\alpha, \beta) = 0$.

(3) The quantity $\max\{\deg_{X_1} R, \deg_{X_2} R\}$ is minimal among all polynomials satisfying properties (1), (2).

(4) Let $i, j \in \{1, 2\}$ be distinct, with i such that $\deg_{X_i} R = \max_k\{\deg_{X_k} R\}$. The quantity $\deg_{X_j} R$ is minimal among all polynomials satisfying properties (1), (2), (3).

(5) The quantity $H(R)$ is minimal among all polynomials satisfying properties (1) – (4).

Example 5.2. If R is a minimal polynomial of (α, β) then $-R$ is also a minimal polynomial of (α, β). This already demonstrates that a minimal polynomial is not unique. Furthermore, uniqueness is not guaranteed even if we take multiplication by -1 into account. Indeed, let

$$\alpha = 2\cos\left(\frac{2\pi}{17}\right), \quad \beta = 2\cos\left(\frac{8\pi}{17}\right).$$

Then both

$$R_1(X, Y) = (X^2 - 1)Y^2 + (X + 1)Y - X^2 + X + 1$$

and

$$R_2(X, Y) = XY^2 + (X^2 + X - 1)Y - X + 1$$

are minimal polynomials for (α, β).

Definition 5.3. Let α, β be algebraic numbers over $\mathbb{Q}$ and put

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha, \beta)\colon \mathbb{Q}(\alpha)].$$

Let μ be a real number such that $0 < \mu \leq d_1$. If R is a minimal polynomial of (α, β) of bi-degree (r_1, r_2), we call it *μ-special* if

$$\left\lfloor \frac{d_1 d_2 + r_1 r_2}{r_1 + r_2} \right\rfloor \geq \frac{\mu^2}{d_1 + \mu}.$$

If, in addition, the curve $R(X_1, X_2) = 0$ contains infinitely many rational points, we call it *μ-exceptional*.

Example 5.4. By Definition 5.1, for any algebraic number α over $\mathbb{Q}$ the polynomials $X_1 - X_2$ and $X_2 - X_1$ are minimal polynomials of (α, α). Furthermore, by Definition 5.3, for any μ such that $0 < \mu \leq d_1$ they are μ-exceptional. More generally, if F is an irreducible binary form, the same observation applies to any polynomial

$$R(X_1, X_2) = vX_1 - u + X_2(tX_1 - s),$$

which corresponds to an automorphism $|sv-tu|^{-1/2}\left(\begin{smallmatrix} s & u \\ t & v \end{smallmatrix}\right) \in \mathrm{Aut}'\,|F|$ (see Proposition 2.7). In Theorems 2.2, 2.3 we exploited the fact that every μ-special polynomial has to arise from an automorphism, provided that we change the definition of a minimal polynomial so that $\deg_{X_2} R \leq 1$ and assume $\beta \in \mathbb{Q}(\alpha)$, $\mu > d/2+1$. In this chapter, we weaken both of these restrictions.

Our first result is given in Theorem 5.5. Note that, unlike in Theorems 2.2, 2.3, all of the numbers in its statement are explicitly computable.

Theorem 5.5. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d_1 \geq 37$ and content one. Let $\alpha_1, \ldots, \alpha_{d_1}$ be the roots of $F(X,1)$, put*

$$d_2 = \max_{1 \leq i \leq d_1} \{[\mathbb{Q}(\alpha_1, \alpha_i) \colon \mathbb{Q}(\alpha_1)]\},$$

and suppose that $d_2 < d_1/36$. Let λ, μ, γ be such that

$$0 \leq \lambda < \frac{d_1}{2\sqrt{d_1 d_2}} - 2, \quad \mu = \frac{d_1}{2+\lambda},$$

$$\gamma = \#\{i \colon \text{a minimal polynomial of } (\alpha_1, \alpha_i) \text{ is } \mu\text{-special}\}.$$

Note that $\gamma \geq 1$, because minimal polynomials of (α_1, α_1) are always μ-special.

For a prime p and a positive integer k, consider the Diophantine equation

$$|F(x,y)| = tp^k. \tag{5.0.1}$$

There exists an explicitly computable number $C_1 = C_1(\mu, F)$ such that if $p^k \geq C_1$ then the number of solutions to (5.0.1) in integers (x,y,t) satisfying

$$\gcd(x,y) = 1, \quad 1 \leq t \leq (p^k)^\lambda$$

is less than

$$2\gamma + 4\gamma \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu+3) - 2\log\left(\mu/\sqrt{2d_1d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1d_2} - 1\right)} \right\rfloor.$$

The usage of minimal and μ-special polynomials deserves explanation. Consider Theorem 5.5 in the case $d_2 = 1$, which was studied in detail in Chapter 2. Let R_i be a minimal

polynomial of (α_1, α_i) of bi-degree (r_1, r_2). Setting $\lambda = 0$ implies that $\mu = d_1/2$, and so a polynomial R_i is considered to be μ-special if

$$\left\lfloor \frac{d_1 + r_1 r_2}{r_1 + r_2} \right\rfloor \geq \frac{d_1}{6}.$$

Applying $x \geq \lfloor x \rfloor$ and rearranging, we obtain

$$\frac{r_1 + r_2}{6} - 1 \leq \frac{r_1 r_2}{d_1} \leq 1,$$

where the last inequality follows from the fact that $\max\{r_1, r_2\} \leq \lfloor \sqrt{d_1 d_2} \rfloor$ (see Proposition 5.10). We conclude that $\max\{r_1, r_2\} \leq 12$, so the maximum of the degrees of R_i is small in comparison to d_1. For d_1 sufficiently large, we expect that among the minimal polynomials $R_1, R_2, \ldots, R_{d_1}$ only few will satisfy such a restrictive condition.

Our second result is as follows.

Theorem 5.6. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be an irreducible binary form of degree $d_1 \geq 8$ and content one. Let $\alpha_1, \ldots, \alpha_{d_1}$ be the roots of $F(X, 1)$, put*

$$d_2 = \max_{1 \leq i \leq d_1} \{[\mathbb{Q}(\alpha_1, \alpha_i) \colon \mathbb{Q}(\alpha_1)]\},$$

and suppose that $d_2 < (d_1 - 2.05)^2/4d_1$. Let λ, μ, γ be such that

$$0 \leq \lambda < \frac{d_1 - 2.05}{2\sqrt{d_1 d_2}} - 1, \quad \mu = \frac{d_1 - 2.05}{1 + \lambda},$$

$$\gamma = \#\{i \colon \text{a minimal polynomial of } (\alpha_1, \alpha_i) \text{ is } \mu\text{-exceptional}\}.$$

Note that $\gamma \geq 1$, because minimal polynomials of (α_1, α_1) are always μ-exceptional.

For a prime p, consider the Diophantine equation

$$|F(x, y)| = t p^z. \tag{5.0.2}$$

The number of solutions to (5.0.2) in integers (x, y, z, t) such that

$$\gcd(x, y) = 1, \quad z \geq 1, \quad 1 \leq t \leq (p^z)^\lambda$$

does not exceed

$$2\gamma \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu + 3) - 2\log\left(\mu/\sqrt{2d_1 d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1 d_2} - 1\right)} \right\rfloor.$$

5.1 Preliminary Results

Lemma 5.7. *Let $P, Q \in \mathbb{Z}[X_1, X_2]$ be polynomials of bi-degrees (r_1, r_2), (s_1, s_2), respectively. Denote $h(X_1) = \mathrm{Res}_{X_2}(P, Q) \in \mathbb{Z}[X_1]$, where polynomials P, Q inside the resultant are viewed as polynomials in X_2 with coefficients in $\mathbb{Z}[X_1]$. Then*

$$\deg h \le r_1 s_2 + r_2 s_1,$$

$$H(h) \le (r_1+1)^{s_2-1}(r_2+1)^{s_2}(s_1+1)^{r_2-1}(s_2+1)^{r_2}(r_1 s_2+1)H(P)^{s_2}H(Q)^{r_2}.$$

Proof. Write

$$P(X_1, X_2) = \sum_{i=0}^{r_2} p_i(X_1)X_2^i, \;\; Q(X_1, X_2) = \sum_{j=0}^{s_2} q_j(X_1)X_2^j.$$

By definition, $\mathrm{Res}_{X_2}(P, Q)$ is equal to the determinant of a $(r_2 + s_2) \times (r_2 + s_2)$ matrix. In turn, this determinant is equal to the summation of at most $(r_2+1)^{s_2}(s_2+1)^{r_2}$ terms of the form

$$\pm p_{i_1}(X_1) \cdots p_{i_{s_2}}(X_1) q_{j_1}(X_1) \cdots q_{j_{r_2}}(X_1),$$

where $i_1, \ldots, i_{s_2} \in \{0, \ldots, r_2\}$ and $j_1, \ldots, j_{r_2} \in \{0, \ldots, s_2\}$. Since $\deg p_j \le r_1$ and $\deg q_j \le s_1$, we conclude that $\deg h \le r_1 s_2 + r_2 s_1$.

It remains to estimate $H(h)$. Recall that for any polynomials $f, g \in \mathbb{Z}[X_1]$ the inequality $H(f) \le (\deg f + 1)H(f)H(g)$ holds. Therefore

$$\begin{aligned}
&H\left(p_{i_1}(X_1) \cdots p_{i_{s_2}}(X_1) q_{j_1}(X_1) \cdots q_{j_{r_2}}(X_1)\right) \\
&\le (r_1 s_2 + 1) H\left(p_{i_1}(X_1) \cdots p_{i_{s_2}}(X_1)\right) H\left(q_{j_1}(X_1) \cdots q_{j_{r_2}}(X_1)\right) \\
&\le (r_1 s_2 + 1)\left((r_1+1)^{s_2-1} H(P)^{s_2}\right)\left((s_1+1)^{r_2-1} H(Q)^{r_2}\right).
\end{aligned}$$

Since there are at most $(r_2+1)^{s_2}(s_2+1)^{r_2}$ terms in total, the result follows. □

Lemma 5.8. *Let α, β be algebraic numbers over $\mathbb{Q}$ and put*

$$d_1 = [\mathbb{Q}(\alpha) \colon \mathbb{Q}], \;\; d_2 = [\mathbb{Q}(\alpha, \beta) \colon \mathbb{Q}(\alpha)].$$

Denote the leading coefficients of the minimal polynomials of α, β by c_α, c_β, respectively. Let $R \in \mathbb{Z}[X_1, X_2]$ be a non-zero polynomial of bi-degree (r_1, r_2) such that $R(\alpha, \beta) \ne 0$.

If $\alpha, \beta \in \mathbb{C}$ then

$$|R(\alpha, \beta)| \ge \frac{\max\{1, |\alpha|\}^{r_1} \max\{1, |\beta|\}^{r_2}}{\left(c_\alpha^{r_1} c_\beta^{r_2}(r_1+1)(r_2+1)H(R)\right)^{d_1 d_2 - 1} M(\alpha)^{r_1 d_2} M(\beta)^{r_2 d_1}}. \tag{5.1.1}$$

If $\alpha, \beta \in \mathbb{Q}_p$ then

$$|R(\alpha,\beta)|_p \geq \frac{\left|c_\alpha^{r_1} c_\beta^{r_2}\right|_p^{-1}}{(c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2 - 1} \left((r_1+1)(r_2+1)H(R)\right)^{d_1 d_2} M(\alpha)^{r_1} M(\beta)^{r_2}}. \tag{5.1.2}$$

Proof. Since $c_\alpha \alpha, c_\beta \beta$ are algebraic integers, the number $c_\alpha^{r_1} c_\beta^{r_2} R(\alpha,\beta)$ is also an algebraic integer.

Suppose that $\alpha, \beta \in \mathbb{C}$. Denote the complex conjugates of α over $\mathbb{Q}$ by $\alpha = \alpha_1, \ldots, \alpha_{d_1}$ and the complex conjugates of β over $\mathbb{Q}(\alpha)$ by $\beta_1, \ldots, \beta_{d_2}$. Then the number

$$N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}\left(c_\alpha^{r_1} c_\beta^{r_2} R(\alpha,\beta)\right) = \prod_{i=1}^{d_1}\prod_{j=1}^{d_2} c_\alpha^{r_1} c_\beta^{r_2} R(\alpha_i,\beta_j)$$

is a non-zero rational integer. We can bound this quantity from above as follows:

$$\begin{aligned}
&\left|N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}\left(c_\alpha^{r_1} c_\beta^{r_2} R(\alpha,\beta)\right)\right| \\
&= \prod_{i=1}^{d_1}\prod_{j=1}^{d_2} \left|c_\alpha^{r_1} c_\beta^{r_2} R(\alpha_i,\beta_j)\right| \\
&\leq (c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2} \prod_{i=1}^{d_1}\prod_{j=1}^{d_2} (r_1+1)(r_2+1)H(R)\max\{1,|\alpha_i|\}^{r_1}\max\{1,|\beta_j|\}^{r_2} \\
&= \left(c_\alpha^{r_1} c_\beta^{r_2}(r_1+1)(r_2+1)H(R)\right)^{d_1 d_2}\left(\prod_{i=1}^{d_1}\max\{1,|\alpha_i|\}\right)^{r_1}\cdot\left(\prod_{j=1}^{d_2}\max\{1,|\beta_j|\}\right)^{r_2} \\
&\leq (c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2 - 1}\left((r_1+1)(r_2+1)H(R)\right)^{d_1 d_2} M(\alpha)^{r_1} M(\beta)^{r_2}.
\end{aligned} \tag{5.1.3}$$

Since $R(\alpha,\beta) \neq 0$, we conclude that

$$\begin{aligned}
|R(\alpha,\beta)|^{-1} &\leq (c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2} \prod_{(i,j)\neq(1,1)} |R(\alpha_i,\beta_j)| \\
&\leq (c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2} \prod_{(i,j)\neq(1,1)} (r_1+1)(r_2+1)H(R)\max\{1,|\alpha_i|\}^{r_1}\max\{1,|\beta_j|\}^{r_2} \\
&= (c_\alpha^{r_1} c_\beta^{r_2})^{d_1 d_2}\left((r_1+1)(r_2+1)H(R)\right)^{d_1 d_2 - 1}\frac{\left(\prod_{i=1}^{d_1}\max\{1,|\alpha_i|\}\right)^{r_1}\cdot\left(\prod_{j=1}^{d_2}\max\{1,|\beta_j|\}\right)^{r_2}}{\max\{1,|\alpha|\}^{r_1}\max\{1,|\beta|\}^{r_2}} \\
&\leq \frac{\left(c_\alpha^{r_1} c_\beta^{r_2}(r_1+1)(r_2+1)H(R)\right)^{d_1 d_2 - 1} M(\alpha)^{r_1} M(\beta)^{r_2}}{\max\{1,|\alpha|\}^{r_1}\max\{1,|\beta|\}^{r_2}}.
\end{aligned}$$

Suppose that $\alpha, \beta \in \mathbb{Q}_p$. Put $\gamma_1 = c_\alpha^{r_1} c_\beta^{r_2} R(\alpha, \beta)$. Then γ_1 is an algebraic integer, so its p-adic absolute value, as well as p-adic absolute values of its conjugates $\gamma_2, \dots, \gamma_{\deg \gamma_1}$, does not exceed one. Therefore

$$\left|N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}(\gamma_1)\right|_p = |\gamma_1|_p \cdot \prod_{i=2}^{\deg \gamma_1} |\gamma_i|_p \le \left|c_\alpha^{r_1} c_\beta^{r_2} R(\alpha, \beta)\right|_p.$$

Since $N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}(\gamma_1)$ is a non-zero integer, it follows from the product formula and (5.1.3) that

$$\begin{aligned}\left|N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}(\gamma_1)\right|_p &\ge \left|N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}(\gamma_1)\right|^{-1} \\ &\ge (c_\alpha^{r_1} c_\beta^{r_2})^{-(d_1 d_2 - 1)} \left((r_1+1)(r_2+1)H(R)\right)^{-d_1 d_2} M(\alpha)^{-r_1} M(\beta)^{-r_2}.\end{aligned}$$

The result follows once we combine upper and lower bounds on $|N_{\mathbb{Q}(\alpha,\beta)/\mathbb{Q}}(\gamma_1)|_p$.

□

Lemma 5.9. *Let α, β be algebraic numbers over $\mathbb{Q}$ and put*

$$d_1 = [\mathbb{Q}(\alpha) \colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha, \beta) \colon \mathbb{Q}(\alpha)].$$

Let c_α, c_β denote the leading coefficients of the minimal polynomials of α, β, respectively. Suppose that, for some ξ, η, the polynomial $R \in \mathbb{Z}[X_1, X_2]$ of bi-degree (r_1, r_2), $r_1 \ge 1$ satisfies

$$R(\alpha, \beta) = R(\xi, \eta) = 0, \quad D_{1,0} R(\alpha, \beta) \ne 0.$$

Then the following two results hold.

- *If $\alpha, \beta, \xi, \eta \in \mathbb{C}$ and*

$$\max\{|\alpha - \xi|, |\beta - \eta|\} \le C_1,$$

then

$$|\alpha - \xi| < \kappa |\beta - \eta|,$$

where

$$C_1 = \min\left\{1, \frac{\max\{1, |\alpha|\} \max\{1, |\beta|\}}{2^{r_1+r_2+3} \left(c_\alpha^{r_1-1} c_\beta^{r_2} r_1^2 (r_2+1)\right)^{d_1 d_2 - 1} M(\alpha)^{r_1-1} M(\beta)^{r_2} \max\{1, |\alpha|, |\beta|\} H(R)^{d_1 d_2}}\right\},$$

$$\kappa = 2^{r_2+2} (r_1+1)^{2d_1 d_2 - 1} \left((r_2+1) c_\alpha^{r_1-1} c_\beta^{r_2}\right)^{d_1 d_2 - 1} \max\{1, |\alpha|\} \max\{1, |\beta|\}^{-1} H(R)^{d_1 d_2}. \tag{5.1.4}$$

- *If* $\alpha, \beta, \xi, \eta \in \mathbb{Q}_p$ *and*

$$\max\{|\alpha - \xi|_p, |\beta - \eta|_p\} \leq C_1,$$

then

$$|\alpha - \xi|_p \leq \kappa |\beta - \eta|_p,$$

where

$$C_1 = 2^{-1} \left(c_\alpha^{r_1 - 1} c_\beta^{r_2}\right)^{-(d_1 d_2 - 1)} \left(r_1^2 (r_2 + 1) H(R)\right)^{-d_1 d_2} M(\alpha)^{-(r_1 - 1)} M(\beta)^{-r_2},$$

$$\kappa = 2 c_\alpha (c_\alpha^{r_1 - 1} c_\beta^{r_2})^{d_1 d_2 - 1} \left(r_1^2 (r_2 + 1) H(R)\right)^{d_1 d_2} M(\alpha)^{r_1 - 1} M(\beta)^{r_2}. \tag{5.1.5}$$

Proof. By Taylor's theorem,

$$\begin{aligned} 0 &= R(\xi, \eta) \\ &= \sum_{i=0}^{r_1} \sum_{j=0}^{r_2} D_{i,j} R(\alpha, \beta)(\xi - \alpha)^i (\eta - \beta)^j \\ &= (\eta - \beta) \sum_{j=1}^{r_2} D_{0,j} R(\alpha, \beta)(\eta - \beta)^{j-1} + (\xi - \alpha) \sum_{i=1}^{r_1} \sum_{j=0}^{r_2} D_{i,j} R(\alpha, \beta)(\xi - \alpha)^{i-1} (\eta - \beta)^j. \end{aligned}$$

Therefore

$$\alpha - \xi = (\eta - \beta) \frac{\sum_{j=1}^{r_2} D_{0,j} R(\alpha, \beta)(\eta - \beta)^{j-1}}{\sum_{i=1}^{r_1} \sum_{j=0}^{r_2} D_{i,j} R(\alpha, \beta)(\xi - \alpha)^{i-1} (\eta - \beta)^j}. \tag{5.1.6}$$

At this point, we distinguish two cases.

Case 1. Suppose that $\alpha, \beta, \xi, \eta \in \mathbb{C}$. Since $D_{1,0} R(\alpha, \beta) \neq 0$, we may apply Lemma 5.8 to the polynomial $D_{1,0} R$. Then it follows from the triangle inequality, (5.1.1) and (4.1.5)

that

$$\begin{aligned}
&\left|\sum_{i=1}^{r_1}\sum_{j=0}^{r_2} D_{i,j}R(\alpha,\beta)(\xi-\alpha)^{i-1}(\eta-\beta)^j\right| \\
&\geq |D_{1,0}R(\alpha,\beta)| - \max\{|\alpha-\xi|,|\beta-\eta|\}\left(\sum_{i=2}^{r_1}|D_{i,0}R(\alpha,\beta)| + \sum_{i=1}^{r_1}\sum_{j=1}^{r_2}|D_{i,j}R(\alpha,\beta)|\right) \\
&\geq |D_{1,0}R(\alpha,\beta)| - C_1\left(H(R)(r_2+1)\max\{1,|\alpha|\}^{r_1-2}\max\{1,|\beta|\}^{r_2}\sum_{i=2}^{r_1}\tbinom{r_1+1}{i+1} + \right. \\
&\left. + H(R)\max\{1,|\alpha|\}^{r_1-1}\max\{1,|\beta|\}^{r_2-1}\sum_{i=1}^{r_1}\sum_{j=1}^{r_2}\tbinom{r_1+1}{i+1}\tbinom{r_2+1}{j+1}\right) \\
&\geq \frac{\max\{1,|\alpha|\}^{r_1-1}\max\{1,|\beta|\}^{r_2}}{\left(c_\alpha^{r_1-1}c_\beta^{r_2}r_1(r_2+1)H(D_{1,0}R)\right)^{d_1d_2-1}M(\alpha)^{r_1-1}M(\beta)^{r_2}} - \\
&- C_1 2^{r_1+r_2+2}H(R)\max\{1,|\alpha|\}^{r_1-2}\max\{1,|\beta|\}^{r_2-1}\max\{1,|\alpha|,|\beta|\} \\
&\geq \frac{\max\{1,|\alpha|\}^{r_1-1}\max\{1,|\beta|\}^{r_2}}{2\left(c_\alpha^{r_1-1}c_\beta^{r_2}r_1^2(r_2+1)H(R)\right)^{d_1d_2-1}M(\alpha)^{r_1-1}M(\beta)^{r_2}},
\end{aligned} \tag{5.1.7}$$

where the last inequality follows from the definition of C_1 and the inequality

$$H(D_{i,j}R) \leq \binom{r_1}{i}\binom{r_2}{j}H(R).$$

Also, by the triangle inequality, $|\eta-\beta| \leq 1$, and (4.1.5),

$$\begin{aligned}
\left|\sum_{j=1}^{r_2} D_{0,j}R(\alpha,\beta)(\eta-\beta)^{j-1}\right| &\leq \sum_{j=1}^{r_2}|D_{0,j}R(\alpha,\beta)| \\
&\leq \sum_{j=1}^{r_2} H(R)(r_1+1)\binom{r_2+1}{j+1}\max\{1,|\alpha|\}^{r_1}\max\{1,|\beta|\}^{r_2-j} \\
&\leq H(R)(r_1+1)\max\{1,|\alpha|\}^{r_1}\max\{1,|\beta|\}^{r_2-1}\sum_{j=1}^{r_2}\binom{r_2+1}{j+1} \\
&< H(R)(r_1+1)2^{r_2+1}\max\{1,|\alpha|\}^{r_1}\max\{1,|\beta|\}^{r_2-1}.
\end{aligned} \tag{5.1.8}$$

Combining (5.1.6), (5.1.7) and (5.1.8), we obtain

$$\begin{aligned}
|\alpha-\xi| &= \left|\sum_{j=1}^{r_2} D_{0,j}R(\alpha,\beta)(\eta-\beta)^{j-1}\right| \cdot \left|\sum_{i=1}^{r_1}\sum_{j=0}^{r_2} D_{i,j}R(\alpha,\beta)(\xi-\alpha)^{i-1}(\eta-\beta)^j\right|^{-1} \cdot |\beta-\eta| \\
&< \kappa\,|\beta-\eta|\,.
\end{aligned}$$

Case 2. Suppose that $\alpha, \beta, \xi, \eta \in \mathbb{Q}_p$. Since $D_{1,0}R(\alpha,\beta) \neq 0$, we may apply Lemma 5.8 to the polynomial $D_{1,0}R$. Then it follows from non-Archimedean triangle inequality, (5.1.2) and (4.1.6) that

$$\begin{aligned}
&\left|\sum_{i=1}^{r_1}\sum_{j=0}^{r_2} D_{i,j}R(\alpha,\beta)(\xi-\alpha)^{i-1}(\eta-\beta)^j\right|_p \\
&\geq |D_{1,0}R(\alpha,\beta)|_p - \max\{|\alpha-\xi|_p, |\beta-\eta|_p\} \cdot \max_{\substack{i\geq 1, j\geq 0 \\ (i,j)\neq(1,0)}} \{|D_{i,j}R(\alpha,\beta)|_p\} \\
&\geq \frac{\left|c_\alpha^{r_1-1}c_\beta^{r_2}\right|_p^{-1}}{(c_\alpha^{r_1-1}c_\beta^{r_2})^{d_1d_2-1}(r_1(r_2+1)H(D_{1,0}R))^{d_1d_2}M(\alpha)^{r_1-1}M(\beta)^{r_2}} - C_1 \max_{\substack{i\geq 1, j\geq 0 \\ (i,j)\neq(1,0)}} \left\{|c_\alpha|_p^{-(r_1-i)}|c_\beta|_p^{-(r_2-j)}\right\} \\
&\geq \frac{\left|c_\alpha^{r_1-1}c_\beta^{r_2}\right|_p^{-1}}{2(c_\alpha^{r_1-1}c_\beta^{r_2})^{d_1d_2-1}\left(r_1^2(r_2+1)H(R)\right)^{d_1d_2}M(\alpha)^{r_1-1}M(\beta)^{r_2}}.
\end{aligned} \tag{5.1.9}$$

Also, by non-Archimedean triangle inequality, $|\eta-\beta|_p \leq 1$ and (4.1.6),

$$\begin{aligned}
\left|\sum_{j=1}^{r_2} D_{0,j}R(\alpha,\beta)(\eta-\beta)^{j-1}\right|_p &\leq \max_{1\leq j\leq r_2}\{|D_{0,j}R(\alpha,\beta)|_p\} \\
&\leq \max_{1\leq j\leq r_2}\left\{|c_\alpha|_p^{-r_1}|c_\beta|_p^{-(r_2-j)}\right\} \\
&\leq |c_\alpha|_p^{-r_1}|c_\beta|_p^{-(r_2-1)}.
\end{aligned} \tag{5.1.10}$$

Combining (5.1.6), (5.1.9) and (5.1.10), the result follows. □

5.2 Minimal and Supplementary Polynomials

In this section, we explore properties of minimal polynomials and introduce the notion of a *supplementary polynomial*.

Proposition 5.10. *Let α, β be algebraic numbers over $\mathbb{Q}$ and put $d = [\mathbb{Q}(\alpha,\beta)\colon \mathbb{Q}]$. Let R be a minimal polynomial of (α,β) and put $r = r(\alpha,\beta) = \max\{\deg_{X_1} R, \deg_{X_2} R\}$. Then R possesses the following properties.*

1. $1 \leq r \leq \lfloor\sqrt{d}\rfloor$.
2. *R is irreducible.*

Proof. Let us prove both of the above statements.

1. If $r = 0$ then R is a constant polynomial, and since $R(\alpha, \beta) = 0$ it must be the case that R is identically equal to zero, which contradicts the definition of a minimal polynomial. Therefore $r \geq 1$.

   Next, let $P \in \mathbb{Q}[X_1, X_2]$ be a polynomial such that

   $$\max_{k=1,2} \left\{ \deg_{X_k} P \right\} \leq t_1,$$

   where $t_1 = \lfloor \sqrt{d} \rfloor$. Then P has at most $(t_1 + 1)^2$ rational coefficients, which we view as variables. The equation $R(\alpha, \beta) = 0$ corresponds to d linear equations over $\mathbb{Q}$. Since $(t_1 + 1)^2 > d$, we conclude that the number of variables exceeds the number of equations, and so there exists a non-zero polynomial P such that $P(\alpha, \beta) = 0$. Thus, if R is a minimal polynomial of (α, β), then by definition

   $$r = \max_{k=1,2} \{ \deg_{X_k} R \} \leq \max_{k=1,2} \{ \deg_{X_k} P \} \leq t_1.$$

2. Let $i, j \in \{1, 2\}$ be distinct, with i such that $\deg_{X_i} R = r$. Then by Definition 5.1,

   $$\deg_{X_j} R \leq \deg_{X_i} R. \tag{5.2.1}$$

   Suppose that $R(X_1, X_2)$ is reducible. Then $R(X_1, X_2) = S(X_1, X_2)T(X_1, X_2)$ for some non-constant polynomials $S, T \in \mathbb{Z}[X_1, X_2]$. Without loss of generality, $S(\alpha, \beta) = 0$. Since the value $\max_k \{ \deg_{X_k} R \}$ is minimal among all non-zero polynomials that vanish at (α, β), it must be the case that

   $$\max_{k=1,2} \{ \deg_{X_k} S \} = r.$$

   Therefore S satisfies Properties (1), (2), (3) in Definition 5.1.

   If $\deg_{X_i} S = \deg_{X_i} R$, by Property (4) in Definition 5.1,

   $$\deg_{X_j} R \leq \deg_{X_j} S.$$

   Since $S \mid R$, we conclude that $\deg_{X_j} S = \deg_{X_j} R$. But then T would have to be a constant polynomial, so we reach a contradiction.

   If $\deg_{X_j} S = \deg_{X_i} R$, by Property (4) in Definition 5.1,

   $$\deg_{X_j} R \leq \deg_{X_i} S.$$

It follows from $S \mid R$ and (5.2.1) that

$$\deg_{X_i} R = \deg_{X_j} S \leq \deg_{X_j} R \leq \deg_{X_i} R,$$

so $\deg_{X_i} R = \deg_{X_j} R$ and $\deg_{X_j} S = \deg_{X_j} R$. Since $S \mid R$,

$$\deg_{X_j} R \leq \deg_{X_i} S \leq \deg_{X_i} R = \deg_{X_j} R,$$

so $\deg_{X_i} S = \deg_{X_j} R = \deg_{X_i} R$. We conclude that T is a constant polynomial, so once again we reach a contradiction. Therefore R is irreducible.

□

Proposition 5.11. *Let α, β be algebraic numbers over $\mathbb{Q}$ and put $d = [\mathbb{Q}(\alpha, \beta) \colon \mathbb{Q}]$. Let R_1 be a minimal polynomial of (α, β) of bi-degree (r_1, r_2). There exists an irreducible polynomial $R_2 \in \mathbb{Z}[X_1, X_2]$ such that $R_2(\alpha, \beta) = 0$, $R_1 \nmid R_2$ and*

$$\max_{k=1,2}\{\deg_{X_k} R_2\} \leq \left\lfloor \frac{d + r_1 r_2}{r_1 + r_2} \right\rfloor.$$

Proof. Let

$$t_1 = \lfloor \sqrt{d} \rfloor, \quad t_2 = \left\lfloor \frac{d + r_1 r_2}{r_1 + r_2} \right\rfloor.$$

We claim that

$$t_1 \leq t_2. \tag{5.2.2}$$

Indeed, the inequality

$$0 \leq \left(\sqrt{d} - r_1\right) \cdot \left(\sqrt{d} - r_2\right)$$

always holds, because by construction of R_1 we have $r_1 \leq \sqrt{d}$ and $r_2 \leq \sqrt{d}$. It is easy to see that the above inequality is equivalent to

$$\sqrt{d} \leq \frac{d + r_1 r_2}{r_1 + r_2}.$$

Thus t_1 does not exceed $(d + r_1 r_2)/(r_1 + r_2)$. Now, suppose for a contradiction that $t_1 > t_2$. Then we end up in the situation

$$\left\lfloor \frac{d + r_1 r_2}{r_1 + r_2} \right\rfloor < \lfloor \sqrt{d} \rfloor \leq \frac{d + r_1 r_2}{r_1 + r_2},$$

which is nonsensical, because for every $x \in \mathbb{R}$ a half-open interval $(\lfloor x \rfloor, x]$ does not contain any integers. Therefore $t_1 \leq t_2$, as claimed.

For an arbitrary positive real number t define the vector spaces

$$\mathcal{P}(\alpha, \beta; t) = \left\{ P \in \mathbb{Q}[X_1, X_2] \colon \substack{P(\alpha,\beta)=0, \\ \max\{\deg_{X_k} P\} \leq t} \right\}, \tag{5.2.3}$$

$$\mathcal{P}(\alpha, \beta; R_1, t) = \left\{ P \in \mathbb{Q}[X_1, X_2] \colon \substack{R_1 \,|\, P, \\ \max\{\deg_{X_k} P\} \leq t} \right\}.$$

Then it follows from Proposition 5.10 and (5.2.2) that $R_1 \in \mathcal{P}(\alpha, \beta; t_2)$. Therefore

$$\mathcal{P}(\alpha, \beta; R_1, t_2) \subseteq \mathcal{P}(\alpha, \beta; t_2).$$

Let us analyze the dimensions of these vector spaces. Certainly, we have

$$\dim \mathcal{P}(\alpha, \beta; R_1, t_2) = (t_2 + 1 - r_1)(t_2 + 1 - r_2). \tag{5.2.4}$$

On the other hand, if $P \in \mathcal{P}(\alpha, \beta; t_2)$, the equation $P(\alpha, \beta) = 0$ corresponds to d linear equations over $\mathbb{Q}$, while the number of coefficients of $P(X, Y)$ (viewed as variables) is equal to $(t_2 + 1)^2$. Therefore

$$\dim \mathcal{P}(\alpha, \beta; t_2) \geq (t_2 + 1)^2 - d. \tag{5.2.5}$$

Since $\lfloor x \rfloor > x - 1$ for any x, we have

$$t_2 > \frac{d + r_1 r_2}{r_1 + r_2} - 1,$$

which is equivalent to

$$(t_2 + 1 - r_1)(t_2 + 1 - r_2) < (t_2 + 1)^2 - d.$$

Combining this inequality with (5.2.4), (5.2.5), we obtain

$$\begin{aligned} \dim \mathcal{P}(\alpha, \beta; R_1, t_2) &= (t_2 + 1 - r_1)(t_2 + 1 - r_2) \\ &< (t_2 + 1)^2 - d \\ &\leq \dim \mathcal{P}(\alpha, \beta; t_2). \end{aligned}$$

Therefore there exists $\hat{R}_2 \in \mathcal{P}(\alpha, \beta; t_2) \setminus \mathcal{P}(\alpha, \beta; R_1, t_2)$.

If $\hat{R}_2$ is irreducible, then we set $R_2 = \hat{R}_2$, and the result follows. Otherwise there exists an irreducible polynomial $R_2 \in \mathbb{Z}[X_1, X_2]$ such that $R_2(\alpha, \beta) = 0$ and $R_2 \mid \hat{R}_2$. Since $\max_k\{\deg_{X_k} R_k\} \leq \max_k\{\deg_{X_k} \hat{R}_k\}$, the result follows. □

Definition 5.12. A polynomial R_2 from Proposition 5.11 is called a *supplementary polynomial* for R_1 and (α, β).

We conclude this section by briefly addressing the question of computation of a minimal polynomial. Let (α, β) be a fixed pair of algebraic numbers and consider the vector space $\mathcal{P}(\alpha, \beta; t_1)$, where $t_1 = \lfloor \sqrt{d} \rfloor$ and $\mathcal{P}(\alpha, \beta; t)$ is defined in (5.2.3). This is a finite-dimensional vector space over $\mathbb{Q}$. Furthermore, in view of Proposition 5.10 part 1, it contains a minimal polynomial R of (α, β). By analogy with Algorithm 1 outlined in Section 2.3, it is possible to determine a minimal polynomial of (α, β) by using row reduction and LLL reduction. We leave the problem of writing an explicit algorithm for its computation for future work.

5.3 A Generalized Gap Principle

Lemma 5.13. (A generalized Archimedean gap principle) *Let $\alpha, \beta \in \mathbb{C}$ be distinct algebraic numbers over $\mathbb{Q}$ such that $d_2 < d_1/4$, where*

$$d_1 = [\mathbb{Q}(\alpha) \colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha, \beta) \colon \mathbb{Q}(\alpha)].$$

Let c_α, c_β denote the leading coefficients of the minimal polynomials of α and β, respectively. Let R_1 be a minimal polynomial for (α, β) of bi-degree (r_1, r_2), and R_2 a supplementary polynomial for R_1 and (α, β). Define

$$t_1 = \left\lfloor \sqrt{d_1 d_2} \right\rfloor, \quad t_2 = \left\lfloor \frac{d_1 d_2 + r_1 r_2}{r_1 + r_2} \right\rfloor. \tag{5.3.1}$$

Let $2t_1 < \mu \leq d_1$, $C_0 > 0$,

$$C_1 = \max \left\{ C_0^{1/\mu}, ((t_2+1)^4 H(R_1) H(R_2))^{t_2}, \left(\frac{C_0 2^{r_1+r_2+3} \left(c_\alpha^{r_1-1} c_\beta^{r_2} r_1^2 (r_2+1)\right)^{d_1 d_2 - 1} M(\alpha)^{r_1-1} M(\beta)^{r_2} \max\{1,|\alpha|,|\beta|\} H(R_1)^{d_1 d_2}}{\max\{1,|\alpha|\} \max\{1,|\beta|\}} \right)^{1/\mu} \right\}, \tag{5.3.2}$$

$$C_{1+\ell} = 4 \max\{1, |\alpha|\} \max\{1, |\beta|\} \cdot (4 C_0 H(R_\ell))^{1/t_\ell}, \quad \ell = 1, 2. \tag{5.3.3}$$

If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms, $H(x_2, y_2) \geq H(x_1, y_1) \geq C_1$ and

$$\left| \alpha - \frac{x_1}{y_1} \right| < \frac{C_0}{H(x_1, y_1)^\mu}, \quad \left| \beta - \frac{x_2}{y_2} \right| < \frac{C_0}{H(x_2, y_2)^\mu},$$

then one of the following holds.

1. $H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu/t_1 - 1}$.

2. $t_2 \geq \mu^2/(d_1 + \mu)$ *(i.e., R_1 is μ-special) and* $R_1(x_1/y_1, x_2/y_2) = 0$.

3. $t_2 < \mu^2/(d_1 + \mu)$ *and* $H(x_1, y_1) < C_4$, *where*

$$C_4 = \left(2^{d_1+1} H(\alpha) \max\{1, |\alpha|\}^{d_1 - 1} \kappa C_0 C_3^{\mu}\right)^{t_2/(\mu^2 - \mu t_2 - d_1 t_2)} \tag{5.3.4}$$

and $\kappa = \kappa(\alpha, \beta, R_1)$ *is defined in (5.1.4).*

Proof. By Propositions 5.10 and 5.11,

$$\max_{k=1,2}\{\deg_{X_k} R_1\} \leq t_1, \ \ \max_{k=1,2}\{\deg_{X_k} R_2\} \leq t_2.$$

We claim that the inequality $H(x_1, y_1) \geq C_1$ implies the existence of $\ell \in \{1, 2\}$ such that $R_\ell(x_1/y_1, x_2/y_2) \neq 0$. Indeed, suppose that $R_1(x_1/y_1, x_2/y_2) = R_2(x_1/y_1, x_2/y_2) = 0$. By Proposition 5.10, R_1 is irreducible. Further, by Proposition 5.11, R_1 does not divide R_2. Then it follows from Bézout's theorem that the algebraic curves $R_1(X, Y) = 0$ and $R_2(X, Y) = 0$ have only finitely many points in common. Let $g(X_1) = \mathrm{Res}_{X_2}(R_1, R_2)$, where R_1, R_2 are viewed as polynomials in one variable X_2 with coefficients in $\mathbb{Z}[X_1]$. Then g is non-zero, and it must be the case that $g(x_1/y_1) = 0$. Since x_1, y_1 are coprime, it follows from the rational roots theorem that x_1, y_1 divide the constant and the leading coefficients of g, respectively. Thus $H(x_1, y_1) \leq H(g)$. By Lemma 5.7 and (5.2.2),

$$\begin{aligned} H(x_1, y_1) &\leq H(g) \\ &\leq (t_2 + 1)^{2t_1 - 1}(t_1 + 1)^{2t_2 - 1}(t_1 t_2 + 1) H(R_1)^{t_2} H(R_2)^{t_1} \\ &< \left((t_2 + 1)^4 H(R_1) H(R_2)\right)^{t_2} \\ &\leq C_1, \end{aligned}$$

which leads us to a contradiction. Hence there exists $\ell \in \{1, 2\}$, which we choose to be the smallest, such that $R_\ell(x_1/y_1, x_2/y_2) \neq 0$.

Let $\ell \in \{1, 2\}$ be as defined above. Since $H(x_1, y_1) \geq C_1$, the inequalities $|\alpha - x_1/y_1| < 1$,

$|\beta - x_2/y_2| < 1$ hold. By Taylor's theorem, triangle inequality and (4.1.5),

$$\begin{aligned}
\frac{1}{H(x_1,y_1)^{t_\ell} H(x_2,y_2)^{t_\ell}} &\le \left| R_\ell\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right| \\
&\le \sum_{i=0}^{t_\ell}\sum_{j=0}^{t_\ell} |D_{i,j}R_\ell(\alpha,\beta)| \left|\alpha - \frac{x_1}{y_1}\right|^i \left|\beta - \frac{x_2}{y_2}\right|^j \\
&< \frac{C_0}{H(x_1,y_1)^\mu} \sum_{i=0}^{t_\ell}\sum_{j=0}^{t_\ell} H(R_\ell)\binom{t_\ell+1}{i+1}\binom{t_\ell+1}{j+1} \max\{1,|\alpha|\}^{t_\ell-i}\max\{1,|\beta|\}^{t_\ell-j} \\
&< \frac{C_0}{H(x_1,y_1)^\mu} 4^{t_\ell+1} H(R_\ell)\max\{1,|\alpha|\}^{t_\ell}\max\{1,|\beta|\}^{t_\ell} \\
&= \frac{C_{1+\ell}^{t_\ell}}{H(x_1,y_1)^\mu}.
\end{aligned}$$

In particular, if $\ell = 1$, we see that $H(x_2,y_2) > C_2^{-1} H(x_1,y_1)^{\mu/t_1-1}$, so case 1 holds.

If, however, $\ell = 2$, then we end up in the situation

$$R_1\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) = 0, \quad R_2\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) \neq 0.$$

If $t_2 \ge \mu^2/(d_1+\mu)$ then case 2 holds.

If $t_2 < \mu^2/(d_1+\mu)$, then $\mu \le d_1$ implies $t_2 \le \mu/2$, which means that the gap principle $H(x_2,y_2) > C_3^{-1} H(x_1,y_1)^{\mu/t_2-1}$ holds. Since $H(x_1,y_1) \ge C_1$, Lemma 5.9 applies, so we obtain the upper bound

$$\left|\alpha - \frac{x_1}{y_1}\right| \le \kappa \left|\beta - \frac{x_2}{y_2}\right| < \frac{\kappa C_0}{H(x_2,y_2)^\mu} \le \frac{\kappa C_0 C_3^\mu}{H(x_1,y_1)^{\mu(\mu/t_2-1)}}.$$

On the other hand, by Lemma 2.11, the lower bound (2.2.1) holds. Combining upper and lower bounds, we obtain

$$\frac{1}{2^{d_1+1}H(\alpha)\max\{1,|\alpha|\}^{d_1-1}H(x_1,y_1)^{d_1}} < \left|\alpha - \frac{x_1}{y_1}\right| < \frac{\kappa C_0 C_3^\mu}{H(x_1,y_1)^{\mu(\mu/t_2-1)}}.$$

Therefore

$$H(x_1,y_1) < \left(2^{d_1+1}H(\alpha)\max\{1,|\alpha|\}^{d_1-1}\kappa C_0 C_3^\mu\right)^{1/(\mu(\mu/t_2-1)-d_1)},$$

so case 3 holds. □

Lemma 5.14. (A generalized non-Archimedean gap principle) *Let $\alpha, \beta \in \mathbb{Q}_p$ be distinct algebraic numbers over $\mathbb{Q}$ such that $d_2 < d_1/4$, where*

$$d_1 = [\mathbb{Q}(\alpha)\colon \mathbb{Q}], \quad d_2 = [\mathbb{Q}(\alpha,\beta)\colon \mathbb{Q}(\alpha)].$$

Let c_α, c_β denote the leading coefficients of the minimal polynomials of α and β, respectively. Let R_1 be a minimal polynomial for (α,β) of bi-degree (r_1, r_2), and R_2 a supplementary polynomial for (α,β). Define

$$t_1 = \left\lfloor \sqrt{d_1 d_2} \right\rfloor, \quad t_2 = \left\lfloor \frac{d_1 d_2 + r_1 r_2}{r_1 + r_2} \right\rfloor. \tag{5.3.5}$$

Let $2t_1 < \mu \le d_1$, $C_0 > 0$,

$$\begin{aligned} C_1 = \ & \max\Big\{ C_0^{1/\mu}, \left((t_2+1)^4 H(R_1)H(R_2)\right)^{t_2}, \\ & \left(2\left(c_\alpha^{r_1-1} c_\beta^{r_2}\right)^{d_1 d_2 - 1} \left(r_1^2 (r_2+1) H(R)\right)^{d_1 d_2} M(\alpha)^{r_1-1} M(\beta)^{r_2}\right)^{1/\mu} \Big\}, \end{aligned} \tag{5.3.6}$$

$$C_{1+\ell} = c_\alpha c_\beta \left(C_0 (t_\ell+1)^2 H(R_\ell)\right)^{1/t_\ell}, \quad \ell = 1, 2. \tag{5.3.7}$$

If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms, $H(x_2,y_2) \ge H(x_1,y_1) \ge C_1$ and

$$|y_1\alpha - x_1|_p < \frac{C_0}{H(x_1,y_1)^\mu}, \quad |y_1\alpha - x_1|_p < \frac{C_0}{H(x_2,y_2)^\mu},$$

then one of the following holds.

1. $H(x_2,y_2) > C_2^{-1} H(x_1,y_1)^{\mu/t_1 - 1}$.
2. $t_2 \ge \mu^2/(d_1+\mu)$ *(i.e., R_1 is μ-special) and* $R_1(x_1/y_1, x_2/y_2) = 0$.
3. $t_2 < \mu^2/(d_1+\mu)$ *and* $H(x_1,y_1) < C_4$, *where*

$$C_4 = \left(c_\alpha^{d_1-1}(d+1)H(\alpha)\kappa C_0 C_3^\mu\right)^{t_2/(\mu^2 - \mu t_2 - d_1 t_2)} \tag{5.3.8}$$

and $\kappa = \kappa(\alpha,\beta,R_1)$ is defined in (5.1.5).

Proof. As in Lemma 5.13, the inequality $H(x_1,y_1) \ge C_1$ implies the existence of $\ell \in \{1,2\}$, which we choose to be the smallest, such that $R_\ell(x_1/y_1, x_2/y_2) \ne 0$. Consequently, the following trivial lower bound holds:

$$\begin{aligned} \left|(y_1y_2)^{t_\ell} R_\ell\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right|_p &\ge \left|(y_1y_2)^{t_\ell} R_\ell\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right)\right|^{-1} \\ &\ge \left((t_\ell+1)^2 H(R_\ell) H(x_1,y_1)^{t_\ell} H(x_2,y_2)^{t_\ell}\right)^{-1}. \end{aligned}$$

Note that for each $(i,j) \in \{0,\dots,t_\ell\}^2$ the p-adic number $c_\alpha^{t_\ell - i} c_\beta^{t_\ell - j} D_{i,j} R_\ell(\alpha,\beta)$ is an algebraic integer. Thus its p-adic absolute value does not exceed one. Via the application of Taylor's theorem we obtain the following upper bound:

$$\begin{aligned}
\left|(y_1y_2)^{t_\ell} R_\ell\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right)\right|_p &\le \max_{(i,j)\neq(0,0)} \left\{ |D_{i,j}R(\alpha,\beta)|_p \cdot |y_1\alpha - x_1|_p^i \cdot |y_2\beta - x_2|_p^j \right\} \\
&= \max_{(i,j)\neq(0,0)} \left\{ \left| \frac{c_\alpha^{t_\ell - i} c_\beta^{t_\ell - j} D_{i,j} R(\alpha,\beta)}{c_\alpha^{t_\ell - i} c_\beta^{t_\ell - j}} \right|_p \cdot |y_1\alpha - x_1|_p^i \cdot |y_2\beta - x_2|_p^j \right\} \\
&\le \left| c_\alpha^{t_\ell} c_\beta^{t_\ell} \right|_p^{-1} \max_{(i,j)\neq(0,0)} \left\{ |y_1 c_\alpha \alpha - c_\alpha x_1|_p^i \cdot |y_2 c_\beta \beta - c_\beta x_2|_p^j \right\} \\
&\le (c_\alpha c_\beta)^{t_\ell} \max\{|y_1 c_\alpha \alpha - c_\alpha x_1|_p, |y_2 c_\beta \beta - c_\beta x_2|_p\} \\
&\le (c_\alpha c_\beta)^{t_\ell} \max\{|c_\alpha|_p, |c_\beta|_p\} \max\{|y_1\alpha - x_1|_p, |y_2\beta - x_2|_p\} \\
&< \frac{C_0 (c_\alpha c_\beta)^{t_\ell}}{H(x_1,y_1)^\mu}.
\end{aligned}$$

Combining upper and lower bounds,

$$\frac{1}{(t_\ell+1)^2 H(R_\ell) H(x_1,y_1)^{t_\ell} H(x_2,y_2)^{t_\ell}} < \frac{C_0 (c_\alpha c_\beta)^{t_\ell}}{H(x_1,y_1)^\mu}.$$

In particular, if $\ell = 1$, we see that $H(x_2,y_2) > C_2^{-1} H(x_1,y_1)^{\mu/t_1 - 1}$, so case 1 holds.

If, however, $\ell = 2$, then we end up in the situation

$$R_1\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right) = 0, \quad R_2\left(\frac{x_1}{y_1},\frac{x_2}{y_2}\right) \neq 0.$$

If $t_2 \ge \mu^2/(d_1+\mu)$ then case 2 holds.

If $t_2 < \mu^2/(d_1+\mu)$, then $\mu \le d_1$ implies $t_2 \le \mu/2$, which means that the gap principle $H(x_2,y_2) > C_3^{-1} H(x_1,y_1)^{\mu/t_2 - 1}$ holds. Since $H(x_1,y_1) \ge C_1$, Lemma 5.9 applies, so we obtain the upper bound

$$\left|\alpha - \frac{x_1}{y_1}\right| \le \kappa \left|\beta - \frac{x_2}{y_2}\right| < \frac{\kappa C_0}{H(x_2,y_2)^\mu} \le \frac{\kappa C_0 C_3^\mu}{H(x_1,y_1)^{\mu(\mu/t_2 - 1)}}.$$

On the other hand, by Lemma 2.12, the lower bound (2.2.4) holds. Combining upper and lower bounds, we obtain

$$\frac{1}{c_\alpha^{d_1 - 1}(d+1) H(\alpha) H(x_1,y_1)^{d_1}} \le \left|\alpha - \frac{x_1}{y_1}\right| < \frac{\kappa C_0 C_3^\mu}{H(x_1,y_1)^{\mu(\mu/t_2 - 1)}}.$$

Therefore

$$H(x_1, y_1) < \left(c_\alpha^{d_1-1}(d+1)H(\alpha)\kappa C_0 C_3^\mu\right)^{1/(\mu(\mu/t_2-1)-d_1)},$$

so case 3 holds. □

5.4 Counting Approximations of Large Height

Theorem 5.15. *Let $K = \mathbb{C}$ or $\mathbb{Q}_p$, where p is a rational prime, and denote the standard absolute value on K by $|\ |$. Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree $d_1 \geq 5$, and $\alpha_1, \ldots, \alpha_{d_1}$ be the roots of f. Define*

$$d_2 = \max_{1 \leq i \leq d} \{[\mathbb{Q}(\alpha_1, \alpha_i) \colon \mathbb{Q}(\alpha_1)]\}$$

and suppose that $d_2 < d_1/4$. Let μ be such that $2.31 \leq \mu \leq 3$ if $d_1 = 3$ and

$$2\sqrt{d_1 d_2} < \mu \leq d_1$$

otherwise. Let C_0 be a positive real number. There exists an explicitly computable positive number C_1, which depends on C_0, μ, f, but not on p in the case $K = \mathbb{Q}_p$, with the following property.

The total number of rationals x/y in lowest terms, which satisfy $H(x, y) \geq C_1$ and

$$\left|\alpha_j - \frac{x}{y}\right| < \frac{C_0}{H(x,y)^\mu} \tag{5.4.1}$$

for some $j \in \{1, 2, \ldots, d\}$ is less than

$$\gamma \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu+3) - 2\log\left(\mu/\sqrt{2d_1d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1d_2} - 1\right)} \right\rfloor,$$

where

$$\gamma = \#\{i \colon \text{a minimal polynomial of } (\alpha_1, \alpha_i) \text{ is } \mu\text{-special}\}.$$

Proof. Throughout the proof we will be adjusting our choice of C_1 several times. We begin by choosing C_1 so that

$$C_1 \geq \left(C_0(d+1)^{3d/2}H(f)^{d-1}\right)^{1/\mu}.$$

Then it follows from Corollary 2.28 that for each x/y satisfying (5.4.1) the index $j \in \{1, 2, \ldots, d\}$ is unique.

Let R_i denote a minimal polynomial for (α_1, α_i) and $R_{i,j} = R_{\sigma(j)}$, where σ is some element of $\mathrm{Gal}(\mathbb{Q}(\alpha_1, \dots, \alpha_d)/\mathbb{Q})$ such that $\sigma\colon \alpha_i \mapsto \alpha_1$. Then by construction $R_{i,j}(\alpha_i, \alpha_j) = 0$. Further, let $t_2(i), C_1(i), C_4(i)$ denote the constants (5.3.1), (5.3.2), (5.3.4) if $K = \mathbb{C}$ or (5.3.5), (5.3.6), (5.3.8) if $K = \mathbb{Q}_p$, with $(\alpha, \beta) = (\alpha_1, \alpha_i)$. We adjust C_1 by choosing it so that

$$C_1 \geq \max\left\{ \max_{1\leq i\leq d} \{C_1(i)\}, \max_{\substack{1\leq i\leq d \\ t_2(i)<\mu^2/(d_1+\mu)}} \{C_4(i)\} \right\}.$$

With such a choice of C_1, case 3 in Lemmas 5.13 or 5.14 cannot hold.

Let $\mathcal{X}$ denote the collection of all rational numbers x/y in lowest terms that satisfy $H(x,y) \geq C_1$ and (5.4.1). Suppose that $\mathcal{X}$ contains at least two elements (otherwise the result holds trivially). Let $\mathcal{X}' = \{x_1/y_1, \dots, x_\ell/y_\ell\}$ be a subset of $\mathcal{X}$, constructed according to the following procedure:

1. Let $x_1/y_1 \in \mathcal{X}$ be such that $H(x_1, y_1)$ is the smallest among elements of $\mathcal{X}$, and let i_1 denote the index such that $|\alpha_{i_1} - x_1/y_1| < C_0 H(x_1, y_1)^{-\mu}$.
2. For $j \geq 2$, let $x_j/y_j \in \mathcal{X}$ be such that the following three properties are satisfied:
   (a) $H(x_j, y_j) > H(x_{j-1}, y_{j-1})$;
   (b) $R_{i_{j-1}, i_j}(x_{j-1}/y_{j-1}, x_j/y_j) \neq 0$, where i_j is such that $|\alpha_{i_j} - x_j/y_j| < C_0 H(x_j, y_j)^{-\mu}$;
   (c) $H(x_j, y_j)$ is the smallest among all $x/y \in \mathcal{X}$ satisfying (a) and (b).

Our first goal is to bound $\mathcal{X}$ in terms of $\ell = |\mathcal{X}'|$. To do so, define

$$\mathcal{X}_j = \{x/y \in \mathcal{X} \setminus \mathcal{X}' \colon H(x_j, y_j) \leq H(x,y) < H(x_{j+1}, y_{j+1})\}, \quad j = 1, \dots, \ell - 1,$$

$$\mathcal{X}_\ell = \{x/y \in \mathcal{X} \setminus \mathcal{X}' \colon H(x_\ell, y_\ell) \leq H(x,y)\}.$$

Then by construction $\mathcal{X} = \mathcal{X}' \cup \mathcal{X}_1 \cup \cdots \cup \mathcal{X}_\ell$. Furthermore, for any $x/y \in \mathcal{X}_j$ with $|\alpha_k - x/y| < C_0 H(x,y)^{-\mu}$ we have

$$R_{i_j,k}\left(\frac{x_j}{y_j}, \frac{x}{y}\right) = 0.$$

Due to our choice of C_1, the above equality and the inequalities $H(x,y) \geq H(x_j, y_j) \geq C_1$ can be satisfied simultaneously only if $R_{i_j,k}$ is a μ-special polynomial (this situation

corresponds to case 2 in Lemmas 5.13, 5.14). Since there always exists an automorphism σ in the Galois group of f such that $\sigma\colon i_j \mapsto 1$, we have

$$R_{i_j,k}\left(\alpha_{i_j},\alpha_k\right) = R_{1,\sigma(k)}\left(\alpha_1,\alpha_{\sigma(k)}\right) = 0,$$

meaning that a minimal polynomial $R_{1,\sigma(k)}$ of $(\alpha_1,\alpha_{\sigma(k)})$ is a μ-special polynomial. By definition, the total number of such polynomials is γ, leading us to a conclusion that there are at most γ possible values for k. Since the minimal polynomial for $(\alpha_{i_j},\alpha_{i_j})$ is $R(X_1,X_2) = X_1 - X_2$, which is always μ-special, one of the possible values is $k = i_j$. However, since

$$R_{i_j,i_j}\left(\frac{x_j}{y_j},\frac{x}{y}\right) = \frac{x_j}{y_j} - \frac{x}{y} = 0,$$

it must be the case that $x/y = x_j/y_j$. This is impossible due to the fact that $x_j/y_j \notin \mathcal{X}_j$, and so we conclude that there are at most $\gamma - 1$ possible values for k.

Now, we claim that the proper choice of C_1 implies $|\mathcal{X}_j| \le \gamma - 1$ for all j, where γ is the total number of μ-special polynomials. Indeed, let us choose C_1 so that

$$C_1 \ge \left(C_0^{\mu+1}2^{\mu+2}(d+1)H(f)\max\{2\overline{|f|},c_d\}^{d-1}\max_{i,j}\{\kappa(i,j)\}\right)^{1/(\mu^2-\mu-d)},$$

where $\kappa(i,j) = \kappa(\alpha_i,\alpha_j,R_{i,j})$ is defined in (5.1.4) if $K = \mathbb{C}$ or (5.1.5) if $K = \mathbb{Q}_p$. Note that due to our choice of μ we have $\mu^2 - \mu - d_1 > 0$. Let $x/y \in \mathcal{X}_j$ and k be such that $|\alpha_k - x/y| < C_0H(x,y)^{-\mu}$. If $|\mathcal{X}_j| \ge \gamma$ then it follows from the Dirichlet's box principle that there exist $x/y, x'/y' \in \mathcal{X}_j$, with $H(x',y') \ge H(x,y)$, which correspond to the same value of k. In other words,

$$R_{i_j,k}\left(\frac{x_j}{y_j},\frac{x}{y}\right) = R_{i_j,k}\left(\frac{x_j}{y_j},\frac{x'}{y'}\right) = 0.$$

Since

$$\left|\alpha_k - \frac{x}{y}\right| < \frac{C_0}{H(x,y)^\mu} \quad\text{and}\quad \left|\alpha_k - \frac{x'}{y'}\right| < \frac{C_0}{H(x',y')^\mu},$$

the standard gap principle applies:

$$\frac{1}{H(x,y)H(x',y')} \le \left|\frac{x}{y} - \frac{x'}{y'}\right| \le \left|\alpha_k - \frac{x}{y}\right| + \left|\alpha_k - \frac{x'}{y'}\right| < \frac{2C_0}{H(x,y)^\mu}.$$

We conclude that

$$H(x',y') > (2C_0)^{-1}H(x,y)^{\mu-1} \ge (2C_0)^{-1}H(x_j,y_j)^{\mu-1}.$$

On the other hand, since $H(x_j, y_j) \geq C_1$, Lemma 5.9 applies to $(\alpha_{i_j}, \alpha_k, R_{i_j,k})$, so we obtain the upper bound

$$\left|\alpha_{i_j} - \frac{x_j}{y_j}\right| \leq \kappa(i_j, k) \left|\beta - \frac{x'}{y'}\right| < \frac{\kappa(i_j, k) C_0}{H(x', y')^{\mu}} < \frac{\kappa(i_j, k) C_0 (2C_0)^{\mu}}{H(x_j, y_j)^{\mu(\mu-1)}}.$$

Applying Lemma 2.11 if $K = \mathbb{C}$ or Lemma 2.12 if $K = \mathbb{Q}_p$, we are also able to produce a lower bound on $|\alpha_{i_j} - x_j/y_j|$:

$$\frac{1}{4(d+1)H(f)\max\{2, 2|\alpha_{i_j}|, c_d\}^{d-1} H(x_j, y_j)^d} \leq \left|\alpha_{i_j} - \frac{x_j}{y_j}\right| < \frac{\kappa(i_j, k) C_0 (2C_0)^{\mu}}{H(x_j, y_j)^{\mu(\mu-1)}}.$$

Therefore

$$H(x_j, y_j) < \left(C_0^{\mu+1} 2^{\mu+2} (d+1) H(f) \max\{2, 2|\alpha_{i_j}|, c_d\}^{d-1} \kappa(i_j, k)\right)^{1/(\mu^2-\mu-d)} \leq C_1,$$

so we reach a contradiction. We conclude that distinct rationals $x/y, x'/y' \in \mathcal{X}_j$ cannot correspond to the same value of k. From this fact we deduce that $|\mathcal{X}_j| \leq \gamma - 1$ for all j, and so

$$|\mathcal{X}| = |\mathcal{X}'| + \sum_{j=1}^{\ell} |\mathcal{X}_j| \leq \gamma \ell.$$

It remains to obtain an upper bound on ℓ. We begin by applying a generalized gap principle to the ordered pair $(\alpha_{i_k}, \alpha_{i_{k+1}})$. Let $C_2(i)$ denote the constant (5.3.3) if $K = \mathbb{C}$ or (5.3.7) if $K = \mathbb{Q}_p$, with $(\alpha, \beta) = (\alpha_1, \alpha_i)$, and define

$$C_2 = \max_{1 \leq i \leq d} \{C_2(i)\},$$

$$E = \mu / \sqrt{d_1 d_2} - 1.$$

Since $\mu > 2\sqrt{d_1 d_2}$, we see that $E > 1$. Note that if $K = \mathbb{Q}_p$ then for all $j \in \{1, \ldots, \ell\}$ we have $|y_j| \leq 1$, and so

$$|y_j \alpha_{i_j} - x_j| = |y_j| \cdot \left|\alpha_{i_j} - \frac{x_j}{y_j}\right| < \frac{C_0}{H(x_j, y_j)^{\mu}}.$$

With the choice of C_1 and $\mathcal{X}'$ as above, neither case 2 nor case 3 hold in Lemmas 2.24, 2.25. Therefore only case 1 remains, i.e., it is possible to apply our generalized gap principle to the ordered pair $(\alpha, \beta) = (\alpha_{i_k}, \alpha_{i_{k+1}})$:

$$\log H(x_{k+1}, y_{k+1}) > E \log H(x_k, y_k) - \log C_2(i_k, i_{k+1}) \geq E \log H(x_k, y_k) - \log C_2$$

for any $k \in \{1, 2, \ldots, \ell - 1\}$, where E is defined in (2.6.6). Consequently,

$$\begin{aligned}\log H(x_\ell, y_\ell) &> E \log H(x_{\ell-1}, y_{\ell-1}) - \log C_2 \\ &> E^2 \log H(x_{\ell-2}, y_{\ell-2}) - (1 + E) \log C_2 \\ &> \ldots \\ &> E^{\ell-1} \log H(x_1, y_1) - (1 + E + \cdots + E^{\ell-2}) \log C_2.\end{aligned}$$

Thus we obtain the following lower bound on $\log H(x_\ell, y_\ell)$:

$$\log H(x_\ell, y_\ell) > E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C_2. \tag{5.4.2}$$

Next, we apply the Thue-Siegel principle from Theorem 4.1 to the pair $(\alpha, \beta) = (\alpha_{i_1}, \alpha_{i_\ell})$. We adjust our definition of C_1 by choosing it so that

$$C_1 \geq C,$$

where $C = C(\alpha, \beta, \mu)$ is the constant from Theorem 4.1 (for the exact definition, see (4.2.2)). Since $H(x_\ell, y_\ell) \geq H(x_1, y_1) \geq C_1$, it follows from Theorem 4.1 that

$$\log H(x_\ell, y_\ell) < \frac{86(\mu + 3)^4}{(\mu/\sqrt{2d_1d_2} - 1)^2} \log H(x_1, y_1).$$

We combine the above upper bound on $\log H(x_\ell, y_\ell)$ with the lower bound given in (5.4.2):

$$E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C_2 < \frac{86(\mu + 3)^4}{(\mu/\sqrt{2d_1d_2} - 1)^2} \log H(x_1, y_1).$$

Reordering the terms yields

$$\left(E^{\ell-1} - \frac{86(\mu + 3)^4}{(\mu/\sqrt{2d_1d_2} - 1)^2} \right) \log H(x_1, y_1) < \frac{E^{\ell-1} - 1}{E - 1} \log C_2. \tag{5.4.3}$$

Let us assume that

$$\ell \geq 1 + \log \left(\frac{86(\mu + 3)^4}{(\mu/\sqrt{2d_1d_2} - 1)^2} \right) / \log E,$$

for otherwise the statement of our theorem holds. Then

$$E^{\ell-1} \geq \frac{86(\mu + 3)^4}{(\mu/\sqrt{2d_1d_2} - 1)^2},$$

so we may use the inequality $H(x_1, y_1) \geq C_1$ to replace $H(x_1, y_1)$ with C_1 in (5.4.3):

$$\left(E^{\ell-1} - \frac{86(\mu+3)^4}{(\mu/\sqrt{2d_1d_2}-1)^2}\right)\log C_1 < \frac{E^{\ell-1}-1}{E-1}\log C_2.$$

We make a final adjustment to C_1 by choosing it so that

$$C_1 \geq C_2^{2/(E-1)}.$$

Then $\log C_2 \leq ((E-1)/2)\log C_1$, so

$$E^{\ell-1} - \frac{86(\mu+3)^4}{(\mu/\sqrt{2d_1d_2}-1)^2} < \frac{E^{\ell-1}-1}{2},$$

leading us to a conclusion that

$$\ell < 1 + \log\left(\frac{172(\mu+3)^4}{(\mu/\sqrt{2d_1d_2}-1)^2}\right)/\log E < 1 + \frac{5.15 + 4\log(\mu+3) - 2\log\left(\mu/\sqrt{2d_1d_2}-1\right)}{\log\left(\mu/\sqrt{d_1d_2}-1\right)}.$$

□

Corollary 5.16. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be an irreducible binary form of degree $d_1 \geq 5$. Let $\alpha_1, \ldots, \alpha_{d_1}$ denote the roots of $F(X,1)$,*

$$d_2 = \max_{1\leq i\leq d_1}\{[\mathbb{Q}(\alpha_1,\alpha_i)\colon \mathbb{Q}(\alpha_1)]\},$$

and suppose that $d_2 < d_1/4$. For a prime p and a positive integer k, consider

$$|F(x,y)| = tp^k, \tag{5.4.4}$$

where t is a positive integer variable. Let λ be a non-negative real number and μ such that

$$2\sqrt{d_1d_2} < \mu \leq d_1.$$

Let C be the maximum between the constant C_1 from Theorem 5.15 and

$$\left(\frac{2^d d^{(d-1)/2} M(F)^{d-2}(p^k)^{1+\lambda}}{|D(F)|^{1/2}}\right)^{\frac{1}{d-\mu}}.$$

Then the number of solutions to (5.4.4) such that

$$\gcd(x,y) = 1,\ \ t \leq (p^k)^\lambda,\ \ H(x,y) \geq C$$

does not exceed

$$4\gamma \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu+3) - 2\log\left(\mu/\sqrt{2d_1d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1d_2} - 1\right)} \right\rfloor,$$

where

$$\gamma = \#\{i\colon \text{a minimal polynomial of } (\alpha_1, \alpha_i) \text{ is } \mu\text{-special}\}.$$

Proof. Let $\alpha_1, \alpha_2, \dots, \alpha_d$ be the roots of $F(X,1)$. By Lemma 2.26, there exists an index $j \in \{1,2,\dots,d\}$ such that

$$\begin{aligned}\min\left\{\left|\alpha_j - \frac{x}{y}\right|, \left|\alpha_j^{-1} - \frac{y}{x}\right|\right\} &< \frac{2^d d^{(d-1)/2} M(F)^{d-2}|F(x,y)|}{|D(F)|^{1/2} H(x,y)^d} \\ &\le \frac{2^d d^{(d-1)/2} M(F)^{d-2}(p^k)^{1+\lambda}}{|D(F)|^{1/2} H(x,y)^d} \\ &\le \frac{1}{H(x,y)^\mu},\end{aligned}$$

where the last inequality follows from $H(x,y) \ge C$. To count the number of solutions to this inequality, we apply Theorem 5.15 twice, first to $C_0 = 1, \mu, f(X) = F(X,1)$, and then to $C_0 = 1, \mu, f(X) = F(1,X)$. The result follows once we recall that the solutions (x,y) and $(-x,-y)$ are regarded as the same. □

5.5 Proof of Theorem 5.5

Throughout the proof we will be adjusting our choice of p^k several times. Let

$$C_0 = \left(\frac{2^{d_1} d^{(d_1-1)/2} M(F)^{d_1-2}}{|D(F)|^{1/2}}\right)^{1/(d_1-\mu)}.$$

Recall the definitions of explicitly computable constants $C_1 = C_1(1, \mu, F(X,1))$ and $C = C(1, \lambda, \mu, F(X,1))$ from Theorem 5.15 and Corollary 5.15, respectively. We begin by choosing p^k so that

$$p^k \ge \left(\frac{|D(F)|^{1/2}}{2^{d_1} d_1^{(d_1-1)/2} M(F)^{d_1-2}}\right)^{\frac{\mu}{d_1-\mu}} C_1^\mu.$$

Then by definition of C,

$$C = C_0 (p^k)^{\frac{1}{\mu}}.$$

By Corollary 5.16, the number of solutions (x, y, t) to (5.0.1) with $H(x, y) \geq C$ does not exceed

$$4\gamma \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu + 3) - 2\log\left(\mu/\sqrt{2d_1d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1d_2} - 1\right)} \right\rfloor.$$

It remains to count the number of solutions (x, y, t) such that

$$H(x, y) < C_0(p^k)^{1/\mu}. \tag{5.5.1}$$

Note that

$$H(x, y)^\mu < C_0^\mu p^k \leq C_0^\mu |F(x, y)|_p^{-1}.$$

Therefore

$$|F(x, y)|_p < \frac{C_0^\mu}{H(x, y)^\mu}. \tag{5.5.2}$$

Let us further adjust our choice of p^k as follows:

$$p^k > |D(F)|.$$

Then

$$|F(x, y)|_p \leq p^{-k} < |D(F)|^{-1} \leq |D(F)|_p.$$

By Lemma 2.16 there exists a unique p-adic root $\alpha \in \mathbb{Q}_p$ of $F(X, 1)$ such that

$$|y\alpha - x|_p \leq \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x, y)|_p.$$

Combining this inequality with (5.5.2), we obtain

$$\begin{aligned} |y\alpha - x|_p &\leq \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x, y)|_p \\ &< \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} \cdot \frac{C_0^\mu}{H(x, y)^\mu} \\ &\leq \frac{C_1}{H(x, y)^\mu}, \end{aligned}$$

where

$$C_1 = C_0^\mu c_d |D(F)|^{1/2}.$$

Now, suppose that there exist two distinct solutions $(x_1, y_1, t_1), (x_2, y_2, t_2)$ to (3.0.1) such that $H(x_2, y_2) \geq H(x_1, y_1)$, with $H(x_1, y_1)$ as small as possible. Let $\alpha_1, \dots, \alpha_t \in \mathbb{Q}_p$ denote the p-adic roots of $F(X, 1)$. Then there exist indices i_1 and i_2 such that

$$|y_1\alpha_{i_1} - x_1|_p < \frac{C_1}{H(x,y)^\mu}, \quad |y_2\alpha_{i_2} - x_2|_p < \frac{C_1}{H(x,y)^\mu}.$$

We would like to apply Lemma 5.14 to $(\alpha_{i_1}, \alpha_{i_2})$. For this purpose, we need to further adjust our choice of p^k. Denote a minimal polynomial for (α_1, α_i) by R_i, and define $R_{i,j} = R_{\sigma(j)}$, where σ is some element of $\mathrm{Gal}(\mathbb{Q}(\alpha_1, \dots, \alpha_d)/\mathbb{Q})$ such that $\sigma\colon \alpha_i \mapsto \alpha_1$. By construction, $R_{i,j}(\alpha_i, \alpha_j) = 0$. Further, let $t_2(i), C_1(i), C_4(i)$ denote the constants (5.3.5), (5.3.6), (5.3.8), with $(\alpha, \beta) = (\alpha_1, \alpha_i)$. We choose p^k so that

$$p^k \geq (d+1)H(F) \max\left\{ \max_{1\leq i\leq d} \{C_1(i)\}, \max_{\substack{1\leq i\leq d \\ t_2(i)<\mu^2/(d_1+\mu)}} \{C_4(i)\} \right\}^d.$$

Since $p^k \leq |F(x_i, y_i)| \leq (d+1)H(F)H(x_i, y_i)^d$, we conclude that

$$H(x_i, y_i) \geq \left(\frac{p^k}{(d+1)H(F)}\right)^{1/d} \geq \max\left\{ \max_{1\leq i\leq d} \{C_1(i)\}, \max_{\substack{1\leq i\leq d \\ t_2(i)<\mu^2/(d_1+\mu)}} \{C_4(i)\} \right\}.$$

It follows from the above inequality and $\mu > 2\sqrt{d_1 d_2}$ that Lemma 5.14 can be applied to $(C_0, \alpha, \beta, \mu) = (C_1, \alpha_{i_1}, \alpha_{i_2}, \mu)$. Furthermore, our choice of m implies that case 3 in Lemma 5.14 does not hold. Therefore we either have

$$H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu/\sqrt{d_1 d_2}-1}, \tag{5.5.3}$$

where C_2 is defined in (5.3.7), or the polynomial $R_{i_1,i_2}(X, Y)$ is μ-special and

$$R_{i_1,i_2}\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) = 0. \tag{5.5.4}$$

Suppose that (5.5.3) holds. We combine this inequality with (5.5.1):

$$\begin{aligned} C_2^{-1} H(x_1, y_1)^{\mu/\sqrt{d_1 d_2}-1} &< H(x_2, y_2) \\ &< C_0 (p^k)^{1/\mu} \\ &\leq C_0 (|F(x_1, y_1)|)^{1/\mu} \\ &\leq C_0 ((d_1+1)H(F)H(x_1, y_1)^{d_1})^{1/\mu}. \end{aligned}$$

Therefore

$$H(x_1, y_1)^{\mu/\sqrt{d_1 d_2}-1-d_1/\mu} < C_0 C_2 \left((d_1+1)H(F)\right)^{1/\mu}.$$

Since $\mu = d_2/(2+\lambda)$, it follows from our choice of λ that

$$\mu/\sqrt{d_1 d_2} - 1 - d_1/\mu > 0.$$

Thus, if we choose p^k so that

$$p^k \geq (d+1)H(F)\left(\max\{2C_1, C_2\}C_0((d+1)H(F))^{1/\mu}\right)^{1/(\mu/\sqrt{d_1 d_2}-1-d_1/\mu)}, \tag{5.5.5}$$

then the inequalities $p^k \leq |F(x_1, y_1)| \leq (d_1+1)H(F)H(x_1, y_1)^{d_1}$ imply

$$H(x_1, y_1) \geq \left(\frac{p^k}{(d_1+1)H(F)}\right)^{1/d_1} \geq \left(C_0 C_2((d_1+1)H(F))^{1/\mu}\right)^{1/(\mu/\sqrt{d_1 d_2}-1-d_1/\mu)},$$

so we obtain a contradiction. Consequently, by choosing p^k sufficiently large we are able to ensure that the inequality (5.5.3) does not hold. But then it follows from Lemma 5.14 that the polynomial $R_{i_1,i_2}(X, Y)$ is μ-special and the equality (5.5.4) holds.

With (x_1, y_1, i_1) fixed, it remains to estimate the number of solutions (x_2, y_2, i_2) to (5.5.4). We claim that it does not exceed 2γ. Indeed, suppose that there are at least $2\gamma+1$ solutions. Certainly, if (x_2, y_2, i_2) is a solution, then so is $(-x_2, -y_2, i_2)$. By the Dirichlet's box principle, there exists some minimal μ-special polynomial R_{i_1,i_2} such that

$$R_{i_1,i_2}\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) = R_{i_1,i_2}\left(\frac{x_1}{y_1}, \frac{x_2'}{y_2'}\right) = 0,$$

where x_2/y_2, x_2'/y_2' are distinct rationals in lowest terms such that

$$|y_2\alpha_{i_2} - x_2|_p < \frac{C_1}{H(x_2, y_2)^\mu}, \quad |y_2'\alpha_{i_2} - x_2'|_p < \frac{C_1}{H(x_2', y_2')^\mu}$$

for some root α_{i_2} of $F(X, 1)$, ordered so that $H(x_2', y_2') \geq H(x_2, y_2)$. By the product formula and the non-Archimedean triangle inequality,

$$\begin{aligned}
\frac{1}{2H(x_2, y_2)H(x_2', y_2')} &\leq |x_2 y_2' - y_2 x_2'|^{-1} \\
&\leq |x_2 y_2' - y_2 x_2'|_p \\
&\leq \max\left\{|y_2\alpha_{i_2} - x_2|_p, |y_2'\alpha_{i_2} - x_2'|_p\right\} \\
&< \frac{C_1}{H(x_2, y_2)^\mu},
\end{aligned}$$

so

$$H(x_2', y_2') > (2C_1)^{-1} H(x_2, y_2)^{\mu-1}. \tag{5.5.6}$$

Combining this inequality with (5.5.1),

$$\begin{aligned}(2C_1)^{-1} H(x_2, y_2)^{\mu(\mu-1)} &< H(x_2', y_2') \\ &< C_0 (p^k)^{1/\mu} \\ &< C_0 (|F(x_2, y_2)|)^{1/\mu} \\ &\leq C_0 ((d_1+1) H(F) H(x_2, y_2)^{d_1})^{1/\mu}.\end{aligned}$$

Therefore

$$H(x_2, y_2)^{\mu(\mu-1)-d_1/\mu} < 2C_0C_1 \left((d+1)H(F)\right)^{1/\mu}.$$

However, since

$$\mu(\mu-1) - d_1/\mu \geq \mu/\sqrt{d_1 d_2} - d_1/\mu > \mu/\sqrt{d_1 d_2} - 1 - d_1/\mu > 0,$$

our choice of p^k made in (5.5.5) implies that

$$\begin{aligned}H(x_2, y_2) &\geq \left(\frac{p^k}{(d_1+1)H(F)}\right)^{1/d_1} \\ &\geq \left(2C_0C_1 \left((d+1)H(F)\right)^{1/\mu}\right)^{1/(\mu/\sqrt{d_1 d_2} - 1 - d_1/\mu)} \\ &> \left(2C_0C_1 \left((d+1)H(F)\right)^{1/\mu}\right)^{1/(\mu(\mu-1)-d_1/\mu)},\end{aligned}$$

so we are able to ensure that the inequality (5.5.6) does not hold. Thus we reach a contradiction, and so there are at most 2γ solutions (x, y, t) to (5.0.1), including (x_1, y_1, t_1), such that $H(x, y) < C_0 (p^k)^{1/\mu}$.

5.6 Proof of Theorem 5.6

By Roth's theorem [29], for every root α of $F(X, 1)$ there exist only finitely many non-zero integers x, y such that $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x, y)^{-2.05}$. Now let (x, y, z, t) be a solution of (2.0.3). Since $|F(x, y)| \leq (d+1)H(F)H(x, y)^d$ and $|F(x, y)| = tp^z$, we have

$$\frac{p}{(d+1)H(F)} \leq \frac{tp^z}{(d+1)H(F)} = \frac{|F(x, y)|}{(d+1)H(F)} \leq H(x, y)^d.$$

Hence by choosing a large enough p we can increase $H(x,y)$ and make it so large that the inequality $\min\{|\alpha - x/y|, |\alpha^{-1} - y/x|\} \leq H(x,y)^{-2.05}$ is no longer satisfied for every root α of $F(X,1)$.

Let (x,y,z,t) be a solution to (5.0.2). On one hand, it follows from our choice of p that

$$\frac{1}{H(x,y)^{2.05}} < \max\left\{\left|\alpha_j - \frac{x}{y}\right|, \left|\alpha_j^{-1} - \frac{y}{x}\right|\right\}.$$

On the other hand, it follows from Lemma 2.26 that there exists an index j such that

$$\max\left\{\left|\alpha_j - \frac{x}{y}\right|, \left|\alpha_j^{-1} - \frac{y}{x}\right|\right\} \leq \frac{C_0|F(x,y)|}{H(x,y)^{d_1}}, \tag{5.6.1}$$

where

$$C_0 = \frac{2^{d_1-1}d^{(d_1-1)/2}M(F)^{d_1-2}}{|D(F)|^{1/2}}.$$

Combining this inequality with (5.6.1),

$$H(x,y) < (C_0 t p^z)^{1/(d_1-2.05)}.$$

Since $t \leq p^\lambda$, we see that

$$tp^z \leq p^{1+\lambda} \leq (p^z)^{1+\lambda} \leq |F(x,y)|_p^{-(1+\lambda)}.$$

Therefore

$$H(x,y) < (C_0 t p^z)^{1/(d_1-2.05)} \leq \left(C_0|F(x,y)|_p^{-(1+\lambda)}\right)^{1/(d_1-2.05)},$$

and so we conclude that

$$|F(x,y)|_p < \frac{C_0^{1/(1+\lambda)}}{H(x,y)^\mu}, \tag{5.6.2}$$

where

$$\mu = \frac{d_1 - 2.05}{1+\lambda}.$$

Since $d_1 \geq 8$, $\lambda < (d_1 - 2.05)/2\sqrt{d_1 d_2} - 1$ and $d_2 < (d_1 - 2.05)^2/(4d_1)$, we conclude that

$$\mu > 2\sqrt{d_1 d_2}.$$

Let us further adjust our choice of p as follows:

$$p > |D(F)|.$$

Then

$$|F(x,y)|_p \le p^{-1} < |D(F)|^{-1} \le |D(F)|_p.$$

By Lemma 2.16 there exists a unique p-adic root $\alpha \in \mathbb{Q}_p$ of $F(X,1)$ such that

$$|y\alpha - x|_p \le \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x,y)|_p.$$

Combining this inequality with (5.6.2), we obtain

$$|y\alpha - x|_p \le \frac{\max\{1, |\alpha|_p\}}{|D(F)|_p^{1/2}} |F(x,y)|_p < \frac{C_1}{H(x,y)^\mu},$$

where

$$C_1 = C_0^{1/(1+\lambda)} c_d |D(F)|^{1/2}.$$

Note that C_1 is independent of p. Further, we choose p so that $p \nmid c_0 c_d$. Such a choice of p implies $p \nmid x$ and $p \nmid y$. Indeed, if $p \mid y$, then p does not divide x, because x and y are coprime. Since $z \ge 1$, it is evident from the equation

$$c_d x^d + y(c_{d-1}x^{d-1} + \cdots + c_0 y^{d-1}) = \pm t p^z$$

that p divides c_d, in contradiction to our choice of p. Then $|y|_p = 1$, and so for any $\alpha \in \mathbb{Q}_p$ we have

$$\left|\alpha - \frac{x}{y}\right|_p = |y\alpha - x|_p.$$

Analogously, we can show that $|x|_p = 1$. Therefore

$$\left|\alpha - \frac{x}{y}\right|_p < \frac{C_1}{H(x,y)^\mu}.$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_d$ be the roots of $F(X,1)$. Applying Theorem 5.15 to $C_1, \mu, f(X) = F(X,1)$, we conclude that there exists a positive number C_2, which depends on C_1, μ, F, but not on p, such that the number of rationals x/y in lowest terms satisfying $H(x,y) \ge C_2$, $|x|_p = 1$, and

$$\left|\alpha_j - \frac{x}{y}\right|_p < \frac{C_1}{H(x,y)^\mu}$$

for some $j \in \{1, 2, \ldots, d\}$ is less than

$$\hat{\gamma} \cdot \left\lfloor 1 + \frac{5.15 + 4\log(\mu+3) - 2\log\left(\mu/\sqrt{2d_1d_2} - 1\right)}{\log\left(\mu/\sqrt{d_1d_2} - 1\right)} \right\rfloor.$$

where $\hat{\gamma}$ is the number of μ-special polynomials. If we choose p so that $p \geq (d_1+1)H(F)C_2^{d_1}$, then

$$C_2^{d_1} \leq \frac{p}{(d_1+1)H(F)} \leq \frac{tp^z}{(d_1+1)H(F)} = \frac{|F(x,y)|}{(d_1+1)H(F)} \leq H(x,y)^{d_1},$$

so the inequality $H(x,y) \geq C_2$ is satisfied.

Suppose that R_i is a μ-special polynomial such that the curve $R_i(X_1, X_2) = 0$ has only finitely many rational points. Then there exists a number $C_3 = C_3(R)$ such that if $H(x_2, y_2) \geq H(x_1, y_1) \geq C_3$ then $R(x_1/y_1, x_2/y_2) \neq 0$. By choosing p sufficiently large we can always ensure that $H(x,y) \geq C_3$, and so we can replace $\hat{\gamma}$, the number of μ-special polynomials, with γ, the number of μ-exceptional polynomials. The result follows once we recall that the solutions (x,y,z,t) and $(-x,-y,z,t)$ are regarded as distinct.

Chapter 6

Conclusion

In this thesis we developed generalizations of two principles from the theory of Diophantine approximation, namely the gap principle and the Thue-Siegel principle. These generalizations enabled us to prove Theorems 2.2 and 2.3, where we established absolute bounds on the number of solutions of certain equations of Thue and Thue-Mahler type.

In our studies we were primarily motivated by Conjecture 1.1, which was stated by Stewart in 1991 [31]. Theorem 2.2 serves as a theoretical evidence in support of Stewart's conjecture, but in order for the conjecture to be established, at least with the techniques investigated in this manuscript, more theory has to be developed. In particular, the following two problems have to be resolved:

1. In Theorem 2.2 irreducible binary forms $F(X, Y)$ have a rather restrictive property that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$, where α is a root of $F(X, 1)$, is a Galois extension. Such forms are only a drop in the ocean of other forms for which we believe that Stewart's conjecture is true. It is a challenging problem to generalize our results to larger families of binary forms, ideally to all forms of degree at least three with non-zero discriminant. When proving Theorem 5.5 we had this goal in mind, but our bounds ended up not being absolute. To be more precise, they depend on a parameter γ, which does not exceed the degree of F. Perhaps, a careful investigation of a function $\gamma = \gamma(\mu, F)$ would lead to further progress in this direction.

2. In equation (2.0.2) we considered positive integers tp^k, where t is small relative to a prime power p^k. However, we believe that Stewart's conjecture should hold for all integers, not just those that are divisible by a large prime power. This suggests another direction of research, namely to generalize Theorem 2.2 to all possible integers.

Finally, it would be interesting to see generalizations of Theorems 2.3 and 5.6, where we established bounds on the number of solutions of certain equations of Thue-Mahler type. In both settings we restricted our attention to only one prime p, but the question what happens if we consider more than one prime remains open. We believe that ideas present in this thesis have a lot of potential for further development and could lead to an answer to this question, as well as to the resolution of Stewart's conjecture.

Bibliography

[1] E. Bombieri. On the Thue-Siegel-Dyson theorem. *Acta Math.*, 148:255–296, 1982.

[2] E. Bombieri and W. Gubler. *Heights in Diophantine Geometry*. Cambridge University Press, 2006.

[3] E. Bombieri and J. Mueller. On effective measures of irrationality for $\sqrt[r]{\frac{a}{b}}$ and related numbers. *J. Reine Agnew. Math.*, 342:173–196, 1983.

[4] E. Bombieri and W. M. Schmidt. On Thue's equation. *Invent. Math.*, 88:69–81, 1987.

[5] E. Bombieri and J. Vaaler. On Siegel's lemma. *Invent. Math.*, 73:11–32, 1983.

[6] Y. Bugeaud and M. Mignotte. On the distance between roots of integer polynomials. *Proc. Edinburgh Math. Soc.*, 47:553–556, 2004.

[7] J. W. S. Cassels. *An Introduction to Diophantine Approximation*. Cambridge University Press, 1957.

[8] C. Curtis and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras*, pages 258–262. John Wiley & Sons, 1962.

[9] M. Demirci and I. N. Cangül. The constant term of the minimal polynomial $\cos(2\pi/n)$ over $\mathbb{Q}$. *Fixed Point Theory and Applications*, 77, 2013.

[10] D. S. Dummit and R. M. Foote. *Abstract Algebra*. John Wiley & Sons, third edition, 2004.

[11] F. Dyson. The approximation of algebraic numbers by rationals. *Acta Math.*, 79:225–240, 1947.

[12] P. Erdős and K. Mahler. On the number of integers which can be represented by a binary form. *J. London Math. Soc.*, 13:134–139, 1938.

[13] P. Erdős, C. L. Stewart, and R. Tijdeman. Some Diophantine equations with many solutions. *Comp. Math.*, 66:37–56, 1988.

[14] H. Esnault and E. Viehweg. Dyson's lemma for polynomials in several variables (and the Theorem of Roth). *Invent. Math.*, 78:445–490, 1984.

[15] J.-H. Evertse. On equations in S-units and the Thue-Mahler equation. *Invent. Math.*, 75:561–584, 1984.

[16] W. Gautschi. Norm estimates for inverses of Vandermonde matrices. *Numer. Math.*, 23:337–347, 1975.

[17] K. Győry. Thue inequalities with a small number of primitive solutions. *Period. Math. Hungar.*, 42(1-2):239–246, 2001.

[18] H. Iwaniec. On the problem of Jacobsthal. *Demonstratio Mathematica*, 11(1):225–231, 1978.

[19] H. Kanold. Über eine zahlentheoretische Funktion von Jacobsthal. *Math. Ann.*, 170(4):314–326, 1967.

[20] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[21] D. Lewis and K. Mahler. Representation of integers by binary forms. *Acta Arith.*, 6:333–363, 1961.

[22] J. J. Liang. On the integral basis of the maximal real subfield of a cyclotomic field. *J. Reine Angew. Math.*, 286-287:223–226, 1976.

[23] J. Liouville. Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationelles algébriques. *J. Math. pures appl.*, 16:133–142, 1951.

[24] K. Mahler. Zur Approximation algebraischer Zahlen. II. Über die Anzahl der Darstellungen ganzer Zahlen durch Binärformen. *Math. Ann.*, 108:37–55, 1933.

[25] M. Newman. *Integral Matrices*, volume 45 of *Pure and Appl. Math.* Academic Press, New York, 1972.

[26] T Pejkovic. *Polynomial root separation and applications*. PhD thesis, Université de Strasbourg, 2012.

[27] V. P. Prasolov. *Polynomials*, volume 11 of *Algorithms and Computation in Mathematics*. Springer-Verlag Berlin Heidelberg, 2004.

[28] G. Robin. Estimation de la fonction de Tchebychef θ sur le k-ième nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n. *Acta Arith.*, 42(4):367–389, 1983.

[29] K. Roth. Rational approximations to algebraic numbers. *Mathematika*, 2(3):1–20, 1955.

[30] W. M. Schmidt. *Diophantine Approximations and Diophantine Equations*. Springer-Verlag, 1991.

[31] C. L. Stewart. On the number of solutions of polynomial congruences and Thue equations. *J. Amer. Math. Soc.*, 4:793–835, 1991.

[32] C. L. Stewart. Cubic Thue equations with many solutions. *International Mathematics Research Notices*, page rnn040, 2008.

[33] A. Thue. Über Annäherungswerte algebraischer Zahlen. *J. Reine Angew. Math.*, 135:284–305, 1909.

[34] J. L. Thunder. Thue equations and lattices. *Illinois J. Math.*, 59(4):999–1023, 2015.

[35] R. C. Vaughan. On the order of magnitude of Jacobsthal's function. *Proc. Edinburgh Math. Soc.*, 20:329–331, 1976–77.

[36] U. Zannier. *Lecture Notes on Diophantine Analysis*. Scuola Normale Superiore Pisa, 2014.