

On the Generalization of the Gap Principle

Anton Mosunov

University of Waterloo

Abstract

Let α be a real algebraic number of degree $d \geq 3$ and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. Let μ be a real number such that $(d/2) + 1 < \mu < d$ and let C_0 be a positive real number. We prove that there exist positive real numbers C_1 and C_2 , which depend only on α , β , μ and C_0 , with the following property. If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms such that

$$H(x_2, y_2) \geq H(x_1, y_1) \geq C_1$$

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 either $H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu-d/2}$, or there exist integers s, t, u, v , with $sv - tu \neq 0$, such that

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

or both. Here $H(x, y) = \max(|x|, |y|)$ is the height of x/y . Since $\mu - d/2$ exceeds one, our result demonstrates that, unless α and β are connected by means of a linear fractional transformation with integer coefficients, the heights of x_1/y_1 and x_2/y_2 have to be exponentially far apart from each other. An analogous result is established in the case when α and β are p -adic algebraic numbers.

1 Introduction

The theory of Diophantine approximation concerns the question of how well real numbers can be approximated by rationals, and its variations. If α is a real number and x/y is a rational number, with $x, y \in \mathbb{Z}$ and $y \geq 1$, then the quality of approximation of α by x/y can be measured by means of a quantity μ such that the inequality

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

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) 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.

It is not difficult to count distinct x/y satisfying (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), then their denominators must be exponentially far apart from each other.

Unfortunately, as the quantity C_2 can be arbitrarily large, the gap principle itself does not allow us to count the number of distinct solutions to (1). However, it was established by Thue [15] that, when α is an irrational algebraic number of degree $d \geq 3$ and $(d/2) + 1 < \mu < d$, then there exist computable positive real numbers C_1 and $\eta > 1$, which depend only on α and μ , such that every solution x_i/y_i with $C_1 \leq y_1 < \dots < y_\ell$ satisfies $y_i < y_1^\eta$. This phenomenon is known as the *Thue-Siegel principle* and it was vastly generalized by Bombieri and Mueller [2]. When combined with the gap principle, the Thue-Siegel principle enables us to count the number of solutions x/y to (1) such that $y \geq C_1$.

For a rational number x/y in lowest terms, let $H(x, y) = \max(|x|, |y|)$ denote the *height* of x/y . In this article, we generalize the gap principle as follows. Notice that the positive real numbers C_1, C_2, \dots occurring throughout the article are all computable.

Theorem 1.1. (A generalized Archimedean gap principle) *Let α be a real algebraic number of degree $d \geq 3$ over \mathbb{Q} and let β be irrational and in $\mathbb{Q}(\alpha)$. Let μ be a real number such that $(d/2) + 1 < \mu < d$ and let C_0 be a positive real number. There exist positive real numbers C_1 and C_2 , which depend only on α, β, μ and C_0 , with the following property. If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms such that $H(x_2, y_2) \geq H(x_1, y_1) \geq C_1$ 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 at least one of the following holds:

- $H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu-d/2}$;
- *There exist integers s, t, u, v , with $sv - tu \neq 0$, such that*

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

Since the exponent $\mu - d/2$ exceeds one, our result demonstrates that, unless α and β are connected by means of a linear fractional transformation with integer coefficients, the heights of x_1/y_1 and x_2/y_2 have to be exponentially far apart from each other.

Next, let $|\cdot|_p$ denote the p -adic absolute value on the field of p -adic numbers \mathbb{Q}_p , normalized so that $|p|_p = p^{-1}$. An analogous result for p -adic algebraic numbers is as follows.

Theorem 1.2. (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 μ be a real number such that $(d/2) + 1 < \mu < d$ and let C_0 be a positive real number. There exist positive real numbers C_3 and C_4 , which depend only on α , β , μ and C_0 , with the following property. If x_1/y_1 and x_2/y_2 are rational numbers in lowest terms such that $H(x_2, y_2) \geq H(x_1, y_1) \geq C_3$ 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 at least one of the following holds:

- $H(x_2, y_2) > C_4^{-1} H(x_1, y_1)^{\mu-d/2}$;
- *There exist integers s, t, u, v , with $sv - tu \neq 0$, such that*

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

We apply our result to establish an absolute bound on the number of large primitive solutions of certain Thue inequalities. A *Thue inequality* is an inequality of the form

$$0 < |F(x, y)| \leq m, \tag{2}$$

where m is a positive integer and $F \in \mathbb{Z}[x, y]$ is an irreducible binary form of degree $d \geq 3$. A solution $(x, y) \in \mathbb{Z}^2$ to the above inequality is called *primitive* when x and y are coprime. In [8, Theorem 1], Györy proved that there exists a positive real number Y_0 , which depends only on m and F , such that the number of primitive solutions (x, y) to (2) with $H(x, y) \geq Y_0$ does not exceed $25d$ (here the solutions (x, y) and $(-x, -y)$ are regarded as the same). Using Theorem 1.1, we can improve Györy's result in the case when F is irreducible and the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois, where α is a root of $F(x, 1)$.

To state our result, we need to introduce the notion of *enhanced automorphism group* of F . For a 2×2 matrix $M = \begin{pmatrix} s & u \\ t & v \end{pmatrix}$, with complex entries, define the binary form F_M by

$$F_M(x, y) = F(sx + uy, tx + vy).$$

Let $\overline{\mathbb{Q}}$ denote the algebraic closure of the rationals and let K be a field containing \mathbb{Q} . We say that a matrix $M = \begin{pmatrix} s & u \\ t & v \end{pmatrix} \in M_2(K)$ is a *K-automorphism* of F (resp.,

$|F|$) if $F_M = F$ (resp., $F_M = \pm F$). The set of all K -automorphisms of F (resp., $|F|$) is denoted by $\text{Aut}_K F$ (resp., $\text{Aut}_K |F|$). We define

$$\text{Aut}' |F| = \left\{ \frac{1}{\sqrt{|sv - tu|}} \begin{pmatrix} s & u \\ t & v \end{pmatrix} : s, t, u, v \in \mathbb{Z} \right\} \cap \text{Aut}_{\overline{\mathbb{Q}}} |F| \quad (3)$$

and refer to it as the *enhanced automorphism group* of F . One can verify that $\text{Aut}' |F|$ is a group. In Section 7 we will show that, under the conditions on F specified above, it is finite and contains at most 24 elements. We prove the following.

Theorem 1.3. *Let $F \in \mathbb{Z}[x, y]$ be an irreducible binary form of degree $d \geq 3$. Let α be a root of $F(x, 1)$ and assume that the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois. For a positive integer m consider the Thue inequality (2). Let μ be a real number such that $(d/2) + 1 < \mu < d$. There exists a positive real number C_5 , which depends only on m , F and μ , such that the number of primitive solutions (x, y) to (2) with $H(x, y) \geq C_5$ does not exceed*

$$\# \text{Aut}' |F| \cdot \left\lceil 1 + \frac{11.51 + 1.5 \log d + \log \mu}{\log(\mu - d/2)} \right\rceil.$$

Here the solutions (x, y) and $(-x, -y)$ are regarded as the same.

Let $\mu = (3d + 2)/4$. Then the function

$$f(d) = 1 + \frac{11.51 + 1.5 \log d + \log((3d + 2)/4)}{\log((d + 2)/4)}$$

is monotonously decreasing on the interval $[3, \infty)$. To see that this is the case, it is sufficient to prove that $g(x) = \frac{\log x}{\log((x+2)/4)}$ and $h(x) = \frac{\log((3x+2)/4)}{\log((x+2)/4)}$ are monotonously decreasing on the specified interval. We leave it as an exercise to the reader to prove that the derivatives of $g(x)$ and $h(x)$ take negative values when evaluated at any $x_0 \geq 3$. Since $f(3) \approx 64.5$, we can use the upper bound $\# \text{Aut}' |F| \leq 24$ established in Lemma 7.2 as well as Theorem 1.3 to conclude that the number of primitive solutions (x, y) to (2) such that $H(x, y) \geq C_5$ does not exceed $24 \cdot \lfloor f(3) \rfloor = 1536$ when $d \geq 3$. Furthermore, since $f(10^{14}) < 4$ and $\lim_{d \rightarrow \infty} f(d) = 3.5$, we can also conclude that it does not exceed $24 \cdot \lfloor f(10^{14}) \rfloor = 72$ when $d \geq 10^{14}$. While it is an interesting task to compare the value of C_5 to the quantity Y_L in [10] (see equation 2.9) or to the quantity Y_0 in [8, Theorem 1], it lies outside the scope of this article.

The article is structured as follows. In Section 2 we outline a number of auxiliary results, which are used in the later sections. We recommend the reader to skip this section and use it as a reference. In Section 3 we introduce the notion of a *minimal pair* $P, Q \in \mathbb{Z}[x]$ for a tuple of algebraic numbers $(\alpha, \beta) \in \overline{\mathbb{Q}} \times (\mathbb{Q}(\alpha) \setminus \mathbb{Q})$, and summarize the properties of minimal pairs. Minimal pairs enable us to construct a nonzero polynomial $R(x, y) = P(x) + yQ(x)$,

which vanishes at the point (α, β) . When $R(x, y)$ does not vanish at the rational point $(x_1/y_1, x_2/y_2)$, establishing the gap principle is a rather easy task. In Section 4, we prove that *despite the vanishing* of $R(x, y)$ at $(x_1/y_1, x_2/y_2)$, it is still possible to prove that the heights of (x_1, y_1) and (x_2, y_2) are exponentially far apart, provided that α and β are not connected by means of a linear fractional transformation with integer coefficients. In Sections 5 and 6 we prove Theorems 1.1 and 1.2, respectively. In Section 7 we investigate the properties of the enhanced automorphism group $\text{Aut}'[F]$. Finally, in Section 8 we prove Theorem 1.3.

2 Auxiliary Results

This section contains several definitions and results, which we utilize in the remaining part of the article. We recommend the reader to skip this section and refer to it when reading the proofs outlined in sections that follow it.

Let us begin with the number of definitions. For an arbitrary polynomial $R \in \mathbb{Z}[x_1, x_2, \dots, x_n]$, we let $H(R)$ denote the maximum of Archimedean absolute values of its coefficients, and refer to this quantity as a *height* of R . For an algebraic number α with the minimal polynomial f , we write $H(\alpha) = H(f)$. For a point $(x_1, x_2, \dots, x_n) \in \mathbb{C}^n$, we define

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

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

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

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

Lemma 2.1. (Liouville's Theorem) *Let $\alpha \in \mathbb{C}$ be an algebraic number of degree d over \mathbb{Q} . There exists a positive number C_6 , which depends only on α , such that for all integers x and y , with $y \neq 0$ and $x \neq y\alpha$, the inequality*

$$\left| \alpha - \frac{x}{y} \right| \geq \frac{C_6}{H(x, y)^d} \tag{4}$$

holds.

Proof. See [13, Theorem 1E]. □

Lemma 2.2. (*p*-adic Liouville's Theorem) *Let p be a rational prime and $\alpha \in \mathbb{Q}_p$ a *p*-adic algebraic number of degree d over \mathbb{Q} . There exists a positive number C_7 , which depends only on α , such that for all integers x and y , with $x \neq y\alpha$, the inequality*

$$|y\alpha - x|_p \geq \frac{C_7}{H(x, y)^d} \tag{5}$$

holds.

Proof. Let

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

be the minimal polynomial of α and let $F(x, y) = y^d f(x/y)$ be its associated binary form. 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

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

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 $|F(x, y)|_p$ and use the inequality $|c_d|_p \geq c_d^{-1}$. \square

Lemma 2.3. (Siegel's lemma, [4]) *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 $i = 1, \dots, N$ and $j = 1, \dots, M$. Then there exist integers t_1, \dots, t_N , not all zero, such that*

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

Proof. See, for example, [16, Lemma 2.7]. \square

Lemma 2.4. *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

$$C_8 = 1 + \max_{0 \leq i \leq d-1} \{|a_{d,i}|\},$$

then $c_\alpha^{\max\{0, r-d+1\}} a_{r,i} \in \mathbb{Z}$ and $|a_{r,i}| \leq C_8^{\max\{0, r-d+1\}}$ for all i such that $0 \leq i \leq d-1$.

Proof. See [16, Proposition 2.6]. \square

Let $P \in \mathbb{C}[x]$ be a polynomial of degree $d \geq 1$. The *house* of P , denoted \overline{P} , is defined to be

$$\overline{P} = \max \{|\alpha_1|, \dots, |\alpha_d|\},$$

where $\alpha_1, \dots, \alpha_d \in \mathbb{C}$ are the roots of P . For an algebraic number α , we define the *house* of α as $\overline{\alpha} = \overline{f}$, where f is the minimal polynomial of α .

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

$$\theta_\alpha = [\mathcal{O} : \mathbb{Z}[c_\alpha \alpha]]. \quad (6)$$

That is, θ_α is equal to the index of $\mathbb{Z}[c_\alpha \alpha]$ in the additive group \mathcal{O} .

Lemma 2.5. *Let α be an algebraic number of degree d over \mathbb{Q} and*

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

where $b_0, b_1, \dots, b_{d-1} \in \mathbb{Q}$. There exists a positive number C_9 , which depends only on α and β , such that

$$\max_{1 \leq i \leq d} \{|b_i|\} \leq C_9.$$

Furthermore,

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

where θ_α is defined in (6). In particular, $\theta_\alpha c_\beta b_i \in \mathbb{Z}$ for all $i = 0, 1, \dots, d-1$.

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

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{d-1} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{d-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_d & \alpha_d^2 & \dots & \alpha_d^{d-1} \end{pmatrix} \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{d-1} \end{pmatrix}. \quad (7)$$

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 [7] that

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

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

Now, let $V^{-1} = (v_{ij})$. Then it follows from (7) that $b_i = \sum_{j=1}^d v_{ij} \beta_j$, so

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

Next, let c_α and c_β denote the leading coefficients of the minimal polynomials of α and β , respectively. 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. \square

Let $P \in \mathbb{C}[x]$ be a polynomial that is not identically equal to zero, with leading coefficient c_P . The *Mahler measure* of P , denoted $M(P)$, is defined to be $M(P) = |c_P|$ if P is the constant polynomial and

$$M(P) = |c_P| \prod_{i=1}^d \max\{1, |\alpha_i|\}$$

otherwise, where $\alpha_1, \dots, \alpha_d \in \mathbb{C}$ are the roots of P . For a binary form $Q \in \mathbb{C}[x, y]$, we define the Mahler measure of Q as $M(Q) = M(Q(x, 1))$. Finally, for an algebraic number α , we define the Mahler measure of α to be $M(\alpha) = M(f)$, where f is the minimal polynomial of α .

The following lemma is a reformulation of a well-known result of Lewis and Mahler [9].

Lemma 2.6. *Let*

$$F(x, y) = c_d x^d + c_{d-1} x^{d-1} y + \dots + c_0 y^d$$

be a binary form of degree $d \geq 2$ with integer coefficients such that $c_0 c_d \neq 0$. Let x_1 and y_1 be nonzero integers. There exists a root α of $F(x, 1)$ such that

$$\min \left\{ \left| \alpha - \frac{x_1}{y_1} \right|, \left| \alpha^{-1} - \frac{y_1}{x_1} \right| \right\} \leq \frac{C_{10} |F(x_1, y_1)|}{|H(x_1, y_1)|^d},$$

where

$$C_{10} = \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2}}{|D(F)|^{1/2}}.$$

Proof. Let α be a root of $F(x, 1)$ that minimizes $|\alpha - x/y|$. By [14, Lemma 3],

$$\left| \alpha - \frac{x_1}{y_1} \right| \leq \frac{C_{10} |F(x_1, y_1)|}{|y_1|^d}.$$

If $|y_1| \geq |x_1|$, then $H(x_1, y_1) = |y_1|$, 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 nonzero, 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 $|\beta^{-1} - y_1/x_1|$, then it follows from [14, Lemma 3] that

$$\left| \beta^{-1} - \frac{y_1}{x_1} \right| \leq \frac{C_{10}|F(x_1, y_1)|}{|x_1|^d}.$$

Since $|x_1| > |y_1|$, the result follows. \square

Lemma 2.7. *Let $K = \mathbb{R}$ 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 $|\cdot|$. Let α and β be distinct numbers in \overline{K} . Let μ and C_0 be positive real numbers.*

If x_1/y_1 is a rational number such that $H(x_1, y_1) \geq (2C_0/|\alpha - \beta|)^{1/\mu}$ and

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

then

$$\left| \beta - \frac{x_1}{y_1} \right| \geq \frac{C_0}{H(x_1, y_1)^\mu}.$$

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

$$|\alpha - \beta| \leq \left| \alpha - \frac{x_1}{y_1} \right| + \left| \beta - \frac{x_1}{y_1} \right| < \frac{2C_0}{H(x_1, y_1)^\mu},$$

and so $H(x_1, y_1) < (2C_0/|\alpha - \beta|)^{1/\mu}$, leading us to a contradiction. \square

Corollary 2.8. *Let $K = \mathbb{R}$ 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 $|\cdot|$. Let $f(x) \in \mathbb{Z}[x]$ be an irreducible polynomial of degree $d \geq 2$ with roots $\alpha_1, \dots, \alpha_d \in \overline{K}$. Let μ and C_0 be positive real numbers. There exists a positive number C_{11} , which depends only on f , μ and C_0 , with the following property. If x_1/y_1 is a rational number such that $H(x_1, y_1) \geq C_{11}$ and*

$$\left| \alpha_i - \frac{x_1}{y_1} \right| < \frac{C_0}{H(x_1, y_1)^\mu}$$

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

$$\left| \alpha_j - \frac{x_1}{y_1} \right| \geq \frac{C_0}{H(x_1, y_1)^\mu}$$

for all $j \neq i$.

Lemma 2.9. (The Thue-Siegel Principle [2]) *Let $K = \mathbb{R}$ or \mathbb{Q}_p , where p is a rational prime, and denote the standard absolute value on K by $|\cdot|$. Let*

$\alpha_1 \in K$ be an algebraic number of degree $d \geq 3$ over \mathbb{Q} and let $\alpha_2 \in \mathbb{Q}(\alpha_1)$ have degree d . Let t and τ be such that

$$\frac{2 + \sqrt{2d^3 + 2d^2 - 4d}}{d(d+1)} < t < \sqrt{\frac{2}{d}}, \quad \sqrt{2 - dt^2} < \tau < t - \frac{2}{d}, \quad (8)$$

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

$$A_i = \frac{t^2}{2 - dt^2} \left(\log M(\alpha_i) + \frac{d}{2} \right) \quad \text{for } i = 1, 2.$$

Let x_1/y_1 and x_2/y_2 be rational numbers in lowest terms that satisfy the inequalities

$$\left| \alpha_i - \frac{x_i}{y_i} \right| < \frac{1}{(4e^{A_i} H(x_i, y_i))^\lambda} \quad \text{for } i = 1, 2.$$

Then

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

where

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

Proof. Since $d \geq 3$, the intervals in (8) are guaranteed to be non-empty, so the statement is not vacuously true. Since

$$\left| \alpha_i - \frac{x_i}{y_i} \right| < \frac{1}{(4e^{A_i} H(x_i, y_i))^\lambda} < \frac{1}{(3e^{A_i} H(x_i, y_i))^\lambda} \quad \text{for } i = 1, 2,$$

the case when $|\alpha_1| \leq 1$ and $|\alpha_2| \leq 1$ follows directly from [2, Section II]. More precisely, the comment on p. 184 of [2] states that the triple (A_1, A_2, τ) is admissible for the data $(\alpha_1, \alpha_2, x_1/y_1, x_2/y_2, t, \vartheta, \delta)$, where we take $\vartheta = 1$. Note also that the comments on p. 74 of [3] apply in our situation:

- (i) the hypothesis $K_{\bar{v}} = k_v$ is not used in the proof and therefore may be omitted;
- (ii) $c(\vartheta t) \leq \log 3$;
- (iii) the chosen value for A_i implies a fortiori $\left| \alpha_i - \frac{x_i}{y_i} \right| < \vartheta(t - \tau)$ for $i = 1, 2$;
- (iv) $h(x_i/y_i) = H(x_i, y_i)$ for $i = 1, 2$;
- (v) the exponent in (5A), p. 179 of [2] should be $2\vartheta/(t - \tau)$, not $2\vartheta^{-1}/(t - \tau)$.

Next, we consider the case when $|\alpha_i| > 1$ for some $i \in \{1, 2\}$. If $K = \mathbb{R}$, then $|x_i/y_i| > 1$, and so

$$\left| \alpha_i^{-1} - \frac{y_i}{x_i} \right| < |\alpha_i|^{-1} |y_i/x_i| (4e^{A_i} H(x_i, y_i))^{-\lambda} < (3e^{A_i} H(x_i, y_i))^{-\lambda}.$$

If $K = \mathbb{Q}_p$, then $|y_i| \leq 1$ and we claim that $|x_i| = 1$. For suppose not and $|x_i| < 1$. Since x_i/y_i is in lowest terms it must be the case that $|y_i| = 1$, so

$$|x_i \alpha_i^{-1}| = |x_i| \cdot |\alpha_i^{-1}| < 1 = |y_i|.$$

Since $|x_i \alpha_i^{-1}| \neq |y_i|$, it follows from the strong triangle inequality that

$$|x_i \alpha_i^{-1} - y_i| = \max\{|x_i \alpha_i^{-1}|, |y_i|\} = \max\{|x_i \alpha_i^{-1}|, 1\} \geq 1.$$

Thus,

$$1 \leq |x_i \alpha_i^{-1} - y_i| = |y_i| |\alpha_i^{-1}| \left| \alpha_i - \frac{x_i}{y_i} \right| < (4e^{A_i} H(x_i, y_i))^{-\lambda},$$

which is impossible. Hence $|x_i| = 1$, so

$$\left| \alpha_i^{-1} - \frac{y_i}{x_i} \right| < |\alpha_i|^{-1} |y_i/x_i| (4e^{A_i} H(x_i, y_i))^{-\lambda} < (3e^{A_i} H(x_i, y_i))^{-\lambda}.$$

We conclude that, as long as $\left| \alpha_i - \frac{x_i}{y_i} \right| < (4e^{A_i} H(x_i, y_i))^{-\lambda}$ for $i = 1, 2$, the inequalities

$$\left| \alpha_i - \frac{x_i}{y_i} \right| < \frac{1}{(3e^{A_i} H(x_i, y_i))^\lambda} \quad \text{and} \quad \left| \alpha_i^{-1} - \frac{y_i}{x_i} \right| < \frac{1}{(3e^{A_i} H(x_i, y_i))^\lambda}$$

hold whenever $|\alpha_i| > 1$ for some $i \in \{1, 2\}$. Consequently, we can always choose $r, s \in \{-1, 1\}$ so that $|\alpha_1^r| \leq 1$, $|\alpha_2^s| \leq 1$ and (A_1, A_2, τ) is admissible for the data $(\alpha_1^r, \alpha_2^s, (x_1/y_1)^r, (x_2/y_2)^s, t, \vartheta, \delta)$. The result now follows from [2, Section II]. \square

3 Minimal Pairs

Let α be an algebraic number of degree d over \mathbb{Q} and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. With a pair (α, β) we associate two polynomials $P, Q \in \mathbb{Z}[x]$, which possess certain minimal properties listed in Definition 3.1. The properties of minimal pairs summarized in Proposition 3.2 will play a crucial role in proofs of Archimedean and non-Archimedean gap principles, which are outlined in Sections 5 and 6, respectively.

Definition 3.1. *Let α be an algebraic number of degree d and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. We say that two univariate polynomials, P and 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 properties (1) and (2).
- (4) The quantity $\max\{H(P), H(Q)\}$ is minimal among all polynomials satisfying properties (1), (2) and (3).

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

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

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

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

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) .

If P, Q is a minimal pair for (α, β) , then we can define a polynomial

$$R(x, y) = P(x) + yQ(x).$$

Polynomials of such form were used by Thue [15] for the purpose of establishing the first instance of the Thue-Siegel principle [2]. More precisely, they were constructed as to achieve high vanishing at the point (α, α) , i.e., $D_i R(\alpha, \alpha) = 0$ for $i = 0, 1, \dots, \ell$ for some large ℓ (see the exposition of Thue's method in [16, Chapter 2]). In turn, we construct $R(x, y)$ so to achieve $R(\alpha, \beta) = 0$ for arbitrary irrational $\beta \in \mathbb{Q}(\alpha)$ for the purpose of obtaining a generalized gap principle. The following proposition summarizes various properties of minimal pairs.

Proposition 3.2. *Let α be an algebraic number of degree d over \mathbb{Q} and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. 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.

$$1 \leq r \leq \lfloor d/2 \rfloor. \tag{9}$$

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. There exists a positive number C_{12} , which depends only on α and β , such that

$$\max\{H(P), H(Q)\} \leq C_{12}. \quad (10)$$

5. If $\alpha \in \mathbb{C}$, there exists a positive real number C_{13} , which depends only on α and β , such that

$$|W(\alpha)| \geq C_{13}. \quad (11)$$

Similarly, if $\alpha \in \mathbb{Q}_p$ for some rational prime p , there exists a positive real number C_{14} , which depends only on α and β , such that

$$|W(\alpha)|_p \geq C_{14}. \quad (12)$$

Proof. Let us prove each of the above statements.

1. First, we prove that $r \geq 1$. If not, then $r = \max\{\deg P, \deg Q\} = 0$, which means that $P = p$ and $Q = q$ for some integers p and q , not both equal to zero. If $q \neq 0$, then $P(\alpha) + \beta Q(\alpha) = 0$ implies $\beta = -p/q$, which contradicts the fact that β is irrational. If $q = 0$, then $p = 0$, which is impossible, since we assumed that both p and q cannot be equal to zero. Thus, $r \geq 1$.

Next, we prove that $r \leq s$, where $s = \lfloor d/2 \rfloor$. Write

$$\hat{P}(x) = \sum_{i=0}^s a_i x^i \quad \text{and} \quad \hat{Q}(x) = \sum_{i=0}^s a_{s+1+i} x^i. \quad (13)$$

We view the $2s+2$ integer coefficients a_0, \dots, 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} , which we will define in the proof of Part 4. Since $2s+2 > d$, the existence of a non-trivial integer solution to the system of d linear equations over \mathbb{Q} in $2s+2$ variables is guaranteed by Lemma 2.3. Therefore, there exist polynomials \hat{P}, \hat{Q} , not both zero, 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. This means that $\deg G = 0$, and so P and 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

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

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 (13). Then,

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

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

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

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

Put

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

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

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

By Lemma 2.5 we have $\theta_\alpha c_\beta b_i \in \mathbb{Z}$ for all i and $B \leq C_9$. Further, by Lemma 2.4 we have $c_\alpha^{\max\{0, k-d+1\}} c_{k,i} \in \mathbb{Z}$ for all i, k and $C \leq C_8^s$. 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

$$A = \theta_\alpha c_\beta c_\alpha^s C_9 (1 + sC_8^s).$$

By Lemma 2.3,

$$\begin{aligned} \max\{H(\hat{P}), H(\hat{Q})\} &= \max_{0 \leq i \leq 2s+1} \{|a_i|\} \\ &\leq ((2s+2)A)^{d/(2s+2-d)} \\ &\leq ((2s+2)\theta_\alpha c_\beta c_\alpha^s C_9 (1 + sC_8^s))^{d/(2s+2-d)}. \end{aligned}$$

Now that we know an upper bound on $\max\{H(\hat{P}), H(\hat{Q})\}$, we can determine an 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 $\max\{\deg \hat{P}, \deg \hat{Q}\} \leq \lfloor d/2 \rfloor$ that

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

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 3.1. By Property (4), the polynomials P and Q satisfy

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

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 nonzero, we have $H(G) \geq 1$. By Gelfond's Lemma [1, 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 and 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

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

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

With the basic properties of heights listed in [16, Section 2.4.1], we find the following upper bound on $H(W)$:

$$\begin{aligned} H(W) &\leq H(PQ') + H(QP') \\ &\leq r(H(P)H(Q') + H(Q)H(P')) \\ &\leq 2r^2H(P)H(Q) \\ &\leq 2r^2 \max\{H(P), H(Q)\}^2 \\ &\leq 2(d/2)^2 C_{12}^2. \end{aligned}$$

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

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

is a nonzero rational integer. Thus,

$$\begin{aligned}
|W(\alpha)|^{-1} &\leq c_\alpha^{d \deg W} \prod_{i=2}^d |W(\alpha_i)| \\
&\leq c_\alpha^{d \deg W} \prod_{i=2}^d (\deg W + 1) H(W) \max\{1, |\alpha_i|\}^{\deg W} \\
&\leq (2rH(W))^{d-1} \left(\frac{c_\alpha^{d-1} M(\alpha)}{\max\{1, |\alpha|\}} \right)^{2r-1} \\
&\leq ((d^3/2)C_{12}^2)^{d-1} \left(\frac{c_\alpha^{d-1} M(\alpha)}{\max\{1, |\alpha|\}} \right)^{d-1}.
\end{aligned}$$

Suppose that $\alpha \in \mathbb{Q}_p$. Let $f(x)$ denote the minimal polynomial of α with the leading coefficient c_α . By [12, Theorem 1.3.2], there exist polynomials $\varphi, \psi \in \mathbb{Z}[x]$ such that $\deg \varphi < \deg W$, $\deg \psi < d$, and

$$\varphi(x)f(x) + \psi(x)W(x) = \text{Res}(f, W).$$

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

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

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

$$\begin{aligned}
|\text{Res}(f, W)| &\leq (\deg f + 1)^{\deg W/2} (\deg W + 1)^{\deg f/2} H(\alpha)^{\deg W} H(W)^{\deg f} \\
&\leq (d+1)^{(2r-1)/2} (2r)^{d/2} H(\alpha)^{2r-1} (2(d/2)^2 C_{12}^2)^d \\
&\leq (d+1)^{(d-1)/2} d^{d/2} H(\alpha)^{d-1} ((d^2/2)C_{12}^2)^d.
\end{aligned}$$

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

$$\begin{aligned}
|W(\alpha)|_p &\geq |c_\alpha^{d-1} \text{Res}(f, W)|_p \\
&\geq |c_\alpha^{d-1} \text{Res}(f, W)|^{-1} \\
&\geq H(\alpha)^{-(d-1)} |\text{Res}(f, W)|^{-1} \\
&\geq (d+1)^{-(d-1)/2} d^{-d/2} H(\alpha)^{-2d+2} ((d^2/2)C_{12}^2)^{-d}.
\end{aligned}$$

□

4 A Gap Principle in the Presence of Vanishing

Let α be an algebraic number over \mathbb{Q} of degree $d \geq 2$ and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. Let $P, Q \in \mathbb{Z}[x]$ be polynomials such that

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

In this section we prove Proposition 4.1, which states that *despite the vanishing* of $P(x) + yQ(x)$ at the point $\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) \in \mathbb{Q}^2$, it is still possible to produce a gap principle, provided that the quantity $r = \max\{\deg P, \deg Q\}$ exceeds one.

Proposition 4.1. *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

$$P\left(\frac{x_1}{y_1}\right) + \frac{x_2}{y_2}Q\left(\frac{x_1}{y_1}\right) = 0. \quad (14)$$

Then

$$H(x_2, y_2) \geq \frac{H(x_1, y_1)^r}{C_{15} \max\{H(P), H(Q)\}^{2r^2+3r}},$$

where

$$C_{15} = 2^{r^2} (r+1)^{(3r^2+2r)/2}.$$

The proof of Proposition 4.1 is given at the end of the section, and it follows directly from the results established below.

Lemma 4.2. *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 ,*

$$P(x, y) = y^r P(x/y) \quad \text{and} \quad Q(x, y) = y^r Q(x/y).$$

Then for all coprime integers a and b the number $g = \gcd(P(a, b), Q(a, b))$ divides

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

where $\text{Res}(P, Q)$ denotes the resultant of P and Q . Furthermore,

$$1 \leq \varrho \leq (r+1)^r \max\{H(P), H(Q)\}^{2r}.$$

Proof. Let a and b be coprime integers and suppose that a prime power p^n exactly divides $g = \gcd(P(a, b), Q(a, b))$. Since a and b are coprime, either a or b is not divisible by p . Suppose that p does not divide b . By [12, Theorem 1.3.2], there exist polynomials $\varphi, \psi \in \mathbb{Z}[x]$ such that

$$\varphi(x)P(x) + \psi(x)Q(x) = \text{Res}(P, Q).$$

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

$$b^t \varphi(a/b)P(a, b) + b^t \psi(a/b)Q(a, b) = \text{Res}(P, Q)b^{r+t}$$

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

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

$$P(1, x) = \mathcal{R}(P) \quad \text{and} \quad Q(1, x) = x^{r-s} \mathcal{R}(Q),$$

so

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

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

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

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

□

Lemma 4.3. *Let*

$$P(x, y) = \prod_{i=1}^r (\alpha_i x + \beta_i y) \quad \text{and} \quad Q(x, y) = \prod_{j=1}^r (\gamma_j x + \delta_j y)$$

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

$$C = C(P, Q) = \frac{\min_{i,j} \{|\alpha_i \delta_j - \beta_i \gamma_j|\}}{\max_{i,j} \{\max\{|\alpha_i| + |\gamma_j|, |\beta_i| + |\delta_j|\}\}}. \quad (15)$$

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

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

Proof. We claim that either

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

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

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

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

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

Analogously, either

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

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

For the proof of the following result, recall the definition of the *Mahler measure* and the *house* of a polynomial introduced in Section 2.

Corollary 4.4. *Let $P, Q \in \mathbb{Z}[x]$ be coprime polynomials of degrees r and s , respectively, such that $r \geq \max\{1, s\}$. Define*

$$P(x, y) = y^r P(x/y) \quad \text{and} \quad Q(x, y) = y^r Q(x/y).$$

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

$$\max\{|P(a, b)|, |Q(a, b)|\} \geq \frac{H(a, b)^r}{2r^2(r+1)^{3r^2/2} \max\{H(P), H(Q)\}^{2r^2+r}}.$$

Proof. Let c_P and c_Q be the leading coefficients of P and Q , respectively. Then

$$|c_P| \cdot \max\{1, |\overline{P}|\} \leq M(P) \quad \text{and} \quad |c_Q| \cdot \max\{1, |\overline{Q}|\} \leq M(Q),$$

and so it follows from [1, Lemma 1.6.7] that

$$|c_P| \cdot \max\{1, |\overline{P}|\} \leq (r+1)^{1/2} H(P) \quad \text{and} \quad |c_Q| \cdot \max\{1, |\overline{Q}|\} \leq (s+1)^{1/2} H(Q). \quad (16)$$

Let μ_1, \dots, μ_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 $\gamma_j = 0$ and $\delta_j = c_Q^{1/r}$. Using (16), the constant C in (15) can be estimated from below as follows:

$$\begin{aligned}
C &= \frac{|c_P c_Q|^{1/r}}{\max_i \{ \max\{|c_P|^{1/r}, |c_P|^{1/r} |\mu_i| + |c_Q|^{1/r}\} \}} \\
&= \frac{|c_Q|^{1/r}}{\max_i \{ \max\{1, |\mu_i| + \left| \frac{c_Q}{c_P} \right|^{1/r} \} \}} \\
&= \frac{1}{\max_i \{ \max\{ \left| \frac{1}{c_Q} \right|^{1/r}, \left| \frac{\mu_i}{c_Q} \right|^{1/r} + \left| \frac{1}{c_P} \right|^{1/r} \} \}} \\
&= \frac{1}{\max\{ \left| \frac{1}{c_Q} \right|^{1/r}, \frac{|P|}{|c_Q|^{1/r}} + \left| \frac{1}{c_P} \right|^{1/r} \}} \\
&\geq \frac{1}{|c_P| \cdot \max\{1, |P|\} + |c_Q| \cdot \max\{1, |Q|\}} \\
&\geq \frac{1}{2(r+1)^{1/2} \max\{H(P), H(Q)\}}.
\end{aligned}$$

Case 2. Suppose that $s \geq 1$. Let ν_1, \dots, ν_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

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

Using (16), the constant C in (15) can be estimated from below as follows:

$$\begin{aligned}
C &= \frac{\min_{i,j} \{ |\alpha_i \delta_j - \beta_i \gamma_j| \}}{\max_{i,j} \{ \max\{ |\alpha_i| + |\gamma_j|, |\beta_i| + |\delta_j| \} \}} \\
&\geq \frac{|c_P c_Q|^{1/r} \min\{1, \min_{i,j} \{ |\mu_i - \nu_j| \} \}}{|c_P|^{1/r} \cdot \max\{1, |P|\} + |c_Q|^{1/r} \cdot \max\{1, |Q|\}} \\
&\geq \frac{\min\{1, \min_{i,j} \{ |\mu_i - \nu_j| \} \}}{2(r+1)^{1/2} \max\{H(P), H(Q)\}}.
\end{aligned}$$

By [5, 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

the lower bound on $\min_{i,j}\{|\mu_i - \nu_j|\}$ with the lower bound on C established above, we obtain

$$C \geq 2^{-r}(r+1)^{-3r/2} \max\{H(P), H(Q)\}^{-2r-1}.$$

The result now follows from Lemma 4.3. \square

Proof of Proposition 4.1. From equation (14) it follows that $Q(x_1/y_1) \neq 0$, for otherwise $P(x_1/y_1) = 0$, which means that P and Q are not coprime. Let

$$P(x, y) = y^r P(x/y) \quad \text{and} \quad Q(x, y) = y^r Q(x/y).$$

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 $P(x_1, y_1)$ and $Q(x_1, y_1)$ are integers, we see that

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

where $g = \gcd(P(x_1, y_1), Q(x_1, y_1))$. By Lemma 4.2, $g \leq (r+1)^r \max\{H(P), H(Q)\}^{2r}$. Thus,

$$H(x_2, y_2) = \frac{\max\{|P(x_1, y_1)|, |Q(x_1, y_1)|\}}{g} \geq \frac{\max\{|P(x_1, y_1)|, |Q(x_1, y_1)|\}}{(r+1)^r \max\{H(P), H(Q)\}^{2r}}.$$

Finally, since P and Q are coprime, Corollary 4.4 applies:

$$\begin{aligned} H(x_2, y_2) &\geq \frac{\max\{|P(x_1, y_1)|, |Q(x_1, y_1)|\}}{(r+1)^r \max\{H(P), H(Q)\}^{2r}} \\ &\geq \frac{H(x_1, y_1)^r}{2^{r^2}(r+1)^{(3r^2+2r)/2} \max\{H(P), H(Q)\}^{2r^2+3r}}. \end{aligned}$$

\square

5 A Generalized Archimedean Gap Principle

In this section we prove Theorem 1.1. Recall that, for any $h \in \mathbb{Z}[x]$ and i such that $0 \leq i \leq \deg h$, the inequality

$$|D_i h(\alpha)| \leq H(h) \binom{\deg h + 1}{i + 1} \max\{1, |\alpha|\}^{\deg h - i} \quad (17)$$

holds.

Let P, Q be a minimal pair for (α, β) , and define $R(x, y) = P(x) + yQ(x)$, so that $R(\alpha, \beta) = 0$. Choose C_1 so that

$$C_1 \geq C_0^{1/\mu}.$$

Then

$$\left| \alpha - \frac{x_1}{y_1} \right| < 1 \quad \text{and} \quad \left| \beta - \frac{x_2}{y_2} \right| < 1.$$

If $R(x_1/y_1, x_2/y_2) \neq 0$, then it follows from the triangle inequality, (17), and the two inequalities established above that

$$\begin{aligned} \frac{1}{H(x_1, y_1)^r H(x_2, y_2)} &\leq \left| R\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) \right| \\ &\leq \sum_{i=0}^r \sum_{j=0}^1 |D_{i,j} R(\alpha, \beta)| \left| \alpha - \frac{x_1}{y_1} \right|^i \left| \beta - \frac{x_2}{y_2} \right|^j \\ &< \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^r (|D_{i,0} R(\alpha, \beta)| + |D_{i,1} R(\alpha, \beta)|) \\ &\leq \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^r (|D_i P(\alpha)| + (1 + |\beta|) \cdot |D_i Q(\alpha)|) \\ &\leq \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^r \left(\binom{r+1}{i+1} H(P) + (1 + |\beta|) \binom{r+1}{i+1} H(Q) \right) \max\{1, |\alpha|\}^{r-i} \\ &\leq \frac{C_0}{H(x_1, y_1)^\mu} (2 + |\beta|) \max\{H(P), H(Q)\} \max\{1, |\alpha|\}^r \sum_{i=0}^r \binom{r+1}{i+1} \\ &< \frac{C_0}{H(x_1, y_1)^\mu} 2^{r+1} (2 + |\beta|) C_{12} \max\{1, |\alpha|\}^r \\ &\leq \frac{C_0}{H(x_1, y_1)^\mu} 2^{(d/2)+1} (2 + |\beta|) C_{12} \max\{1, |\alpha|\}^{d/2} \\ &= \frac{C_2}{H(x_1, y_1)^\mu}, \end{aligned}$$

where the second-to-last inequality follows from (10). By (9),

$$H(x_2, y_2) > C_2^{-1} H(x_1, y_1)^{\mu-r} \geq C_2^{-1} H(x_1, y_1)^{\mu-d/2},$$

which means that 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 and v . Note that $sv - tu \neq 0$, for otherwise the number β would have to be rational. Since $R(\alpha, \beta) = 0$ and $R(x_1/y_1, x_2/y_2) = 0$, case 2 holds.

It remains to consider the case when $R(x_1/y_1, x_2/y_2) = 0$ and $r \geq 2$. We will prove that $H(x_1, y_1) < C$ for some positive real number C , which depends only on α, β, μ and C_0 . By choosing C_1 so that $C_1 \geq C$, we then arrive to a contradiction. Note that

$$\left| \beta - \frac{x_2}{y_2} \right| = \left| \frac{P(\alpha)}{Q(\alpha)} - \frac{P(x_1/y_1)}{Q(x_1/y_1)} \right| = \frac{|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|}{|Q(\alpha)Q(x_1/y_1)|}. \quad (18)$$

Further,

$$|Q(\alpha)| \leq (r+1) \max\{H(P), H(Q)\} \max\{1, |\alpha|\}^r \leq (r+1)C_{12} \max\{1, |\alpha|\}^r, \quad (19)$$

$$\begin{aligned} \left| Q\left(\frac{x_1}{y_1}\right) \right| &\leq \sum_{i=0}^r |D_i Q(\alpha)| \cdot \left| \alpha - \frac{x_1}{y_1} \right|^i \\ &< \sum_{i=0}^r |D_i Q(\alpha)| \\ &\leq H(Q) \sum_{i=0}^r \binom{r+1}{i+1} \max\{1, |\alpha|\}^{r-i} \\ &< 2^{r+1} C_{12} \max\{1, |\alpha|\}^r. \end{aligned} \quad (20)$$

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, (17) and (11),

$$\begin{aligned} &\left| P(\alpha)Q\left(\frac{x_1}{y_1}\right) - Q(\alpha)P\left(\frac{x_1}{y_1}\right) \right| \\ &= \left| P(\alpha) \sum_{i=0}^r D_i Q(\alpha) \left(\frac{x_1}{y_1} - \alpha\right)^i - Q(\alpha) \sum_{i=0}^r D_i P(\alpha) \left(\frac{x_1}{y_1} - \alpha\right)^i \right| \\ &= \left| \alpha - \frac{x_1}{y_1} \right| \cdot \left| \sum_{i=0}^{r-1} (P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)) \left(\frac{x_1}{y_1} - \alpha\right)^i \right| \\ &= \left| \alpha - \frac{x_1}{y_1} \right| \cdot \left| W(\alpha) + \left(\frac{x_1}{y_1} - \alpha\right) \sum_{i=1}^{r-1} (P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)) \left(\frac{x_1}{y_1} - \alpha\right)^{i-1} \right| \\ &> \left| \alpha - \frac{x_1}{y_1} \right| \left(|W(\alpha)| - \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=1}^{r-1} |P(\alpha)D_{i+1}Q(\alpha) - Q(\alpha)D_{i+1}P(\alpha)| \right) \\ &\geq \left| \alpha - \frac{x_1}{y_1} \right| \left(|W(\alpha)| - \frac{C_0}{H(x_1, y_1)^\mu} 2(r+1) \max\{H(P), H(Q)\}^2 \sum_{i=1}^{r-1} \binom{r+1}{i+2} \max\{1, |\alpha|\}^{2r-i-1} \right) \\ &\geq \left| \alpha - \frac{x_1}{y_1} \right| (C_{13} - C_1^{-\mu} C_0 2^{r+2} (r+1) C_{12}^2 \max\{1, |\alpha|\}^{2r}), \end{aligned}$$

where the last inequality follows from $H(x_1, y_1) \geq C_1$, (10) and (11). Thus, if we choose C_1 so that

$$C_1^\mu \geq 2^{(d/2)+3} ((d/2) + 1) C_0 C_{12}^2 C_{13}^{-1} \max\{1, |\alpha|\}^d,$$

then it follows from $r \leq d/2$ that

$$\begin{aligned} \left| P(\alpha)Q\left(\frac{x_1}{y_1}\right) - Q(\alpha)P\left(\frac{x_1}{y_1}\right) \right| &\geq \left| \alpha - \frac{x_1}{y_1} \right| (C_{13} - C_1^{-\mu} C_0 2^{r+2} (r+1) C_{12}^2 \max\{1, |\alpha|\}^{2r}) \\ &\geq \left| \alpha - \frac{x_1}{y_1} \right| (C_{13} - C_1^{-\mu} C_0 2^{(d/2)+2} ((d/2) + 1) C_{12}^2 \max\{1, |\alpha|\}^d) \\ &\geq \frac{C_{13}}{2} \left| \alpha - \frac{x_1}{y_1} \right|. \end{aligned}$$

Combining the above result with (18), (19) and (20) yields

$$\left| \beta - \frac{x_2}{y_2} \right| = \frac{|P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)|}{|Q(\alpha)Q(x_1/y_1)|} > \frac{C_{13}}{2^{r+2}(r+1)C_{12}^2 \max\{1, |\alpha|\}^{2r}} \left| \alpha - \frac{x_1}{y_1} \right|.$$

By Proposition 4.1,

$$H(x_2, y_2) \geq \frac{H(x_1, y_1)^r}{C_{15} \max\{H(P), H(Q)\}^{2r^2+3r}} \geq \frac{H(x_1, y_1)^r}{C_{15} C_{12}^{2r^2+3r}},$$

where

$$C_{15} = 2^{r^2}(r+1)^{(3r^2+2r)/2} \leq 2^{d^2/4}((d/2)+1)^{(3d^2+4d)/8}. \quad (21)$$

Consequently,

$$\begin{aligned} \left| \alpha - \frac{x_1}{y_1} \right| &< \frac{2^{r+2}(r+1)C_{12}^2 \max\{1, |\alpha|\}^{2r}}{C_{13}} \left| \beta - \frac{x_2}{y_2} \right| \\ &< \frac{2^{r+2}(r+1)C_{12}^2 \max\{1, |\alpha|\}^{2r}}{C_{13}} \cdot \frac{C_0}{H(x_2, y_2)^\mu} \\ &\leq \frac{2^{r+2}(r+1)C_0 C_{12}^{(2r^2+3r)\mu+2} C_{15}^\mu \max\{1, |\alpha|\}^{2r}}{C_{13} H(x_1, y_1)^{r\mu}}. \end{aligned}$$

Thus, we obtain an upper bound on $|\alpha - x_1/y_1|$. On the other hand, by Lemma 2.1, the lower bound (4) holds. Combining upper and lower bounds,

$$\frac{C_6}{H(x_1, y_1)^d} \leq \left| \alpha - \frac{x_1}{y_1} \right| < \frac{2^{r+2}(r+1)C_0 C_{12}^{(2r^2+3r)\mu+2} C_{15}^\mu \max\{1, |\alpha|\}^{2r}}{C_{13} H(x_1, y_1)^{r\mu}}$$

Since $\mu > (d/2) + 1$ and $r \geq 2$, we see that $r\mu - d > 0$, so

$$\begin{aligned} H(x_1, y_1) &< \left(2^{r+2}(r+1)C_0(C_6 C_{13})^{-1} C_{12}^{(2r^2+3r)\mu+2} C_{15}^\mu \max\{1, |\alpha|\}^{2r} \right)^{1/(r\mu-d)} \\ &\leq \left(2^{d^2\mu/4}((d/2)+1)^{(3d^2+4d)\mu/8} C_0(C_6 C_{13})^{-1} C_{12}^{(d^2+3d)\mu/2+2} \max\{1, |\alpha|\}^d \right)^{1/(2\mu-d)} \\ &= C. \end{aligned}$$

Notice how in the second-to-last inequality we have utilized the upper bound on C_{15} given in (21). Thus, if we choose C_1 so that $C_1 \geq C$, then $H(x_1, y_1) \geq C_1 \geq C$, and so we arrive to a contradiction.

6 A Generalized Non-Archimedean Gap Principle

In this section we prove Theorem 1.2. Let P, Q be a minimal pair for (α, β) , and 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:

$$\begin{aligned}
\left| y_1^r y_2 R\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}\right) \right|_p &\geq \frac{1}{|y_1^r y_2 R(x_1/y_1, x_2/y_2)|} \\
&\geq \frac{1}{2(r+1) \max\{H(P), H(Q)\} H(x_1, y_1)^r H(x_2, y_2)} \\
&\geq \frac{1}{2((d/2)+1) C_{12} H(x_1, y_1)^{d/2} H(x_2, y_2)}.
\end{aligned}$$

Let c_α and c_β denote the leading coefficients of the minimal polynomials of α and β , respectively. Note that for each $(i, j) \in \{0, \dots, r\} \times \{0, 1\}$ 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

$$\frac{1}{2((d/2)+1) C_{12} H(x_1, y_1)^{d/2} H(x_2, y_2)} < \frac{C_0 c_\alpha^r c_\beta}{H(x_1, y_1)^\mu}.$$

If we now set $C_4 = (d+2) C_0 C_{12} c_\alpha^{d/2} c_\beta$, then

$$H(x_2, y_2) > C_4^{-1} H(x_1, y_1)^{\mu-d/2},$$

which means that 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 $R(\alpha, \beta) = 0$ and $R(x_1/y_1, x_2/y_2) = 0$, case 2 holds.

It remains to consider the case when $R(x_1/y_1, x_2/y_2) = 0$ and $r \geq 2$. We will prove that $H(x_1, y_1) < C$ for some positive number C , which depends only on α and β . By choosing C_3 so that $C_3 \geq C$, we then arrive to a contradiction.

First, we claim that by choosing C_3 so that

$$C_3 \geq C_0^{1/\mu}$$

we can ensure that $|y_1|_p \geq c_\alpha^{-1}$. This inequality clearly holds when p does not divide y_1 , so we assume that $p \mid y_1$. Since x_1 and y_1 are coprime, it must be the case that p does not divide x_1 . Suppose that $|y_1\alpha|_p \neq |x_1|_p$. Then it follows from the strong triangle inequality that

$$|y_1\alpha - x_1|_p = \max\{|y_1\alpha|_p, |x_1|_p\} = \max\{|y_1\alpha|_p, 1\} \geq 1.$$

Since $|y_1\alpha - x_1|_p < C_0 H(x_1, y_1)^{-\mu}$, we find that $H(x_1, y_1) < C_0^{1/\mu} \leq C_3$, so we reach a contradiction. Thus, $|y_1\alpha|_p = |x_1|_p = 1$. Since $c_\alpha\alpha$ is an algebraic integer, it must be the case that $|c_\alpha\alpha|_p \leq 1$, so

$$|y_1|_p = |\alpha^{-1}|_p \geq |c_\alpha|_p \geq c_\alpha^{-1},$$

as claimed.

Now that we have chosen C_3 so that $|y_1|_p \geq c_\alpha^{-1}$, we turn our attention to the equation $R(x_1/y_1, x_2/y_2) = 0$, which implies that

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

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

$$|y_2\beta - x_2|_p = \frac{|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p}{|gQ(\alpha)|_p}. \quad (22)$$

Since g is an integer, we have

$$|g|_p \leq 1. \quad (23)$$

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

$$|Q(\alpha)|_p \leq |c_\alpha|_p^{-r} \leq c_\alpha^r \leq c_\alpha^{d/2}. \quad (24)$$

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,

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

Now, let $W = PQ' - QP'$ denote the Wronskian of P and Q . By Taylor's Theorem and (12),

$$\begin{aligned}
& |P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p \\
&= \left| P(\alpha) \sum_{i=0}^r D_i Q(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} - Q(\alpha) \sum_{i=0}^r D_i P(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} \right|_p \\
&= |y_1 \alpha - x_1|_p \left| \sum_{i=0}^{r-1} (P(\alpha) D_{i+1} Q(\alpha) - Q(\alpha) D_{i+1} P(\alpha)) (x_1 - \alpha y_1)^i y_1^{r-1-i} \right|_p \\
&\geq |y_1 \alpha - x_1|_p \left(|W(\alpha) y_1^{r-1}|_p - |y_1 \alpha - x_1|_p \max_{i=0, \dots, r-1} \left\{ \left| \frac{P(\alpha) D_{i+1} Q(\alpha) - Q(\alpha) D_{i+1} P(\alpha)}{c_\alpha^i} \right|_p \right\} \right) \\
&> |y_1 \alpha - x_1|_p \left(|W(\alpha) y_1^{r-1}|_p - \frac{C_0}{H(x_1, y_1)^\mu} c_\alpha^{2r-1} \right) \\
&\geq |y_1 \alpha - x_1|_p \left(C_{14} c_\alpha^{-(r-1)} - C_3^{-\mu} C_0 c_\alpha^{2r-1} \right),
\end{aligned}$$

where the last inequality follows from $H(x_1, y_1) \geq C_3$, (12) and $|y_1|_p \geq c_\alpha^{-1}$. Thus, if we choose C_3 so that

$$C_3^\mu \geq 2C_0 C_{14}^{-1} c_\alpha^{(3d-4)/2},$$

then

$$|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p > |y_1 \alpha - x_1|_p \left(C_{14} c_\alpha^{-(r-1)} - C_3^{-\mu} C_0 c_\alpha^{2r-1} \right) \geq \frac{C_{14}}{2c_\alpha^{(d-2)/2}} |y_1 \alpha - x_1|_p.$$

Combining this observation with (22), (23) and (24),

$$|y_2 \beta - x_2|_p = \frac{|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p}{|gQ(\alpha)|_p} > \frac{C_{14}}{2c_\alpha^{d-1}} |y_1 \alpha - x_1|_p.$$

By Proposition 4.1,

$$H(x_2, y_2) \geq \frac{H(x_1, y_1)^r}{C_{15} \max\{H(P), H(Q)\}^{2r^2+3r}} \geq \frac{H(x_1, y_1)^r}{C_{15} C_{12}^{2r^2+3r}},$$

where C_{15} satisfies the upper bound (21). Consequently,

$$|y_1 \alpha - x_1|_p < \frac{2c_\alpha^{d-1}}{C_{14}} |y_2 \beta - x_2|_p < \frac{2c_\alpha^{d-1} C_0}{C_{14} H(x_2, y_2)^\mu} \leq \frac{2c_\alpha^{d-1} C_0 C_{12}^{(2r^2+3r)\mu} C_{15}^\mu}{C_{14} H(x_1, y_1)^{r\mu}}$$

Thus, we obtain an upper bound on $|y_1 \alpha - x_1|_p$. On the other hand, by Lemma 2.2 we have the lower bound (5). Combining upper and lower bounds,

$$\frac{C_7}{H(x_1, y_1)^d} \leq |y_1 \alpha - x_1|_p < \frac{2c_\alpha^{d-1} C_0 C_{12}^{(2r^2+3r)\mu} C_{15}^\mu}{C_{14} H(x_1, y_1)^{r\mu}}$$

Since $\mu > (d/2) + 1$ and $r \geq 2$, we have

$$\begin{aligned}
H(x_1, y_1) &< \left(2c_\alpha^{d-1} C_0 C_7^{-1} C_{12}^{(2r^2+3r)\mu} C_{14}^{-1} C_{15}^\mu \right)^{1/(r\mu-d)} \\
&\leq \left(2^{d^2\mu/4} ((d/2) + 1)^{(3d^2+4d)\mu/8} c_\alpha^{d-1} C_0 C_7^{-1} C_{12}^{(d^2+3d)\mu/2} C_{14}^{-1} \right)^{1/(2\mu-d)} \\
&= C.
\end{aligned}$$

Notice how in the second-to-last inequality we have utilized the upper bound on C_{15} given in (21). Thus, if we choose C_3 so that $C_3 \geq C$, then $H(x_1, y_1) \geq C_3 \geq C$, and so we arrive to a contradiction.

Group	Generators	Group	Generators
C_1	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	D_1	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
C_2	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	D_2	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
C_3	$\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$	D_3	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$
C_4	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	D_4	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
C_6	$\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$	D_6	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$

Table 1: Representatives of equivalence classes of finite subgroups of $\mathrm{GL}_2(\mathbb{Q})$ under $\mathrm{GL}_2(\mathbb{Q})$ -conjugation.

7 The Enhanced Automorphism Group

In this section we establish several results about the enhanced automorphism group $\mathrm{Aut}'|F|$ of a binary form F . At the end we prove Proposition 7.3, where we explain the relation between automorphisms of F and the roots of $F(x, 1)$.

Lemma 7.1. *Let $F \in \mathbb{Z}[x, y]$ be a binary form of degree $d \geq 3$ and nonzero discriminant $D(F)$. Then $\mathrm{Aut}_{\mathbb{Q}}|F|$ is $\mathrm{GL}_2(\mathbb{Q})$ -conjugate to one of the groups from Table 1.*

Proof. See [11]. □

Lemma 7.2. *Let $F \in \mathbb{Z}[x, y]$ be a binary form of degree $d \geq 3$ and nonzero discriminant $D(F)$. Let $\mathrm{Aut}'|F|$ be as in (3). Then $\mathrm{Aut}'|F| \cong C_n$ or $\mathrm{Aut}'|F| \cong D_n$, 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 7.1, $\mathrm{Aut}_{\mathbb{Q}}|F|$ 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 $\mathrm{GL}_2(\mathbb{R})$, including $\mathrm{Aut}'|F|$, is isomorphic to either a cyclic group C_n of order n or a dihedral group D_n of order $2n$.

Now suppose that $\mathrm{Aut}'|F|$ contains at least 25 distinct elements M_1, \dots, M_{25} . By Schur's Theorem [6], any finitely generated torsion subgroup of $\mathrm{GL}_n(\mathbb{C})$ is finite. Hence $\langle M_1, \dots, M_{25} \rangle$ is a finite subgroup of $\mathrm{GL}_2(\mathbb{R})$, so it is isomorphic to either C_n or 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. □

Let us give an example of a group of the form (3) that is not a subgroup of $\text{GL}_2(\mathbb{Q})$. 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 $G \cong \mathbf{D}_{12}$. If we choose coprime integers a and 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 $F_M = F$ for any $M \in G$. Consequently, if (x, y) is a solution to the Thue equation $F(x, y) = m$, 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 [14, Section 6] with respect to binary forms invariant under \mathbf{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 7.3. *Let $F(x, y) = c_d x^d + c_{d-1} x^{d-1} y + \cdots + c_0 y^d \in \mathbb{Z}[x, y]$ be an irreducible binary form of degree $d \geq 3$. Let $\alpha_1, \dots, \alpha_d$ be the roots of $F(x, 1)$. There exists an index $j \in \{1, \dots, d\}$ such that*

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

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

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

is an element of $\text{Aut}'|F|$. Furthermore, if $M \in \text{Aut}'|F|$, then $|sv - tu| = \left| \frac{F(s, t)}{c_d} \right|^{2/d}$.

Proof. Suppose that there exists an index $j \in \{1, \dots, d\}$ such that $\alpha_j = \frac{v\alpha_1 - u}{-t\alpha_1 + s}$ for some integers s, t, u and v . Since $F(x, 1)$ is irreducible, its Galois group acts transitively on the roots $\alpha_1, \alpha_2, \dots, \alpha_d$. Therefore,

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

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

$$\begin{aligned}
F(x, y) &= c_d \prod_{i=1}^d \left(x - \frac{v\alpha_i - u}{-t\alpha_i + s} y \right) \\
&= \frac{c_d}{\prod_{i=1}^d (s - t\alpha_i)} \prod_{i=1}^d ((-t\alpha_i + s)x - (v\alpha_i - u)y) \\
&= \frac{c_d}{F(s, t)} F(sx + uy, tx + vy) \\
&= \pm \eta^d F(sx + uy, tx + vy) \\
&= \pm F_M(x, y),
\end{aligned}$$

where $\eta = |c_d/F(s, t)|^{1/d}$ and $M = \eta \begin{pmatrix} s & u \\ t & v \end{pmatrix}$. Since

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

and $F_M = \pm F$, we see that $|D(F_M)| = |D(F)|$, so $|\det M| = 1$. Hence $|\eta|^2 \cdot |sv - tu| = 1$, which leads us to the conclusion that $\eta = |\eta| = |sv - tu|^{-1/2}$ and $M \in \text{Aut}' |F|$.

Conversely, suppose that $M = |sv - tu|^{-1/2} \begin{pmatrix} s & u \\ t & v \end{pmatrix}$ is in $\text{Aut}' |F|$. Then

$$\begin{aligned}
\pm F(x, y) &= F_M(x, y) \\
&= \frac{c_d}{|sv - tu|^{d/2}} \prod_{i=1}^d (sx + uy - \alpha_i(tx + vy)) \\
&= \frac{F(s, t)}{|sv - tu|^{d/2}} \prod_{i=1}^d \left(x - \frac{v\alpha_i - u}{-t\alpha_i + s} y \right).
\end{aligned}$$

We see that the polynomial $F_M(x, 1)$ vanishes at $\frac{v\alpha_i - u}{-t\alpha_i + s}$ for $i = 1, \dots, d$. Since $F_M = \pm F$, the polynomials $F_M(x, 1)$ and $F(x, 1)$ have the same roots, so there exists some index j such that $\alpha_j = \frac{v\alpha_1 - u}{-t\alpha_1 + s}$. Furthermore, the leading coefficients of $F(x, 1)$ and $F_M(x, 1)$ must equal up to a sign, i.e., $|c_d| = \frac{|F(s, t)|}{|sv - tu|^{d/2}}$. Of course, this is the same as $|sv - tu| = \left| \frac{F(s, t)}{c_d} \right|^{2/d}$. \square

8 Counting Primitive Solutions of Large Height to Certain Thue Inequalities

In this section we prove Theorem 1.3. It follows from a more general result stated in Theorem 8.1, where we count approximations x/y of large height to distinct algebraic numbers $\alpha_1, \dots, \alpha_n$ such that $\mathbb{Q}(\alpha_i) = \mathbb{Q}(\alpha_1)$ for all $i = 1, 2, \dots, n$. In order to state the main result of this section we need to introduce the notion of an *orbit*. For an irrational number α , the *orbit* of α is the set

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

Theorem 8.1. *Let $K = \mathbb{C}$ or \mathbb{Q}_p , where p is a rational prime, and denote the standard absolute value on K by $|\cdot|$. Let $\alpha_1 \in K$ be an algebraic number of degree $d \geq 3$ over \mathbb{Q}*

and $\alpha_2, \alpha_3, \dots, \alpha_n$ be distinct elements of $\mathbb{Q}(\alpha_1)$, different from α_1 , each of degree d . Let μ be such that $(d/2) + 1 < \mu < d$. Let C_0 be a real number such that $C_0 > (4e^A)^{-1}$, where

$$A = 500^2 \left(\log \max_{i=1, \dots, n} \{M(\alpha_i)\} + \frac{d}{2} \right). \quad (25)$$

There exists a positive real number C_{16} , which depends on $\alpha_1, \alpha_2, \dots, \alpha_n, \mu$ and C_0 , with the following property. The total number of rationals x/y in lowest terms, which satisfy $H(x, y) \geq C_{16}$ and

$$\left| \alpha_j - \frac{x}{y} \right| < \frac{C_0}{H(x, y)^\mu} \quad (26)$$

for some $j \in \{1, 2, \dots, n\}$ is less than

$$\gamma \left[1 + \frac{11.51 + 1.5 \log d + \log \mu}{\log(\mu - d/2)} \right],$$

where

$$\gamma = \max\{\gamma_1, \dots, \gamma_n\}, \quad \gamma_i = \#\{j : \alpha_j \in \text{orb}(\alpha_i)\}. \quad (27)$$

Let us see why Theorem 1.3 follows from Theorem 8.1.

Proof of Theorem 1.3. Let $\alpha_1, \alpha_2, \dots, \alpha_d$ be the roots of $F(x, 1)$. Notice that, since $F(x, y)$ is irreducible, the roots of $F(1, x)$ are given by $\alpha_1^{-1}, \dots, \alpha_d^{-1}$. Furthermore, since the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is Galois, we have $\mathbb{Q}(\alpha_i) = \mathbb{Q}(\alpha_1)$ for all $i = 1, \dots, d$.

Choose C_5 so that

$$C_5^{d-\mu} > \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2} m}{|D(F)|^{1/2}}.$$

Let (x, y) be a primitive solution to (2) such that $H(x, y) \geq C_5$. Then it follows from the result of Lewis and Mahler stated in Lemma 2.6 that there exists an index $j \in \{1, 2, \dots, d\}$ such that

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

Next, adjust the choice of C_5 so that Theorem 8.1 applies:

$$C_5 \geq \max \{C_{16}(\alpha_1, \dots, \alpha_d, \mu, C_0), C_{16}(\alpha_1^{-1}, \dots, \alpha_d^{-1}, \mu, C_0)\},$$

where $C_0 = 1$. If we let γ be as in (27), then it follows from Theorem 8.1 that x/y is one of at most

$$2\gamma \left[1 + \frac{11.51 + 1.5 \log d + \log \mu}{\log(\mu - d/2)} \right]$$

rationals in lowest terms that satisfy either of the two inequalities

$$\left| \alpha_j - \frac{x}{y} \right| < \frac{C_0}{H(x, y)^\mu}, \quad \left| \alpha_j^{-1} - \frac{y}{x} \right| < \frac{C_0}{H(x, y)^\mu}.$$

It now follows from Proposition 7.3 that $\gamma \leq \frac{\text{Aut}'[F]}{2}$. The division by 2 appears due to the presence of the matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ in $\text{Aut}'[F]$, which maps (x, y) to $(-x, -y)$. \square

We conclude this section with the proof of Theorem 8.1.

Proof of Theorem 8.1. Throughout the proof we will be adjusting our choice of C_{16} four times. First, let $C_{16} \geq C_{11}$, where the positive real number C_{11} is defined in Corollary 2.8. Then it follows from Lemma 2.8 that for each x/y satisfying (26) the index $j \in \{1, 2, \dots, n\}$ is unique.

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

1. $C_{16} \leq H(x_1, y_1) \leq H(x_2, y_2) \leq \dots \leq H(x_\ell, y_\ell)$.
2. $\gcd(x_j, y_j) = 1$ for all $j = 1, 2, \dots, \ell$.
3. For each $j \in \{1, 2, \dots, \ell\}$, there exists the index $i_j \in \{1, 2, \dots, n\}$ 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 $j, k \in \{1, 2, \dots, \ell\}$, if $\alpha_{i_k} \in \text{orb}(\alpha_{i_j})$, i.e.,

$$\alpha_{i_k} = \frac{s\alpha_{i_j} + t}{u\alpha_{i_j} + v}$$

for some integers s, t, u and v , then

$$\frac{x_k}{y_k} \neq \frac{sx_j + ty_j}{ux_j + vy_j}.$$

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 rational approximation in the list

$$\frac{x_1}{y_1}, \quad \frac{x_2}{y_2}, \quad \dots, \quad \frac{x_\ell}{y_\ell}$$

correspond several rational 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 $\alpha \in \text{orb}(\alpha_{i_j})$ 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 and v . Then

$$\alpha - \frac{x'_j}{y'_j} = \frac{tu - sv}{(u\alpha_{i_j} + v)(u(x_j/y_j) + v)} \left(\alpha_{i_j} - \frac{x_j}{y_j} \right),$$

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 $x_1/y_1, x_2/y_2, \dots, x_\ell/y_\ell$.

In order to account for the presence of derived rational approximations, we introduce the value γ_i defined in (27). 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,

$$N \leq \sum_{j=1}^\ell \gamma_{i_j} \leq \gamma^\ell,$$

where γ is defined in (27). 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 $(\alpha_{i_k}, \alpha_{i_{k+1}})$. Choose C_{16} and define C as follows:

$$C_{16} \geq \max_{j,k} \{C_1(\alpha_j, \alpha_k, \mu, C_0), C_3(\alpha_j, \alpha_k, \mu, C_0)\},$$

$$C = \max_{j,k} \{C_2(\alpha_j, \alpha_k, \mu, C_0), C_4(\alpha_j, \alpha_k, \mu, C_0)\},$$

where the positive real numbers C_1, C_2, C_3 and C_4 are taken from Theorems 1.1 and 1.2, respectively. Note that if $K = \mathbb{Q}_p$, then $|y_k| \leq 1$, and so

$$|y_k \alpha_{i_k} - x_k| = |y_k| \cdot \left| \alpha_{i_k} - \frac{x_k}{y_k} \right| < \frac{C_0}{H(x_k, y_k)^\mu}.$$

Analogously,

$$|y_{k+1} \alpha_{i_{k+1}} - x_{k+1}| < \frac{C_0}{H(x_{k+1}, y_{k+1})^\mu}.$$

It follows from Theorems 1.1 and 1.2 that, for every $k \in \{1, 2, \dots, \ell - 1\}$,

$$H(x_{k+1}, y_{k+1}) > C^{-1} H(x_k, y_k)^E,$$

where

$$E = \mu - d/2. \quad (28)$$

Notice that case 2 in the aforementioned theorems is impossible due to the fact that the list $x_1/y_1, \dots, x_\ell/y_\ell$ does not contain derived rational approximations. Consequently,

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

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

$$\log H(x_\ell, y_\ell) > E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C. \quad (29)$$

Next, we apply the Thue-Siegel principle from Lemma 2.9 to the pair $(\alpha, \beta) = (\alpha_{i_1}, \alpha_{i_\ell})$. Observe that, since all α_i 's 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,

$$\begin{aligned} \frac{t^2}{2 - dt^2} &= \frac{1}{a^2} = 500^2, \\ A_1 &= 500^2 \left(\log M(\alpha_{i_1}) + \frac{d}{2} \right), \quad A_\ell = 500^2 \left(\log M(\alpha_{i_\ell}) + \frac{d}{2} \right), \end{aligned}$$

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

Note that

$$\delta^{-1} < 41667d^2. \quad (30)$$

We further adjust our definition of C_{16} by choosing it so that

$$C_{16} \geq C_0^{\frac{1}{\mu-1.42\sqrt{d}}} \left(4e^A\right)^{\frac{1.42\sqrt{d}}{\mu-1.42\sqrt{d}}}, \quad (31)$$

where A is defined in (25). Now with the help of inequalities $\lambda < 1.42\sqrt{d}$ and $H(x_j, y_j) \geq C_{16}$ we obtain

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

so that the hypothesis of Lemma 2.9 is satisfied. Thus, we arrive at the conclusion that

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

where the last inequality follows from (30). Thus

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

We combine the above upper bound on $\log H(x_\ell, y_\ell)$ with the lower bound given in (29):

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

Reordering the terms yields

$$\left(E^{\ell-1} - 41667d^2 \right) \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C < 41667d^2 \log(4e^{A_1}). \quad (32)$$

Let us assume that

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

for otherwise the statement of our theorem holds. Then $E^{\ell-1} \geq 41667d^2$, so we may use the inequality $H(x_1, y_1) \geq C_{16}$ to replace $H(x_1, y_1)$ with C_{16} in (32):

$$\left(E^{\ell-1} - 41667d^2 \right) \log C_{16} - \frac{E^{\ell-1} - 1}{E - 1} \log C < 41667d^2 \log(4e^{A_1}).$$

Since $E = \mu - d/2$,

$$(\mu - d/2)^{\ell-1} \left(\log C_{16} - \frac{\log C}{E - 1} \right) < 41667d^2 \log C_{16} + 41667d^2 \log(4e^{A_1}) + \frac{\log C}{E - 1}.$$

We make a final adjustment to C_{16} by choosing it so that

$$C_{16} \geq C^{2/(E-1)}. \quad (33)$$

Then

$$(\mu - 0.5d)^{\ell-1} \frac{\log C_{16}}{2} < \left(41667d^2 + \frac{1}{2}\right) \log C_{16} + 41667d^2 \log(4e^{A_1}),$$

leading us to a conclusion

$$(\mu - 0.5d)^{\ell-1} < 1 + 83334d^2 \left(1 + \frac{\log(4e^{A_1})}{\log C_{16}}\right). \quad (34)$$

By our choice of C_{16} ,

$$\log C_{16} \geq \frac{1}{\mu - 1.42\sqrt{d}} \log C_0 + \frac{1.42\sqrt{d}}{\mu - 1.42\sqrt{d}} \log(4e^A),$$

which means that

$$\frac{\log(4e^A)}{\log C_{16}} \leq \frac{\mu - 1.42\sqrt{d}}{1.42\sqrt{d} + \log C_0 / \log(4e^A)} < \frac{\mu - 1.42\sqrt{d}}{1.42\sqrt{d} - 1},$$

where the last inequality follows from the fact that $C_0 > (4e^A)^{-1}$. Plugging the above inequality into (34), 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

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

The result follows once we multiply the right-hand side by the constant γ defined in (27). \square

Acknowledgements

The author is grateful to Prof. Cameron L. Stewart for his wise supervision as well as to the anonymous referee for their excellent suggestions on how to improve the article.

References

- [1] E. Bombieri and W. Gubler. *Heights in Diophantine Geometry*. Cambridge University Press, 2006.
- [2] E. Bombieri and J. Mueller. On effective measures of irrationality for $\sqrt[r]{\frac{a}{b}}$ and related numbers. *J. Reine Angew. Math.*, 342:173–196, 1983.
- [3] E. Bombieri and W. M. Schmidt. On Thue’s equation. *Invent. Math.*, 88:69–81, 1987.
- [4] E. Bombieri and J. Vaaler. On Siegel’s lemma. *Invent. Math.*, 73:11–32, 1983.
- [5] Y. Bugeaud and M. Mignotte. On the distance between roots of integer polynomials. *Proc. Edinburgh Math. Soc.*, 47:553–556, 2004.

- [6] C. Curtis and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras*, pages 258–262. John Wiley & Sons, 1962.
- [7] W. Gautschi. Norm estimates for inverses of Vandermonde matrices. *Numer. Math.*, 23:337–347, 1975.
- [8] K. Györy. Thue inequalities with a small number of primitive solutions. *Period. Math. Hungar.*, 42(1-2):239–246, 2001.
- [9] D. Lewis and K. Mahler. Representation of integers by binary forms. *Acta Arith.*, 6:333–363, 1961.
- [10] J. Mueller and W. M. Schmidt. Thue’s equation and a conjecture of siegel. *Acta Math.*, 160:207–247, 1988.
- [11] M. Newman. *Integral Matrices*, volume 45 of *Pure and Appl. Math.* Academic Press, New York, 1972.
- [12] V. P. Prasolov. *Polynomials*, volume 11 of *Algorithms and Computation in Mathematics*. Springer-Verlag Berlin Heidelberg, 2004.
- [13] W. M. Schmidt. *Diophantine Approximations and Diophantine Equations*. Springer-Verlag, 1991.
- [14] C. L. Stewart. On the number of solutions of polynomial congruences and Thue equations. *J. Amer. Math. Soc.*, 4:793–835, 1991.
- [15] A. Thue. Über Annäherungswerte algebraischer Zahlen. *J. Reine Angew. Math.*, 135:284–305, 1909.
- [16] U. Zannier. *Lecture Notes on Diophantine Analysis*. Scuola Normale Superiore Pisa, 2014.