On the generalization of the gap principle

Anton Mosunov

Accepted: 27 June 2022 / Published online: 10 February 2023
© Akadémiai Kiadó, Budapest, Hungary 2023

Abstract

Let $\alpha$ be a real algebraic number of degree $d \geq 3$ and let $\beta \in \mathbb{Q}(\alpha)$ be irrational. Let $\mu$ 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 $\alpha$, $\beta$, $\mu$ 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

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

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

or both. Here $H(x, y) = \max(|x|, |y|)$ is the height of $x/y$. Since $\mu - d/2$ exceeds 1, our result demonstrates that, unless $\alpha$ and $\beta$ 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 $\alpha$ and $\beta$ 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 $\alpha$ 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 $\alpha$ by $x/y$ can be measured by means of a quantity $\mu$ such that the inequality

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

(1.1)

1 University of Waterloo, Waterloo, ON, Canada
is satisfied. The larger \( \mu \) is, the better the approximation of \( x/y \) with respect to \( \alpha \) 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 \( \alpha \) is real and irrational. On the other hand, Liouville pointed out that if \( \alpha \) is an irrational algebraic number of degree \( d \) and \( \mu > d \), then (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.

It is not difficult to count distinct \( x/y \) satisfying (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

\[
\frac{1}{y_1y_2} \leq \left| \frac{x_1}{y_1} - \frac{x_2}{y_2} \right| \leq \left| \alpha - \frac{x_1}{y_1} \right| + \left| \alpha - \frac{x_2}{y_2} \right| < \frac{1}{y_1^\mu} + \frac{1}{y_2^\mu} < \frac{2}{y_1^\mu},
\]

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), 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 (1.1). However, it was established by Thue [16] that when \( \alpha \) 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 \( \alpha \) and \( \mu \), such that every solution \( x_i/y_i \) with \( C_1 \leq y_1 < \cdots < y_e \) satisfies \( y_i < y_1^{\eta} \). This phenomenon is known as the Thue–Siegel principle and it was vastly generalized by Bombieri [1], with further improvements made by Bombieri and Mueller [3]. When combined with the gap principle, the Thue–Siegel principle enables us to count the number of solutions \( x/y \) to (1.1) 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 \( C_1, C_2, \ldots \) occurring throughout the article are all computable.

**Theorem 1.1** (a generalized Archimedean gap principle) Let \( \alpha \) be a real algebraic number of degree \( d \geq 3 \) over \( \mathbb{Q} \) and let \( \beta \) be irrational and in \( \mathbb{Q}(\alpha) \). Let \( \mu \) 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 \( \alpha, \beta, \mu, \) 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

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

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 1, our result demonstrates that, unless \( \alpha \) and \( \beta \) 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 \(|x|_p\) 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 \( \beta \) be irrational and in \( \mathbb{Q}(\alpha) \). Let \( \mu \) 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 \( \alpha, \beta, \mu \) 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

\[
0 < |F(x, y)| \leq m,
\]

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 [9, Theorem 1], Győry 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 (1.2) 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 Győry’s result in the case when \( F \) is irreducible and the field extension \( \mathbb{Q}(\alpha)/\mathbb{Q} \) is Galois, where \( \alpha \) 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 \( 2 \times 2 \) 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

\[
\text{Aut}' |F| = \left\{ \frac{1}{\sqrt{|sv - tu|}} \begin{pmatrix} s & u \\ t & v \end{pmatrix} : s, t, u, v \in \mathbb{Z} \right\} \cap \text{Aut}_{\overline{\mathbb{Q}}} |F|\quad (1.3)
\]

and refer to it as the enhanced automorphism group of \( F \). One can verify that \( \text{Aut}' |F| \) is a group. In Sect. 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 \( \alpha \) 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 (1.2). Let \( \mu \) 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 \( \mu \), such that the number of primitive solutions \((x, y)\) to (1.2) with \( H(x, y) \geq C_5 \) does not exceed

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

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 (1.2) 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 \( \lim_{d \to \infty} f(d) = 3.5 \), 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 [11] (see equation 2.9) or to the quantity \( Y_0 \) in [9, Theorem 1], it lies outside the scope of this article.

The article is structured as follows. In Sect. 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 Sect. 3 we introduce the notion of a minimal pair \( P, Q \in \mathbb{Z}[x] \) for a tuple of algebraic numbers \((\alpha, \beta) \in \mathbb{Q} \times (\mathbb{Q}(\alpha) \setminus \mathbb{Q})\), 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 \((\alpha, \beta)\). 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 Sect. 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 \( \alpha \) and \( \beta \) are not connected by means of a linear fractional transformation with integer coefficients. In Sects. 5 and 6 we prove Theorems 1.1 and 1.2, respectively. In Sect. 7 we investigate the properties of the enhanced automorphism group \( \text{Aut}^+ |F| \). Finally, in Sect. 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 a number of definitions. 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 the height of \( R \). For an algebraic number \( \alpha \) with minimal polynomial \( f \), we write \( H(\alpha) = H(f) \). For a point \((x_1, x_2, \ldots, x_n) \in \mathbb{C}^n \), 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 \((x_1, x_2, \ldots, x_n)\).

In this section, as well as in all 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 \( \alpha \), such that for all integers \( x \) and...
y, with \( y \neq 0 \) and \( x \neq y \alpha \), the inequality
\[
|\alpha - \frac{x}{y}| \geq \frac{C_6}{H(x, y)^d}
\]
holds.

**Proof** See [14, Theorem 1E]. \( \square \)

**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 \( \alpha \), such that for all integers \( x \) and \( y \), with \( x \neq y \alpha \), the inequality
\[
|y \alpha - x|_p \geq \frac{C_7}{H(x, y)^d}
\]
holds.

**Proof** Let
\[
f(x) = c_d x^d + \cdots + c_1 x + c_0
\]
be the minimal polynomial of \( \alpha \) 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
\[
F(x, y) = (x - \alpha y) \sum_{i=1}^d D_i f(\alpha)(x - \alpha y)^{i-1} y^{d-i}
\]
\[
= (x - \alpha y) \sum_{i=1}^d \frac{D_i f(\alpha)}{c_d^{i-1}} (c_d x - c_d \alpha y)^{i-1} y^{d-i}.
\]
Since \( c_d \alpha \) and \( c_d^{d-i} D_i f(\alpha) \) are algebraic integers, their \( p \)-adic absolute values do not exceed 1, so
\[
|F(x, y)|_p \leq |y \alpha - x|_p \cdot \max_{i=1, \ldots, d} \left\{ \frac{|D_i f(\alpha)|}{c_d^{i-1}} \right\}
\]
\[
= |y \alpha - x|_p \cdot \max_{i=1, \ldots, d} \left\{ \frac{c_d^{d-i} D_i f(\alpha)}{c_d^{d-1}} \right\}
\]
\[
\leq |y \alpha - x|_p \cdot |c_d|_p^{-d+1}.
\]
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, [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, [17, Lemma 2.7]. \qed

Lemma 2.4 (see [17, Proposition 2.6]) Let $\alpha$ be an algebraic number of degree $d$ over $\mathbb{Q}$. Then for every nonnegative 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 $c_\alpha$ denotes the leading coefficient of the minimal polynomial of $\alpha$ and we 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 $0 \leq i \leq d - 1$.

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

$$[P] = \max \{|\alpha_1|, \ldots, |\alpha_d|\},$$

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

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

$$\theta_\alpha = [\mathcal{O}: \mathbb{Z}[c_\alpha \alpha]]. \quad (2.3)$$

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

Lemma 2.5 Let $\alpha$ 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 $b_0, b_1, \ldots, b_{d-1} \in \mathbb{Q}$. There exists a positive number $C_9$, which depends only on $\alpha$ and $\beta$, 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 $\theta_\alpha$ is defined in (2.3). 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 $\alpha$. For $j = 1, \ldots, d$, let

$$\beta_j = \sum_{i=0}^{d-1} b_i \alpha_j^i.$$
Then each $\beta_j$ is a conjugate of $\beta = \beta_1$. Further,

\[
\begin{pmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_d
\end{pmatrix} =
\begin{pmatrix}
1 & \alpha_1^2 & \ldots & \alpha_1^{d-1} \\
1 & \alpha_2^2 & \ldots & \alpha_2^{d-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \alpha_d^2 & \ldots & \alpha_d^{d-1}
\end{pmatrix}
\begin{pmatrix}
b_0 \\
b_1 \\
\vdots \\
b_{d-1}
\end{pmatrix}.
\] (2.4)

Let $V$ denote the Vandermonde matrix on the right-hand side of the above expression. Then it follows from (4.1) in [8] that

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

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

Now, let $V^{-1} = (v_{ij})$. Then it follows from (2.4) that $b_i = \sum_{j=1}^{d} v_{ij} \beta_j$, so

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

Next, let $c_\alpha$ and $c_\beta$ denote the leading coefficients of the minimal polynomials of $\alpha$ and $\beta$, respectively. Note that $\theta_\alpha c_\beta \beta = (\#O/\mathbb{Z}[c_\alpha \alpha]) c_\beta \beta \in \mathbb{Z}[c_\alpha \alpha]$ due to the fact that $c_\beta \beta \in O$. Finally, observe that

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

Hence 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 $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ 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 $\alpha$, we define the Mahler measure of $\alpha$ to be $M(\alpha) = M(f)$, where $f$ is the minimal polynomial of $\alpha$.

The following lemma is a reformulation of a well-known result of Lewis and Mahler [10].

**Lemma 2.6** 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 $x_1$ and $y_1$ be nonzero integers. There exists a root $\alpha$ of $F(x, 1)$ such that

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

\(\square\)
where

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

**Proof** Let \( \alpha \) be a root of \( F(x, 1) \) that minimizes \( |\alpha - x/y| \). By [15, Lemma 3],

\[ |\alpha - \frac{x_1}{y_1}| \leq \frac{C_{10}|F(x_1, y_1)|}{|y_1|^d}. \]

If \(|y_1| \geq |x_1|\), then \( H(x_1, y_1) = |y_1| \), and so the result holds. Otherwise, since \( c_0c_d \neq 0 \), 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 \( \alpha^{-1} \), where \( \alpha \) is a root of \( F(x, 1) \). If we let \( \beta^{-1} \) be a root of \( F(1, x) \) that minimizes \( |\beta^{-1} - y_1/x_1| \), then it follows from [15, Lemma 3] that

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

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 \( \alpha \) and \( \beta \) be distinct numbers in \( \overline{K} \). Let \( \mu \) 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

\[ |\alpha - \frac{x_1}{y_1}| < \frac{C_0}{H(x_1, y_1)^\mu}, \]

then

\[ |\beta - \frac{x_1}{y_1}| \geq \frac{C_0}{H(x_1, y_1)^\mu}. \]

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

\[ |\alpha - \beta| \leq |\alpha - \frac{x_1}{y_1}| + |\beta - \frac{x_1}{y_1}| \leq \frac{2C_0}{H(x_1, y_1)^\mu}, \]

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 \( \alpha_1, \ldots, \alpha_d \in \overline{K} \). Let \( \mu \) and \( C_0 \) be positive real numbers. There exists a positive number \( C_{11} \), which depends only on \( f \), \( \mu \), 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

\[ |\alpha_i - \frac{x_1}{y_1}| < \frac{C_0}{H(x_1, y_1)^\mu} \]

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

\[ |\alpha_j - \frac{x_1}{y_1}| \geq \frac{C_0}{H(x_1, y_1)^\mu} \]

for all \( j \neq i \).
Lemma 2.9 (The Thue–Siegel Principle [3]) 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 $\tau$ be such that

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

(2.5)

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

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

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

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

Then

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

where

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

Proof Since $d \geq 3$, the intervals in (2.5) are guaranteed to be nonempty, so the statement is not vacuously true. Since

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

the case when $|\alpha_1| \leq 1$ and $|\alpha_2| \leq 1$ follows directly from [3, Section II]. More precisely, the comment on p. 184 of [3] states that the triple $(A_1, A_2, \tau)$ 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 [4] apply in our situation:

(i) the hypothesis $K_\varnothing = k_\varnothing$ 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 [3] 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

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

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

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

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

$$|x_i\alpha_i^{-1} - y_i| = \max \left\{|x_i\alpha_i^{-1}|, |y_i|\right\} = \max \left\{|x_i\alpha_i^{-1}|, 1\right\} \geq 1.$$
Thus,
\[ 1 \leq |x_i \alpha_i^{-1} - y_i| = |y_i ||\alpha_i^{-1} - \frac{x_i}{y_i}| < \left( 4e^{A_i} H(x_i, y_i) \right)^{-\lambda}, \]
which is impossible. Hence \(|x_i| = 1\), so
\[ \left| \alpha_i - \frac{y_i}{x_i} \right| < \left( 4e^{A_i} H(x_i, y_i) \right)^{-\lambda}, \]
We conclude that, as long as \( \left| \alpha_i - \frac{y_i}{x_i} \right| < \left( 4e^{A_i} H(x_i, y_i) \right)^{-\lambda} \) for \( i = 1, 2 \), the inequalities
\[ \left| \alpha_i - \frac{x_i}{y_i} \right| < \frac{1}{\left( 3e^{A_i} H(x_i, y_i) \right)^{\lambda}} \quad \text{and} \quad \left| \alpha_i^{-1} - \frac{y_i}{x_i} \right| < \frac{1}{\left( 3e^{A_i} H(x_i, y_i) \right)^{\lambda}} \]
hold whenever \( |\alpha_i| > 1 \) for some \( i \in \{1, 2\} \). Consequently, we can always choose \( r, s \in \{-1, 1\} \) so that \( |\alpha_r^r| \leq 1, |\alpha_s^s| \leq 1 \) and \((A_1, A_2, \tau)\) is admissible for the data \((\alpha_1^r, \alpha_2^s, (x_1/y_1)^r, (x_2/y_2)^s, t, \theta, \delta)\). The result now follows from [3, Section II]. \( \Box \)

3 Minimal pairs

Let \( \alpha \) be an algebraic number of degree \( d \) over \( \mathbb{Q} \) and let \( \beta \in \mathbb{Q}(\alpha) \) be irrational. With a pair \((\alpha, \beta)\) 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 Sects. 5 and 6, respectively.

Definition 3.1 Let \( \alpha \) 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 \((\alpha, \beta)\) 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 \((\alpha, \beta)\), we write
\[ r(\alpha, \beta) = \max\{ \deg P, \deg Q \}. \]

If \( P, Q \) is a minimal pair for \((\alpha, \beta)\) then \(-P, -Q\) is also a minimal pair for \((\alpha, \beta)\). 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 1. 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 \((\alpha, \beta)\).

If \(P, Q\) is a minimal pair for \((\alpha, \beta)\), then we can define a polynomial
\[
R(x, y) = P(x) + yQ(x).
\]
Polynomials of such form were used by Thue [16] for the purpose of establishing the first instance of the Thue–Siegel principle [3]. More precisely, they were constructed so as to achieve high vanishing at the point \((\alpha, \alpha)\), i.e.,
\[
\text{Di}_{R}(\alpha, \alpha) = 0 \quad \text{for some large} \quad \ell
\]
(see the exposition of Thue’s method in [17, Chapter 2]). In turn, we construct \(R(x, y)\) so that \(R(\alpha, \beta) = 0\) 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 \(\alpha\) 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 \((\alpha, \beta)\) and put \(r = r(\alpha, \beta)\). Then the polynomials \(P, Q,\) and their Wronskian \(W = PQ' - QP'\) possess the following properties.

1. \(1 \leq r \leq \lfloor d/2 \rfloor\). \hfill (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. There exists a positive number \(C_{12}\), which depends only on \(\alpha\) and \(\beta\), such that
\[
\max\{H(P), H(Q)\} \leq C_{12}. \hfill (3.2)
\]

5. If \(\alpha \in \mathbb{C}\), then there exists a positive real number \(C_{13}\), which depends only on \(\alpha\) and \(\beta\), such that
\[
|W(\alpha)| \geq C_{13}. \hfill (3.3)
\]
Similarly, if \(\alpha \in \mathbb{Q}_p\) for some rational prime \(p\), then there exists a positive real number \(C_{14}\), which depends only on \(\alpha\) and \(\beta\), such that
\[
|W(\alpha)|_p \geq C_{14}. \hfill (3.4)
\]

**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 \(\beta\) is irrational. If \(q = 0\), then \(p = 0\), which is impossible, since we assumed that both \(p\) and \(q\) are different from zero. Thus, \(r \geq 1\).

Next, we prove that \(r \leq s\), where \(s = \lfloor d/2 \rfloor\). Write
\[
\hat{P}(x) = \sum_{i=0}^{s} a_i x^i \quad \text{and} \quad \hat{Q}(x) = \sum_{i=0}^{s} a_{s+1+i} x^i. \hfill (3.5)
\]
We view the $2s + 2$ integer coefficients $a_0, \ldots, a_{2s+1}$ as variables. Since $\alpha$ 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 nontrivial 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 $\alpha$ 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 $\alpha$ has degree $d$ and
\[
\deg (P \hat{Q} - Q \hat{P}) \leq \max\{\deg P, \deg Q\} + \max\{\deg \hat{P}, \deg \hat{Q}\} \leq r + (d - 1 - r) < d,
\]
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:
\[
\alpha^k = c_{k,d-1}\alpha^{d-1} + \cdots + c_{k,1}\alpha + c_{k,0},
\]
\[
\beta = b_{d-1}\alpha^{d-1} + \cdots + b_1\alpha + b_0.
\]
Let $\hat{P}, \hat{Q}$ be as in (3.5). Then,
\[
\hat{P}(\alpha) + \beta \hat{Q}(\alpha) = \sum_{i=0}^{s} a_i \alpha^i + \left( \sum_{i=0}^{d-1} b_i \alpha^i \right) \cdot \left( \sum_{j=0}^{s} a_{s+1+j} \alpha^j \right)
\]
\[
= \sum_{i=0}^{s} a_i \alpha^i + \sum_{j=0}^{s} a_{s+1+j} \sum_{i=0}^{d-1} b_i \alpha^{i+j}
\]
\[
= \sum_{i=0}^{s} a_i \alpha^i + \sum_{j=0}^{s} a_{s+1+j} \left( \sum_{i=j}^{d-1} b_i \alpha^{i-j} + \sum_{k=d}^{d-1+j} b_{k-j} \alpha^k \right).
\]
On the generalization of the gap principle 131

Put

By Lemma 2.3,

Then we can bound the (rational) coefficients of \( \bar{L} \), where

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

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

Put

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

By Lemma 2.5 we have \( \theta_\alpha c_\beta b_i \in \mathbb{Z} \) for all \( i \) and \( B \leq C_0 \). 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 \( C \leq C_0^{s} \). Hence the linear forms

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

By Lemma 2.3,

\[
\max\{H(\hat{P}), H(\hat{Q})\} = \max_{0 \leq i \leq 2s+1} \{|a_i|\}
\leq ((2s + 2)A)^{d/(2s+2-d)}
\leq ((2s + 2)\theta_\alpha c_\beta c_\alpha^s C_0 (1 + sC_0^s))^{d/(2s+2-d)}.
\]
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 \( \max\{\deg \hat{P}, \deg \hat{Q}\} \leq \lceil d/2 \rceil \) that

\[
d \leq r + \max\{\deg \hat{P}, \deg \hat{Q}\} \leq 2[d/2].
\]

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 \( \text{[2, Lemma 1.6.11]} \),

\[
H(P) \leq H(G)H(P) \leq 2^{\deg(P)}H(G)P \leq 2d^2H(\hat{P}).
\]

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 \( \alpha \) has degree \( d \) and

\[
\deg W = \deg(PQ' - QP') \\
\leq \max\{\deg P, \deg Q\} + \max\{\deg P', \deg Q'\} \\
\leq d/2 + (d/2 - 1) \\
< d,
\]

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

With the basic properties of heights listed in \( \text{[17, Section 2.4.1]} \), we find the following upper bound on \( H(W) \):

\[
H(W) \leq H(PQ') + H(QP') \\
\leq r(H(P)H(Q') + H(Q)H(P')) \\
\leq 2r^2H(P)H(Q) \\
\leq 2r^2 \max\{H(P), H(Q)\}^2 \\
\leq 2(d/2)^2C_{12}^2.
\]

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

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

is a nonzero rational integer. Thus,

\[
|W(\alpha)|^{-1} \leq c_\alpha^{d \deg W} \prod_{i=2}^{d} |W(\alpha_i)|
\]

\( \copyright \) Springer
Let $\alpha \in \mathbb{Q}_p$. Let $f(x)$ denote the minimal polynomial of $\alpha$ with leading coefficient $c_\alpha$. By [13, 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 $\alpha$ is a root of $f(x)$, we see that $\psi(\alpha)W(\alpha) = \text{Res}(f, W)$. Since $c_\alpha^{-1}\psi(\alpha)$ is an algebraic integer, its $p$-adic absolute value does not exceed 1, so

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

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

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

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

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

\[\Box\]

4 A gap principle in the presence of vanishing

Let $\alpha$ 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 1.
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

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

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

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

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 

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

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 [13, 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 the 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 $\text{Res}(P(x, y), Q(x, y))$. Let $R(f) = x^{\deg f} f(1/x)$ denote the reciprocal of a polynomial $f(x)$. Then

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

so

$$\text{Res}(P(1, x), Q(1, x)) = \text{Res}\left(R(P), x^{r-s}R(Q)\right)$$

$$= \text{Res}(R(P), x)^{r-s} \text{Res}(R(P), R(Q)).$$
Applying Hadamard’s inequality and
Then
\[ |c_p^{r-s} \operatorname{Res}(P, Q)| \leq |c_p|^r (r + 1)^{r/2} (s + 1)^{r/2} H(P)^r H(Q)^r \]
\[ \leq (r + 1)^r \max \{ H(P), H(Q) \}^{2r}. \]

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

Finally, since \( P(x) \) and \( Q(x) \) 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
\[ |c_p^{r-s} \operatorname{Res}(P, Q)| \leq |c_p|^r (r + 1)^{r/2} (s + 1)^{r/2} H(P)^r H(Q)^r \]
\[ \leq (r + 1)^r \max \{ H(P), H(Q) \}^{2r}. \]

\[ \square \]

**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

\[ C = C(P, Q) = \frac{\min_{i,j} \{ |\alpha_i \delta_j - \beta_i \gamma_j| \}}{\max_{i,j} \{ \max( |\alpha_i|, |\gamma_j|, |\beta_i|, |\delta_j|) \}}. \]

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

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

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

\[ |(\alpha_i \delta_j - \beta_i \gamma_j) b| = |\alpha_i (\gamma_j a + \delta_j b) - \gamma_j (\alpha_i a + \beta_i b)| \]
\[ \leq (|\alpha_i| + |\gamma_j|) \max \{ |\alpha_i a + \beta_i b|, |\gamma_j a + \delta_j b| \} \]
\[ < (|\alpha_i| + |\gamma_j|) C |b| \]
\[ \leq \min \{ |\alpha_i \delta_j - \beta_i \gamma_j| \} |b|, \]

so we obtain 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

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

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.

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

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

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

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

and so it follows from [2, Lemma 1.6.7] that

$$|c_P| \cdot \max\{1, |P|\} \leq (r + 1)^{1/2} H(P) \quad \text{and} \quad |c_Q| \cdot \max\{1, |Q|\} \leq (s + 1)^{1/2} H(Q). \quad (4.3)$$

Let $\mu_1, \ldots, \mu_r$ 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 $\alpha_i = c_P^{1/r} \mu_i$ and $\beta_i = -c_P^{1/r}$. 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 (4.3), the constant $C$ in (4.2) can be estimated from below as follows:

$$C = \frac{|c_P c_Q|^{1/r}}{\max_j \left\{ \max\{|c_P|^{1/r}, |c_P|^{1/r} |\mu_i| + |c_Q|^{1/r} \} \right\}}$$

$$= \frac{|c_Q|^{1/r}}{\max_j \{ \max\{1, |\mu_i| + \left(\frac{|c_Q|}{|c_P|}\right)^{1/r} \} \}}$$

$$= \frac{1}{\max_j \{ \max\{ \left(\frac{1}{|c_Q|}\right)^{1/r}, \left(\frac{\mu_i}{c_Q}\right)^{1/r} + \left(\frac{1}{|c_P|}\right)^{1/r} \} \}}$$

$$\geq \frac{1}{\max\{ \left(\frac{1}{|c_Q|}\right)^{1/r}, \left(\frac{|P|}{|c_Q|^{1/r}} + \left(\frac{1}{|c_P|}\right)^{1/r} \} \}}$$

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

**Case 2.** Suppose that $s \geq 1$. Let $\nu_1, \ldots, \nu_s$ 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

\[ y_j = \begin{cases} 
   c_Q^{1/r}, & \text{if } 1 \leq i \leq s, \\
   0, & \text{if } s + 1 \leq i \leq r, 
\end{cases} \]

and

\[ \delta_j = \begin{cases} 
   -c_Q^{1/r} v_i, & \text{if } 1 \leq i \leq s, \\
   c_Q^{1/r}, & \text{if } s + 1 \leq i \leq r. 
\end{cases} \]

Using (4.3), the constant \( C \) in (4.2) can be estimated from below as follows:

\[
C = \frac{\min_{i,j} |\alpha_i \delta_j - \beta_i y_j|}{\max_{i,j} \{ |\alpha_i| + |\gamma_j|, |\beta_i| + |\delta_j| \}} 
\geq \frac{|c_p c_Q|^{1/r} \min \{1, \min_{i,j} |\mu_i - v_j| \}}{2(r + 1)^{1/2} \max \{H(P), H(Q)\}}. 
\]

By [6, Theorem A],

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

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 1. Combining the lower bound on \( \min_{i,j} |\mu_i - v_j| \) 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 Eq. (4.1) 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

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

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,

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

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

\[
H(x_2, y_2) \geq \frac{\max \{ |P(x_1, y_1)|, |Q(x_1, y_1)| \} / (r + 1)^r \max \{H(P), H(Q)\}^{2r}}{\max \{H(P), H(Q)\}^{2r}} \geq \frac{2^r (r + 1)^{(3r^2 + 2r)}/2} \max \{H(P), H(Q)\}^{2r+3r}. 
\]

\( \square \)
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

\[
|D_i h(\alpha)| \leq H(h) \left( \frac{\deg h + 1}{i + 1} \right) \max \{ 1, |\alpha| \}^{\deg h - i}
\]

(5.1)

holds.

Let \( P, Q \) be a minimal pair for \((\alpha, \beta)\), and define \( R(x, y) = P(x) + y Q(x) \), 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, (5.1), and the two inequalities established above that

\[
\frac{1}{H(x_1, y_1) H(x_2, y_2)} \leq \left| \frac{R(x_1/y_1, x_2/y_2)}{R(x_1/y_1, x_2/y_2)} \right|
\]

\[
\leq \sum_{i=0}^{r} \sum_{j=0}^{1} |D_{i,j} R(\alpha, \beta)| \left| \alpha - \frac{x_1}{y_1} \right|^i \left| \beta - \frac{x_2}{y_2} \right|^j
\]

\[
< \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^{r} \left( |D_{i,0} R(\alpha, \beta)| + |D_{i,1} R(\alpha, \beta)| \right)
\]

\[
\leq \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^{r} \left( |D_{i} P(\alpha)| + (1 + |\beta|) \cdot |D_{i} Q(\alpha)| \right)
\]

\[
\leq \frac{C_0}{H(x_1, y_1)^\mu} \sum_{i=0}^{r} \left( \left( \begin{array}{c} r + 1 \\ i + 1 \end{array} \right) H(P) + (1 + |\beta|) \left( \begin{array}{c} r + 1 \\ i + 1 \end{array} \right) H(Q) \right) \max \{ 1, |\alpha| \}^{r-i}
\]

\[
\leq \frac{C_0}{H(x_1, y_1)^\mu} (2 + |\beta|) \max \{ H(P), H(Q) \} \max \{ 1, |\alpha| \}^{r} \sum_{i=0}^{r} \left( \begin{array}{c} r + 1 \\ i + 1 \end{array} \right)
\]

\[
< \frac{C_0}{H(x_1, y_1)^\mu} 2^{r+1}(2 + |\beta|) C_{12} \max \{ 1, |\alpha| \}^{r}
\]

\[
\leq \frac{C_0}{H(x_1, y_1)^\mu} 2^{(d/2)+1}(2 + |\beta|) C_{12} \max \{ 1, |\alpha| \}^{d/2}
\]

\[
= \frac{C_2}{H(x_1, y_1)^\mu},
\]

where the second-to-last inequality follows from (3.2). By (3.1),

\[
H(x_2, y_2) > C_{2}^{-1} H(x_1, y_1)^{\mu-d/2} \geq C_{2}^{-1} H(x_1, y_1)^{\mu-d/2},
\]

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 \( \beta \) 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 $\alpha$, $\beta$, $\mu$ and $C_0$. By choosing $C_1$ so that $C_1 \geq C$, we then arrive at a contradiction.

Note that

$$\left| \beta - \frac{x_2}{y_2} \right| = \left| \frac{P(\alpha)}{Q(\alpha)} - \frac{P(x_1/y_1)}{Q(x_1/y_1)} \right| = \left| \frac{P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)}{Q(x_1/y_1)} \right|.$$  \hspace{1cm} (5.2)

Further,

$$\left| Q(\alpha) \right| \leq (r + 1) \max\{H(P), H(Q)\} \max\{1, |\alpha|\}' \leq (r + 1)C_{12} \max\{1, |\alpha|\}' \hspace{1cm} (5.3)$$

$$\left| Q \left( \frac{x_1}{y_1} \right) \right| \leq \sum_{i=0}^{r} |D_i Q(\alpha)| \cdot \left| \alpha - \frac{x_1}{y_1} \right|$$

$$\leq H(Q) \sum_{i=0}^{r} (r + 1) \max\{1, |\alpha|\}'^{r-i}$$

$$\leq 2^{r+1}C_{12} \max\{1, |\alpha|\}'.$$ \hspace{1cm} (5.4)

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, (5.1) and (3.3),

$$\left| P(\alpha)Q \left( \frac{x_1}{y_1} \right) \right| - Q(\alpha)P \left( \frac{x_1}{y_1} \right)$$

$$= \left| P(\alpha) \sum_{i=0}^{r} D_i Q(\alpha) \left( \frac{x_1}{y_1} - \alpha \right)^i - Q(\alpha) \sum_{i=0}^{r} D_i P(\alpha) \left( \frac{x_1}{y_1} - \alpha \right)^i \right|$$

$$= \left| \alpha - \frac{x_1}{y_1} \right| \cdot \sum_{i=0}^{r-1} \left( P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha) \right) \left( \frac{x_1}{y_1} - \alpha \right)^i$$

$$= \left| \alpha - \frac{x_1}{y_1} \right| \left( W(\alpha) + \left( \frac{x_1}{y_1} - \alpha \right) \sum_{i=1}^{r-1} \left( P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha) \right) \left( \frac{x_1}{y_1} - \alpha \right)^{i-1} \right)$$

$$\geq \left| \alpha - \frac{x_1}{y_1} \right| \left( W(\alpha) - \frac{C_0}{H(x_1, y_1)^{\mu}} \sum_{i=1}^{r-1} \left| P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha) \right| \left( \frac{x_1}{y_1} - \alpha \right)^{i-1} \right)$$

$$\geq \left| \alpha - \frac{x_1}{y_1} \right| \left( W(\alpha) - \frac{C_0}{H(x_1, y_1)^{\mu}} 2(r + 1) \max\{H(P), H(Q)\}^2 \sum_{i=1}^{r-1} \left( \frac{r+1}{i+2} \right) \max\{1, |\alpha|\}^{2r-i-1} \right)$$

$$\geq \left| \alpha - \frac{x_1}{y_1} \right| \left( C_{13} - C_1^{\mu} C_0 2^{r+2} (r + 1) C_{12}^2 \max\{1, |\alpha|\}^{2r} \right),$$

where the last inequality follows from $H(x_1, y_1) \geq C_1$, (3.2) and (3.3). 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

$$\left| P(\alpha)Q \left( \frac{x_1}{y_1} \right) - Q(\alpha)P \left( \frac{x_1}{y_1} \right) \right| \geq \left| \alpha - \frac{x_1}{y_1} \right| \left( C_{13} - C_1^{\mu} C_0 2^{r+2} (r + 1) C_{12}^2 \max\{1, |\alpha|\}^{2r} \right)$$
By Proposition 4.1, 

\[ \delta - \frac{x_2}{y_2} = \left| \frac{P(\alpha)Q(x_1/y_1) - Q(\alpha)P(x_1/y_1)}{Q(\alpha)Q(x_1/y_1)} \right| \]

\[ > \frac{C_{13}}{2(r+2)} \left| \alpha - x_1 \right| \left| \beta - x_2 \right| \]

Combining the above result with (5.2), (5.3) and (5.4) yields

\[ \left| \alpha - x_1 \right| < \frac{2^{r+2}(r+1)C_{12}^2 \max[1, |\alpha|]^{2r}}{H(x_1, y_1)^{2r} + 3r} \left| \beta - x_2 \right| \leq \frac{C_0}{C_{13}H(x_1, y_1)^{2r}} \]

By Proposition 4.1,

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

where

\[ C_{15} = 2^r (r+1)^{(3r+2)} \approx 2^{d/4}((d/2) + 1)^{(d^2/4d)}/8. \] (5.5)

Consequently,

\[ \frac{C_6}{H(x_1, y_1)^d} \leq \left| \alpha - x_1 \right| < \frac{2^{r+2}(r+1)C_0C_{12}^{(2r+3)\mu+2}C_{15}^\mu \max[1, |\alpha|]^{2\mu}}{C_{13}H(x_1, y_1)^{2r}}. \]

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

\[ H(x_1, y_1) < \left( \frac{2^{r+2}(r+1)C_0C_{12}^{(2r+3)\mu+2}C_{15}^\mu \max[1, |\alpha|]^{2\mu}}{C_{13}H(x_1, y_1)^{2r}} \right)^{1/(r\mu - d)} \]

\[ \leq \left( \frac{2^{d\mu/4}((d/2) + 1)^{3d^2+4d})}{2^{d^2/4d})}C_0C_{12}^{(2r+3)\mu+2}C_{15}^\mu \max[1, |\alpha|]^{2\mu} \right)^{1/(d\mu - d)} \]

Notice how in the second-to-last inequality we have utilized the upper bound on \( C_{15} \) given in (5.5). 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 at 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 \( (\alpha, \beta) \), and define \( R(x, y) = P(x) + yQ(x) \), so that \( R(\alpha, \beta) = 0 \). Suppose \( R(x_1/y_1, x_2/y_2) \neq 0 \). Then the
following trivial lower bound holds:

\[
\left| y_1^j y_2^k R \left( \frac{x_1}{y_1}, \frac{x_2}{y_2} \right) \right|_p \geq \frac{1}{|y_1^j y_2^k R(x_1/y_1, x_2/y_2)|} \geq \frac{1}{2(r+1) \max(H(P), H(Q)) H(x_1, y_1) H(x_2, y_2)} \geq \frac{1}{2((d/2) + 1)C_{12} H(x_1, y_1)^{d/2} H(x_2, y_2)}.
\]

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

\[
\left| y_1^j y_2^k R \left( \frac{x_1}{y_1}, \frac{x_2}{y_2} \right) \right|_p \leq \max_{(i,j)\neq(0,0)} \left| D_{ij} R(\alpha, \beta) \right|_p \cdot \left| y_1^{i} - x_1^{i} \right|_p \cdot \left| y_2^{j} - x_2^{j} \right|_p \leq \frac{c_\alpha^i c_\beta^j}{H(x_1, y_1)^{\mu}}.
\]

Upon combining the upper and lower bounds, we obtain

\[
\frac{1}{2((d/2) + 1)C_{12} H(x_1, y_1)^{d/2} H(x_2, y_2)} \leq \frac{C_0 c_\alpha^i c_\beta^j}{H(x_1, y_1)^{\mu}}.
\]

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 \( \beta \) 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 \( \alpha \) and \( \beta \). By choosing \( C_3 \) so that \( C_3 \geq C \), we then arrive at 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 \( |y_1/\alpha|_p = |x_1|_p \). 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 obtain 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,

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

(6.1)

Since $g$ is an integer, we have

$$|g|_p \leq 1.$$

(6.2)

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

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

(6.3)

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

$$c_\alpha^{2r-1-i} (P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha))$$

is an algebraic integer, so its $p$-adic absolute value does not exceed 1. Consequently,

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

Now, let $W = PQ' - QP'$ denote the Wronskian of $P$ and $Q$. By Taylor’s Theorem and (3.4),

$$|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p$$

$$= \left| P(\alpha) \sum_{i=0}^r D_i Q(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} - Q(\alpha) \sum_{i=0}^r D_i P(\alpha) (x_1 - \alpha y_1)^i y_1^{r-i} \right|_p$$

$$= \left| y_1 \alpha - x_1 \right|_p \sum_{i=0}^{r-1} \left| (P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha)) (x_1 - \alpha y_1)^i y_1^{r-1-i} \right|_p$$

$$\geq \left| y_1 \alpha - x_1 \right|_p \left( |W(\alpha)y_1^{r-1}|_p - |y_1 \alpha - x_1|_p \max_{i=0, \ldots, r-1} \left| \frac{P(\alpha)D_{i+1} Q(\alpha) - Q(\alpha)D_{i+1} P(\alpha)}{c_\alpha^i} \right|_p \right)$$

$$> |y_1 \alpha - x_1|_p \left( |W(\alpha)y_1^{r-1}|_p - \frac{C_0}{H(x_1, y_1)^\mu c_\alpha^{2r-1}} \right).$$
where the last inequality follows from $H(x_1, y_1) \geq C_3$, (3.4) and $|y_1|_p \geq c^{-1}_\alpha$. Thus, if we choose $C_3$ so that

$$C_3 \mu \geq 2C_0C_{14}^{-1}c^{(3d-4)/2}_\alpha,$$

then

$$|P(\alpha)Q(x_1, y_1) - Q(\alpha)P(x_1, y_1)|_p > |y_1\alpha - x_1|_p \left(C_{14}^{-1} - C_3^{-\mu}C_0c^{-2r-1}_\alpha\right)$$

Combining this observation with (6.1), (6.2) and (6.3),

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

By Proposition 4.1,

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

where $C_{15}$ satisfies the upper bound (5.5). Consequently,

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

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

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

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

$$H(x_1, y_1) \leq \left(2d^{-1} C_0 C_{12}^{-1} C_{14}^{-1} C_{15}^{-\mu}\right)^{1/(r\mu-d)}$$

$$\leq \left(2d^{d/4}((d/2) + 1)(3d^2+4d)^{\mu/8}c^{d-1}_\alpha C_0 C_{12}^{-1} C_{14}^{-1} C_{15}^{-\mu}\right)^{1/(2\mu-d)} = C.$$

Notice how in the second-to-last inequality we have utilized the upper bound on $C_{15}$ given in (5.5). 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 at a contradiction.

### 7 The enhanced automorphism group

In this section we establish several results about the enhanced automorphism group $\text{Aut}'|F|$ 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** (see [12].) If $F \in \mathbb{Z}[x, y]$ is a binary form of degree $d \geq 3$ and nonzero discriminant $D(F)$, then $\text{Aut}_1 |F|$ is $\text{GL}_2(\mathbb{Q})$-conjugate to one of the groups from Table 1.
Lemma 7.2 If $F \in \mathbb{Z}[x, y]$ is a binary form of degree $d \geq 3$ with nonzero discriminant $D(F)$ and $\text{Aut'} |F|$ is as in (1.3), then $\text{Aut'} |F| \cong C_n$ or $\text{Aut'} |F| \cong D_n$, where $n \in \{1, 2, 3, 4, 6, 8, 12\}$.

Proof Note that $\text{Aut}_Q |F|$ is a subgroup of $\text{Aut'} |F|$. Furthermore, for any $M \in \text{Aut'} |F|$ we have $M^2 \in \text{Aut}_Q |F|$. By Lemma 7.1, $\text{Aut}_Q |F|$ is finite, and so any $M \in \text{Aut'} |F|$ has finite order. In fact, since the orders of elements in $\text{Aut}_Q |F|$ are $\{1, 2, 3, 4, 6\}$, the only possible orders of elements in $\text{Aut'} |F|$ are $\{1, 2, 3, 4, 6, 8, 12\}$.

Next, recall a classical result that any finite subgroup of $\text{GL}_2(\mathbb{R})$ is $\text{GL}_2(\mathbb{R})$-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 $\text{GL}_2(\mathbb{R})$, including $\text{Aut'} |F|$, is isomorphic to either a cyclic group $C_n$ of order $n$ or a dihedral group $D_n$ of order $2n$.

Now suppose that $\text{Aut'} |F|$ contains at least 25 distinct elements $M_1, \ldots, M_{25}$. By Schur’s Theorem [7], any finitely generated torsion subgroup of $\text{GL}_n(\mathbb{C})$ is finite. Hence $(M_1, \ldots, M_{25})$ is a finite subgroup of $\text{GL}_2(\mathbb{R})$, 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 $\text{Aut'} |F|$ can have is 12. Therefore $\text{Aut'} |F|$ contains at most 24 elements. \hfill \Box

Let us give an example of a group of the form (1.3) that is not a subgroup of $\text{GL}_2(\mathbb{Q})$. Consider

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

Then $G \cong D_{12}$. If we choose coprime integers $a$ and $b$ so that $a \equiv 3b \pmod{10}$, then any (reciprocal) binary form

$$F(x, y) = a(x^{12} + y^{12}) - 6axy(x^{10} + y^{10}) + \frac{231a + 2b}{5} x^2 y^2(x^8 + y^8) - (176a + 2b)x^3 y^3(x^6 + y^6) + \frac{495a + 5b}{2} x^4 y^4(x^4 + y^4) + 2bx^5 y^5(x^2 + y^2) - \frac{1122a + 29b}{5} x^6 y^6$$

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

| Group | Generators | Group | Generators |
|-------|------------|-------|------------|
| $C_1$ | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ | $D_1$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ |
| $C_2$ | $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ | $D_2$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ |
| $C_3$ | $\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ | $D_3$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ |
| $C_4$ | $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ | $D_4$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ |
| $C_6$ | $\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$ | $D_6$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$ |
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 [15, Section 6] with respect to binary forms invariant under $D_6$, which is a subgroup of $G$. In addition to these 12 solutions, we have $F(x', y') = 729m$ for any $(x', y') \in \{(x+y, -x+2y), (-x+2y, -2x+y), (-2x+y, -x-y), (-x-y, x-2y), (x-2y, 2x-y), (2x-y, x+y), (-x+2y, x+y), (-2x+y, -x+2y), (-x-2y, -x-y), (2x-y, x-2y), (x+y, 2x-y)\}$. 

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 $\alpha_1, \ldots, \alpha_d$ be the roots of $F(x, 1)$. There exists an index $j \in \{1, \ldots, d\}$ 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 $j \in \{1, \ldots, d\}$ 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 $\alpha_1, \alpha_2, \ldots, \alpha_d$. 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,

$$F(x, y) = c_d \prod_{i=1}^d \left(x - \frac{v\alpha_i - u}{-t\alpha_i + s} y\right)$$

$$= \frac{c_d}{\prod_{i=1}^d (s - t\alpha_i)} \prod_{i=1}^d ((-t\alpha_i + s)x - (v\alpha_i - u)y)$$

$$= \frac{c_d}{F(s, t)} F(sx + uy, tx + vy)$$

$$= \pm \eta^d F(sx + uy, tx + vy)$$

$$= \pm F_M(x, y),$$

where $\eta = |c_d/F(s, t)|^{1/d}$ and $M = \eta \begin{pmatrix} \vphantom{v} u \\ t v \end{pmatrix}$. 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 $M = |sv - tu|^{-1/2} \begin{pmatrix} \vphantom{v} u \\ t v \end{pmatrix}$ is in $\text{Aut}' |F|$. Then

$$\pm F(x, y) = F_M(x, y)$$
Theorem 8.1
Let $K$ be the following property. The total number of rationals $x/y$ such that $F(x, y) = 0$ is less than $C_0 > (4eA)^{-1}$, where $A = 500^2 \left( \log \max_{1 \leq i \leq n} (M(\alpha_i)) + \frac{d}{2} \right)$. (8.1)

There exists a positive real number $C_{16}$, which depends on $\alpha_1, \alpha_2, \ldots, \alpha_n, \mu$ 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

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

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

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

where

$$\gamma = \max\{\gamma_1, \ldots, \gamma_n\}, \quad \gamma_i = \#\{j : \alpha_j \in \text{orb}(\alpha_i)\}. \quad (8.3)$$

Let us see why Theorem 1.3 follows from Theorem 8.1.

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

$\square$
Choose $C_5$ so that
\[ C_5^{d-\mu} = \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2} m}{\vert D(F) \vert^{1/2}}. \]
Let $(x, y)$ be a primitive solution to (1.2) 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 $j \in \{1, 2, \ldots, d\}$ such that
\[
\min \left\{ \left| \alpha_j - \frac{x}{y} \right|, \left| \alpha_j^{-1} - \frac{y}{x} \right| \right\} \leq \frac{2^{d-1} d^{(d-1)/2} M(F)^{d-2} m}{\vert D(F) \vert^{1/2} H(x, y)^d} < \frac{1}{H(x, y)^\mu}.
\]
Next, adjust the choice of $C_5$ so that Theorem 8.1 applies:
\[ C_5 \geq \max \left\{ C_{16}(\alpha_1, \ldots, \alpha_d, \mu, C_0), C_{16}(\alpha_1^{-1}, \ldots, \alpha_d^{-1}, \mu, C_0) \right\}, \]
where $C_0 = 1$. If we let $\gamma$ be as in (8.3), 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 $\gamma \leq \frac{\text{Aut}^\ell \vert F 
\ell} {2}$. The division by 2 appears due to the presence of the matrix \((-1 0 \quad 0 1\)) in Aut $\vert F$, which maps $(x, y)$ to $(-x, -y)$. \hfill \Box

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 (8.2) 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 in lowest terms that satisfy the following conditions.

1. \[ C_{16} \leq H(x_1, y_1) \leq H(x_2, y_2) \leq \cdots \leq H(x_\ell, y_\ell). \]
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 an 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 $\alpha_{i_k} \in \text{orb}(\alpha_{i_j})$, 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
   \[
   \frac{x_k}{y_k} \neq \frac{s x_j + t y_j}{u x_j + v y_j}.
   \]
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

\[
\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots, \frac{x_{\ell}}{y_{\ell}}
\]
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 \(\alpha \in \text{orb}(\alpha_{ij})\) as follows. Let

\[
\alpha = \frac{s\alpha_{ij} + t}{u\alpha_{ij} + v} \quad \text{and} \quad \frac{x_j'}{y_j'} = \frac{sx_j + ty_j}{ux_j + vy_j}
\]

for some integers \(s, t, u\) and \(v\). Then

\[
\alpha - \frac{x_j'}{y_j'} = \frac{tu - sv}{(u\alpha_{ij} + v)(u(x_j/y_j) + v)} \left( \alpha_{ij} - \frac{x_j}{y_j} \right),
\]

so rational approximations to \(\alpha\) and \(\alpha_{ij}\) 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 \(\gamma_i\) defined in (8.3). Note that the value \(\gamma_{ij}\) 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 \(\sum_{j=1}^{\ell} \gamma_{ij}\). Therefore,

\[
N \leq \sum_{j=1}^{\ell} \gamma_{ij} \leq \gamma \ell,
\]

where \(\gamma\) is defined in (8.3). Thus, it remains to estimate \(\ell\).

To derive an upper bound on \(\ell\), we begin by applying a generalized gap principle to the ordered pair \((\alpha_{ik}, \alpha_{ik+1})\). Choose \(C_{16}\) and define \(C\) as follows:

\[
C_{16} = \max_{j,k} \{C_1(\alpha_j, \alpha_k, \mu, C_0), C_3(\alpha_j, \alpha_k, \mu, C_0)\},
\]

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

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

\[
|y_k \alpha_{ik} - x_k| = |y_k| \cdot \left| \alpha_{ik} - \frac{x_k}{y_k} \right| \leq \frac{C_0}{H(x_k, y_k)^{\mu}}.
\]

Analogously,

\[
|y_{k+1} \alpha_{ik+1} - x_{k+1}| \leq \frac{C_0}{H(x_{k+1}, y_{k+1})^\mu}.
\]

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

\[
H(x_{k+1}, y_{k+1}) > C^{-1} H(x_k, y_k)^E,
\]
where
\[ E = \mu - d/2. \] (8.4)

Notice that Case 2 in the aforementioned theorems is impossible due to the fact that the list \( x_1/y_1, \ldots, x_\ell/y_\ell \) does not contain derived rational approximations. Consequently,
\[
\begin{align*}
\log H(x_\ell, y_\ell) &> E \log H(x_{\ell-1}, y_{\ell-1}) - \log C \\
    &> E^2 \log H(x_{\ell-2}, y_{\ell-2}) - (1 + E) \log C \\
    &> \cdots \\
    &> E^{\ell-1} \log H(x_1, y_1) - (1 + E + \cdots + E^{\ell-2}) \log C.
\end{align*}
\]

Thus, we obtain the following lower bound on \( \log H(x_\ell, y_\ell) \):
\[
\log H(x_\ell, y_\ell) > E^{\ell-1} \log H(x_1, y_1) - \frac{E^{\ell-1} - 1}{E - 1} \log C. \] (8.5)

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 \( \alpha_i^i \)'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
\[
\lambda = 2t - \tau = \frac{2}{(1 - 2a)t} < 1.42\sqrt{d}.
\]

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
\[
\delta^{-1} < 41667d^2. \] (8.6)

We further adjust our definition of \( C_{16} \) by choosing it so that
\[
C_{16} \geq C_{16}' \geq C_{0}^\frac{1}{1-42\sqrt{d}} \left( 4e^A \right)^{\frac{1.42\sqrt{d}}{1-42\sqrt{d}}}, \] (8.7)
where \( A \) is defined in (8.1). 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
\[
\log H(x_\ell, y_\ell) \leq \delta^{-1} \left( \log(4e^{A_1}) + \log H(x_1, y_1) \right) - \log(4e^{A_1}).
\]
\[ < 41667d^2 \left( \log(4e^{A_1}) + \log H(x_1, y_1) \right), \]

where the last inequality follows from (8.6). 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 (8.5):

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

Reordering the terms yields

\[ \left( E_\ell^{-1} - 41667d^2 \right) \log H(x_1, y_1) - \frac{E_\ell^{-1} - 1}{E - 1} \log C < 41667d^2 \log(4e^{A_1}). \quad (8.8) \]

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 (8.8):

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

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

\[ C_{16} \geq C^{2/(E-1)}. \quad (8.9) \]

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 the conclusion

\[ (\mu - 0.5d)_\ell^{-1} < 1 + 83334d^2 \left( 1 + \frac{\log(4e^{A_1})}{\log C_{16}} \right). \quad (8.10) \]

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 (8.10), 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

$$\ell < 1 + \frac{\log(98897d^{3/2}\mu)}{\log(\mu - d/2)} < 1 + \frac{11.51 + 1.51\log d + \log \mu}{\log(\mu - d/2)}.$$

The result follows once we multiply the right-hand side by the constant $\gamma$ defined in (8.3). $\square$

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, On the Thue- Siegel-Dyson theorem. Acta Math. 148, 255–296 (1982)
2. E. Bombieri, W. Gubler, Heights in diophantine geometry (Cambridge University Press, Cambridge, 2006)
3. E. Bombieri, J. Mueller, On effective measures of irrationality for $\sqrt{a/b}$ and related numbers. J. Reine Angew. Math. 342, 173–196 (1983)
4. E. Bombieri, W.M. Schmidt, On Thue’s equation. Invent. Math. 88, 69–81 (1987)
5. E. Bombieri, J. Vaaler, On Siegel’s lemma. Invent. Math. 73, 11–32 (1983)
6. Y. Bugeaud, M. Mignotte, On the distance between roots of integer polynomials. Proc. Edinburgh Math. Soc. 47, 553–556 (2004)
7. C. Curtis, I. Reiner, Representation theory of finite groups and associative algebras (Wiley, New York, 1962), pp. 258–262
8. W. Gautschi, Norm estimates for inverses of Vandermonde matrices. Numer. Math. 23, 337–347 (1975)
9. K. Györy, Thue inequalities with a small number of primitive solutions. Period. Math. Hungar. 42(1–2), 239–246 (2001)
10. D. Lewis, K. Mahler, Representation of integers by binary forms. Acta Arith. 6, 333–363 (1961)
11. J. Mueller, W.M. Schmidt, Thue’s equation and a conjecture of siegel. Acta Math. 160, 207–247 (1988)
12. M. Newman, Integral Matrices Pure and Appl Math. (Academic Press, New York, 1972)
13. V.P. Prasolov, Polynomials Algorithms and Computation in Mathematics. (Springer-Verlag, Berlin Heidelberg, 2004)
14. W.M. Schmidt, Diophantine approximations and diophantine equations (Springer-Verlag, Berlin, 1991)
15. C.L. Stewart, On the number of solutions of polynomial congruences and Thue equations. J. Amer. Math. Soc. 4, 793–835 (1991)
16. A. Thue, Über Annäherungswerte algebraischer Zahlen. J. Reine Angew. Math. 135, 284–305 (1909)
17. U. Zannier, Lecture notes on diophantine analysis (Scuola Normale Superiore Pisa, Italy, 2014)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.