APPROXIMATE REALS BY ALGEBRAIC NUMBERS

BENJAMIN LI

ABSTRACT. Wirsing showed that we can approximate reals by algebraic numbers α of degree $\leq n$ with an error bounded by $O\big(\|\alpha\|^{-(n+3)/2+\epsilon}\big)$. In this paper, we will strengthen Wirsing's result by removing the ϵ in his statement.

CONTENTS

1. INTRODUCTION

The classical answer to the question of how well can one approximate irrational real numbers by rational numbers is given by Dirichlet's theorem.

Theorem 1.1 (Dirichlet). *For all irrational numbers $\xi \in \mathbb{R}-\mathbb{Q}$, there are infinitely many coprime pairs (p, q) such that $|\xi - p/q| < 1/|q|^2$.*

Proof. See [4, Chapter 1-3]. □

Dirichlet theorem tells us that we can approximate reals by rational numbers with error bounded by the square of the denominator. It is a natural question to ask how well we can approximate reals $\xi \in \mathbb{R}$ by algebraic numbers of degree $\leq n$. Obviously, when ξ is algebraic with degree $\leq n$ then we can approximate ξ perfectly by itself, so the real interesting case is when ξ is either transcendental or algebraic of degree $> n$.

Example 1.2. Consider the transcendental number $\xi = 1/e$. Then the best rational approximate p/q of ξ such that $q \leq 10$ is $3/8$, with difference $|\xi - 3/8| \approx 0.007$. However, if we approximate ξ by quadratic numbers then after letting $\alpha = (\sqrt{3}-1)/2$ we have $|\xi - \alpha| \approx 0.001$, so we see using algebraic numbers can give a better approximation.

To formalize our statement, we have to measure how "big" an algebraic number α is, which will be used to replace $|q|$ in Theorem 1.1. In order to measure how big α is, we look at $F_\alpha \in \mathbb{Z}[x]$, the minimal polynomial of α. If we can measure how big F_α is, then we can simply set the size of α to be the size of F_α.

Date: May 13, 2026.

There are two common choices for how big F_α is. Let us define them here.

Definition 1.3. Let $F(x) = c_r x^r + \cdots + c_0 \in \mathbb{Z}[x]$, then the (polynomial) height of F is $\|F\| = \max(|c_0|, |c_1|, \ldots, |c_r|)$. For $\alpha \in \overline{\mathbb{Q}}$, we define $\|\alpha\| = \|F_\alpha\|$.

Example 1.4. Let $\alpha = (\sqrt{3}-1)/2$. Then $F_\alpha(x) = 2x^2 + 2x - 1$, so that $\|\alpha\| = \|F_\alpha\| = \max(|2|, |2|, |-1|) = 2$.

Definition 1.5. For a complex number $z = x + iy$, we define its absolute value to be $|z| = \sqrt{x^2 + y^2}$.

Definition 1.6. Let $F(x) = c_r x^r + \cdots + c_0 = c_r \prod_{i=1}^r (x - \alpha_i) \in \mathbb{Z}[x]$. Then the Mahler measure of F is $\mathrm{Mah}(f) = c_r \prod_{|\alpha_i| \geq 1} |\alpha_i| = c_r \prod_{i=1}^r \max(1, |\alpha_i|)$.

Example 1.7. Let $F(x) = 2x^2 + 2x - 1 = 2(x - (\sqrt{3}-1)/2)(x - (-\sqrt{3}-1)/2)$. Then $\mathrm{Mah}(F) = 2\max(1, \left|(\sqrt{3}-1)/2\right|)\max(1, \left|(-\sqrt{3}-1)/2\right|) = 2 \cdot 1 \cdot (\sqrt{3}+1)/2 = \sqrt{3}+1$.

In fact, using either the polynomial height or the Mahler measure does not make a distinction.

Definition 1.8. Let f, g be two non-negative real functions that depend on some variables. We write $f \ll g$ if there exists $C \in \mathbb{R}_{>0}$ such that $f \leq Cg$. We write $f \asymp g$ if $f \ll g$ and $g \ll f$.

We will see later that after fixing n, we have $\|F\| \asymp \mathrm{Mah}(F)$ when $\deg F \leq n$, so that we can use polynomial height or Mahler measure interchangably.

In his paper [3], Wirsing proved the following theorem.

Theorem 1.9 (Wirsing). *Fix $\epsilon > 0, n \in \mathbb{N}$, and $\xi \in \mathbb{R}$ that is either transcendental or algebraic of degree $> n$, there are infinitely many $\alpha \in \overline{\mathbb{Q}}$ such that $\deg \alpha \leq n$ and*

$$|\xi - \alpha| < \|\alpha\|^{-\frac{n+3}{2}+\epsilon}.$$

In this paper, we are going to reproduce and strengthen his theorem. In particular, we will prove the following theorem.

Theorem 1.10 (Wirsing, Strengthened). *Fix $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$ that is either transcendental or algebraic of degree $> n$. Then there are infinitely many $\alpha \in \overline{\mathbb{Q}}$ such that $\deg \alpha \leq n$ and*

$$|\xi - \alpha| \ll \|\alpha\|^{-\frac{n+3}{2}}.$$

Notice that when $n = 1$, we get the following statement: for all irrational $\xi \in \mathbb{R}$, there are infinitely many coprime pairs (p, q) such that

$$|\xi - p/q| \ll \max(p, q)^{-2},$$

which is Theorem 1.1.

Remark 1.11. The exponent in Theorem 1.10 has undergone various improvement. Theorem 1.10 shows $|\xi - \alpha| \ll \|\alpha\|^{-n/2+O(1)}$; in [1], Poëls shows that $|\xi - \alpha| \ll \|\alpha\|^{-n/(2-\log 2)+O(1)}$, which is stronger than Theorem 1.10.

The structure of this paper is as follows. In Section 2 we will show that $\|F\| \asymp_n \mathrm{Mah}(f)$ along with some of its applications. In Section 3 we will show that we can find $P \in \mathbb{Z}[x]$ such that $|P(\xi)|$ is very small, using geometry of numbers. In Section 4 we will show that we can even take $P \in \mathbb{Z}[x]$ to be irreducible and $|P(\xi)|$ is still very small; the irreducibility will be crucial in the proof of Wirsing's theorem. Finally, in Section 5 we will give a proof for Theorem 1.10.

2. Heights

This section deals with the two notions of the size of polynomials. We will prove that $\mathrm{Mah}(f) \asymp \|f\|$ when $\deg f \le n$. One fundamental axiom for all sorts of height functions is the Northcott property:

Proposition 2.1 (Northcott Property, [2])**.** *Fix $n \in \mathbb{N}, B \in \mathbb{R}_{>0}$. Then there exist only finitely many $f \in \mathbb{Z}[x]$ such that $\deg f \le n$ and $\|f\| \le B$.*

Proof. Write $f(x) = c_n x^n + \cdots + c_0$ (with possibly $c_n = 0$). Then we see $|c_i| \le B$, so there are at most $(2\lfloor B \rfloor + 1)^{n+1}$ such f's. □

The Northcott property is important because it tells us that after fixing a degree bound, we can enumerate all algebraic numbers according to their heights.

As we promised, we will show that $\|f\| \asymp \mathrm{Mah}(f)$.

Proposition 2.2. *Fix $n \in \mathbb{N}$. Then for polynomial $f \in \mathbb{Z}[x]$ with $\deg f \le n$, we have $\|f\| \asymp \mathrm{Mah}(f)$.*

In the following lemmas, we will separately prove that $\|f\| \ll \mathrm{Mah}(f)$ and $\mathrm{Mah}(f) \ll \|f\|$, hence finish the proof for Proposition 2.2.

Lemma 2.3. *Fix $n \in \mathbb{N}$. Then for polynomial $f \in \mathbb{Z}[x]$ with $\deg f \le n$, we have $\|f\| \le 2^n \mathrm{Mah}(f)$.*

Proof. Let $f(x) = c_r x^r + \cdots + c_0 = c_r \prod_{i=1}^{r}(x - \alpha_i)$. Let $S = \{1, 2, \ldots, r\}$. Then for $1 \le u \le r$, we have

$$\begin{aligned}
|a_u| &= \left| c_r \sum_{\substack{I \subset S \\ |I|=u}} \prod_{s \in S} \alpha_s \right| && \text{(Vieta's formula)} \\
&\le |c_r| \sum_{\substack{I \subset S \\ |I|=u}} \prod_{s \in S} |\alpha_s| && \text{(properties of absolute values)} \\
&\le |c_r| \sum_{\substack{I \subset S \\ |I|=u}} \prod_{s \in S} \max(1, |\alpha_s|) && (\text{Since } |\alpha_s| \le \max(1, |\alpha_s|)) \\
&\le \sum_{\substack{I \subset S \\ |I|=u}} |c_r| \prod_{s=1}^{r} \max(1, |\alpha_s|) && (\text{Since } \max(1, |\alpha_s|) \ge 1) \\
&= \sum_{\substack{I \subset S \\ |I|=u}} \mathrm{Mah}(f) \\
&= \binom{r}{u} \mathrm{Mah}(f)
\end{aligned}$$

Finally, note that $2^n \ge 2^r = \sum_{i=0}^{r} \binom{r}{i}$, hence

$$\le 2^n \mathrm{Mah}(f),$$

which implies $\|f\| \le 2^n \mathrm{Mah}(f)$. □

For the reverse $\ll$, we will use the following identification of the Mahler measure. Since its proof involves complex analysis, we will omit the proof and instead give a reference.

Lemma 2.4. *For $f \in \mathbb{Z}[x]$, we have*

$$\mathrm{Mah}(f) = \exp\left(\int_0^1 \log\left|f(e^{2\pi it})\right| dt\right).$$

Proof. See [5]. $\square$

Thanks to Lemma 2.4, we are ready to prove $\mathrm{Mah}(f) \ll \|f\|$.

Lemma 2.5. *Fix $n \in \mathbb{N}$. Then for polynomial $f \in \mathbb{Z}[x]$ with $\deg f \leq n$, we have $\mathrm{Mah}(f) \leq (n+1)\|f\|$.*

Proof. Let $f(x) = c_r x^r + \cdots + c_0$. By Lemma 2.4 we have

$$\mathrm{Mah}(f) = \exp\left(\int_0^1 \log\left|f(e^{2\pi it})\right| dt\right)$$

By Jensen's inequality,

$$\begin{aligned}
&\leq \int_0^1 \exp\left(\log\left|f(e^{2\pi it})\right|\right) dt \\
&= \int_0^1 \left|f(e^{2\pi it})\right| dt \\
&\leq \int_0^1 \sum_{j=0}^{r} |c_j| dt \\
&\leq \int_0^1 (n+1)\|f\|\, dt \\
&= (n+1)\|f\|,
\end{aligned}$$

as desired. $\square$

With the help from Lemma 2.3 and Lemma 2.5, the proof of Proposition 2.2 is straightforward.

Proof of Proposition 2.2. By Lemma 2.3 we have $\|f\| \ll \mathrm{Mah}(f)$, and by Lemma 2.5 we have $\mathrm{Mah}(f) \ll \|f\|$. Thus, $\|f\| \asymp \mathrm{Mah}(f)$, as desired. $\square$

As an application of Proposition 2.2, let's show that the height of a polynomial is translational invariant (up to a scalar). First note that we can extend the definition of $\|P\|$ and $\mathrm{Mah}(P)$ for all $P \in \mathbb{C}[x]$ as follows. When $P(x) = c_r x^r + \cdots + c_0 = c_r \prod_{j=1}^{r}(x - \alpha_j) \in \mathbb{C}[x]$, we define $\|P\| = \max|c_j|$ and $\mathrm{Mah}(P) = c_r \prod_{|\alpha_j| \geq 1} |\alpha_j|$. Then Lemma 2.3, Lemma 2.4, and Lemma 2.5 still hold for the case of complex polynomials.

Corollary 2.6. *Fix $n \in \mathbb{N}$ and $z \in \mathbb{C}$. Then for $P \in \mathbb{C}[x]$ with $\deg P \leq n$, we have $\|P(x)\| \asymp \|P(x - z)\|$.*

Proof. By Proposition 2.2 it suffices to show $\mathrm{Mah}(P) \ll \mathrm{Mah}(P(x-z))$ (then we replace z with $-z$ to get the other side). After normalization we can assume $P(x) = \prod_{j=1}^{r}(x-\alpha_j)$. Then

$$\begin{aligned}\mathrm{Mah}(P) &= \prod_{|\alpha_j|\geq 1} |\alpha_j| \\ &= \prod_{1\leq|\alpha_j|<1+|z|} |\alpha_j| \prod_{|\alpha_j|\geq 1+|z|} |\alpha_j|\end{aligned}$$

Since $|\alpha_j - z| \geq |\alpha_j| - |z|$, we get

$$\begin{aligned}&\leq (1+|z|)^n \prod_{|\alpha_j|\geq 1+|z|} (1+|z|)|\alpha_j+z| \\ &\leq (1+|z|)^{2n} \prod_{|\alpha_j|\geq 1+|z|} |\alpha_j+z|\end{aligned}$$

Note that if $|\alpha_j| \geq 1+|z|$, then $|\alpha_j+z| \geq 1$, so

$$\leq (1+|z|)^{2n} \prod_{|\alpha_j+z|\geq 1} |\alpha_j+z|$$

Now observe $P(x-z) = \prod_{j=1}^{r}(x-(\alpha_j+z))$, so

$$= (1+|z|)^{2n}\,\mathrm{Mah}(P(x-z)),$$

as desired. ☐

Note that $\mathrm{Mah}(f)$ is multiplicative. This is because if $f(x) = c\prod_{j=1}^{n}(x-\alpha_j), g(x) = d\prod_{k=1}^{m}(x-\beta_k)$, then $(fg)(x) = cd\prod_{j=1}^{n}(x-\alpha_j)\prod_{k=1}^{m}(x-\beta_k)$, so $\mathrm{Mah}(fg) = cd\prod_{|\alpha_j|\geq 1}|\alpha_j|\prod_{|\beta_k|\geq 1}|\beta_k| = \left(c\prod_{|\alpha_j|\geq 1}|\alpha_j|\right)\left(d\prod_{|\beta_k|\geq 1}|\beta_k|\right) = \mathrm{Mah}(f)\,\mathrm{Mah}(g)$. As a result, $\|f\|$ is also multiplicative (up to some bounded scales).

Corollary 2.7. *Fix $n \in \mathbb{N}$. For $P, Q \in \mathbb{Z}[x]$ with $\deg P + \deg Q \leq n$, we have*

$$8^{-n}\|P\|\|Q\| \leq \|PQ\| \leq 8^n \|P\|\|Q\|.$$

Proof. By Lemma 2.3 and Lemma 2.5 we have

$$\begin{aligned}\|PQ\| &\geq (n+1)^{-1}\mathrm{Mah}(PQ) = (n+1)^{-1}\mathrm{Mah}(P)\,\mathrm{Mah}(Q) \\ &\geq (n+1)^{-1}(2^{-n}\|P\|)(2^{-n}\|Q\|) = \|P\|\|Q\|/((n+1)4^n) \geq 8^{-n}\|P\|\|Q\| \\ \|PQ\| &\leq 2^n\,\mathrm{Mah}(PQ) = 2^n\,\mathrm{Mah}(P)\,\mathrm{Mah}(Q) \leq 2^n((n+1)\|P\|)((n+1)\|Q\|) \\ &= (n+1)^2 2^n \|P\|\|Q\| \leq 8^n\|P\|\|Q\|.\end{aligned}$$

☐

Corollary 2.6 and Corollary 2.7 will be needed later to prove Theorem 1.10.

3. Polynomials with Small Evaluation at ξ

The proof of Wirsing's theorem can be roughly divided into 2 steps. In the first step, Wirsing claimed that we can find polynomials P such that $|P(\xi)|$ is small. Then he uses such polynomials to find α's that approximate ξ. The motivation is that if $|P(\xi)|$ is small, then after writing $P(x) = c\prod_{i=1}^{d}(x-\alpha_i)$ we have at least one $|\xi - \alpha_i|$ has to be small.

The main goal of this section is to exhibit $P \in \mathbb{Z}[x]$ such that $|P(\xi)|$ is small. This is achieved by Corollary 3.5. First let us recall Minkowski's theorem.

Theorem 3.1 (Minkowski)**.** *Let $m \in \mathbb{N}$ and $\Omega \subset \mathbb{R}^m$ be a convex bounded symmetric (i.e., $x \in \Omega$ implies $-x \in \Omega$) measurable set with volume $> 2^m$. Then $\Omega \cap \mathbb{Z}^m \neq \{0\}$.*

Proof. See, for example, [8, Theorem B.1]. □

Remark 3.2. [6] shows that every convex set in $\mathbb{R}^m$ is measurable, so the measurability assumption in Theorem 3.1 is unnecessary.

The next lemma tells us that we can find a polynomial P of degree $\leq n$ such that $|P(\xi)|$ is small.

Lemma 3.3. *Fix $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$. Then there exists $C \in \mathbb{R}_{>0}$ such that for all $H \in \mathbb{R}_{\geq 1}$, there exists $P \in \mathbb{Z}[x] - \{0\}$ with the property that $\deg P \leq n$, $|P(\xi)| \leq CH^{-n}$, and $\|P\| \leq H$.*

Proof. Take $D > 0$ and consider $\Omega_D \subset \mathbb{R}^{n+1}$ consists of $(c_0, \ldots, c_n)$ with constraints $|c_0 + c_1\xi + \cdots + c_n\xi^n| < 2D^n H^{-n}, |c_1| < H/D, |c_2| < H/D, \ldots, |c_n| < H/D$. Then $|\Omega_D| = 2^{n+2} > 2^{n+1}$ (because the matrix defining Ω_D has determinant 1, so the volumn of Ω_D is simply the volume of the rectangle defined by $|c_0'| < 2D^n H^{-n}$ and $|c_i'| < H/D$ for $1 \leq i \leq n$), so by Theorem 3.1 there exists a nonzero integral point that lies in Ω_D. We now observe that if $(c_0, \ldots, c_n) \in \Omega_D$ then $|c_0| \leq |c_0 + c_1\xi + \cdots + c_n\xi^n| + |c_1\xi| + |c_2\xi^2| + \cdots + |c_n\xi^n| \leq 2D^n H^{-n} + (|\xi| + |\xi^2| + \cdots + |\xi^n|)H/D$.

We claim that $C = 2(8n \max(1, |\xi|^n))^n$ works. If $H \leq 8n \max(1, |\xi|^n)$, then we just have to take $P = 1$, so we assume $H > 8n \max(1, |\xi|^n)$. However, in this case, if we take $D = 2n \max(1, |\xi|^n)$, then $|c_0| \leq 2 \cdot 4^{-n} + H/2 < H$, so if we take $P(x) = c_0 + c_1 x + \cdots + c_n x^n$, where $(c_0, \ldots, c_n) \in \Omega_D \cap (\mathbb{Z}^{n+1} - \{0\})$, then $|P(\xi)| < 2D^n H^{-n} < CH^{-n}$, as desired. □

Remark 3.4. Lemma 3.3 uses the condition that $\xi \in \mathbb{R}$; otherwise $c_0 + c_1\xi + \cdots + c_n\xi^n$ may not be real, so that the volume of Ω_D will not be 2^{n+2}.

A direct consequence of Lemma 3.3 is

Corollary 3.5. *Fix $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$ that is either transcendental or algebraic with degree $> n$. Then there exist $C \in \mathbb{R}_{>0}$ and infinitely many $P \in \mathbb{Z}[x]$ of degree $\leq n$ such that $|P(\xi)| \leq C\|P\|^{-n}$.*

Remark 3.6. The constant C in Corollary 3.5 can be computed efficiently. More explicitly, by the proof of Lemma 3.3 we see $C = 2(8n \max(1, |\xi|^n))^n$ works.

4. Gauss' Lemma and Irreducibility

We want to upgrade Corollary 3.5 to restrict P further to irreducible polynomials. This is because later we want to pick infinitely many coprime pairs of P, Q such that $|P|$ and $|Q|$ are small, so irreducibility will be the key. However, before doing that we need to deal with some technicalities of irreducibility in $\mathbb{Q}[x]$ versus irreducibility in $\mathbb{Z}[x]$.

Definition 4.1. Let R be a ring and $f(x) = c_q x^q + \cdots + c_0 \in R[x]$. We say f is primitive if $(c_0, c_1, \ldots, c_q) = R$ as ideals.

Lemma 4.2 (Gauss)**.** *Let R be a ring. If $f, g \in R[x]$ are primitive, then so is fg.*

Proof. Suppose the coefficients of fg generate the ideal $I \subset R$. If $I \neq R$, then there exists a maximal ideal $\mathfrak{m}$ containing I. Then after quotienting out by $\mathfrak{m}$ we see $fg = 0$ in $(R/\mathfrak{m})[x]$, but neither $f = 0$ nor $g = 0$ in $(R/\mathfrak{m})[x]$ (since they are primitive), contradicts the fact that $(R/\mathfrak{m})[x]$ is an integral domain. Thus, $I = R$ and fg is primitive. □

The next two propositions are standard applications of Gauss' lemma.

Proposition 4.3. *Let $f, g \in \mathbb{Z}[x] - \{0\}$ such that f is primitive. If $f \mid g$ in $\mathbb{Q}[x]$, then $f \mid g$ in $\mathbb{Z}[x]$.*

Proof. It suffices to show the case where g is primitive, since if $g = ag_0$ for some $a \in \mathbb{N}$ and $g_0 \in \mathbb{Z}[x]$ being primitive then $f \mid g_0$ in $\mathbb{Q}[x]$; so that if $f \mid g_0$ in $\mathbb{Z}[x]$ then $f \mid g$ in $\mathbb{Z}[x]$.

Let $g = fh$ with $h \in \mathbb{Q}[x]$. Write $h = bh_0$, where $b \in \mathbb{Q}_{>0}$ and $h_0 \in \mathbb{Z}[x]$ is primitive. Then fh_0 is primitive by Lemma 4.2. Thus, looking at the $\mathbb{Z}$-module generated by the coefficients from both sides of $g = bfh_0$ we see $b = 1$, i.e., $h \in \mathbb{Z}[x]$. □

Proposition 4.4. *Let $f \in \mathbb{Z}[x] - \{0\}$ be primitive. Then f is irreducible in $\mathbb{Z}[x]$ if and only if f is irreducible in $\mathbb{Q}[x]$.*

Proof. Proposition 4.4 is equivalent to say f is reducible in $\mathbb{Z}[x]$ if and only if f is reducible in $\mathbb{Q}[x]$. If f is reducible in $\mathbb{Z}[x]$, then because f is primitive we see each factor has degree ≥ 1, hence f is reducible in $\mathbb{Q}[x]$.

Converse, assume $f = gh$ is reducible in $\mathbb{Q}[x]$, with $\deg g \geq 1, \deg h \geq 1$. Then write $g = ag_0$ with $a \in \mathbb{Q}$ and $g_0 \in \mathbb{Z}[x]$ being primitive. We have $f = g_0(ah)$, so $g_0 \mid f$ in $\mathbb{Q}[x]$. Thus, by Proposition 4.3 we see $ah \in \mathbb{Z}[x]$, hence f is reducible in $\mathbb{Z}[x]$, as desired. □

We can now upgrade Corollary 3.5 to irreducible polynomials.

Proposition 4.5. *Fix $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$ that is either transcendental or algebraic with degree $> n$. Then there exist $C \in \mathbb{R}_{>0}$ and infinitely many $P \in \mathbb{Z}[x]$, irreducible over $\mathbb{Z}[x]$, such that $\deg P \leq n$ and $|P(\xi)| \leq C \|P\|^{-n}$.*

Proof. Define the set $\mathscr{P} = \{P \in \mathbb{Z}[x] - \mathbb{Z} : \deg P \leq n, P \text{ irreducible in } \mathbb{Z}[x]\}$ and consider the map $\varphi : \mathscr{P} \to \mathbb{R}_{>0}, P \mapsto \|P\|^n |P(\xi)|$.

Assume the contrary, then for all $D > 0$, the set $\{P \in \mathscr{P} : \varphi(P) \leq D\}$ is finite. Now Corollary 3.5 tells us that there exist $C > 0$ and infinitely many $Q \in \mathbb{Z}[x] - \{0\}$ such that $C \|Q\|^{-n} \geq |Q(\xi)|$. Write $Q = aQ_0$, where $a \in \mathbb{N}$ and Q_0 is primitive. Then $Q_0 = \prod_{i=1}^r P_i$ with $P_i \in \mathscr{P}$. We now have

$$\begin{aligned} C \|Q\|^{-n} &\geq |Q(\xi)| = a|Q_0(\xi)| \geq |Q_0(\xi)| \\ &= \prod_{i=1}^r |P_i(\xi)| = \prod_{i=1}^r (\|P_i\|^n |P_i(\xi)|) \prod_{i=1}^r \|P_i\|^{-n} \end{aligned}$$

By definition of φ,

$$= \prod_{i=1}^r \varphi(P_i) \prod_{i=1}^r \|P_i\|^{-n}$$

By Corollary 2.7,

$$\geq \prod_{i=1}^{r} \varphi(P_i) \cdot 8^{-n^2} \left\| \prod_{i=1}^{r} P_i \right\|^{-n} = 8^{-n^2} \prod_{i=1}^{r} \varphi(P_i) \left\| Q_0 \right\|^{-n}$$
$$= a^n 8^{-n^2} \prod_{i=1}^{r} \varphi(P_i) \left\| Q \right\|^{-n} \geq 8^{-n^2} \prod_{i=1}^{r} \varphi(P_i) \left\| Q \right\|^{-n}.$$

As a result, $\prod_{i=1}^{r} \varphi(P_i) \leq 8^{n^2} C$. Now let $\delta = \min(1, \{\varphi(P) : P \in \mathscr{P}\})$ (from our assumption $S = \{P \in \mathscr{P} : \varphi(P) < 1\}$ is finite, so δ exists), then

$$8^{n^2} C \geq \prod_{i=1}^{r} \varphi(P_i) = \prod_{\substack{1 \leq i \leq r \\ \varphi(P_i) < 1}} \varphi(P_i) \prod_{\substack{1 \leq i \leq r \\ \varphi(P_i) \geq 1}} \varphi(P_i) \geq \delta^n \prod_{\substack{1 \leq i \leq r \\ \varphi(P_i) \geq 1}} \varphi(P_i).$$

Thus, for each j such that $\varphi(P_j) \geq 1$, we must have

$$\varphi(P_j) \leq \prod_{\substack{1 \leq i \leq r \\ \varphi(P_i) \geq 1}} \varphi(P_i) \leq 8^{n^2} C \delta^{-n}.$$

Now because the set $\left\{ P \in \mathscr{P} : \varphi(P) \leq 8^{n^2} C \delta^{-n} \right\}$ is finite, we see only finitely many $P \in \mathscr{P}$ can appear in the decomposition of Q_0. Now because in the decomposition of Q_0, there can be at most n factors, we see the set of Q_0 is finite. Finally, for a certain Q_0, there can only be finitely many Q's with $|Q(\xi)| \leq C \left\| Q \right\|^{-n}$. This is because $|Q(\xi)| = a|Q_0(\xi)|$ and $C \left\| Q \right\|^{-n} = a^{-n} C \left\| Q_0 \right\|^{-n}$. Therefore, we conclude that there can only be finitely many $Q \in \mathbb{Z}[x]$ such that $|Q(\xi)| \leq C \left\| Q \right\|^{-n}$, which contradicts Corollary 3.5. □

Remark 4.6. The constant C in Proposition 4.5 is non-computable. This is because we do not know what δ is.

Remark 4.7. Proposition 4.5 is the reason why we are able to strengthen Wirsing's theorem. In his paper [3] Wirsing only showed that $|P(\xi)| \ll \|P\|^{-n+\epsilon}$, so that Theorem 1.9 carries a ϵ in its statement. Here we kill this ϵ, so that we are able to produce Theorem 1.10.

5. The Proof of Wirsing's Theorem

In this section we will collect what we have in Section 2 and Section 4 to give a proof for Theorem 1.10. Our proof strategy is as follows: pick polynomials P, Q such that both $|P(\xi)|$ and $|Q(\xi)|$ are small, then use resultant of P and Q to give a bound on how close the roots of P or Q to ξ are, and finally split into 4 cases and analyze them separately. Because we are going to use resultant to give a certain bound, we will first define what resultant is.

Definition 5.1. Let $P(x) = a \prod_{\nu=1}^{s} (x - \alpha_\nu)$ and $Q(x) = b \prod_{\mu=1}^{t} (x - \beta_\mu)$ be two polynomials. Then their <u>resultant</u> is $\mathrm{Res}(P, Q) = a^t b^s \prod_{\nu,\mu} (\alpha_\nu - \beta_\mu)$.

The key property of resultant we will use is the following proposition.

Proposition 5.2. *For $P, Q \in \mathbb{Z}[x]$, we have* $\mathrm{Res}(P, Q) \in \mathbb{Z}$.

Proof. See [7, Definition 1.1 and Theorem 1.6]. □

We provide a further lemma that controls certain product of difference with ξ.

Lemma 5.3. *Fix $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$. Then for $P = c\prod_{i=1}^{r}(x-\alpha_i) \in \mathbb{Z}[x]$ of degree $r \leq n$, we have*

(1) $\prod_{|\xi-\alpha_i|\geq 1} |\xi - \alpha_i| \asymp \|P\|/c$.
(2) $\prod_{|\xi-\alpha_i|<1} |\xi - \alpha_i| \asymp |P(\xi)|/\|P\|$.

Proof. Part (1): this is exactly Corollary 2.6.

Part (2): we have $\prod |\xi - \alpha_i| = |P(\xi)|/c$, and this equation divides Part (1) we get the desired statement. ☐

Finally we are at a position to prove our main theorem.

Proof of Theorem 1.10. We use Proposition 4.5 to pick an irreducible $P \in \mathbb{Z}[x]$ such that $|P(\xi)| \ll \|P\|^{-n}$. We then use Lemma 3.3 to pick $Q \in \mathbb{Z}[x]$ such that $\|Q\| \leq 9^{-n}\|P\|$ and $|Q(\xi)| \ll \|P\|^{-n}$. Then we claim P and Q share no common root. To show the claim, suppose P and Q share a common root, then they are not coprime in $\mathbb{Q}[x]$, so by Proposition 4.3 and Proposition 4.4 we see $P \mid Q$ in $\mathbb{Z}[x]$. Write $Q = PR$ for some $R \in \mathbb{Z}[x]$, then $\|Q\| = \|PR\| \geq 8^{-n}\|P\|\|R\| \geq 8^{-n}\|P\|$ by Corollary 2.7, contradicts our assumption that $\|Q\| \leq 9^{-n}\|P\|$. Thus, we have shown that P and Q share no common root.

Let $P(x) = a\prod_{\nu=1}^{s}(x-\alpha_\nu), Q(x) = b\prod_{\mu=1}^{t}(x-\beta_\mu), p_\nu = |\xi-\alpha_\nu|, q_\mu = |\xi - \beta_\mu|$, and we order α_ν, β_μ in a way so that $p_1 \leq p_2 \leq \cdots \leq p_s, q_1 \leq q_2 \leq \cdots \leq q_t$. Because $P, Q \in \mathbb{Z}[x]$ we see $\mathrm{Res}(P,Q) = a^t b^s \prod_{\nu,\mu}(\alpha_\nu - \beta_\mu)$ is an integer by Proposition 5.2, and is nonzero because P, Q share no common root. Also, after taking $\|P\|$ big enough (so $|P(\xi)|$ and $|Q(\xi)|$ small enough) we can assume $p_1 < 1, q_1 < 1$.

We now have

$$\begin{aligned} 1 &\leq |\mathrm{Res}(P,Q)| \\ &= a^t b^s \prod_{\nu,\mu} |\alpha_\nu - \beta_\mu| \end{aligned}$$

Using $s, t \leq n$ and $|p-q| \leq 2\max(p,q)$ (when $p, q > 0$) we see

$$\ll a^n b^n \prod_{\nu,\mu} \max(p_\nu, q_\mu)$$

Now if we let $e_\nu = \#\{\mu : q_\mu < p_\nu\}, f_\mu = \#\{\nu : p_\nu \leq q_\mu\}$, then

$$\begin{aligned} &= a^n b^n \prod_{\nu} p_\nu^{e_\nu} \prod_{\mu} q_\mu^{f_\mu} \\ &\leq a^n b^n \prod_{p_\nu \geq 1} p_\nu^n \prod_{q_\mu \geq 1} q_\mu^n \prod_{p_\nu < 1} p_\nu^{e_\nu} \prod_{q_\mu<1} q_\mu^{f_\mu} \\ &\ll \prod_{p_\nu<1} p_\nu^{e_\nu} \prod_{q_\mu<1} q_\mu^{f_\mu} \|P\|^n \|Q\|^n \qquad \text{(by Lemma 5.3).} \end{aligned}$$

We claim that $e_1 = f_1 = 1$. If $e_1 \geq 2$ then $e_\nu \geq 2$, so

$$\begin{aligned}
1 &\ll \prod_{p_\nu<1} p_\nu^{e_\nu} \prod_{q_\mu<1} q_\mu^{f_\mu} \|P\|^n \|Q\|^n \\
&\leq \prod_{p_\nu<1} p_\nu^{e_\nu} \|P\|^n \|Q\|^n && \text{(by ignoring the second term)}, \\
&\leq \prod_{p_\nu<1} p_\nu^{2} \|P\|^n \|Q\|^n && \text{(since } e_v \geq 2\text{)}, \\
&\ll \frac{|P(\xi)|^2}{\|P\|^2} \|P\|^n \|Q\|^n && \text{(by Lemma 5.3)},
\end{aligned}$$

Using assumptions $|P(\xi)| \ll \|P\|^{-n}$ and $\|Q\| \leq 9^{-n} \|P\|$ we get

$$\ll \|P\|^{-2},$$

which is impossible as $\|P\| \to +\infty$.

Similarly, if $f_1 \geq 2$ then $f_\mu \geq 2$, so

$$\begin{aligned}
1 &\ll \prod_{p_\nu<1} p_\nu^{e_\nu} \prod_{q_\mu<1} q_\mu^{f_\mu} \|P\|^n \|Q\|^n \\
&\leq \prod_{q_\mu<1} q_\mu^{f_\mu} \|P\|^n \|Q\|^n && \text{(by ignoring the first term)}, \\
&\leq \prod_{q_\mu<1} q_\mu^{2} \|P\|^n \|Q\|^n && \text{(since } f_\mu \geq 2\text{)}, \\
&\ll \frac{|Q(\xi)|^2}{\|Q\|^2} \|P\|^n \|Q\|^n && \text{(by Lemma 5.3)},
\end{aligned}$$

Using assumptions $\|Q\| \leq 9^{-n} \|P\|$ and $|Q(\xi)| \ll \|P\|^{-n}$ we get

$$\ll \|Q\|^{-2},$$

which is impossible as $|Q(\xi)| \to 0^+$ (so that $\|Q\| \to +\infty$ by Proposition 2.1).

Thus, $e_1 = f_1 = 1$, and we see the order of (p_1, q_1, p_2, q_2) must fall into one of the 4 categories:

(1) $p_1 \leq q_1 < p_2 \leq q_2$.
(2) $p_1 \leq q_1 \leq q_2 < p_2$.
(3) $q_1 < p_1 \leq p_2 \leq q_2$.
(4) $q_1 < p_1 \leq q_2 < p_2$.

Case (1): $p_1 \leq q_1 < p_2 \leq q_2$. Then

$$\begin{aligned}
p_1^2 \leq p_1 q_1 &\ll \|P\|^n \|Q\|^n \prod_{p_\nu<1} p_\nu \prod_{q_\mu<1} q_\mu^2 \\
&\ll \|P\|^{n-1} \|Q\|^{n-2} \left|P(\xi)Q(\xi)^2\right| && \text{(by Lemma 5.3)},
\end{aligned}$$

Use $|P(\xi)| \ll \|P\|^{-n}, |Q(\xi)| \ll \|P\|^{-n}, \|Q\| \ll \|P\|$ we get

$$\begin{aligned}
&\ll \|P\|^{2n-3} \|P\|^{-3n} \\
&= \|P\|^{-n-3}.
\end{aligned}$$

Case (2): $p_1 \le q_1 \le q_2 < p_2$. Then

$$\begin{aligned} p_1^2 &\ll \|P\|^n \|Q\|^n \prod_{p_\nu<1} p_\nu^2 \prod_{q_\mu<1} q_\mu \\ &\ll \|P\|^{n-2} \|Q\|^{n-1} \left|P(\xi)^2 Q(\xi)\right| \qquad \text{(by Lemma 5.3)}, \end{aligned}$$

Use $|P(\xi)| \ll \|P\|^{-n}, |Q(\xi)| \ll \|P\|^{-n}, \|Q\| \ll \|P\|$ we get

$$\begin{aligned} &\ll \|P\|^{2n-3} \|P\|^{-3n} \\ &= \|P\|^{-n-3}. \end{aligned}$$

Case (3): $q_1 < p_1 \le p_2 \le q_2$. Then

$$\begin{aligned} q_1^2 &\ll \|P\|^n \|Q\|^n \prod_{p_\nu<1} p_\nu \prod_{q_\mu<1} q_\mu^2 \\ &\ll \|P\|^{n-1} \|Q\|^{n-2} \left|P(\xi) Q(\xi)^2\right| \qquad \text{(by Lemma 5.3)}, \end{aligned}$$

Use $|P(\xi)| \ll \|P\|^{-n}, |Q(\xi)| \ll \|P\|^{-n}$ we get

$$\begin{aligned} &\ll \|P\|^{-2n-1} \|Q\|^{n-2} \\ &\ll \|Q\|^{-n-3} \qquad \text{(since } \|P\| \gg \|Q\|). \end{aligned}$$

Case (4): $q_1 < p_1 \le q_2 < p_2$. Then

$$\begin{aligned} q_1^2 \le q_1 p_1 &\ll \|P\|^n \|Q\|^n \prod_{p_\nu<1} p_\nu^2 \prod_{q_\mu<1} q_\mu \\ &\ll \|P\|^{n-2} \|Q\|^{n-1} \left|P(\xi)^2 Q(\xi)\right| \qquad \text{(by Lemma 5.3)}, \end{aligned}$$

Use $|P(\xi)| \ll \|P\|^{-n}, |Q(\xi)| \ll \|P\|^{-n}$ we get

$$\begin{aligned} &\ll \|P\|^{-2n-2} \|Q\|^{n-1} \\ &\ll \|Q\|^{-n-3} \qquad \text{(since } \|P\| \gg \|Q\|). \end{aligned}$$

As a result, we see either $p_1^2 \ll \|P\|^{-n-3}$ or $q_1^2 \ll \|Q\|^{-n-3}$. In the former case, $|\xi - \alpha_1| = p_1 \ll \|\alpha_1\|^{-(n+3)/2}$ since P is irreducible. In the latter case, $|\xi - \beta_1| = q_1 \ll \|Q\|^{-(n+3)/2}$. Let $F \in \mathbb{Z}[x]$ be the minimal polynomial of β_1 and let $Q = FG$. Then $G \in \mathbb{Z}[x]$ by Proposition 4.3, and $\|Q\| \ge 8^{-n} \|F\| \|G\| \ge 8^{-n} \|F\|$ by Corollary 2.7, so $|\xi - \beta_1| \ll \|Q\|^{-(n+3)/2} \ll \|F\|^{-(n+3)/2} = \|\beta_1\|^{-(n+3)/2}$.

Now for any irreducible P, we have found γ_P such that $|\xi - \gamma_P| \le C \|\gamma_P\|^{-(n+3)/2}$ for some fixed constant C. Note that $p_1^n \le |P(\xi)| \ll \|P\|^{-n}$ and $q_1^n \le |Q(\xi)| \ll \|P\|^{-n}$, so because we are taking γ_P to be either α_1 or β_1 we see that $|\xi - \gamma_P|^n \ll \|P\|^{-n}$, or $|\xi - \gamma_P| \cdot \|P\| \ll 1$. As a result, when we let $\|P\| \to +\infty$ we see that $|\xi - \gamma_P| \to 0^+$, hence exhibits infinitely many γ such that $|\xi - \gamma| \ll \|\gamma\|^{-(n+3)/2}$. This finishes our proof. □

References

[1] A. Poëls: *On approximation to a real number by algebraic numbers of bounded degree*, Ann. Math. **201** (2025), 307-330.

[2] D. Northcott: *Periodic Points on an Algebraic Variety*, Ann. Math. **51** (1950), no. 1, 167-177.

[3] E. Wirsing: *Approximation mit algebraischen Zahlen beschränkten Grades*, J. Reine Angew. Math. **206** (1961), 67–77.

[4] J. F. Koksma: *Diophantische Approximationen*, Ergebn. Math. IV, 4 (1935).
[5] K. Mahler: *An application of Jensen's formula to polynomials*, Mathematika **7** (1960), 98-100.
[6] L. Robert: *A note on the measurability of convex sets*, Arch. Math. (Basel) **47** (1986), no. 1, 90–92.
[7] S. Janson: *Resultant and discriminant of polynomials*, 2007.
[8] Y. Bugeaud: *Approximation by Algebraic Numbers*, Cambridge Uniersity Press (2009).

Massachusetts Institute of Technology
Email address: liben@mit.edu