Polynomials Disguised in Different Senarios

李浩然　学号：13354154

中山大学　数学学院

摘要

在本论文里，我首先讨论了有限域上的Nikodym问题，有限域上的Kakeya问题，空间直线交点问题，空间点距问题[1]。这些问题之前人们一直都在不停的尝试，使用了各式各样的方法，拓扑学的方法，分析学的方法等等，但却没能取得什么大的进展。而这里使用多项式以及一些简单的线性代数，就得以使这些问题获得解答，其核心思想在论文中有所总结，也就是消失引理
接下来，我有讨论一个代数几何中基本结果，Bézout定理，我选择了一个使用较少的证明方法[2]，但我却认为这个方法更加简洁，美妙并且自然，这里仅仅用到一个简单的概念，结式，定理用两条互素并且既约的代数曲线的次数给出了它们交点个数的上界
最后是一个丢番图逼近中的重要定理，Roth定理[3][4]，此定理主要依赖于有理系数多项式在有理点和一般点上的不同表现
[**关键词**]组合几何学; Bézout定理; Roth定理

Abstract

In this thesis, first I talk about the finite field Nikodym problem, the finite field Kakeya problem, the joints problem and the distinct distance problem[1], all these problems have been tried in many ways before, topology, analysis, etc. It was very hard to make some progress. However, using polynomials and some basic linear algebra, these problems could be solved easily, all of them are based on a crucial fact summrized in the thesis as vanishing lemma
Next I present a fundamental result in algebraic geometry, Bézout's theorem[2], I chose a less used proof method, which I think is more succint, elegant and natural, using a simple notion, resultant, which roughly states a upper bound for intersections of two coprime irreducible algebraic curves
In the final part it's about the renowned Roth's theorem[3][4], rely mainly on the different behaviors of polynomials of rational coefficients on rational points and general points
[**Keywords**] Combinatorial geometry; Bézout's theorem; Roth's theorem

目录

1 Introduction

Polynomial, a simple notion that everybody think they know everything about it, but in fact we are often so shocked that we hardly know anything about it, especially when the degree or its coefficients changes. Yet it has so many beautiful and elegant structures hidden behind it, driving generations of generations of brilliant mathematicians to explore its wonderful mestery This thesis mainly focuses on only some problems concerning a few combinatorial and algebraic structures of polynomials which otherwise would be very hard to solve. For example, the finite field Kakeya problem has been tried in many ways without succeeding, however it could be easily solved by some arguments in linear algebra. So is many other problems in this thesis, all the proofs are quite elementary, involving only some linear or abstract algebra, including the famous Roth's theorem. In essence, all the proofs are based on some simple algebraic structures(vector spaces, rings, etc.) of polynomials and how they behave on the underlying fields (especially vanishing property).

2 Combinatorial Geometry

2.1 Preliminaries

definition 1

Definition 1. Let $\mathrm{Poly}(\mathbb{F}^n)$ be the set of polynomials in n variables with coefficients in $\mathbb{F}$, and $\mathrm{Poly}_D(\mathbb{F}^n)$ be the set of polynomials in n variables with coefficients in $\mathbb{F}$ and with degree no more than D. Both of them are vector spaces over the field $\mathbb{F}$

proposition 2

Proposition 2. Suppose $S \subset \mathbb{F}^n$, if $\dim \mathrm{Poly}_D(\mathbb{F}^n) > |S|$, then there is a non-zero polynomial $Q \in \mathrm{Poly}_D(\mathbb{F}^n)$ vanishes on S

Proof. Let $S = \{p_1, \cdots, p_{|S|}\}$. Consider the evaluation map

$$\phi : \mathrm{Poly}_D(\mathbb{F}^n) \to \mathbb{F}^{|S|}$$

$$Q \mapsto \big(Q(p_1), \cdots, Q(p_{|S|})\big)$$

which is a linear map with $\ker\phi$ being the set of polynomials in $\mathrm{Poly}_D(\mathbb{F}^n)$ that vanish on S. If $\dim \mathrm{Poly}_D(\mathbb{F}^n) > |S|$, then $\ker\phi$ must be non-trivial. □

So it is a natural question to ask: how to compute $\dim \mathrm{Poly}_D(\mathbb{F}^n)$?

proposition 3

Proposition 3.

$$\dim \mathrm{Poly}_D(\mathbb{F}^n) = \binom{D+n}{n}$$

Proof. Since

$$A = \left\{ x_1^{D_1} \cdots x_n^{D_n} \,\middle|\, D_i \in \mathbb{N}, D_1 + \cdots + D_n \le D \right\}$$

form a basis of $\mathrm{Poly}_D(\mathbb{F}^n)$, define

$$B = \left\{ (j_1, \cdots, j_n) \,\middle|\, j_i \in \mathbb{Z}_+, j_1 < \cdots < j_n \le D + n \right\}$$

then there is a natural bijective correspondence between A and B.

$$x_1^{D_1} \cdots x_n^{D_n} \mapsto (D_1 + 1, \cdots, D_1 + \cdots + D_n + n)$$

$$(j_1, \cdots, j_n) \mapsto x_1^{j_1 - 1} \cdots x_n^{j_n - j_{n-1} - 1}$$

thus $|A| = |B| = \binom{D+n}{n}$ □

Remark. Note that

$$\dim \mathrm{Poly}_D(\mathbb{F}^n) > \frac{D^n}{n!} \ge \frac{(D+1)^n}{n^n}$$

Combining Proposition 2 and 3, we have

corollary 4

Corollary 4. $S \subset \mathbb{F}^n$, if $|S| < \binom{D+n}{n}$, then there is a non-zero polynomial $P \in \text{Poly}_D(\mathbb{F}^n)$ that vanishes on S

For computational convenience, we mostly use another corollary

corollary 5

Corollary 5. Let $S \subset \mathbb{F}^n$ be a finite set, there is a non-zero polynomial that vanishes on S with degree no more than $n|S|^{\frac{1}{n}}$

Proof. Let $D = \left\lfloor n|S|^{\frac{1}{n}} \right\rfloor$, then

$$n|S|^{\frac{1}{n}} < D + 1 \Rightarrow |S| < \frac{(D+1)^n}{n^n} < \binom{D+n}{n}$$

By Corollary 4, there is a non-zero polynomial $P \in \text{Poly}_D(\mathbb{F}^n)$ that vanishes on S □

Now let's consider the behavior of a polynomial on lines

proposition 6

Proposition 6. If $P \in \text{Poly}_D(\mathbb{F}^n)$ vanishes at $D+1$ points on a line l, then P vanishes on l

Proof. There is a map

$$\gamma : \mathbb{F} \to \mathbb{F}^n, t \mapsto at + b$$

Where $a, b \in \mathbb{F}^n, a \neq 0$, and define $Q(t) = P(\gamma(t)) = P(at+b)$ which is a polynomial with degree no more than D, since Q has $D+1$ zeros, $Q = 0$, hence P vanishes on l □

2.2 The finite field Nikodym problem

definition 7

Definition 7. $\mathbb{F}_q$ is the finite field of q elements. A set $N \subset \mathbb{F}_q^n$ is a Nikodym set if, for each point $x \in \mathbb{F}_q^n$, there is a line L through x such that $L - x \subset N$

proposition 8

Proposition 8. If $P \in \text{Poly}_{q-1}(\mathbb{F}_q^n)$ vanishes on all the points of $\mathbb{F}_q^n$, then $P = 0$

Proof. When $n = 1$, this statement is obviously true, we use induction on n
Since

$$P(x_1, \cdots, x_n) = \sum_{i=0}^{q-1} P_i(x_1, \cdots, x_{n-1}) x_n^i, \deg P_i \leq q - 1$$

Fix any values of $x_1, \cdots, x_{n-1}$, this is a polynomial with degree no more than $q-1$ hence a zero polynomial, then by induction we have $P_i = 0$, which concludes that $P = 0$ □

As it turns out, a Nikodym set cannot be too small

theorem 9

Theorem 9. *If $N \subset \mathbb{F}_q^n$ is a Nikodym set, then there is a constant c_n such that $|N| \geq c_n q^n$*

Proof. Suppose $N \subset \mathbb{F}_q^n$ is a Nikodym set such that $|N| < cq^n$, by Corollary 5, there exists a non-zero polynomial P vanishes on N with $\deg P \leq n|N|^{\frac{1}{n}}$, assume

$$\deg P < q - 1$$

then for every $x \in \mathbb{F}_q^n$, P will be vanishing on $q-1$ points and thus vanishes on x according to Proposition 6
We have thus concluded that P vanishes on all the points of $\mathbb{F}_q^n$ which contradicts Proposition 8. All we need to do is to find a constant c that satisfies the inequality

$$\deg P \leq n|N|^{\frac{1}{n}} < nc^{\frac{1}{n}} q \leq q - 1$$

And any number $c_n \leq \left(\frac{q-1}{nq}\right)^n$ should suffice □

2.3 The finite field Kakeya problem

definition 10

Definition 10. A set $K \subset \mathbb{F}_q^n$ is a Kakeya set if it contains a line in every direction.

Just like a Nikodym set, a Kakeya set cannot be too small as well

theorem 11

Theorem 11. *If $K \subset \mathbb{F}_q^n$ is a Kakeya set, then there is a constant c_n such that $|K| \geq c_n q^n$*

Proof. Suppose $K \subset \mathbb{F}_q^n$ is a Kakeya set such that $|K| < cq^n$, by Corollary 5, there exists a non-zero polynomial P vanishes on K with $\deg P \leq n|K|^{\frac{1}{n}}$, assume

$$\deg P < q$$

Let $P = H + Q$, where H is the homogeneous polynomial of highest degree which is also non-zero, every line in $\mathbb{F}_q^n$ is of the form $\left\{at + b \middle| t \in \mathbb{F}_q^n, a \neq 0\right\}$, thus $P(at+b)$ is a zero polynomial in t, and this implies that $H(a) = 0$, hence H vanishes on $\mathbb{F}_q^n$ which contradicts Proposition 8. All we need to do is to find a constant c that satisfies the inequality

$$\deg P \leq n|K|^{\frac{1}{n}} < nc^{\frac{1}{n}} q \leq q$$

And any number $c_n \leq \dfrac{1}{n^n}$ should suffice □

2.4 The joints problem

definition 12

Definition 12. $\mathscr{L}$ is a set of lines in $\mathbb{R}^n$, a r-rich point is a point lies in r lines of $\mathscr{L}$ which are linearly independent, denote the set of r-rich points $P_r(\mathscr{L})$.

You may be interested in how many joints that L lines can form

theorem 13

Theorem 13. *$\mathscr{L}$ is a set of L lines in $\mathbb{R}^3$, then there is a constant C such that $|P_r(\mathscr{L})| \leq CL^{\frac{3}{2}}$*

Proof. First note that if $\mathscr{L}$ is a set of lines with J joints, then there is one line that could contain at most $3J^{\frac{1}{3}}$ joints. Otherwise, there is a polynomial P with the least degree $\deg P \leq 3J^{\frac{1}{3}}$ that vanishes on all the joints. But then it vanishes on more than $3J^{\frac{1}{3}}$ points on each line, thus P vanishes on $\mathscr{L}$, at each joint, all the directional derivatives would be zero. Hence any derivative of P is a polynomial with degree less than $\deg P$ but still vanishes on all the joints, which is a contradiction
Let $J(L)$ be the maximal number of joints that L lines can form, clearly we have

$$J(L-1) \leq J(L)$$

then there is one line that contains at most $3J(L)^{\frac{1}{3}}$ joints, thus we have

$$J(L) \leq J(L-1) + 3J(L)^{\frac{1}{3}} \leq J(L-2) + 2 \cdot 3J(L)^{\frac{1}{3}} \leq \cdots \leq L \cdot 3J(L)^{\frac{1}{3}}$$

which give us $J(L) \leq CL^{\frac{3}{2}}$ □

2.5 Generalization of Vanishing lemma

lemma 14

Lemma 14. Suppose $P \in \mathrm{Poly}_D(\mathbb{R}^n)$ vanishes on $D+1$ distinct $(n-1)$-dimensional planes, then $P = 0$

Proof. For any point x not on any one of the planes, we can draw a line that intersects $D+1$ distinct points with these planes, thus $P(x) = 0$, therefore $P = 0$ □

Corollary 5 could be captured as a vanishing lemma,the Following Theorem could be seen as a generalization of Corollary 5

theorem 15

Theorem 15. *Given N distinct k-dimensional planes in $\mathbb{R}^n$, Then there is a polynomial P with degree $\deg P \leq CN^{\frac{1}{n-k}}$ that vanishes on all the planes, where C depends only on k, n*

Proof. Restrict on a k-dimensional plane, P would be a polynomial in $\mathrm{Poly}_D(\mathbb{R}^{n-k})$, If we could find $D+1$ distinct $(k-1)$-dimensional planes on which P vanishes, then P would vanish on this k-dimensional plane by Lemma 14. Carrying over this procecedure, If we could find $N(D+1)^k$ points on which P vanishes, then P vanishes on all these intermediate affine subspaces. Accroding to Corollary 5, If we could choose D so that

$$D \leq nN^{\frac{1}{n}}(D+1)^{\frac{k}{n}}$$

then there exists a polynomial P satisfies all these requirements

$$\begin{aligned}
&D \leq nN^{\frac{1}{n}}(D+1)^{\frac{k}{n}} \leq n2^{\frac{k}{n}}N^{\frac{1}{n}}D^{\frac{k}{n}} \\
&\Rightarrow D^{\frac{n-k}{n}} \leq n2^{\frac{k}{n}}N^{\frac{1}{n}} \\
&\Rightarrow D^{n-k} \leq n^n 2^k N \\
&\Rightarrow D \leq n^{\frac{n}{n-k}}2^{\frac{k}{n-k}}N^{\frac{1}{n-k}} \leq CN^{\frac{1}{n-k}}
\end{aligned}$$

□

theorem 16

Theorem 16. *If $P \in \mathrm{Poly}_D(\mathbb{R}^n)$ is a non-zero polynomial, then $Z(P)$, the zero set of P, has Lebesgue measure zero.*

Proof. We prove this by induction on n, we denote Lebesgue measure as μ
Assume

$$P = (x_n - a_1)^{n_1} \cdots (x_n - a_k)^{n_k} P_1$$

Restricts on hyperplane $x_n = a$, P_1 would be a non-zero polynomial in $\mathrm{Poly}_D(\mathbb{R}^{n-1})$, by induction hypothesis, $\mu(Z(P_1)) = 0$, Since $Z(P)$ is a closed set in $\mathbb{R}^n$, $\chi_{Z(P)}$ is a nonnegative measurable function, according to Tonelli's Theorem, we have

$$\begin{aligned}
\mu(Z(P)) &= \int_{\mathbb{R}^n} \chi_{Z(P)} \\
&\leq \int_{\mathbb{R}^n} \sum_{j=1}^{k} \chi_{Z(x_n - a_j)} + \chi_{Z(P_1)} \\
&= \int_{\mathbb{R}^n} \sum_{j=1}^{k} \chi_{Z(x_n - a_j)} + \int_{\mathbb{R}^n} \chi_{Z(P_1)} \\
&= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \sum_{j=1}^{k} \chi_{Z(x_n - a_j)} \right) + \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} \chi_{Z(P_1)} \right) \\
&= 0
\end{aligned}$$

□

2.6 Distinct distance problem

definition 17

Definition 17. Let $P \subset \mathbb{R}^n$, define

$$\mathrm{dist}(P) := \left\{ |p-q| \middle| p, q \in P \right\}$$

theorem 18

Theorem 18. *$P \subset \mathbb{R}^n$ satisfies $|\mathrm{dist}(P)| \leq s$, Then $|P| \leq \binom{n+s+1}{s}$*

Proof. Let $P = \{p_1, \cdots, p_N\}, \mathrm{dist}(P) = \{d_1, \cdots, d_s\}$, then for each $p_j \in P$, define

$$f_j(x) = \prod_{r=1}^{s} \left(|x - p_j|^2 - d_r^2 \right)$$

then

$$f_j(p_i) = \delta_{ij}(-1)^s \prod_{r=1}^{s} d_r^2$$

where δ_{ij} is the Kronecker delta
hence f_j are linearly independent in $\mathrm{Poly}(\mathbb{R}^n)$. On the other hand,

$$|x - p_j|^2 - d_r^2 = |x|^2 - 2p_j \cdot x + |p_j|^2 - d_r^2$$

is generated by $1, x_1, \cdots, x_n, |x|^2$, thus

$$f_j(x) = \prod_{r=1}^{s} \left(|x|^2 - 2p_j \cdot x + |p_j|^2 - d_r^2\right)$$

is generated by $\binom{n+s+1}{s}$ polynomials in $\mathrm{Poly}(\mathbb{R}^n)$ by a similar argument as Proposition 3, so is any linear combination of f_j, Therefore, we have

$$|P| = N \leq \binom{n+s+1}{s}$$

□

3 Bézout's Theorem

3.1 Sylvester matrix and resultant

Next, we will talk about Bézout's Theorem

definition 19

Definition 19. For $p, q \in \mathrm{Poly}(\mathbb{F})$, Assume

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

$$q(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n$$

Define its Sylvester matrix $S(p, q)$ to be a $(m + n) \times (m + n)$ matrix

$$S(p,q) = \begin{bmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdots & 0 \\ a_2 & a_1 & \ddots & 0 & b_2 & b_1 & \ddots & 0 \\ \vdots & \vdots & \ddots & a_0 & \vdots & \vdots & \ddots & b_0 \\ a_m & a_{m-1} & \cdots & \vdots & b_n & b_{n-1} & \cdots & \vdots \\ 0 & a_m & \ddots & \vdots & 0 & b_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{m-1} & \vdots & \vdots & \ddots & b_{n-1} \\ 0 & 0 & \cdots & a_m & 0 & 0 & \cdots & b_n \end{bmatrix}$$

Define the resultant $\mathrm{Res}(p, q)$ of p, q to be $\det S(p, q)$, and using (p, q) for the greatest common divisor of p, q
Define $\mathrm{Coe}(p)$ to be the vector of coefficients

$$\mathrm{Coe}(p) = [a_0, \cdots, a_n]^{\mathrm{T}}$$

theorem 20

Theorem 20. $\mathrm{Res}(p, q) = 0$ *if and only if* p, q *has common roots.*

Proof. Since

$$(p, q) = d \Leftrightarrow sp + tq = d$$

For unique $s, t \in \mathrm{Poly}(\mathbb{F})$ with $\deg s < \deg q - \deg d, \deg t < \deg p - \deg d$, this result is also known as Bézout's Theorem

Moreover

$$sp + tq = 0, \text{for some } s, t \text{ with } \deg s < \deg q, \deg t < \deg p \Leftrightarrow (p, q) = d \neq 1$$

Where $\Rightarrow$ is true because of the uniqueness of Bézout's Theorem, and $\Leftarrow$ is true because we could take $s = q/d, t = -p/d$
Suppose

$$s(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_k x^k$$
$$t(x) = d_0 + d_1 x + d_2 x^2 + \cdots + d_l x^l$$

Then

$$Coe(sp + tq) = S(p, q) \begin{bmatrix} Coe(p) \\ Coe(q) \end{bmatrix}$$

Hence

$$sp + tq = 0, \text{for some } s, t \text{ with } \deg s < \deg q, \deg t < \deg p \Leftrightarrow Res(p, q) = 0$$

□

Now we introduce a version of Fundamental Theorem of Algebra in $\mathbb{CP}^1$

`theorem 21`

Theorem 21. *Suppose $p \in \text{Poly}(\mathbb{C}^2)$ is a homogenous polynomial of degree n then it can be factored into linear factors.*

Proof. Assume

$$p(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_{n-1} x y^{n-1} + a_n y^n$$

Define

$$q(z) = a_0 + a_1 z + \cdots + a_n z^n$$

Then by the usual Fundamental Theorem of Algebra on $\mathbb{C}$, hence $q(z)$ can be factored into linear factors

$$q(z) = C(z - \lambda_1) \cdots (z - \lambda_n)$$

Therefore we have

$$p(x, y) = C(y - \lambda_1 x) \cdots (y - \lambda_n x)$$

□

3.2 Bézout's Theorem

The next result is formally known as Bézout's Theorem

`theorem 22`

Theorem 22. *Suppose two curves in $\mathbb{C}^2$ defined by polynomials $p, q \in \text{Poly}(\mathbb{C}^2)$ with $(p, q) = 1$ and $\deg p = m, \deg q = n$, then the number of intersections of the two curves is no more than mn*

Remark. $(p, q) = 1$ means that p and q don't have common factors, otherwise p, q would have infinitely many common roots

Proof. Suppose p, q have more than mn common roots, say, $mn+1$ common roots $P_1, \cdots, P_{mn+1}$, than after an affine transformation, we could assume that neither two of lines OP_i will overlap
Suppose the homogenization of p, q are

$$p(x, y, z) = \sum_{i+j+k=m} c_{ijk} x^i y^j z^k = p_0(x, y) z^m + p_1(x, y) z^{m-1} + \cdots + p_m(x, y)$$

$$q(x, y, z) = \sum_{i+j+k=n} d_{ijk} x^i y^j z^k = q_0(x, y) z^n + q_1(x, y) z^{n-1} + \cdots + q_n(x, y)$$

Where $[x, y, z]$ is the homogenous coordinate, and $p_i(x, y)$ are homogenous polynomial of degree i
We could consider $\text{Res}(p, q)$ which would be a homogenous polynomial of degree mn, and the proof of this fact we defer to the next lemma. And $\text{Res}(p, q) \neq 0$, otherwise f, g would have common roots in $k(x, y)[z]$, by Gauss's lemma, f, g would have common factors in $k[x, y, z]$
According to Theorem 21, $\text{Res}(p, q)$ could be factored into mn linear factors, then each P_i should lie on different linear factor, which is a contradiction □

lemma 23

Lemma 23. Res(p, q) is a homogenous polynomial of degree mn

Proof. we would prove that each term in the $(m+n)!$ terms of the expansion of Sylvester matrix $S(p, q)$ is a homogenous polynomial of degree mnassume one term is

$$\pm S_{i_1,1} \cdots S_{i_n,n} S_{i_{n+1},n+1} \cdots S_{i_{m+n},m+n}$$

Where $(i_1 \cdots i_{m+n}) \in S_{m+n}$ is a permutation of $1, \cdots, m+n$, if we want this term not to be zero, we need further to assume

$$j \leq i_j \leq j+m, j = 1, \cdots, n$$

$$j \leq i_{n+j} \leq j+n, j = 1, \cdots, m$$

Due to the arranging of $S(p, q)$, we have the degree d of this term satisfies

$$\sum_{k=1}^{m+n} k = \sum_{j=1}^{m+n} i_j = d + \sum_{k=1}^{n} k + \sum_{k=1}^{m} k$$

Thus we have $d = mn$ □

4 Roth's Theorem

4.1 History and developement

Our last topic would be a famous result in Diophantine approximation, Thue-Siegel-Roth's Theorem. But before we get to Roth's Theorem, I would like to talk about some history and developement

theorem 24

Theorem 24. *For any irrational number α ,there are infinitely many rational numbers $\frac{p}{q}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

Remark. This theorem is known as Dirichlet's theorem, and note that the theorem won't hold if α is rational

Proof. Define $\langle x \rangle = x - \lfloor x \rfloor$, thus

$$\{0, \langle\alpha\rangle, \langle 2\alpha\rangle, \cdots, \langle n\alpha\rangle\} \subset [0, 1)$$

Obviously, no two can be the same, assume

$$0 < \langle i_1\alpha\rangle < \langle i_2\alpha\rangle < \cdots < \langle i_n\alpha\rangle$$

hence there must be two adjacent terms with difference strictly less than $\frac{1}{n}$, thus there exist integer $q \leq n$ and integer p such that

$$|q\alpha - p| < \frac{1}{n} \leq \frac{1}{q} \Rightarrow \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

as n becomes larger, q must also becomes larger, otherwise $|q\alpha - p|$ won't be small enough, if $(p, q) = d$, then $|q\alpha - p| = d\left|\frac{q}{d}\alpha - \frac{p}{d}\right|$, hence if there are only finitely many rationals $\frac{r}{s}$ where $(r, s) = 1$, then $|s\alpha - r|$ must have a lower bound, thus as q becomes larger, $|q\alpha - p|$ will certainly be very large, which is a contradiction □

theorem 25

Theorem 25. *The set S of those $x \in \mathbb{R}$ such that there exist infinitely many fractions $\frac{p}{q}, (p, q) = 1$ such that $\left|x - \frac{p}{q}\right| \leq \frac{1}{q^{2+\epsilon}}$ is of measure zero, where $\epsilon > 0$*

Remark. Note that the condition could be changed into a weaker one

$$\left|x-\frac{p}{q}\right| \leq \frac{1}{q^2\log(q)^{1+\epsilon}}$$

Proof. First, let's concentrate on $x \in [0,1]$, consider

$$E_k = \left\{ x \in S \cap [0,1] \middle| \left|x-\frac{l}{k}\right| \leq \frac{1}{k^{2+\epsilon}}, 0<l<k, (k,l)=1 \right\}$$

Then

$$\mu(E_k) \leq 2 \cdot \frac{1}{k^{2+\epsilon}} \cdot (k+1) \leq \frac{4}{k^{1+\epsilon}}$$

Thus $\sum_{k=1}^{\infty} \mu(E_k) < \infty$

According to Borel-Cantelli lemma, $\mu(S \cap [0,1]) = 0$, thus $\mu(S) = \sum_{n=-\infty}^{+\infty} \mu(S \cap [n-1,n]) = 0$ □

theorem 26

Theorem 26. *If α is an irrational algebraic number of degree n, then there exists $A>0$ such that*

$$\forall p,q \in \mathbb{Z}, q>0, \left|\alpha-\frac{p}{q}\right| > \frac{A}{q^n}$$

Remark. This theorem is known as Liouville's theorem, and note that the theorem won't hold if α is rational

Proof. Assume α is a root of an polynomial $f \in \mathbb{Z}[x]$ of degree n, denote the maximum of $|f'(x)|$ in $[\alpha-1, \alpha+1]$ as M, and assume all the distinct roots of f in $\mathbb{R}$ are $\alpha_1 < \alpha_2 < \cdots < \alpha_m$. Choose A small enough so that

$$A < \min\left(1, \frac{1}{\mathrm{M}}, |\alpha_{\mathrm{i}} - \alpha_{\mathrm{j}}|\right)$$

Suppose there is a fraction $\frac{p}{q}, q>0, (p,q)=1$ such that $\left|x-\frac{p}{q}\right| \leq \frac{A}{q^n}$, then $f\left(\frac{p}{q}\right) \neq 0$ and $\left|f\left(\frac{p}{q}\right)\right| \geq \frac{1}{q^n}$. By the intermediate theorem

$$f(\alpha) - f\left(\frac{p}{q}\right) = f'(x_0)\left(\alpha-\frac{p}{q}\right)$$

For some $x_0 \in [\alpha-1, \alpha+1]$, hence

$$\left|\alpha-\frac{p}{q}\right| = \frac{\left|f(\frac{p}{q})\right|}{|f'(x_0)|} \geq \frac{1}{Mq^n} > \frac{A}{q^n}$$

Which would be a contradiction □

In fact, Liouville use this to construct the first transcendental number

$$\ell = \sum_{k=1}^{\infty} \frac{1}{10^{k!}}$$

Suppose ℓ is an algebraic number of degree n, then by Liouville's theorem

$$\exists A, \left|\ell-\frac{p}{q}\right| \geq \frac{A}{q^n}$$

Consider $\frac{p_m}{q_m} = \sum_{k=1}^{m} \frac{1}{10^{k!}}$ and $q_m = \frac{1}{10^{m!}}$, then

$$\begin{aligned}\left|\ell - \frac{p_m}{q_m}\right| &= \sum_{k=m+1}^{\infty} \frac{1}{10^{k!}} \\ &= \frac{1}{10^{(m+1)!}} \sum_{k=0}^{\infty} \frac{1}{10^{(k+m+1)!-(m+1)!}} \\ &\leq \frac{1}{10^{(m+1)!}} \sum_{k=0}^{\infty} \frac{1}{10^{k}} \\ &= \frac{1}{10^{(m+1)!}} \frac{10}{9}\end{aligned}$$

Thus

$$\frac{A}{q^n} \leq \left|\ell - \frac{p}{q}\right| \leq \frac{1}{10^{(m+1)!}} \frac{10}{9} \Rightarrow$$
$$A \leq \frac{10}{9} 10^{(n-m-1)m!} \to 0 (m \to \infty)$$

Which would be a contradiction, hence ℓ is a transcendental number
Another way of saying Roth's theorem is

$$\forall \kappa > 2, \exists C_{\alpha,\kappa} > 0, \left|\alpha - \frac{p}{q}\right| \geq \frac{C_{\alpha,\kappa}}{q^{\kappa}}$$

However, simply use polynomial of one variable can at best reach Liouville's result in some sense
If we pick $P(x) \in \mathbb{Q}[x]$ with $\deg P = r$, and α is a root of order i, then

$$P\left(\frac{p}{q}\right) = \sum_{j=i}^{r} \frac{P^{(j)}(\alpha)}{j!} \left(\frac{p}{q} - \alpha\right)^j$$

Then we have

$$\frac{1}{q^r} \leq P\left(\frac{p}{q}\right) \leq \frac{C}{q^{i\kappa}}$$

Thus we should have

$$i\kappa \leq r \Leftrightarrow \kappa \leq \left(\frac{i}{r}\right)^{-1}$$

Hence we should make $\frac{i}{r}$ as large as possible, but if α has order i, then $r \geq ni \Leftrightarrow \frac{1}{n} \geq \frac{1}{n}$, thus we didn't get a better result than Liouville's. That is a motivation for using polynomials of multiple variables

4.2 Thue-Siegel-Roth's Theorem

Since Thue-Siegel-Roth's Theorem is rather complicated, we will using a flexible proof technique, postpone the details and the technical part of the theorem, first present its basic idea, make it more readable, but before that, let's introduce some notations first

`definition 27`

Definition 27. Let $R = (r_1, \cdots, r_m), I = (i_1, \cdots, i_m) \in \mathbb{N}^m$ be a multi-index, and define $\text{Poly}_R(\mathbb{F}^m)$ to be the set of polynomials with degree at x_i no more than r_i, and $|I| = i_1 + \cdots + i_m$, $\frac{I}{R} := (\frac{i_1}{r_1}, \cdots, \frac{i_m}{r_m})$

`theorem 28`

Theorem 28. *If α is an irrational algebraic number, then for any $\kappa > 2$, $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{\kappa}}$ has only finitely many solutions*

Remark. This theorem is known as Thue-Siegel-Roth's Theorem, which could be seen as a generalization of Dirichlet's Theorem and Liouville's theorem, and this result is sharp in the sense that the solutions of $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{\kappa}}$ is infinitely many if $\kappa = 2$ and finitely many if $\kappa > 2$

Proof. **Step 1(Simplification):**
First we simplify the question, we could assume α to be an algebraic integer
Suppose the theorem holds for all algebraic integers, if α is an algebraic number of degree n with minimal polynomial

$$m_\alpha(x) = x^n + \frac{s_{n-1}}{r_{n-1}}x^{n-1} + \cdots + \frac{s_1}{r_1}x + \frac{s_0}{r_0}$$

Where $\frac{s_{n-1}}{r_{n-1}}, \cdots, \frac{s_0}{r_0}$ are rational numbers, let $M = r_{n-1}\cdots r_0$, then we have

$$(M\alpha)^n + a_{n-1}(M\alpha)^{n-1} + \cdots + a_1(M\alpha) + a_0 = 0$$

Where $a_{n-1}, \cdots, a_0$ are integers, thus $M\alpha$ is an algebraic integer, if

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^\kappa}$$

has infinitely many solutions, then q could be taken sufficiently large, thus there exists $\kappa' > 2$ so that

$$\left|M\alpha - \frac{Mp}{q}\right| \leq \frac{M}{q^\kappa} \leq \frac{1}{q^{\kappa'}}$$

Where q could be sufficiently large, on the other hand

$$\left|M\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{\kappa'}}$$

Should have only finitely many solutions, thus q should have an upper bound, which would be a contradiction
Second, we could assume all the rational approximations to be in reduced forms, for if $\frac{p}{q}$ satisfies

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^\kappa}$$

$\frac{p}{q}$ could be writen as $\frac{rp_1}{rq_1}$ where $(p_1, q_1) = 1$, then

$$\left|\alpha - \frac{p}{q}\right| = \left|\alpha - \frac{p_1}{q_1}\right| \leq \frac{1}{q^\kappa} \leq \frac{1}{q_1^\kappa}$$

but there could only be finitely many $\frac{rp_1}{rq_1}$ satisfy the inequality, since $\left|\alpha - \frac{p_1}{q_1}\right|$ is a fixed number, thus, there are infinitely many reduced rationals satisfy the inequality as long as there are infinitely many rationals satisfy the inequality
Step 2(Idea):
Assume α is an algebraic integer of degree n with minimal polynomial

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

The basic idea is to find a polynomial $Q \in \mathrm{Poly}_R(\mathbb{Z}^m)$ of somewhat bounded coefficients such that $Q\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) \neq 0$ but vanishes to a very high order at $(\alpha, \cdots, \alpha)$, with $\frac{h_i}{q_i}$ be reduced good approximations that satisfies $\left|\alpha - \frac{h_i}{q_i}\right| \leq \frac{1}{q_i^\kappa}$, this would raise a contradiction
We have

$$Q\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) \geq \frac{1}{q_1^{r_1}\cdots q_m^{r_m}}$$

on one hand, and on the other hand, since

$$Q\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) = \sum_{0 \leq I \leq R} \partial_I Q(\alpha, \cdots, \alpha)\left(\frac{h_1}{q_1} - \alpha\right)^{i_1} \cdots \left(\frac{h_m}{q_m} - \alpha\right)^{i_m}$$

And

$$\left|\left(\frac{h_1}{q_1}-\alpha\right)^{i_1}\cdots\left(\frac{h_m}{q_m}-\alpha\right)^{i_m}\right|\leq\frac{1}{(q_1^{i_1}\cdots q_m^{i_m})^{\kappa}}$$

Pick m,q_1,r_1 to be sufficiently large, pick $q_2,\cdots,q_m$ even larger subsequently, and to simplify things, select $r_2,\cdots,r_m$ so that $q_j^{r_j}\approx q_1^{r_1}$, thus

$$\frac{1}{q_1^{r_1}\cdots q_m^{r_m}}\gtrsim q_1^{-mr_1}$$

And

$$\frac{1}{(q_1^{i_1}\cdots q_m^{i_m})^{\kappa}}=\left(q_1^{-r_1\frac{i_1}{r_1}}\cdots q_m^{-r_m\frac{i_m}{r_m}}\right)^{\kappa}\approx q_1^{-r_1\kappa(\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m})}=q_1^{-r_1\kappa|\frac{I}{R}|}$$

Define

$$\gamma=\min\left\{\left|\frac{I}{R}\right|\middle|\partial_I Q(\alpha,\cdots,\alpha)\neq 0\right\}$$

as a measure how much order does Q vanish at $(\alpha,\cdots,\alpha)$ by the name index of Q at $(\alpha,\cdots,\alpha)$ with respect to R. When q_1 is quite large, then all $\partial_I Q(\alpha,\cdots,\alpha)$ would be somewhat bounded so that

$$\left|Q\left(\frac{h_1}{q_1},\cdots,\frac{h_m}{q_m}\right)\right|\lesssim q_1^{-r_1\kappa\gamma}$$

Hence

$$q_1^{-mr_1}\lesssim q_1^{-r_1\kappa\gamma}\Rightarrow -mr_1\leq -r_1\kappa\gamma$$

But in the construction of Q, we could make sure $\dfrac{\gamma}{m}$ and $\dfrac{1}{2}$ are as close as possible, therefore we have achieved a contradiction

$$\gamma\kappa>m$$

Now let's turn to the construction of Q, choose m to be sufficiently large, then choose $\delta>0$ to be small enough, and then choose q_1 very large, which will be chosen later, and then choose $q_2,\cdots,q_m$, we need to make sure $\frac{r_{j-1}}{r_j}>\delta^{-1}$ for latter construction, since we want

$$q_1^{r_1}\approx q_m^{r_m}\Rightarrow\frac{\log q_j}{\log q_{j-1}}\approx\frac{r_{j-1}}{r_j}$$

We simply let

$$\frac{\log q_j}{\log q_{j-1}}>2\delta^{-1}$$

Then let r_1 be arbitrarily large, and take $r_2,\cdots,r_m$ such that

$$r_1\frac{\log q_1}{\log q_j}\leq r_j<r_1\frac{\log q_1}{\log q_j}+1$$

We have

$$r_{j-1}\geq r_1\frac{\log q_1}{\log q_{j-1}},r_j<r_1\frac{\log q_1}{\log q_j}+1$$

Thus

$$\frac{r_{j-1}}{r_j}>\frac{r_1\frac{\log q_1}{\log q_{j-1}}}{r_1\frac{\log q_1}{\log q_j}+1}>\frac{\frac{\log q_j}{\log q_{j-1}}}{1+\frac{\log q_{j-1}}{r_1\log q_1}}>\frac{2\delta^{-1}}{1+\frac{\log q_m}{r_1\log q_1}}>\delta^{-1}$$

Given r_1 sufficiently large

$$\left[r_1>\frac{\log q_m}{\log q_1}\right]$$

□

theorem 29

Theorem 29. *Following the idea and notations above, we could find an integer polynomial $Q \in \mathbb{Z}^m$ such that*

$$Q\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) \neq 0$$

$$q_1^{-r_1} \cdots q_m^{-r_m} \geq q_1^{-mr_1(1+\delta)}$$

$$\left|Q\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right)\right| \leq q_1^{(3\delta - \kappa \mathrm{Ind}Q)r_1}$$

$$-mr_1(1+\delta) > (3\delta - \kappa \mathrm{Ind}Q)r_1$$

Where $\mathrm{Ind}Q = \dfrac{m}{2} - \xi m - C_m \delta^{\frac{1}{2^{m-1}}}$

Proof. Since

$$\begin{aligned} r_1 \log q_1 + \cdots + r_m \log q_m &\leq m r_1 \log q_1 + m \log q_m \\ &= m r_1 \log q_1 \left(1 + \frac{\log q_m}{r_1 \log q_1}\right) \\ &\leq m r_1 \log q_1 (1+\delta) \end{aligned}$$

Given r_1 sufficiently large

$$\left[r_1 > \frac{\log q_m}{\delta \log q_1}\right]$$

Thus

$$q_1^{-r_1} \cdots q_m^{-r_m} \geq q_1^{-mr_1(1+\delta)}$$

Define

$$\mathrm{Poly}_{R,B}(\mathbb{Z}_+^m) := \left\{P \in \mathrm{Poly}_R(\mathbb{Z}^m) \middle| 0 \leq C_I \leq B\right\}$$

then we have

$$\left|\mathrm{Poly}_{R,B}(\mathbb{Z}_+^m)\right| = (B+1)^{(r_1+1)\cdots(r_m+1)}$$

Given γ, it is hard to find $H \in \mathrm{Poly}_{R,B}(\mathbb{Z}_+^m)$ such that the index of H, $\mathrm{Ind}(H) \geq \gamma$, we will use pigeonhole, by estimate the number of distinct value of $\partial_J W(\alpha, \cdots, \alpha)$ to ensure there exist $W_1, W_2 \in \mathrm{Poly}_{R,B}(\mathbb{Z}_+^m)$ such that

$$\partial_J W_1(\alpha, \cdots, \alpha) = \partial_J W_2(\alpha, \cdots, \alpha)$$

$\forall \left|\frac{J}{R}\right| \leq \gamma$, then $H = W_1 - W_2$ would be a desired polynomial. Consider the remainder $r(x)$ of $\partial_J W(x, \cdots, x)$ modulo $f(x)$, then $r(\alpha) = \partial_J W(\alpha, \cdots, \alpha)$. Define $[W]$ to be the maximum of the absolute value of W, then we have

$$\partial_J(x_1^{r_1}, \cdots, x_m^{r_m}) = \binom{r_1}{j_1} \cdots \binom{r_m}{j_m}$$

But since $\binom{r}{j} \leq 2^r$, thus we have

$$[\partial_J W(x_1, \cdots, x_m)] \leq 2^{r_1 + \cdots + r_m}[W]$$

since $\partial_J W(x_1, \cdots, x_m)$ has no more than $(r_1+1)\cdots(r_m+1)$ terms, hence

$$[\partial_J W(x, \cdots, x)] \leq (r_1+1)\cdots(r_m+1)2^{r_1+\cdots+r_m}[W]$$

Then by the Lemma 30 we have

$$[r] \leq (r_m+1)2^{r_1+\cdots+r_m}(1+[f])^m[W]$$

□

lemma 30

Lemma 30. $P \in \mathbb{Z}[x]$ with $\deg P = m \geq n$, r is the remainder of P modulo f, then $[r] \leq (1+[f])^{m-n+1}[P]$

Proof. Suppose $P = c_0 + \cdots + c_m x_m$, then $P = P_1 + c_m x^{m-n} f$ with $\deg P_1 < m$, $[P_1] \leq [P] + [P][f] = (1 + [f])[P]$, carry on this process, we have $[r] \leq (1 + [f])^{m-n+1}[P]$ □

Hence the number of distinct values of $\partial_J W(\alpha, \cdots, \alpha) = r(\alpha)$ is at most $(2(r_m + 1)2^{r_1+\cdots+r_m}(1 + [f])^m[W] + 1)^n$ for a single J, next we need to estimate the number of distinct J such that $\left|\frac{J}{R}\right| \leq \gamma$, as we analyzed above, we want $\frac{\gamma}{m} \approx \frac{1}{2}$, we might define $\gamma = \frac{m}{2} - \xi m$ with $\xi > 0$, then the number of distinct J is bounded by the Lemma 31

lemma 31

Lemma 31. $S := \left\{0 \leq J \leq R \middle| \left|\frac{J}{R}\right| - \frac{m}{2} \leq -\xi m\right\}$, then $n(\xi) = |S| \leq (r_1 + 1) \cdots (r_m + 1) e^{-\frac{m}{4}\xi^2}$

Remark. The idea is to view $\frac{j_i}{r_i}$ as random variables, then by the law of large numbers, there can't be too many away from the mean value

Proof. $n(\xi) = \sum_{J \in S} 1$, thus

$$\begin{aligned} n(\xi) e^{\frac{m}{2}\xi^2} &\leq \sum_{0 \leq J \leq R} e^{-\frac{\xi}{2}(|\frac{J}{R}| - \frac{m}{2})} \\ &= \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} e^{-\frac{\xi}{2}(|\frac{j_1}{r_1}| - \frac{1}{2})} \cdots e^{-\frac{\xi}{2}(|\frac{j_m}{r_m}| - \frac{1}{2})} \\ &= \prod_{i=1}^{m} \left(\sum_{j_i=0}^{r_i} e^{-\frac{\xi}{2}(|\frac{j_i}{r_i}| - \frac{1}{2})} \right) \end{aligned}$$

Since $1 + x \leq e^x \leq 1 + x + x^2 \quad \forall |x| \leq 1$, thus

$$\begin{aligned} \sum_{j=0}^{r} e^{-\frac{\xi}{2}(|\frac{j}{r}| - \frac{1}{2})} &\leq \sum_{j=0}^{r} \left(1 - \frac{\xi}{2} \left(\left| \frac{j}{r} \right| - \frac{1}{2} \right) + \frac{\xi^2}{4} \left(\left| \frac{j}{r} \right| - \frac{1}{2} \right)^2 \right) \\ &\leq (r + 1) \left(1 + \frac{\xi^2}{4} \right) \\ &\leq (r + 1) e^{\frac{\xi^2}{4}} \end{aligned}$$

Therefore

$$n(\xi) \leq (r_1 + 1) \cdots (r_m + 1) e^{-\frac{m}{4}\xi^2}$$

□

Since

$$2(r_1 + 1) \cdots (r_m + 1) 2^{r_1+\cdots+r_m} (1 + [f])^m [W] + 1 \leq 2^{2+2(r_1+\cdots+r_m)} (1 + [f])^m B$$

We should make sure

$$\left(2^{2+2(r_1+\cdots+r_m)} (1 + [f])^m B \right)^{n(r_1+1)\cdots(r_m+1)e^{-\frac{m}{4}\xi^2}} < B^{(r_1+1)\cdots(r_m+1)}$$

Or simply make sure

$$\left(2^{4mr_1} (1 + [f])^m B \right)^{ne^{-\frac{m}{4}\xi^2}} < B$$

If We take $B = q_1^{\delta r_1}$, for q_1 sufficiently large $\left[\log q_1 > m\delta^{-1} \log\left(16(1 + [f])\right)\right]$, we then have

$$2^{4mr_1} (1 + [f])^m \leq 2^{4mr_1} (1 + [f])^{mr_1} \leq q_1^{\delta r_1}$$

then we only need

$$2ne^{-\frac{m}{4}\xi^2} < 1 \Leftrightarrow \log(2n) < \frac{m}{4}\xi^2$$

But still $\xi \to 0$ as $m \to \infty$ $\left[\text{Suppose } \xi = \left\lfloor \sqrt{\frac{4\log(2n)}{m}} \right\rfloor + 1\right]$

Now that we have H vanishes to high order at $(\alpha, \cdots, \alpha)$, but we couldn't make sure $H\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) \neq 0$. However, we could find J_0 with $\left|\frac{J_0}{R}\right|$ fairly small compare to $\frac{m}{2}$ that $\partial_{J_0} H\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right) \neq 0$ and then simply choose $Q = \partial_{J_0} H$. This relies upon the behavior of integer polynomials on rational points, which is fairly different than on irrational points, and this makes all the difference, in the case where H has just one variable, this is easy, but when H has many variables, we need to seek a way of making inductions on the number of variables, and that would be the most tricky part. First we introduce the notion of generalized Wronskian

definition 32

Definition 32. Assume $\varphi_0, \cdots, \varphi_{n-1} \in \text{Poly}(\mathbb{F}^m)$, differential operator $\partial_J = \frac{1}{J!}\partial_x^J = \frac{1}{j_1! \cdots j_m!}\partial_{x_1}^{j_1} \cdots \partial_{x_m}^{j_m}$, then the generalized Wronskian is defined to be

$$\begin{vmatrix} \partial_{\Delta_0}\varphi_0 & \cdots & \partial_{\Delta_0}\varphi_{n-1} \\ \vdots & \ddots & \vdots \\ \partial_{\Delta_{n-1}}\varphi_0 & \cdots & \partial_{\Delta_{n-1}}\varphi_{n-1} \end{vmatrix}$$

Where Δ_i is some some operator ∂_J with $|J| = i$

lemma 33

Lemma 33. Assume $\varphi_0, \cdots, \varphi_{n-1} \in \mathbb{Q}[x]$, then they are independent if and only if their Wronskian

$$\begin{vmatrix} \varphi_0 & \cdots & \varphi_{n-1} \\ \vdots & \ddots & \vdots \\ \frac{1}{(n-1)!}\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\varphi_0 & \cdots & \frac{1}{(n-1)!}\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\varphi_{n-1} \end{vmatrix} \neq 0$$

Proof. If $\det\left(\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\varphi_l\right) \neq 0$ and $\varphi_0, \cdots, \varphi_{n-1}$ are linearly independent, then $\exists c_0, \cdots, c_{n-1}$ are not all zero, such that

$$c_0\varphi_0 + \cdots + c_{n-1}\varphi_{n-1} = 0$$

But then

$$c_0\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\varphi_0 + \cdots + c_{n-1}\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\varphi_{n-1} = 0$$

Which is a contradiction. The other direction is easy. □

We generalize the lemma above to the generalized Wronskian

lemma 34

Lemma 34. If $\varphi_0, \cdots, \varphi_{n-1} \in \text{Poly}(\mathbb{Q}^m)$ are linearly independent, then there exists a generalized Wronskian $\det(\Delta_k\varphi_l) \neq 0$

Proof. Choose k to be large enough, larger than the degrees of φ_i, then consider

$$\phi_i(t) = \varphi_i\left(t, t^k, \cdots, t^{k^{m-1}}\right) \in \mathbb{Q}[t]$$

Which will still be linearly independent, otherwise, suppose $\varphi_i(x_1, \cdots, x_m) = \sum C_S^{(i)} x_1^{s_1} \cdots x_m^{s_m}$, then

$$\sum_{i=0}^{n-1} A_i\phi_i(t) = \sum_{i=0}^{n-1} A_i \sum C_S^{(i)} t^{s_1 + ks_2 + \cdots + k^{m-1}s_m}$$

But $s_1 + ks_2 + \cdots + k^{m-1}s_m$ will all be different like k-adic numbers by the Lemma 35, therefore $A_i = 0$ □

lemma 35

Lemma 35. If two k-adic number $s_0 + s_1k + \cdots + s_nk^n$ and $t_0 + t_1k + \cdots + t_nk^n$ are equal, then $s_i = t_i$

Proof. Since

$$s_0 + s_1k + \cdots + s_nk^n = t_0 + t_1k + \cdots + t_nk^n$$

We have

$$\begin{aligned}\left|(t_n - s_n)k^n\right| &= \left|(s_0 - t_0) + (s_1 - t_1)k + \cdots + (s_{n-1} - t_{n-1})k^{n-1}\right| \\ &\leq (k-1)(1 + k + \cdots + k^{n-1}) \\ &= k^n - 1\end{aligned}$$

Thus $t_n = s_n$, carry on the process, we have $s_i = t_i$ □

According to Lemma 33 above, we know that the Wronskian

$$\det\left(\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\phi_l\right) \neq 0$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_l = \partial_{x_1}\varphi_l + kt^{k-1}\partial_{x_2}\varphi_l + \cdots + k^{m-1}t^{k^{m-1}-1}\partial_{x_m}\varphi_l$$

And

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}\phi_l = \left(f_1(t)\Delta_k^1 + \cdots + f_r(t)\Delta_k^r\right)\varphi\left(t, t^k, \cdots, t^{k^{m-1}}\right)$$

Where Δ_k^i are of order k. Expand the Wronskian $\det\left(\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\phi_l\right)$ to a sum of generalized Wronskian, $g_0(t)W_0 + \cdots + g_s(t)W_s$, hence at least one of the generalized Wronskian $W_i \neq 0$
In order to make induction on the number of variables, we use the next argument to separate the variables

lemma36

Lemma 36. Given $R \in \mathrm{Poly}(\mathbb{Z}^m)$, $m \geq 2$, $R \neq 0$, $\deg_{x_i} R \leq r_i$, then $\exists l \in \mathbb{Z}$ such that $1 \leq l \leq r_m + 1$ and $\exists \Delta_0, \cdots, \Delta_{l-1}$ of orders $0, \cdots, l-1$, and we can construct

$$F(x_1, \cdots, x_m) = \det\left(\Delta_p \frac{1}{q!}\partial_{x_m}^q R\right)$$

Where $0 \neq F \in \mathrm{Poly}(\mathbb{Z}^m)$, We can separate x_m from $x_1, \cdots, x_{m-1}$

$$F(x_1, \cdots, x_m) = U(x_1, \cdots, x_{m-1})V(x_m)$$

Where $U \in \mathrm{Poly}(\mathbb{Z}^{m-1})$, $V \in \mathbb{Z}[x]$, and $\deg_{x_i} U \leq lr_i$, $\deg_{x_m} V \leq lr_m$

Proof. We can write R in the form of

$$\psi_0(x_m)\phi_0(x_1, \cdots, x_{m-1}) + \cdots + \psi_{l-1}(x_m)\phi_{l-1}(x_1, \cdots, x_{m-1})$$

Where l is the smallest possible, such an expression always exists since

$$R(x_1, \cdots, x_m) = \phi_0(x_1, \cdots, x_{m-1}) + \cdots + x_m^{r_m}\phi_{r_m}(x_1, \cdots, x_{m-1})$$

Which also implies that $1 \leq l \leq r_m + 1$. Also $\psi_0, \cdots, \psi_{l-1}$, $\phi_0, \cdots, \phi_{l-1}$ must be linearly independent otherwise l won't be the smallest, thus according to lemma, there exist a Wronskian

$$W(x_m) = \det\left(\frac{1}{q!}\partial_{x_m}^q \psi_i\right) \neq 0$$

And a generalized Wronskian

$$G(x_1, \cdots, x_{m-1}) = \det\left(\Delta_p \phi_i\right) \neq 0$$

Let

$$F(x_1, \cdots, x_m) = WG = \det\left(\Delta_p \frac{1}{q!}\partial_{x_m}^q R\right) \neq 0$$

By Gauss's lemma, after multiply a integer factor, we have

$$F(x_1, \cdots, x_m) = U(x_1, \cdots, x_{m-1})V(x_m)$$

Where $U \in \mathrm{Poly}(\mathbb{Z}^{m-1})$, $V \in \mathbb{Z}[x]$, and $\deg_{x_i} U \leq lr_i$, $\deg_{x_m} V \leq lr_m$ □

Next, we estimate the coefficients of F

lemma 37

Lemma 37. If $[R] \leq B$, then

$$[F] \leq ((r_1+1)\cdots(r_m+1))^l \, l! B^l 2^{(r_1+\cdots+r_m)l}$$

Proof. $\Delta_p \frac{1}{q!}\partial_{x_m}^q R$ has at most $(r_1+1)\cdots(r_m+1)$ terms and a upper bound of each term would be

$$\binom{r_1}{p_1}\cdots\binom{r_{m-1}}{p_{m-1}}\binom{r_m}{s} B \leq 2^{r_1+\cdots+r_m} B$$

And $\det\left(\Delta_p \frac{1}{q!}\partial_{x_m}^q R\right)$ can be expanded to at most $l!$ terms of product of l terms. In conclusion

$$[F] \leq ((r_1+1)\cdots(r_m+1))^l \, l! B^l 2^{(r_1+\cdots+r_m)l}$$

□

We first dealt the case of just one variable

definition 38

Definition 38. $\Theta(B; q_1, \cdots, q_m; r_1, \cdots, r_m) := \max\{\mathrm{Ind}(P)\}$, where P is taken over $\mathrm{Poly}_{B,R}(\mathbb{Z}^m)$, and h_i is taken over $(h_i, q_i) = 1$, and the $\mathrm{Ind}(P)$ is at $\left(\frac{h_1}{q_1}, \cdots, \frac{h_m}{q_m}\right)$ with respect to R

lemma 39

Lemma 39.

$$\Theta(B; q; r) \leq \frac{\log B}{r \log q}$$

Proof. Assume $P \in \mathrm{Poly}_{B,R}(\mathbb{Z})$, $(h,q) = 1$, $\theta = \mathrm{Ind}(P)$, then we know that $\frac{h}{q}$ is a root of P of order θr, by Gauss's lemma

$$P = (qx - h)^{\theta r} g(x)$$

Where $g \in \mathbb{Z}[x]$, thus we have $q^{\theta r} \leq B$ which gives the estimate

$$\Theta(B; q; r) \leq \frac{\log B}{r \log q}$$

□

Before we jump into the case of multiple variables, we state some basic arithmetic properties of index

lemma 40

Lemma 40.

$$\mathrm{Ind}(P+Q) \geq \min(\mathrm{Ind}(P), \mathrm{Ind}(Q))$$
$$\mathrm{Ind}(PQ) = \mathrm{Ind}(P) + \mathrm{Ind}(Q)$$

Proof. Assume $P = \sum c_I x^I, Q = \sum d_J x^J$, then $P+Q = \sum (c_I + d_I) x^I, PQ = \sum c_I d_I x^{I+J}$, suppose

$$\left|\frac{I_1}{R_1}\right| = \cdots = \left|\frac{I_p}{R_p}\right| = \mathrm{Ind}(P)$$
$$\left|\frac{J_1}{R_1}\right| = \cdots = \left|\frac{J_q}{R_q}\right| = \mathrm{Ind}(Q)$$

If $c_I + d_I \neq 0$, then $c_I \neq 0$ or $d_I \neq 0$, thus

$$\mathrm{Ind}(P+Q) \geq \min(\mathrm{Ind}(P), \mathrm{Ind}(Q))$$

$c_I d_I \neq 0$, if and only if $c_I \neq 0$ and $d_I \neq 0$, thus

$$\mathrm{Ind}(PQ) = \mathrm{Ind}(P) + \mathrm{Ind}(Q)$$

□

Finally, we discuss the case of multiple variables by making induction on the number of variables

lemma 41

Lemma 41.

$$\Theta(q_1^{\delta r_1}; q_1, \cdots, q_m; r_1, \cdots, r_m) \leq C_m \delta^{\frac{1}{2^{m-1}}}$$

Proof. On one hand, since $F = UV$, $[U], [V] \leq [F]$, $\mathrm{Ind}(F) = \mathrm{Ind}(U) + \mathrm{Ind}(V)$, and

$$\begin{aligned}
[F] &\leq ((r_1+1)\cdots(r_m+1))^l \, l! q_1^{l\delta r_1} 2^{(r_1+\cdots+r_m)l} \\
&\leq \left(2^{2(r_1+\cdots+r_m)} l q_1^{\delta r_1}\right)^l \\
&\leq \left(2^{2mr_1} l q_1^{\delta r_1}\right)^l \\
&\leq q_1^{2l\delta r_1}
\end{aligned}$$

Given q_1 large enough, thus

$$\mathrm{Ind}(U) \leq l\Theta(q_1^{2l\delta r_1}; q_1, \cdots, q_{m-1}; lr_1, \cdots, lr_{m-1})$$

$$\mathrm{Ind}(V) \leq l\Theta(q_1^{2l\delta r_1}; q_m; lr_m)$$

Hence

$$\mathrm{Ind}(F) \leq l\left(\Theta(q_1^{2l\delta r_1}; q_1, \cdots, q_{m-1}; lr_1, \cdots, lr_{m-1}) + \Theta(q_1^{2l\delta r_1}; q_m; lr_m)\right)$$

On the other hand, since $F = \det\left(\Delta_p \frac{1}{q!} \partial_{x_m}^q R\right)$, and

$$\begin{aligned}
\mathrm{Ind}\left(\Delta_p \frac{1}{q!} \partial_{x_m}^q R\right) &\geq \mathrm{Ind}(R) - \frac{p_1}{r_1} - \cdots - \frac{p_{m-1}}{r_{m-1}} - \frac{q}{r_{m-1}} \\
&= \mathrm{Ind}(R) - \frac{q}{r_m} - \frac{p}{r_{m-1}} \\
&\geq \mathrm{Ind}(R) - \frac{q}{r_m} - \frac{l\delta}{r_m} \\
&> \mathrm{Ind}(R) - \frac{q}{r_m} - \delta
\end{aligned}$$

But since $\mathrm{Ind}\left(\Delta_p \frac{1}{q!} \partial_{x_m}^q R\right) \geq 0$, thus in case $\mathrm{Ind}(R)$ is to small, we have

$$\mathrm{Ind}\left(\Delta_p \frac{1}{q!} \partial_{x_m}^q R\right) \geq \max\left\{0, \mathrm{Ind}(R) - \frac{q}{r_m}\right\} - \delta$$

And in the expansion of $\det\left(\Delta_p \frac{1}{q!} \partial_{x_m}^q R\right)$, each of the $l!$ terms in the expansion is of the form

$$\pm \left(\Delta_{p_0} R\right)\left(\Delta_{p_0} \partial_{x_m} R\right) \cdots \left(\Delta_{p_{l-1}} \frac{1}{(l-1)!} \partial_{x_m}^{l-1} R\right)$$

Thus

$$\mathrm{Ind}(F) \geq \sum_{q=0}^{l-1} \max\left\{0, \mathrm{Ind}(R) - \frac{q}{r_m}\right\} - l\delta$$

Denote $\Phi = \Theta(q_1^{2l\delta r_1}; q_1, \cdots, q_{m-1}; lr_1, \cdots, lr_{m-1}) + \Theta(q_1^{2l\delta r_1}; q_m; lr_m)$, $\theta = \mathrm{Ind}R$, then there are two cases

Case I: $\theta r_m \geq l$

$$\begin{aligned}
\mathrm{Ind}(F) &\geq \sum_{q=0}^{l-1} \left(\theta - \frac{q}{r_m}\right) - l\delta \\
&= l\theta - \frac{l(l-1)}{2r_m} - l\delta \\
&\geq \frac{l}{2}\theta - l\delta
\end{aligned}$$

Which would implie

$$l\Phi \geq \frac{l}{2}\theta - l\delta \Rightarrow \theta \leq 2(\Phi + \delta)$$

Case II: $\theta r_m < l$

$$\begin{aligned}
\mathrm{Ind}(F) &\geq \sum_{0 \leq q \leq \lfloor \theta r_m \rfloor} \left(\theta - \frac{q}{r_m}\right) - l\delta \\
&= \left(\lfloor \theta r_m \rfloor + 1\right)\theta - \frac{\lfloor \theta r_m \rfloor \left(\lfloor \theta r_m \rfloor + 1\right)}{2r_m} - l\delta \\
&= \left(\lfloor \theta r_m \rfloor + 1\right)\left(\theta - \frac{\lfloor \theta r_m \rfloor}{2r_m}\right) - l\delta \\
&\geq \left(\lfloor \theta r_m \rfloor + 1\right)\frac{\theta}{2} - l\delta \\
&\geq \frac{\theta^2 r_m}{2} - l\delta
\end{aligned}$$

Which would implie

$$\begin{gathered}
l\Phi \geq \frac{\theta^2 r_m}{2} - l\delta \Rightarrow \\
\theta^2 \leq \frac{2l}{r_m}(\Phi + \delta) \leq \frac{2(r_m + 1)}{r_m}(\Phi + \delta) \leq 4(\Phi + \delta) \Rightarrow \\
\theta \leq 2(\Phi + \delta)^{\frac{1}{2}} \leq 2(\Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}})
\end{gathered}$$

In conclusion, we have

$$\theta \leq 2(\Phi + \Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}})$$

Since R is arbitrary, we have

$$\Theta(q_1^{\delta r_1}; q_1, \cdots, q_m; r_1, \cdots, r_m) \leq 2(\Phi + \Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}})$$

Already we have

$$\Theta(q_1^{\delta r_1}; q_1; r_1) \leq \delta$$

Since

$$\begin{cases}
\Theta(q_1^{2l\delta r_1}; q_2; lr_2) \leq \dfrac{2l\delta r_1 \log q_1}{lr_2 \log q_2} \leq 2\delta \\
\Theta(q_1^{2l\delta r_1}; q_1; lr_1) \leq \dfrac{2l\delta r_1 \log q_1}{lr_1 \log q_1} \leq 2\delta
\end{cases}$$

We have

$$\begin{aligned}
\Theta(q_1^{\delta r_1}; q_1, q_2; r_1, r_2) &\leq 2(\delta^{\frac{1}{2}} + 4\delta + (4\delta)^{\frac{1}{2}}) \\
&\leq 2(\delta^{\frac{1}{2}} + 4\delta^{\frac{1}{2}} + 2\delta^{\frac{1}{2}}) \\
&= 14\delta^{\frac{1}{2}}
\end{aligned}$$

Since

$$\begin{cases}
\Theta(q_1^{2l\delta r_1}; q_3; lr_3) \leq \dfrac{2l\delta r_1 \log q_1}{lr_3 \log q_3} \leq 2\delta \\
\Theta(q_1^{2l\delta r_1}; q_1, q_2; lr_1, lr_2) \leq 14\delta^{\frac{1}{2}}
\end{cases}$$

We have

$$\begin{aligned}
\Theta(q_1^{\delta r_1}; q_1, q_2, q_3; r_1, r_2, r_3) &\leq 2(\delta^{\frac{1}{2}} + (2\delta + 14\delta^{\frac{1}{2}}) + (2\delta + 14\delta^{\frac{1}{2}})^{\frac{1}{2}}) \\
&\leq 2(\delta^{\frac{1}{4}} + (2^{\frac{1}{4}} + 14\delta^{\frac{1}{4}}) + (2\delta^{\frac{1}{2}} + 14\delta^{\frac{1}{2}})^{\frac{1}{2}}) \\
&= 42\delta^{\frac{1}{4}}
\end{aligned}$$

Since

$$\begin{cases}
\Theta(q_1^{2l\delta r_1}; q_4; lr_4) \leq \dfrac{2l\delta r_1 \log q_1}{lr_4 \log q_4} \leq 2\delta \\
\Theta(q_1^{2l\delta r_1}; q_1, q_2, q_3; lr_1, lr_2, lr_3) \leq 42\delta^{\frac{1}{4}}
\end{cases}$$

We have

$$\begin{aligned}\Theta(q_1^{\delta r_1};q_1,q_2,q_3,q_4;r_1,r_2,r_3,r_4) &\le 2(\delta^{\frac{1}{2}}+(2\delta+42\delta^{\frac{1}{4}})+(2\delta+42\delta^{\frac{1}{4}})^{\frac{1}{2}})\\ &\le 2(\delta^{\frac{1}{8}}+(2\delta^{\frac{1}{8}}+42\delta^{\frac{1}{8}})+(2\delta^{\frac{1}{4}}+42\delta^{\frac{1}{4}})^{\frac{1}{2}})\\ &\le 104\delta^{\frac{1}{8}}\end{aligned}$$

And so on, we could write

$$\begin{cases}\Theta(q_1^{2l\delta r_1};q_m;lr_m)\le \dfrac{2l\delta r_1\log q_1}{lr_m\log q_m}\le 2\delta\\ \Theta(q_1^{2l\delta r_1};q_1,\cdots,q_{m-1};lr_1,\cdots,lr_{m-1})\le C_{m-1}\delta^{\frac{1}{2^{m-2}}}\end{cases}$$

Thus

$$\begin{aligned}&\Theta(q_1^{\delta r_1};q_1,\cdots,q_m;r_1,\cdots,r_m)\\ &\le 2\left(\delta^{\frac{1}{2}}+\left(2\delta+C_{m-2}\delta^{\frac{1}{2^{m-2}}}\right)+\left(2\delta+C_{m-2}\delta^{\frac{1}{2^{m-2}}}\right)^{\frac{1}{2}}\right)\\ &\le 2\left(\delta^{\frac{1}{2^{m-1}}}+\left(2\delta^{\frac{1}{2^{m-1}}}+C_{m-1}\delta^{\frac{1}{2^{m-1}}}\right)+\left(2\delta^{\frac{1}{2^{m-2}}}+C_{m-1}\delta^{\frac{1}{2^{m-2}}}\right)^{\frac{1}{2}}\right)\\ &\le 2\left(1+(2+C_{m-1})+(2+C_{m-1})^{\frac{1}{2}}\right)\delta^{\frac{1}{2^{m-1}}}\end{aligned}$$

Obviously, we need $2(1+(2+C_{m-1})+(2+C_{m-1})^{\frac{1}{2}})\le C_m$, quick check we can simply take $C_m=10^m$. But still we have

$$\Theta(q_1^{\delta r_1};q_1,\cdots,q_m;r_1,\cdots,r_m)\le C_m\delta^{\frac{1}{2^{m-1}}}\to 0(\delta\to 0)$$

Finally, we get a desired Q with the following property

$$\begin{aligned}|\partial_J Q(\alpha,\cdots,\alpha)| &\le (r_1+1)\cdots(r_m+1)2^{r_1+\cdots+r_m}(1+[f])^m[H]\\ &\le 2^{2(r_1+\cdots+r_m)}(1+[f])^m q_1^{\delta r_1}\\ &\le 2^{2mr_1}(1+[f])^m q_1^{\delta r_1}\\ &\le q_1^{2\delta r_1}\end{aligned}$$

Given q_1 large enough $\left[\text{as above}\log q_1>m\delta^{-1}\log\left(16(1+[f])\right)\right]$, and

$$\mathrm{Ind}Q\ge\gamma-C_m\delta^{\frac{1}{2^{m-1}}}=\frac{m}{2}-\xi m-C_m\delta^{\frac{1}{2^{m-1}}}$$

$$\left|Q\left(\frac{h_1}{q_1},\cdots,\frac{h_m}{q_m}\right)\right|\ge q_1^{-r_1}\cdots q_m^{-r_m}\ge q_1^{-mr_1(1+\delta)}$$

$$\begin{aligned}\left|Q\left(\frac{h_1}{q_1},\cdots,\frac{h_m}{q_m}\right)\right| &\le \sum_{|\frac{I}{R}|\ge\mathrm{Ind}Q}|\partial_I Q(\alpha,\cdots,\alpha)|\left|\left(\frac{h_1}{q_1}-\alpha\right)^{i_1}\cdots\left(\frac{h_m}{q_m}-\alpha\right)^{i_m}\right|\\ &\le (r_1+1)\cdots(r_m+1)|\partial_I Q(\alpha,\cdots,\alpha)|q_1^{-i_1\kappa}\cdots q_m^{-i_m\kappa}\\ &\le 2^{mr_1}q_1^{2\delta r_1}q_1^{-r_1\kappa\frac{i_1}{r_1}}\cdots q_m^{-r_m\kappa\frac{i_m}{r_m}}\\ &\le q_1^{3\delta r_1}q_1^{-r_1\kappa\mathrm{Ind}Q}=q_1^{(3\delta-\kappa\mathrm{Ind}Q)r_1}\end{aligned}$$

Thus we need

$$\begin{aligned}-mr_1(1+\delta)>\left(3\delta-\kappa\left(\frac{m}{2}-\xi m-C_m\delta^{\frac{1}{2^{m-1}}}\right)\right)r_1 &\Leftrightarrow\\ \kappa\left(\frac{m}{2}-\xi m-C_m\delta^{\frac{1}{2^{m-1}}}\right)>m(1+\delta)+3\delta &\Leftrightarrow\\ \kappa\left(\frac{1}{2}-\xi-\frac{C_m\delta^{\frac{1}{2^{m-1}}}}{m}\right)>1+\delta+\frac{3\delta}{m} &\Leftrightarrow\end{aligned}$$

$$\frac{\kappa-2}{2} > \kappa\xi + \frac{\kappa 10^m \delta^{\frac{1}{2^{m-1}}}}{m} + 4\delta$$

Which would be true given m large enough and δ small enough

$$\left[\xi < \frac{\kappa-2}{4\kappa}(m\to\infty), \quad \frac{\kappa 10^m \delta^{\frac{1}{2^{m-1}}}}{m} + 4\delta < \frac{\kappa-2}{4\kappa}(\delta\to 0)\right]$$

□

As a major implication, we state Thue's theorem

theorem 42

Theorem 42. *Suppose $P(x,y) \in \mathbb{Z}[x,y]$ is an irreducible homogeneous polynomial with $d = \deg P \geq 3$, then Thue's equation $P(x,y) = m$ has at most finitely many integer solutions, where $m \in \mathbb{Z}$*

Remark. This theorem is known as Thue's theorem. When $d = 2$, theorem doesn't hold. Consider Diophantine equation $y^2 - 2x^2 = 1$, let $x_1 = y_1 = 1$, $x_2 = 3$, $y_2 = 2$, and

$$\begin{cases} x_{n+1} = 2x_n + x_{n-1} \\ y_{n+1} = 2y_n + y_{n-1} \end{cases}$$

Then we have $y_n^2 - 2x_2^2 = (-1)^n$. Note that (x,y) should be best approximations of $\sqrt{2}$, since if $\left|\frac{y}{x}\right| \neq \sqrt{2}$, then $\left|y^2 - 2x_2\right| \geq 1$, thus we simply pick x_n, y_n from the continued fractions $\frac{y_n}{x_n}$ of $\sqrt{2}$

Proof. Assume $P(x,y) = \sum_{j=0}^{d} a_j x^{d-j} y^j$, then we have $\left|\sum_{j=0}^{d} a_j \left(\frac{y}{x}\right)\right| = |A||x|^{-d}$, define $Q(z) = \sum_{j=0}^{d} a_j z^j$, thus $\left|Q\left(\frac{y}{x}\right)\right| = |A||x|^{-d}$. Let β_j be the roots of Q, β_j would be distinct since Q is irreducible, thus $Q'(\beta_j) \neq 0$, choose some $\delta > 0$ such that $\forall z \in (\beta_j - \delta, \beta_j + \delta), |Q(z)| \geq \frac{1}{2}|Q'(\beta_j)||z - \beta_j|$, since $\frac{Q(z)}{z-\beta_j} = \frac{Q(z)-Q(\beta_j)}{z-\beta_j}$ is continues with $\lim_{z\to\beta_j} \frac{Q(z)}{z-\beta_j} = Q'(\beta_j)$. Suppose Thue's equation has infinitely many solutions, then according to $\left|Q\left(\frac{y}{x}\right)\right| = |A||x|^{-d}$, there would be solution with $|x|$ arbitrarily large, otherwise there would be solution with $\left|\frac{y}{z}\right|$ arbitrarily large, but that's impossible since $\lim_{|z|\to\infty} |Q(z)| = \infty$. Thus there should be solution such that $\left|Q\left(\frac{y}{x}\right)\right|$ arbitrarily small which means there should be infinitely many solutions satisfying $\frac{y}{x} \in (\beta_j - \delta, \beta_j + \delta)$ for some j. On the other hand, if (x,y) is one such solution, since we already have $\left|Q\left(\frac{y}{x}\right)\right| \geq \frac{1}{2}|Q'(\beta_j)|\left|\beta_j - \frac{y}{x}\right|$, thus $\left|\beta_j - \frac{y}{x}\right| \leq \frac{2|A|}{|Q'(\beta_j)|}\frac{1}{|x|^d}$, but by Roth's Theorem, there could only be finitely many solutions, and we have reached a contradiction □

5 Conclusion and Further Developement

Bézout's theorem could be generalized to higher dimensional projectivce space with the inequality replaced with equality
Roth's theorem could be generalized to the following statement:
If α is an irrational algebraic number, then

$$|\alpha - \xi| < \frac{1}{H(\xi)^\kappa}$$

has only finitely many solutions, where height $H(\xi)$ is the maximum of the absolute values of the coefficients of its primitive minimal polynomial, and $\kappa > 2\deg(\xi)$, and infinitely many solutions if $\kappa \leq 2\deg(\xi)$
There are also generalizations by LeVeque to arbitratry algebraic number field, and by Schmidt to higher dimension which called Schmidt's subspace theorem, and an analogue to p-adic metric Roth's theorem
Also, Serge Lang made a stronger conjecture, that if α is an irrational algebraic number, then

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2 \log(q)^{1+\epsilon}}$$

has only finitely many solutions
The proofs in this thesis, they often don't construct a desired polynomial concretely, rather they use a probabilistic method, or pegion hole's principle by certifying only the existence of such a polynomial
I have to say I'm astonished over and over again by how some simple yet powerful structure hidden behind could link and shed light on so many interesting questions

6 Reference

References

[1] *Polynomial Methods in Combinatorics* - Larry Guth (2016). American Mathematical Society.

[2] *Ideals, Varieties, and Algorithms* - David Cox, John Little and Donal O'shea (2010). Springer. pages 422-431.

[3] *Diophantine Approximation* - Wolfgang M. Schmidt (1980). Springer. pages 114-125.

[4] *Rational Approximations to Algebraic Numbers* - K. F. Roth (1955). Mathematika 2. pages 1-20.

[5] *Plane Algebraic Curves* - Harold Hilton (Oxford 1920). pages 10-11.

[6] *Using Algebraic Geometry* - Cox, David; Little, John; O'Shea, Donal (2005). Springer.

[7] *Diophantine Approximations and Diophantine Equations* - Wolfgang M. Schmidt (1991). Springer.

[8] *The Approximation to Algebraic Numbers by Rationals* - Dyson, Freeman J. (1947). Acta Mathematica.

[9] Introduction to Diophantine Approximations - Serge Lang (1995). Springer.

[10] *Diophantine Geometry: An Introduction* - Marc Hindry and Joseph H. Silverman (2000). Graduate Texts in Mathematics.

7 Acknowledgements

First, I have to express my deepest gratitude to my advisor, Prof. Lixin Yan, he recommended this interesting topic and some related books to me, and helped me with a lot of difficulties that I came across. Second, I want to thank Prof. Binglong Chen, Prof. Jianxun Hu and Associate Prof. Xiaoyong Fu for their imparted wisdom and constant caring. And last, I must thank my family for choosing to support me no matter what