Goal: approximate algebraic # by algebraic # of a fixed field K.

First development: Thue (1909) first improvement of Liouville's thm.

Thm (Thue) Let α be a real algebraic # of deg $d \geq 3$. Let $\varepsilon > 0$. Then there're only finitely many rational # p/q, $q > 1$ such that $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{\frac{d}{2}+1+\varepsilon}}$.

As consequence, Thue proved Thue eqn

$F(x,y) = m$ where $F \in \mathbb{Z}[x,y]$ is homogeneous of degree d with at least 3 proportional factors over $\mathbb{C}$ has only finitely many solutions in integer x, y for $\forall\, m \neq 0$ integer fixed.

Drawback: It's ineffective since no bound can be a-priori placed on ht of rational approximations p/q.

We need to assume: at least 1 solution is known, then can deduce things.
(need a solution with ht above given

constant.

Siegel: $\frac{d}{2}+1$ replaced by $\min_{s\in\mathbb{N}}\left(s+\frac{d}{s+1}\right) < 2\sqrt{d}$

$\Rightarrow$ finiteness of # of integral pts on a curve of genus $g \geq 1$.

Lem Thue's theorem $\Rightarrow$ finiteness of int solutions of Thue eqn.

Pf Assume F is irreducible, then

$$F\left(\frac{x}{y},1\right) = a_d\left(\frac{x}{y}-\alpha_1\right)\cdots\left(\frac{x}{y}-\alpha_d\right) = \frac{m}{y^d}.$$

into linear factors. If there are infinitely many integer solutions (x_n, y_n) of Thue equation, then $|y_n| \to \infty$. Now assume by passing to a sequence, $x_n/y_n \to \alpha_j$.

Other factor bounded away from 0, so infinitely many solutions of $|x/y-\alpha_j| \leq C|y|^{-d}$ for some $C>0$. Since $d\geq 3$ contradicts Thue's theorem.

In general, $F_1, \dots, F_r \neq 0$ irreduc. polys dividing F. May assume y is not a divisor of F by change of coordinates.

Now Dirichlet's box principle gives finitely many divisors m_j of m s.t. $F_j(x,y) = m_j$ $(j=1,\dots,r)$ has infinitely-many solutions (x_n, y_n).

$\Rightarrow$ $x_n/y_n \to 0$ of any F_j and hence $r=1$.

since F has at least 3 different linear factors

$\deg(F_1) \geq 3$. So the irreducible cases gives a contradiction with initial assumption of infinitely many solution of $F_1(x,y)=m$,

New: Roth's theorem

<u>Notation</u> v place of # field K, $|\cdot|_v$ is normalized so product formula is 1.

K_v it's extension to completion.

<u>Thm</u> Let K be # field, S finite set of places. For each $v \in S$ let $\alpha_v \in K_v$ be K-algebraic. Let $\kappa > 2$, only finitely many $\beta \in K$ s.t.

$$\prod_{v \in S} \min(1, |\beta - \alpha_v|_v) \leq H(\beta)^{-\kappa}$$

The classical theorem of Roth's is the special case $K = \mathbb{Q}$, $S = \{\infty\}$, $|\cdot|_v$ is ordinary absolute value over $\mathbb{R}$.

Rmk Ineffective, since no upper bound $H(\beta)$
Refinement: α_v vary with β) in 6.5

Rmk allow $\alpha_v = \infty$, $|\infty - \beta|_v$ replace by $|\frac{1}{\beta}|_v$

Consider $K = \mathbb{Q}$, $S = \{\infty, p\}$, with $\alpha_\infty = \infty$, $\alpha_p = \alpha$, and α alge. int. in $\mathbb{Q}_p$. $\beta = n \in \mathbb{Z}$.

then $|\frac{1}{n}|_\infty = \frac{1}{H(n)}$, by thm $|\alpha - n|_p < |n|^{-1-\varepsilon}$
has only finitely many solution in n for $\varepsilon > 0$ fixed

Applications ① Waring's Problem.
② Results for more general field of characteristic 0. E.g. function fields.

Sketch of Proof of Roth's theorem.

Case $K = \mathbb{Q}$, $S = \{\infty\}$
There is only one α to worry about.

Suppose we have infinite many rational approximation p/q to α such that

$$|\alpha - \frac{p}{q}| \le q^{-k}$$

Then for $\forall$ positive integer m and large constant M, we can find m rational approximations to α, namely p_j/q_j, $j=1, \dots, m$
$\log q_1 > L$, $\log q_{j+1} > M \log q_j$, $j = 1, 2, \dots, m-1$

Sequence of approximation —— (L, M) independent

Step I: The auxiliary construction at the algebraic point
$\vec{x} = (x_1, \dots, x_m)$ construct $P(\vec{x}) \in \mathbb{Z}[\vec{x}]$ with partial degree $d_1, \dots, d_m$, vanished to a (weighted)

higher order at $(\alpha, \cdots, \alpha)$

d_j are chosen so that $d_j \log d_j$ are nearly the same for $j = 1, \cdots, m$. Vanishing to higher order means vanishing of ΔP at $(\alpha, \cdots, \alpha)$ with Δ any differential operator, order $(i_1, \cdots, i_m)$ with

$$\frac{i_1}{d_1} + \cdots + \frac{i_m}{d_m} < t.$$

Here t is a parameter, which we want to be large.

To do this, apply Siegel's lemma.

# of equations is asymptotic to $V_m(t) d_1 \cdots d_m$ with $V_m(t)$ the volume of the region $\mathcal{V}_m(t) := \{\vec{x} \in \mathbb{R}^m \mid x_1 + \cdots + x_m < t,\ 0 \le x_j \le 1$ for $j = 1, \cdots, m\}$

# of coefficient is asymptotic to $d_1 \cdots d_m$, and if $[\mathbb{Q}(\alpha) : \mathbb{Q}]\, V_m(t) \le 1 - \delta$, can find such a polynomial P with integer coefficients bounded by $C^{d_1 + \cdots + d_m}$, for a suitable constant C depending only on $H(\alpha)$, m and δ.

<u>Step II</u>: Non-vanishing on a rational point.

P constructed in Step I, or a suitable derivative it of rather small order, does not vanish at the rational point $(p_1/q_1, \cdots, p_m/q_m)$ provided points p_i/q_i are (L, M)-independent and L, M sufficiently large.

Difficult step.

Step 3: The upper bound.

since P vanishes to higher order at algebraic points $(\alpha, \cdots, \alpha)$, Taylor expansion at $(\alpha, \cdots, \alpha)$ shows $|p_j/q_j - \alpha| < c q_j^{-s}$, then

$$|P(p_1/q_1, \cdots, p_m/q_m)| \leq C'^{d_1 + \cdots + d_m} \max_j q_j^{-s t d_j}$$

for C' depending only on α, m, δ.

Step 4: Liouville lower bound

since P does not vanish at rational point, it is bounded away from 0 as

$$q_1^{-d_1} \cdots q_m^{-d_m} \leq |P(p_1/q_1, \cdots, p_m/q_m)|.$$

Step 5: Comparison of upper & lower bound.

have chosen $d_1 \log q_1 \sim \cdots \sim d_m \log q_m$.
since C'^{d_j} is negligible w.r.t. $q_j^{d_j}$ (because $q_j \to \infty$), comparison of upper and lower bounds show that

$K_t \leq m + O(1/M)$
constant involved in the $O(\cdot)$ symbol depends on α, m, and δ.

Thus as $M \to \infty$, find $K_\delta \leq m/t$ provided $[\mathbb{Q}(\alpha):\mathbb{Q}]\, V_m(t) \leq 1-\delta$.

Probability estimate $\Rightarrow$ choose $t = (\frac{1}{2} - \varepsilon)m$, then $V_m(t)$ tend to 0 as m increases, so choice of t is admissible for large m.
$\Rightarrow K_\delta \leq \frac{1}{\frac{1}{2}-\varepsilon}$, but ~~$K$~~ since $\varepsilon > 0$ sufficiently small.

Preliminary Lemmas.

$\vec{x} = (x_1, \ldots, x_m)$, $\vec{\alpha}_\nu = (\alpha_{\nu 1}, \ldots, \alpha_{\nu m})$

write
$$\binom{\vec{m}}{\vec{\mu}} = \prod_{j=1}^{m} \binom{m_j}{\mu_j}$$

and
$$\partial_{\vec{\mu}} = \frac{1}{\mu_1! \cdots \mu_m!} \left(\frac{\partial}{\partial x_1}\right)^{\mu_1} \cdots \left(\frac{\partial}{\partial x_m}\right)^{\mu_m}$$

we have
$$\partial_{\vec{\mu}}\, \vec{x}^{\,m} = \binom{\vec{m}}{\vec{\mu}} \vec{x}^{\,\vec{m}-\vec{\mu}}$$

Work with polys in $F[x_1, \ldots, x_m]$ for a field F vanishing to higher order at a point.
For $P \in F[x_1, \ldots, x_m]$ positive weights $\vec{d} = (d_1, \ldots, d_m)$, define $\boxed{\text{index}}$ of P at $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$ to be

$$\operatorname{ind}(P;\vec{d};\vec{\alpha}) = \min_{\mu}\left\{\frac{\mu_1}{d_1}+\cdots+\frac{\mu_m}{d_m} \;\middle|\; \partial_{\vec{\mu}} P(\vec{\alpha}) \neq 0\right\}.$$

(a) $\operatorname{ind}(P+Q;\vec{d};\vec{\alpha}) \geq \min(\operatorname{ind}(P;\vec{d};\vec{\alpha}), \operatorname{ind}(Q;\vec{d};\vec{\alpha}))$

(b) $\operatorname{ind}(PQ;\vec{d};\vec{\alpha}) = \operatorname{ind}(P;\vec{d};\vec{\alpha}) + \operatorname{ind}(Q;\vec{d};\vec{\alpha})$.

(c) $\operatorname{ind}(\partial_{\vec{\mu}} P;\vec{d};\vec{\alpha}) \geq \operatorname{ind}(P;\vec{d};\vec{\alpha}) - \frac{\mu_1}{d_1} - \frac{\mu_2}{d_2} - \cdots - \frac{\mu_m}{d_m}$

By means of Taylor expansion at α, see easily $\operatorname{ind}(P;\vec{d};\vec{\alpha}) = \infty$ iff $P = 0$.
So "ind" is a valuation.

Def $\mathcal{V}_m(t) := \{\vec{x} \mid \sum_{i=1}^{m} x_i \leq t,\ 0 \leq x_i \leq 1\}$
and $V_m(t) := \operatorname{vol}(\mathcal{V}_m(t))$ for $\vec{t} \in \mathbb{R}_+^n$, define

$$V_m(\vec{t}) := \sum_{i=1}^{n} V_m(t_i).$$

<u>Lem</u> $\vec{\alpha_i}$, $i = 1, \dots, n$ be $\alpha_i = (\alpha_{i1}, \dots, \alpha_{im})$ with α_{ij} in a finite extension F/k of a number field k, of degree $r = [F:k]$.
Suppose $t \in \mathbb{R}_+^n$ satisfy

$$r\, V_m(\vec{t}) < 1.$$

Then for $\forall$ $d_1, \dots, d_m$ sufficiently large, $\exists$ $P \in k[\vec{x}]$ not identically 0 with partial degrees at most $d_1, \dots, d_m$ such that

(a) index bounded below by $\operatorname{ind}(P;\vec{d};\vec{\alpha_i}) \geq t_i$ for $i = 1, \dots, n$

(b) height of P is bounded by

$$h(P) \leq \frac{r}{1 - rV_m(\vec{t})} \sum_{i=1}^{n} \sum_{j=1}^{m} V_m(t_i)\left(h(\alpha_{ij}) + \log 2 + o(1)\right) d_j$$

as $d_j \to \infty$ for $j=1,\dots,m$.

<u>Pf</u> We abbreviate $I=(i_1,\dots,i_m)$, $J=(j_1,\dots,j_m)$
set $P(\vec{x}) = \sum p_J \vec{x}^J$ and consider

$$\partial_I P(\alpha_k)=0 \quad \text{for } \frac{i_1}{d_1}+\dots+\frac{i_m}{d_m}<t_k,\ k=1,\dots,n.$$

Linear system in coefficients to be solved non-trivially on K; coefficients of linear system are in field F, of degree $[F:K]=r$.

N unknowns is $N=(d_1+1)\cdots(d_m+1) \sim d_1\cdots d_m$ as $d_j \to \infty$.

while M is asymptotically $M \sim V_m(\vec{t})\, d_1\cdots d_m$ because # lattice points $(i_1/d_1,\dots,i_m/d_m)$ in $D_m(t)$ is asymptotic to $V_m(t)\, d_1\cdots d_m$.

Let Z be this number. Associate to each lattice point $(i_1/d_1,\dots,i_m/d_m)$ parallelpiped $i_\nu/d_\nu \le x_\nu \le (i_\nu+1)/d_\nu$, $\nu=1,\dots,m$.
$\Rightarrow V_m(t)\, d_1\cdots d_m \le Z$

For upper bound Z, note

$$\frac{i_1+1}{d_1}+\dots+\frac{i_m+1}{d_m} \le t+\frac{1}{d_1}+\dots+\frac{1}{d_m}$$

and $i_\nu+1 \le d_\nu+1$. Follows that if we rescale $D_m(t)$ by $\lambda := 1+\max(1,t^{-1})(1/d_1+\dots+1/d_m)$

rescaled domain contains all parallelpipeds associated to lattice pts in $D_m(t)$. Hence

$$V_m(t)\, d_1 \cdots d_m \leq Z \leq V_m(t)\, \lambda^m\, d_1 \cdots d_m$$

Hence $Z \sim V_m(t)\, d_1, \ldots, d_m$ as $d_j \to \infty$.

so $N > rM$ if $r V_m(t) < 1$ and $d_j \to \infty$.

Now matrix

$$\vec{A} = \left(\binom{J}{I} (\alpha_k)^{J-I} \right)$$

with rows indexed by (I, k) columns indexed by J

Now $\boxed{\text{Siegel's lemma}}$, $\exists\, \vec{x}$ s.t.

$$H(\vec{x}) \leq |D_K|^{\frac{1}{2r}} \left(\sqrt{N} \prod_{(I,k)} H(\vec{A}_{(I,k)}) \right)^{\frac{r}{N-rM}}$$

where $\vec{A}_{(I,k)}$ is the (I,k)-th row of $\vec{A}$

For fixed (I, k) estimate corresponding row vector $\vec{A}_{(I,k)}$ of $\vec{A}$ as follows: The vector $\vec{A}_{(I,k)}$ has entries $\binom{J}{I} (\alpha_k)^{J-I}$, and hence

$$H(\vec{A}_{(I,k)}) \leq \prod_{j=1}^{m} (2H(\alpha_{k_j}))^{d_j}$$

This bound is independent of I.

For fixed k have $\sim V_m(t_0)\, d_1 \ldots d_m$ possibilities of I thus product of heights of rows of $\vec{A}$ bounded by

$$\left(\prod_{k=1}^{n} \prod_{j=1}^{m} (d_j+1)^{V_m(t_k)/2} (2H(\alpha_{kj}))^{d_j V_m(t_k)}\right)^{(1+o(1))d_1 \cdots d_m}$$

Now by Siegel's lemma, $|D_k|^{1/2r}$ and terms $(d_j+1)^{V_m(t_k)/2}$ are negligible w.r.t. $2^{d_j V_m(t_k)}$

Lem If $0 \le \varepsilon \le \frac{1}{2}$, then

$$V_m\left(\left(\tfrac{1}{2}-\varepsilon\right)m\right) \le e^{-6m\varepsilon^2}$$

Pf see pg 159 | 6.3.5 | method of probability is used.

Now simplest way to achieve Step II in proof of Roth's theorem is by Roth's lemma.

Lem $P(X_1, \dots, X_m)$, m variables, partial degrees at most $d_1, \dots, d_m$ with $d_i \ge 1$ coefficients in $\mathbb{Q}$ and not identically 0. Let $(\xi_1, \dots, \xi_m) \in \bar{\mathbb{Q}}^m$ and let $0 < \sigma \le \frac{1}{2}$, suppose that

(a) weights $d_1, \dots, d_m$ are rapidly decreasing, namely $d_{j+1}/d_j \le \sigma$

(b) the point $(\xi_1, \dots, \xi_m)$ has components with large height, in the sense that

Rmk

Pf

R

$$\min_j d_j\, h(\xi_j) \geq \sigma^{-1}\left(h(P) + 4m d_1\right)$$

$\boxed{(1)\ d_1 \cdots d_m}$

Then $\operatorname{ind}(P; \vec{d}; \vec{\xi}) \leq 2m\, \sigma^{1/2^{m-1}}$

Rmk constant $2m$ is not optimal, but for now it doesn't affect statements we are trying to achieve

Pf Induction. Base case: $m=1$

$\operatorname{ind}(P; d_1; \xi_1)\, d_1 h(\xi_1) \leq h(P) + d_1 \log 2$,

from ht of polynomials (1.6.5) and $(x_1 - \xi_1)^{\operatorname{ind}(P) d_1}$ is a factor of P.

Induction: use Wronskian thm.

write $P = \sum_{j=0}^{s} f_j(x_1, \dots, x_{m-1})\, g_j(x_m)$

$s \leq d_m$ and f_js and g_js are linearly independent polynomials defined over $\mathbb{Q}$

Wronskian criterion:

Prop Let K be a field char $= 0$. $x_1, \dots, x_m$ alg. indep $/K$. Let $\varphi_j \in K[x_1, \dots, x_m]$, $j = 1, \dots, n$ be n polynomials. Then $\varphi_1, \dots, \varphi_n$ are lin. indep. over K iff some generalized Wronskian

$$W_{\mu_1, \dots, \mu_n}(x_1, \dots, x_m) := \det \begin{pmatrix} \partial_{\mu_1} \varphi_1 & \cdots & \partial_{\mu_1} \varphi_n \\ \partial_{\mu_2} \varphi_1 & \cdots & \vdots \\ \vdots & & \vdots \\ \partial_{\mu_n} \varphi_1 & \cdots & \partial_{\mu_n} \varphi_n \end{pmatrix}$$

with $|\vec{\mu_i}| = \mu_{i1} + \cdots + \mu_{im} \leq i-1$ not identically 0.

<u>Pf</u> If $\varphi_1, \dots, \varphi_n$ are linearly dep. $/K$
all generalized Wronskians vanish.

Let $c_1\varphi_1 + \cdots + c_n\varphi_n = 0$
be a linear dependence relation among φ_j

If we apply differential operators $\partial_{\vec{\mu_i}}$ $(i = 1, \dots, n)$
to this relation, obtain homogeneous linear
system in the coefficients c_j and its
determinant must vanish.

If $\varphi_1, \dots, \varphi_n$ linearly indep $/K$.
consider Kronecker substitution
$(x_1, x_2, \dots, x_m) \mapsto (t, t^d, \dots, t^{d^{m-1}})$ where
t is a new determinate. This maps
monomials in $x_1, \dots, x_m$ into powers of t
and is injective on the set of monomials
with partial degrees $< d$.
Hence if $d >$ partial degrees of φ_j, then

$\varphi_1, \dots, \varphi_n$ are linearly indep. $/K$ iff

$$\Phi_j(t) := \varphi_j(t, t^d, \dots, t^{d^{m-1}})$$

are linearly indep. $/K$. By Wronski's well known result, this is true iff

(A) $$W(t) := \det\left((d/dt)^{i-1} \Phi_j\right)_{i,j=1,\dots,n}$$

is not identically 0.

we have, for certain $a_{\vec{\mu},i}(t; d, m) \in \mathbb{Q}(t)$

$$(d/dt)^{i-1}\Phi_j = \sum_{|\vec{\mu}| \le i-1} a_{\vec{\mu},i}(t; d, m)\, \partial_{\vec{\mu}} \varphi_j(t, \dots, t^{d^{m-1}})$$

Substituting into A we see $W(t)$ is a linear combination of generalized Wronskians $W_{\vec{\mu}_1, \dots, \vec{\mu}_n}(t, t^d, \dots, t^{d^{m-1}})$ with $|\vec{\mu}_i| \le i-1$. Since $W(t)$ is not identically 0, some generalized Wronskian does not vanish identically

Now by Prop (Wronskian).

there are two Wronskians

$$U(x_1, \dots, x_{m-1}) := \det(\partial_{\vec{\mu}_i} f_j)_{i,j=0,\dots,s}$$

and $V(x_m) := \det(\partial_\nu g_j)_{\nu,j=0,\dots,s}$

they're not identically 0. We have
$\vec{\mu_i} = (\mu_{1i}, \dots, \mu_{m-1,i})$
and $|\vec{\mu_i}| \le s \le d_m$.

$$W(x_1, \dots, x_m) := \det(\partial_{\vec{\mu_i}, \nu} P) = U(x_1, \dots x_{m-1}) V(x_m)$$

Since $d_{j+1}/d_j \le \frac{1}{2}$ for $\forall j$, we have
$d_1 + \dots + d_m \le 2 d_1$.
Partial degrees U and V are bounded by
$((s+1) d_1, \dots, (s+1) d_{m-1})$ and $(s+1) d_m$.
we have

$$\boxed{A} \quad h(U) + h(V) = h(W) \le (s+1) (h(P) + 4 d_1).$$

~~...~~ $\boxed{A}$: $h(W) = h(U) + h(V)$ since U and V involves disjoint sets of variables. See prop. 1.6.2 in $\boxed{\text{Bom}}$.

Now $h(W) \le \sum_{\nu} \max_{\pi} \log \left| \prod_{i=0}^{s} \partial_{\vec{\mu_i}, \pi(i)} P \right|_{\nu} + \log$

Then Gauss's lemma & Gelfond's lemma lead to

$$h(W) \le \sum_{\nu} \max_{\pi} \sum_{i=0}^{s} \log |\partial_{\mu_i, \pi(i)} P|_{\nu}$$
$$+ (s+1)(d_1 + \dots + d_m) \log 2 + \log((s+1)!).$$

Now $d_1 + \dots + d_m \le 2 d_1$ yield:
$h(W) \le \sum_{i=0}^{s} (h(P) + (d_1 + \dots + d_m) \log 2)$

$+(s+1)\{(d_1+\cdots+d_m)\log 2+\log(d_m+1)\}$

$\leq (s+1)(h(P)+4d_1)$.

since $\log(d_m+1)\leq d_m \leq \frac{1}{2}d_1$

<u>Next</u>: obtain a lower bound for ind (W) by expanding determinant for W using properties 6.3.2 (a) (b) (c) of index to estimate from below index of W in terms of index of a typical term in expansion.

$$\begin{aligned}\text{ind}(\partial_{\mu,\nu}P) &\geq \text{ind}(P)-\frac{\mu_1}{d_1}-\cdots-\frac{\mu_{m-1}}{d_{m-1}}-\frac{\nu}{d_m}\\ &\geq \text{ind}(P)-\frac{d_m}{d_{m-1}}-\frac{\nu}{d_m}\\ &\geq \text{ind}(P)-\frac{\nu}{d_m}-\sigma\end{aligned}$$

Now since the index ≥ 0, in fact

$$\text{ind}(\partial_{\mu,\nu}P)\geq \max\left(\text{ind}(P)-\frac{\nu}{d_m},0\right)-\varepsilon.$$

Now expanding W by Laplace expansion.

$$\begin{aligned}\text{ind}(W) &\geq \min_{\pi}\left(\sum_{i=0}^{s}\text{ind}(\partial_{\mu_i,\pi(i)}P)\right)\\ &\geq \min_{\pi}\sum_{i=0}^{s}\left(\max\left(\text{ind}(P)-\frac{\pi(i)}{d_m},0\right)-\sigma\right).\\ &=\sum_{i=0}^{s}\left(\max\left(\text{ind}(P)-\frac{i}{d_m},0\right)-\sigma\right).\\ &\geq (s+1)\min\left(\frac{1}{2}\text{ind}(P),\frac{1}{2}\text{ind}(P)^2\right)-(s+1)\sigma.\end{aligned}$$

where $\displaystyle\sum_{i=0}^{s}\max\left(t-\frac{i}{s},0\right)\geq (s+1)\min\left(\frac{1}{2}t,\frac{1}{2}t^2\right)$.

Now ind (W) = ind(U) + ind(V).

we obtain upper bound by using Roth's lemma inductively on # variables to estimate ind (U) and ind (V).

Suppose true for $l < m$ variables. Now apply inductive assumption to U & V but with $(s+1) d_j$ in place of d_j. Now may verify assumptions are satisfied. Hence

$$\mathrm{ind}(U) \le 2(m-1)(s+1)\,\sigma^{1/2^{m-2}}, \quad \mathrm{ind}(V) \le (s+1)\,\sigma$$

Now $\min(\mathrm{ind}(P), \mathrm{ind}(P)^2) \le 4(m-1)\,\sigma^{1/2^{m-2}} + 4\delta'$

$\Rightarrow \mathrm{ind}(P) \le m$. Hence preceding bound may be simplified to

$$\mathrm{ind}(P)^2 \le 4m(m-1)\,\sigma^{1/2^{m-2}} + 4m\,\varepsilon \le 4m^2\,\sigma^{1/2^{m-2}}$$

this proves Roth's lemma.

Now

Roth's theorem

Thm K #-field. S: finite set of places. F: f.d. extension of K. $v \in S$. Let $\alpha_v \in F$. $|\cdot|_v$ extends to $|\cdot|_{v,K}$ of F. Then for $\kappa > 2$ there are only finitely many $\beta \in K$ s.t.

$$\prod_{v \in S} \min(1, |\beta - \alpha_v|_{v,K}) \le H(\beta)^{-\kappa}$$

The statement implies Roth's theorem by