LINEAR FORMS IN LOGARITHMS AND DIOPHANTINE EQUATIONS
C.L. STEWART

Notes taken by D. Wolczuk

A complex number α is said to be algebraic if it is the root of a non-zero polynomial with integer coefficients. A complex number which is not algebraic is said to be transcendental. Recall: The degree of an algebraic numbers is the degree of its minimal polynomial.

In 1844, Liouville proved,

Theorem 1. *Let α be an algebraic number of degree $d > 1$. There is a positive number C which is effectively computable in terms of α, such that*

$$\left|\alpha - \frac{p}{q}\right| > \frac{C(\alpha)}{q^d},$$

for $p, q \in \mathbb{Z}$, $q > 0$.

Proof:
Suppose that $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ is the minimal polynomial of α.

Note that if α is not real, then we can take $C(\alpha)$ to be the absolute value of the imaginary part of α.

Suppose $\alpha \in \mathbb{R}$. Then for $p/q \in \mathbb{Q}$,

$$|f(p/q)| \geq 1/q^d$$

since $f(p/q) \neq 0$ since f is the minimal polynomial of α and $d > 1$.

By the mean value theorem, $\exists$ a real number θ, with θ between α and p/q such that

$$1/q^d \leq |f(p/q)| = |f(\alpha) - f(p/q)| = |\alpha - p/q||f'(\theta)|. \tag{1}$$

but $f'(x) = da_d x^{d-1} + \cdots + a_1$. Observe that if $|\alpha - p/q| > 1$, we can take $C(\alpha)$ to be any positive number less than 1. Thus we may assume that $|\alpha - p/q| \leq 1$. Therefore

$$|f'(\theta)| \leq (d|a_d|(|\alpha| + 1)^{d-1} + \cdots + |a_1|) = (C(\alpha))^{-1}.$$

The result now follows from (1). □

Liouville proved the existence of transcendental numbers such as

$$\theta = \sum_{n=1}^{\infty} 10^{-n!}$$

by using Thrm 1. For example, put $p_k/q_k = \sum_{n=1}^{k} 10^{-n!}$ thus $q_k = 10^{k!}$ and $p_k = 10^{k!} \cdot \sum_{n=1}^{k} 10^{-n!}$. Obeserve that

$$|\theta - p_k/q_k| = \sum_{n=k+1}^{\infty} 10^{-n!} < \frac{2}{10^{(k+1)!}} = \frac{2}{q_k^{k+1}}. \tag{2}$$

Suppose that θ was algebraic of degree d. Note that $d > 1$, since θ is not a rational number. Then by Thrm 1, $\exists C(\alpha) > 0$ s.t. $|\alpha - p_k/q_k| > C(\alpha)/q_k^d$. But this and (2) imply $q_k^{k+1-d} < 2/C(\alpha)$ but $q_k^{k+1-d} \to \infty$ as $k \to \infty$ so α is not algebraic.

In 1873, Hermite proved that e is transcendental.

In 1874, Cantor proved that the transcendental numbers are dense in $\mathbb{R}$ by making use of the fact that the algebraic numbers are countable.

In 1882, Lindemann proved that π is transcendental.

The Hermite-Lindemann theorem states that if $\beta \in \mathbb{C}$, $\beta \neq 0$ then one at least of $\{\beta, e^\beta\}$ is transcendental.

Consider $\{\pi i, e^{\pi i}\} = \{\pi i, -1\}$ since i is algebraic, it follows that π is transcendental.

Lindemann stated and Weierstrass proved in 1885:

If $\beta_1, \dots, \beta_n$ are algebraic numbers which are linearly independent over $\mathbb{Q}$ then $e^{\beta_1}, \dots, e^{\beta_n}$ are algebraically independent.

In 1900, Hilbert proposed as the 7th of his problems:

If α is algebraic and $\alpha \neq 0, 1$ and β is algebraic and irrational, prove that α^β is transcendental. This was proved by Gelfond and Schneider independentaly in 1934.

In 1967, Baker proved that if $\alpha_1, \dots, \alpha_n$ are algebraic numbers different from 0 and 1 and $\beta_1, \dots, \beta_n$ are algebraic numbers for which $1, \beta_1, \dots, \beta_n$ are linearly independent over $\mathbb{Q}$ then

$$\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$$

is transcendental. He also proved that $e^{\beta_0}\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ is transcendental for all nonzero algebraic numbers $\beta_0, \dots, \beta_n, \alpha_1, \dots, \alpha_n$.

Quantitative Results? One can ask for a measure of how small the quantity

$$|\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n - \log \alpha_{n+1}|$$

when α_{n+1} is an algebraic number.

For our applications we are interested in the degenerate case when $\beta_0 = 0$, and β_i's are integers.

We still get bounds from Baker's argument. Put

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n \tag{3}$$

where $b_i \in \mathbb{Z}$, $i = 1, \dots, n$.

One can prove that $\Lambda = 0$ or $|\Lambda|$ is bounded away from 0 in terms of the size of the $|b_i|'s$, n, the degree of the α_i's and the heights of the α_i's.

For any algebraic number α, we define the height of α, denoted by $H(\alpha)$, by

$$H(\alpha) = \max\{|a_d|, |a_{d-1}|, \cdots, |a_0|\}$$

where $f(x) = a_d x^d + \cdots a_1 x + a_0$ is the minimal polyomial of α.

$H(\alpha)$ is known as the naive height.

Aim: If $\Lambda \neq 0$, then $|\Lambda|$ is not too small in terms of $b_1, \dots, b_n, \alpha_1, \dots, \alpha_n$.

Suppose, in (3) that the logs are always the principal branch. Put $d = [\mathbb{Q}(\alpha_1, \dots, \alpha_n) : \mathbb{Q}]$. Suppose that $A_i = \max(H(\alpha_i), e)$ for $i = 1, \dots, n$ and that $B = \max(|b_1|, \dots, |b_n|, e)$.

Theorem 2. *(1993, Baker and Wustholz) If $\Lambda \neq 0$, then*

$$|\Lambda| > \exp(-(16nd)^{2n+4} \cdot \log A_1 \cdots \log A_n \cdot \log B).$$

Proposition 3. *Suppse that α is an algebraic number, $\alpha \neq 0$ with minimal polynomial $f(x) = a_d x^d + \cdots + a_0$ then*

$$|\alpha| < \frac{H(\alpha)}{|a_d|} + 1.$$

Proof:

If $|\alpha| \leq 1$, the result is immediate. So suppose that $|\alpha| > 1$.

We have $f(\alpha) = 0$ and so

$$a_d\alpha = -a_{d-1} - a_{d-2}\alpha^{-1} - \cdots - a_0\alpha^{-d+1}.$$

Hence

$$\begin{aligned}|a_d||\alpha| &\le (|a_{d-1}| + |a_{d-2}||\alpha^{-1}| + \cdots + |a_0||\alpha^{-d+1}|)\\ &\le H(1 + |\alpha|^{-1} + \cdots + |\alpha|^{-d+1})\\ &< H\left(\frac{1}{1-|\alpha|^{-1}}\right)\\ |\alpha| - 1 &< \frac{H}{|a_d|}\end{aligned}$$

□

Remark. Note that since α^{-1} has minimal polynomial $x^d f(1/x)$, we see from Prop 3 that

$$|\alpha| > \left(\frac{H}{|a_0|} + 1\right)^{-1}.$$

Let $b_1, \ldots, b_n \in \mathbb{Z}$ with absolute value at most $B \ge 2$ and let $\alpha_1, \ldots, \alpha_n$ be nonzero algebraic numbers with heights at most A.

Proposition 4. *If* $\Lambda \neq 0$ *then* $|\Lambda| > (3A)^{-dnB}$

Proof:
Let a_j denote the leading coefficient in the minimal polynomial of α_j when $b_j \ge 0$, and let a_j denote the leading coefficient of the minimal polynomial of $(\alpha_j)^{-1}$ when $b_j < 0$.

Then we put

$$\begin{aligned}w &= a_1^{|b_1|}\cdots a_n^{|b_n|}(\alpha_1^{b_1}\cdots\alpha_n^{b_n} - 1)\\ &= a_1^{|b_1|}\cdots a_n^{|b_n|}(\alpha_1^{\epsilon_1|b_1|}\cdots\alpha_n^{\epsilon_n|b_n|} - 1)\end{aligned}$$

where $\epsilon_i = b_i/|b_i|$ for $i = 1, \ldots, n$.

Notice that w is an algebraic integer of degree at most d as $a_i^{|b_i|}\alpha_i^{\epsilon_i|b_i|}$ is an algebraic integer. This is because $a_d\alpha$ is a root of $y^d + a_{d-1}y^{d-1} + \cdots + a_0a_d^{d-1}$.

Let σ be an embedding of $\mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{C}$ which fixes $\mathbb{Q}$. Each conjugate $\sigma(w)$ of w is of the form

$$\sigma(w) = a_1^{|b_1|}\cdots a_n^{|b_n|}(\sigma(\alpha_1^{\epsilon_1})^{|b_1|}\cdots\sigma(\alpha_n^{\epsilon_n})^{|b_n|} - 1).$$

By Prop 3, $|a_i\sigma(\alpha_i^{\epsilon_i})| < 2A$ and so

$$|\sigma(w)| < 2(2A)^{B_n}. \tag{1}$$

If $w = 0$ and $\Lambda \neq 0$ then Λ is a multiple of $2\pi i$ and the result holds.

Suppose $w \neq 0$, then $|N_{k/\mathbb{Q}}(w)| \ge 1$.

Thus, by (1),

$$|w| \ge \frac{1}{\prod_{\sigma\neq id}|\sigma(w)|} \ge (2\cdot(2A)^{nB})^{-d+1}.$$

¿From the inequality $|e^z - 1| \le |z|e^{|z|}$ for all $z \in \mathbb{C}$, and on setting $z = \Lambda$, we see that

$$|\alpha_1^{b_1}\cdots\alpha_n^{b_n} - 1| \le |\Lambda|e^{|\Lambda|}.$$

If $|\Lambda| \ge 1/2$ we're done, so we may assume $|\Lambda| < 1/2$, hence $e^{|\Lambda|} < 2$. Thus $|\alpha_1^{b_1}\cdots\alpha_n^{b_n} - 1| \le 2|\Lambda|$. Recall that

$$|a_1^{|b_1|}\cdots a_n^{|b_n|}(\alpha_1^{b_1}\cdots\alpha_n^{b_n} - 1)| \ge (2(2A)^{nB})^{-d+1}$$

and so

$$|\Lambda| \geq \frac{1}{a_1^{|b_1|}\cdots a_n^{|b_n|}}\frac{1}{2(2(2A)^{nB})^{d-1}}$$
$$\geq \frac{1}{2^d(2A)^{nBd}} > (3A)^{-ndB}$$

since $nB \geq 2$. □

We can rewrite Prop 4 as if $\Lambda \neq 0$ then $|\Lambda| > \exp(-nd(\log 3A)B)$.

Suppose further that $A \geq e$ then $3A < A^3$ and so $|\Lambda| > \exp(-3nd(\log A)B)$.

Compare this with Thrm 2. Notice that Thrm 2 gives a better lower bound for the modulus of Λ, $|\Lambda|$, when $3nd(\log A)B > (16nd)^{2n+4}\log A_1\cdots\log A_n \cdot \log B$.

WLOG, we may assume that $A = A_n$, so α_n has largest height. Thus Thrm 2 improves on Prop 4 when

$$\frac{B}{\log B} > (16nd)^{2n+4}\log A_1\cdots\log A_{n-1}.$$

We get nontrivial information from Thrm 2 when the b_i's are large relative to $n, d, A_1, \dots, A_{n-1}$. Basically, Thrm 2 tells us that products of large powers of algebraic numbers can't be too close together.

Simpler situation: Suppse $a_1, \dots, a_n \in \mathbb{Q}$, nonzero and let $b_1, \dots, b_n \in \mathbb{Z}$, nonzero. Put $B_j = |b_j|$, $B = \max_j |b_j|$, $A_j = \max(H(a_j), 1)$ and $\Lambda = b_1 \log a_1 + \cdots + b_n \log a_n$.

Conjecture. (Lang + Waldschmidt) Let $\epsilon > 0$. $\exists C(\epsilon) > 0$ such that if $\Lambda \neq 0$, then

$$|\Lambda| > \frac{(C(\epsilon))^n B}{(B_1\cdots B_n A_1^2\cdots A_n^2)^{1+\epsilon}}.$$

Notice if we take $\epsilon = 1/2$ then

$$|\Lambda| > \frac{(C(\epsilon))^n}{B^{n(1+\epsilon)}A^{2n(1+\epsilon)}} > \exp(-3n(\log C(\epsilon)^{-1}) + \log B + \log A).$$

The rationale behind the conjecture:

Let S be the set of linear combinations of the $\log a_i$'s of the form $b_1 \log a_1 + \cdots + b_n \log a_n$ where $|b_j| \leq B_j$ and $H(a_j) \leq A_j$ for $j = 1, \dots, n$.

S has cardinality at most

$$(2B_1 + 1)\cdots(2B_n + 1)(2A_1 + 1)^2\cdots(2A_n + 1)^2.$$

The numbers in S are contained in the interval

$$[-nB\log A, nB\log A].$$

If the numbers are uniformly distributed in the interval we would expect that the distance to 0 from the smallest nonzero element of S in absolute value is about

$$\frac{2nB\log A}{(2B_1 + 1)\cdots(2B_n + 1)(2A_1 + 1)^2\cdots(2A_n + 1)^2}.$$

This motivates their conjecture.

Suppose that $a_1, \dots, a_n$ are positive integers of at most A and suppose that $b_1, \dots, b_n$ are postive integers of size at most B. Fix $a_1, \dots, a_n$ and suppose that $\log a_1, \dots, \log a_n$ are linearly independent over $\mathbb{Q}$.

Then the set T of linear combinations $b_1 \log a_1 + \cdots + b_n \log a_n$ has cardinality at least $(B-1)^n$ (due to linear independence). They all lie in the interval $[D, Bn\log A]$. Thus there is a nonzero difference of two elements of R of the form $b_i' \log a_1 + \cdots b_n' \log a_n$ where $|b_j'| \leq B$ for $j = 1, \dots, n$ of size at most $\frac{nB\log A}{(B-1)^n}$.

Thus, in general, the term $\log B$ which occurs in the lower bound for $|\Lambda|$ as in Thrm 2 can't be improved. Similarly, it can be shown that we need a factor of $\log A$ also.

Suppose that $\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n = 0$ with $\alpha_1, \ldots, \alpha_n$ algebraic numbers and $b_1, \ldots, b_n$ nonzero integers. Suppose that no subset of $n-1$ elements from $\{\log \alpha_1, \ldots, \log \alpha_n\}$ is $\mathbb{Q}$-lineraly dependent. Then, up to ± 1, there is a unique nonzero n-tuple of coprime integers $(k_1, \ldots, k_n)$ such that

$$k_1 \log \alpha_1 + \ldots + k_n \log \alpha_n = 0.$$

Claim: We can bound the $|k_i|$'s from above in terms of n, d and $A_1, \ldots, A_n$ ($A_i = H(\alpha_i)$).

Let α be an algebraic number with minimal polynomial $f(x) = a_d x^d + \cdots a_1 x + a_0$ so $H(\alpha) = \max_i(|a_i|)$.

Suppose that over $\mathbb{C}$, $f(x) = a_d(x-\alpha_1)\cdots(x-\alpha_d)$ so wlog $\alpha = \alpha_1$.

We put $M(\alpha) = |a_d| \prod_{i=1}^d \max(1, |\alpha_i|)$. $M(\alpha)$ is known as the Mahler measure of α and it is a more natural height function for α than $H(\alpha)$.

Jensen's Forumla. Let f be an analytic function in region containing the closed ball centered at the origin of radius $r > 0$. Suppose that $\alpha_1, \ldots, \alpha_n$ are the zeros of f in the ball repeated with multiplicity. If $f(0) \neq 0$ then

$$\log|f(0)| = -\sum_{i=1}^n \log|r/\alpha_i| + \frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})| d\theta.$$

Proposition 5. *(Landau) Let α be an algebraic number of degree d. Then*

$$M(\alpha) \leq (d+1)^{1/2} H(\alpha).$$

Proof:

Let $f(z) = a_d z^d + \cdots + a_1 z + a_0$ be the minimal polynomial of α over $\mathbb{Q}$. Note that by Parseval's equality

$$\frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2 d\theta = a_0^2 + \cdots + a_d^2.$$

On the other hand,

$$|f(e^{i\theta})| = |a_d||e^{i\theta} - \alpha_1| \cdots |e^{i\theta} - \alpha_d|.$$

Apply Jensen's Formula with $r = 1$ and let $\alpha_1, \ldots, \alpha_n$ be the roots of f of modulus at most 1. Then

$$\log|a_0| = -\log\frac{1}{|\alpha_1 \cdots \alpha_n|} + \frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})| d\theta$$

so

$$\log\frac{|a_0|}{|\alpha_1 \cdots \alpha_n|} = \frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})| d\theta.$$

Note that $f(z) = a_d(z-\alpha_1)\cdots(z-\alpha_d)$ so $|a_d \alpha_1 \cdots \alpha_d| = |a_0|$. Thus

$$\frac{|a_0|}{|\alpha_1 \cdots \alpha_n|} = |a_d||\alpha_{n+1} \cdots |\alpha_d| = |a_d| \prod_{i=1}^d \max(1, |\alpha_i|) = M(\alpha).$$

Thus

$$\log M(\alpha) = \frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})| d\theta$$

so

$$M(\alpha) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})| d\theta\right)$$

hence

$$M(\alpha)^2 = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})|^2 d\theta\right).$$

By the arithmetic geometric mean inequality for functions, Thrm 184 of Inequalities by Polya, Hardy and Littlewood,

$$\exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log|f(e^{i\theta})|^2 d\theta\right) \leq \frac{1}{2\pi}\int_0^{2\pi}|f(e^{i\theta})|^2 d\theta$$

hence

$$M(\alpha)^2 \leq a_0^2 + \cdots + a_d^2 \leq (d+1)H(\alpha)^2,$$

as required. □

Theorem 6. *Let $\alpha_1, \ldots, \alpha_n$ be nonzero algebraic numbers and suppose that $\log\alpha_1, \ldots, \log\alpha_n$ are linearly dependent over $\mathbb{Q}$. Suppose that $A_j = \max(M(\alpha_j), e^{\frac{|\log\alpha_j|}{d}}, e)$ for $j = 1, \ldots, n$ where $d = [\mathbb{Q}(\alpha_1, \ldots, \alpha_n) : \mathbb{Q}]$. Then there exist integers $t_1, \ldots, t_n$ not all zero for which $t_1\log\alpha_1 + \cdots + t_n\log\alpha_n = 0$ with*

$$|t_i| \leq \frac{(11(n-1)d^3)^{n-1}\log A_1 \cdots \log A_n}{\log A_i}$$

for $i = 1, \ldots, n$.

For the proof of Thrm 6, we need some ideas from the geometry of numbers.

A set S in $\mathbb{R}^n$ is said to symmetric about the orgin $\overline{0} = (0, \ldots, 0)$ if whenever $\overline{x} \in S$, then $-\overline{x} \in S$.

S is said to be convex if whenever $\overline{x}, \overline{y} \in S$ and λ is a real number with $0 \leq \lambda \leq 1$ then $\lambda\overline{x} + (1-\lambda)\overline{y} \in S$.

The volume of a set S in $\mathbb{R}^n$ is the Riemann integral of the characteristic function of the set, when it is integrable. It can be shown that every bounded convex set in $\mathbb{R}^n$ has a volume.

An integer point $\overline{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ is a vector with $x_i \in \mathbb{Z}$ for $i = 1, \ldots, n$.

Theorem 7. *(Minkowski, 1896) Let A be a set in $\mathbb{R}^n$ which is convex, bounded, symmetric about the origin and has volume $M(A)$.*

If $M(A) > 2^n$ then A contains an integer point different from the origin.

Proof:

Let A_m be the set of rational points in A all of whose coordinates have denominator m, so $A_m = \{(\frac{t_1}{m}, \ldots, \frac{t_n}{m}) \in A \mid t_i \in \mathbb{Z}, i = 1, \ldots, n\}$.

Let $|A_m|$ denote the cardinality of m. Then

$$\lim_{m\to\infty}\frac{|A_m|}{m^n} = M(A).$$

For m sufficiently large

$$|A_m| > (2m)^n.$$

Thus there are two distinct points $\overline{a} = (\frac{a_1}{m}, \ldots, \frac{a_n}{m})$ and $\overline{b} = (\frac{b_1}{m}, \ldots, \frac{b_n}{m}) \in A_m$ with $a_i \equiv b_i($ mod $2m)$ for $i = 1, \ldots, n$.

Then $\frac{1}{2}(\overline{a} - \overline{b})$ is an integer point (claim) which isn't $\overline{0}$.

$\overline{a}$ and $\overline{b} \in A_m$ hence in A. $-\overline{b} \in A$ since A is symmetric about $\overline{0}$. $\frac{1}{2}\overline{a} + \frac{1}{2}(-\overline{b}) = \frac{1}{2}(\overline{a} - \overline{b}) \in A$ since A is convex.

And the result follows. □

Remarks concerning Minkowski's Convex Body Thrm.

(1) 2^n is sharp. Consider the set A in $\mathbb{R}^n$ given by

$$A = \{x_1, \ldots, x_n) \in \mathbb{R}^n \mid |x_i| < 1\}.$$

Note that $\mu(A) = 2^n$ but the only integer point in A is the origin.
(2) Note that in the conclusion of Minkowski's Convex Body Thrm we can claim the existence of 2 non-zero points in the set since if $\underline{\underline{g}}$ is in the set then so is $\underline{\underline{-g}}$ by symmetry.

Theorem 7'. (Minkowski's Linear Forms Thrm) Let $B = (\beta_{ij})$ be an n x n matrix with real entries and non-zero determinant. Let $c_1, \dots, c_n$ be positive real numbers with $c_1 \cdots c_n \geq |\det B|$. Then there exists a non-zero integer point $\underline{\underline{x}} = (x_1, \dots, x_n)$ such that

$$|\beta_{i1}x_1 + \cdots + \beta_{in}x_n| < c_i \textbf{ for } i = 1, \dots, n-1$$

and

$$|\beta_{n1}x_1 + \cdots + \beta_{nn}x_n| \leq c_n.$$

Proof:
Write $L_i(\underline{\underline{x}}) = \beta_{i1}x_1 + \cdots + \beta_{in}x_n < c_i$ for $i = 1, \dots, n-1$ and put $L_i'(\underline{\underline{x}}) = \frac{1}{c_i}L_i(\underline{\underline{x}})$ for $i = 1, \dots, n$. Thus we wish to solve

$$|L_i'(\underline{\underline{x}})| < 1 \text{ for } i = 1, \dots, n-1$$

and $|L_n'(\underline{\underline{x}})| \leq 1$.

The determinant of the matrix determined by $L_i'(\underline{\underline{x}})$ for $i = 1, \dots, n$ is $\frac{\det B}{c_1 \cdots c_n} \leq 1$. Therefore, WLOG we may suppose that $c_1 = \cdots = c_n = 1$ and that $|\det B| \leq 1$.

For each $\epsilon > 0$ define A_ϵ to be the subset of $\underline{\underline{x}} \in \mathbb{R}^n$ for which $|L_i(\underline{\underline{x}})| < 1$ for $i = 1, \dots, n-1$ and $|L_n(\underline{\underline{x}})| < 1 + \epsilon$.

Note that A_ϵ is symmetric about the origin $\underline{\underline{0}}$ and is bounded. Further A_ϵ is convex since if $\underline{\underline{x}}$ and $\underline{\underline{y}}$ are in A_ϵ and $\lambda \in \mathbb{R}$ with $0 \leq \lambda \leq 1$ then

$$\begin{aligned} |L_i(\lambda\underline{\underline{x}}) + (1-\lambda)\underline{\underline{y}})| &\leq \lambda|L_i(\underline{\underline{x}})| + (1-\lambda)|L_i(\underline{\underline{y}})| \\ &< \begin{cases} \lambda + 1 - \lambda = 1 & \text{for } i = 1, \dots, n-1 \\ \lambda(1+\epsilon) + (1-\lambda)(1+\epsilon) = 1 + \epsilon & \text{for } i = n \end{cases} \end{aligned}$$

Since $\mu(A_\epsilon) = (1+\epsilon) \cdot 2^n > 2^n$ and so by Minkowski's Convex Body Thrm there is a non-zero integer point $\underline{\underline{x}}_\epsilon$ in A_ϵ.

Consider $A_{1/k}$ for $k = 1, 2, \dots$ and note that $A_1 \supseteq A_{1/2} \supseteq A_{1/3} \supset \cdots$. We obtain a sequence $\underline{\underline{x}}_{1/k}$ of non-zero integer points in $A_{1/k}$ for $k = 1, 2, \dots$. Since A_1 is bounded and contains only finitely many integer points there must be one point $\underline{\underline{y}}$ of the form $\underline{\underline{x}}_{1/k}$ for infinitely many k. Note that then $|L_i(\underline{\underline{y}})| < 1$ for $i = 1, \dots, n-1$ and $|L_n(\underline{\underline{y}})| \leq 1$. □

Suppose that α is an algebraic number with $M(\alpha) \leq 1$. Note that if $f(x) = a_d x^d + \cdots + a_0$ is the minimal polynomial of α, then, since $M(\alpha) = a_d \prod_{i=1}^d \max(1, |\alpha_i|)$ where $\alpha_1, \dots, \alpha_d$ are the roots of f, we see that $|a_d| = 1$. Therefore if $M(\alpha) \leq 1$ then α is an algebraic integer. Let $\alpha = \alpha_1$ so $\alpha_1, \dots, \alpha_d$ are the conjugates of α.

Note that $|\alpha_i^k| \leq 1$ for $i = 1, \dots, d$ and $k = 1, 2, \dots$. The elementary symmetric polynomials in the variables $x_1, \dots, x_d$ are bounded in absolute value by 2^d when evaluated at points $(y_1, \dots, y_d)$ with $|y_i| \leq 1$, for $i = 1, \dots, d$. Thus when $(y_1, \dots, y_d) = (\alpha_1^k, \dots, \alpha_d^k)$ they are integers of size at most 2^d. In particular, α^k is the root of a non-zero polynomial with integer coefficients of absolute value at most 2^d, hence of a finite set of non-zero polynomials. Therefore $\alpha^k = \alpha^l$ for some $k, l \in \mathbb{Z}^+$. Thus either $\alpha = 0$ or α is a root of unity.

This was first proved by Kronecker in 1857.

$$S = \{M(\alpha) \mid \alpha \text{ algebraic }\}.$$

Countable, is it dense?

In 1933, Lehmer asked if for each $\epsilon > 0$ there exists an algebraic number α for which $1 < M(\alpha) < 1 + \epsilon$.

The smallest value of $M(\alpha)$ larger than 1 which he found was $M(\alpha_0) = 1.17628081\cdots$ where α_0 is a root of $x^{10} + x^9 - x^7 - x^6 - x^5 - x^3 + x + 1$. This is an example of an Salem number. One root inside the circle, one outside the circle and the rest on the unit circle. No smaller example has been found.

Pisot numbers: Algebraic integers α all of whose conjugates, apart from $\alpha > 1$ lie strictly inside the unit circle. Let $\alpha = \alpha_1, \dots, \alpha_d$ be the conjugates of a Pisot number α. Then for each $k \in \mathbb{Z}^+$, $\alpha_1^k + \cdots + \alpha_d^k$ is an integer. Thus if we let $\|x\|$ denote the distance from x to the nearest integer for any $x \in \mathbb{R}$ then

$$\lim_{k\to\infty} \|\alpha_1^k\| = 0.$$

Open Question: If $\theta \in \mathbb{R}$ with $\theta > 1$ and $\lim_{k\to\infty} \|\theta^k\| = 0$, is θ a Pisot number (P.V. number, Pisot-Vijayaraghavan number)?

There is a smallest Pisot number.

In 1979, Dobrowolski proved that if α is an algebraic number of degree $d \geq 3$ and α is not a root of unity then $M(\alpha) > 1 + \frac{1}{1200}\left(\frac{\log\log d}{\log d}\right)^3$.

Theorem 8. *Let $d \in \mathbb{Z}^+$ and let α be an algebraic number of degree at most d, which is not a root of unity. Then*

$$\log M(\alpha) > \frac{1}{11d^2}.$$

We first need:

Proposition 9. *Let p be a prime number and let $f \in \mathbb{Z}[x_1, \dots, x_k]$. Then there exists $g \in \mathbb{Z}[x_1, \dots, x_k]$ such that*

$$f(x_1^p, \dots, x_k^p) - f(x_1, \dots, x_k)^p = pg(x_1, \dots, x_k).$$

Proof:

Put $\underline{\underline{x}} = (x_1, \dots, x_k)$ and $\underline{\underline{x}}^p = (x_1^p, \dots, x_k^p)$. Observe that if $f(\underline{\underline{x}})$ is a monomial, say $f(\underline{\underline{x}}) = a(x_1^{i_1} \cdots x_k^{i_k})$, then

$$f(\underline{\underline{x}}^p) - f(\underline{\underline{x}})^p = (a - a^p)x_1^{i_1 p} \cdots x_k^{i_k p} = pg(\underline{\underline{x}}),$$

where $g(\underline{\underline{x}}) = \frac{a-a^p}{p} x_1^{i_1 p} \cdots x_k^{i_k p} \in \mathbb{Z}[x_1, \dots, x_k]$ since $p \mid a - a^p$ by Fermat's little theorem.

Suppose now that the result holds for $f_1(\underline{\underline{x}})$ and $f_2(\underline{\underline{x}})$. Thus $f_1(\underline{\underline{x}}^p) - f_1(\underline{\underline{x}})^p = pg_1(\underline{\underline{x}})$ and $f_2(\underline{\underline{x}}^p) - f_2(\underline{\underline{x}})^p = pg_2(\underline{\underline{x}})$, with $g_1, g_2 \in \mathbb{Z}[x_1, \dots, x_k]$. Then

$$\begin{aligned}(f_1 + f_2)(\underline{\underline{x}}^p) - (f_1 + f_2)(\underline{\underline{x}})^p &= f_1(\underline{\underline{x}}^p) + f_2(\underline{\underline{x}}^p) - (f_1 + f_2)^p(\underline{\underline{x}}) \\ &= f_1(\underline{\underline{x}})^p + pg_1(\underline{\underline{x}}) + f_2(\underline{\underline{x}})^p + pg_2(\underline{\underline{x}}) - (f_1 + f_2)^p(\underline{\underline{x}})\end{aligned}$$

But

$$(f_1 + f_2)^p = f_1^p + \binom{p}{1} f_1^{p-1} f_2 + \cdots + \binom{p}{p-1} f_1 f_2^{p-1} + f_2^p.$$

Thus

$$\begin{aligned}(f_1 + f_2)(\underline{\underline{x}}^p) - (f_1 + f_2)(\underline{\underline{x}})^p &= p(g_1(\underline{\underline{x}}) + g_2(\underline{\underline{x}}) + \frac{1}{p}\binom{p}{1} f_1^{p-1} f_2 + \cdots + \frac{1}{p}\binom{p}{p-1} f_1 f_2 p - 1) \\ &\in \mathbb{Z}[x_1, \dots, x_k]\end{aligned}$$

since $p \mid \binom{p}{j}$ for $1 \leq j \leq p-1$.

The result now follows by induction on the number of monomials of f. □

Proposition 10. *Let α be a non-zero algebraic number. Suppose that h and l are distinct positive integers for which α^h and α^l are conjugates. Then α is a root of unity.*

Proof:
Let $\alpha = \alpha_1, \dots, \alpha_d$ be the conjugates of α and put $k = \mathbb{Q}(\alpha_1, \dots, \alpha_d)$. Since α^h and α^l are conjugates there is an element σ of $Gal(k/\mathbb{Q})$ for which

$$\sigma(\alpha^h) = \alpha^l.$$

We claim that for $n = 1, 2, \dots$

$$\sigma^n(\alpha^{h^n}) = \alpha^{l^n}.$$

True for $n = 1$. Suppose true for $1 \le k \le n$. Then

$$\begin{aligned}\sigma^{n+1}(\alpha^{h^{n+1}}) &= \sigma(\sigma^n(\alpha^{h^n})^h) = \sigma((\alpha^{l^n})^h)\\ &= \sigma((\alpha^h))^{l^n} = (\alpha^l)^{l^n} = \alpha^{l^{n+1}}.\end{aligned}$$

The claim follows by induction.

Since the Galois group is finite there is finite there is $t \in \mathbb{Z}^+$ such that σ^t is the identity. Then $\sigma^t(\alpha^{h^t}) = \alpha^{l^t}$ so $\alpha^{h^t} = \alpha^{l^t}$. Since α is non-zero α is a root of unity. □

Given an algebraic number α with conjugates $\alpha = \alpha_1, \dots, \alpha_d$ over $\mathbb{Q}$ we define the house of α, denoted $\overline{|\alpha|}$, by

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

Theorem 11. *(Dobrowolski, 1978) If α is a non-zero algebraic integer of degree d which is not a root of unity then*

$$\overline{|\alpha|} > 1 + \frac{1}{4ed^2}.$$

Proof:
Let $\alpha = \alpha_1, \dots, \alpha_d$ be the conjugates of α over $\mathbb{Q}$. Let $f_h(x_1, \dots, x_d) = x_1^h + \cdots + x_d^h$ for $h = 1, 2, \dots$. Put $S_h = f_h(\alpha_1, \dots, \alpha_d)$ for $h = 1, 2, \dots$. Note that S_h is an integer for $h = 1, 2 \dots$ since it is fixed under elements of the Galois group of the splitting field over $\mathbb{Q}$. Let p be a prime number. By Fermat's little thrm,

$$S_h^p \equiv S_h \pmod{p} \text{ for } h = 1, 2, \dots$$

On the other hand, by Prop 9,

$$S_{hp} - S_h^p = pg(\alpha_1, \dots, \alpha_d) \text{ where } g \in \mathbb{Z}[x_1 \dots, x_d].$$

Note that S_{hp} and S_h^p are in $\mathbb{Q}$ so $g(\alpha_1, \dots, \alpha_d)$ is in $\mathbb{Q}$. Since $\alpha_1, \dots, \alpha_d$ are algebraic integers and $g \in \mathbb{Z}[x_1, \dots, x_d]$, $g(\alpha_1, \dots, \alpha_d) \in \mathbb{Z}$. Thus

$$S_{hp} \equiv S_h^p \pmod{p}.$$

Therefore

$$S_{hp} \equiv S_h \pmod{p}, \text{ for } h = 1, 2, \dots$$

For each $h \in \mathbb{Z}^+$

$$|S_h| = |\alpha_1^h + \cdots + \alpha_d^h| \le d\overline{|\alpha|}^h.$$

Suppose that $\overline{|\alpha|} \le 1 + \frac{1}{4ed^2}$. By Bertrand's Postulate there is a prime p with $2ed < p < 4ed$. For $1 \le h \le d$,

$$|S_h| \le d\left(1 + \frac{1}{4ed^2}\right)^h \le de^{d\log(1+\frac{1}{4ed^2})} \le de.$$

and

$$|S_{hp}| \le d\left(1+\frac{1}{4ed^2}\right)^{hp} \le de^{4ed^2\log(1+\frac{1}{4ed^2})} \le de.$$

Therefore, for $1 \le h \le d$,

$$|S_{hp} - S_h| \le 2ed < p.$$

Thus, since $S_{hp} \equiv S_h \pmod p$,

$$S_{hp} = S_h \text{ for } h = 1, \dots, d.$$

But the Newton Sums, for $h = 1, \dots, d$ determine the elementary symmetric polynomials in $x_1, \dots, x_d$ and so the minimal polynomials of α and of α^p are the same. Thus α and α^p are conjugates and so by Prop 10, α is a root of unity. □

Let $c, k \in \mathbb{R}^+$ with $k > c$. Consider the function

$$f(t) = \log\left(1+\frac{1}{ct}\right) - \frac{1}{kt}, \text{ for } t > 0.$$

$$f'(t) = -\frac{1}{1+\frac{1}{ct}}\frac{1}{ct^2} + \frac{1}{kt^2}$$

and so $f'(t) > 0$ for $-\frac{1}{c+\frac{1}{t}} + \frac{1}{k} > 0$
so when $\frac{1}{k} > \frac{1}{c+\frac{1}{t}}$
i.e. when $\frac{1}{t} > k - c$ so when $t < \frac{1}{k-c}$. Futher $f'(t) < 0$ for $t > \frac{1}{k-c}$. In addition $\log(1+\frac{1}{ct}) - \frac{1}{kt}$ is positive for t sufficiently large.

Take $k = 11$ and $c = 4e$. Then

$$\log\left(1+\frac{1}{4et}\right) > \frac{1}{11t} \text{ for } t > \frac{1}{11-4e} = 7.88\cdots.$$

Thus, for $d \ge 3$,

$$\log\left(1+\frac{1}{4ed^2}\right) > \frac{1}{11d^2}.$$

Since

$$\log\left(1+\frac{1}{16e}\right) > .022732\cdots > \frac{1}{44} = .0227272\cdots.$$

Therefore for $d \ge 2, \in \mathbb{Z}$

$$\log\left(1+\frac{1}{4ed^2}\right) > \frac{1}{11d^2}. \tag{1}$$

Proof:
(Theorem 8 continued) If α is not an algebraic integer and $f(x) = a_d x^d + \cdots + a_0$ is the minimal polynomail of α then $|a_d| \ge 2$ and so $M(\alpha) \ge 2$ hence $\log M(\alpha) \ge \log 2$ and the result holds.

Thus we may suppose α is an algebraic integer. Note that the result holds if $d = 1$, so suppose $d \ge 2$. Then, by Thrm 11,

$$\log M(\alpha) \ge \log \overline{|\alpha|} \ge \log\left(1+\frac{1}{4ed^2}\right)$$

and, since $d \ge 2$, by (1)

$$\log M(\alpha) > \frac{1}{11d^2}.$$

□

Further remarks on heights.

Definition. Let K be a field and let $|\ |: K \to \mathbb{R}$. We say that $|\ |$ is a valuation on K if
1) $\forall a \in K, |a| \geq 0$ and $|a| = 0$ iff $a = 0$.
2) $\forall a, b \in K, |ab| = |a||b|$.
3) $\forall a, b \in K, |a+b| \leq |a| + |b|$.

Examples:
1) The ordinary absolute value on $\mathbb{C}$.
2) Let K be any field. The map

$$|a| = \begin{cases} 1 & \text{if } a \neq 0 \\ 0 & \text{if } a = 0 \end{cases}$$

is known as the trivial valuation on K.
3) Let p be a prime and let $K = \mathbb{Q}$. For any non-zero rational a/b we can define the p-adic order of a/b denoted by $\operatorname{ord}_p(a/b)$, by $\operatorname{ord}_p(a/b) = \operatorname{ord}_p a - \operatorname{ord}_p b$ where the p-adic order of an integer is the exact power of p which divides it. i.e. $ord_5(100) = 2$.

We now define $|\ |_p$ on $\mathbb{Q}$ by

$$|a/b|_p = \begin{cases} p^{-\operatorname{ord}_p(a/b)} & \text{if } a/b \neq 0 \\ 0 & \text{if } a/b = 0 \end{cases}.$$

One can check that $|\ |_p$ is a valuation. It is called the p-adic valuation on $\mathbb{Q}$.
4) Let k be any field and let T be a transcendental element over k. Put $K = k(T)$. Let λ be a real number with $0 < \lambda < 1$. Let $p(T)$ be an irreducible element in K. Then every non-zero element a of K can be written in the form $p(T)^j \frac{f(T)}{g(T)}$ where $j \in \mathbb{Z}$, f and g are coprime with p. Define $|\ |$ on K by

$$|a| = \begin{cases} \lambda^j & \text{if } a \neq 0 \\ 0 & \text{if } a = 0 \end{cases}.$$

$|\ |$ is a valuation on K.

Definition. A valuation $|\ |$ on a field on K is said to be non-Archimedean if $\forall a, b \in K$,

$$|a+b| \leq \max(|a|, |b|).$$

The last 3 examples are non-Archimedean. For example $|\ |_p$ on $\mathbb{Q}$ is non-Arch. since if $a/b = p^\alpha a_1/b_1$ with $(a_1, p) = (b_1, p) = 1$ and $c/d = p^\beta c_1/d_1$ with $(c_1, p) = (d_1, p) = 1$ then $|a/b|_p = p^{-\alpha}, |c/d|_p = p^{-\beta}$. WLOG let $\alpha = \min(\alpha, \beta)$

$$\begin{aligned} \left| p^\alpha \frac{a_1}{b_1} + p^\beta \frac{c_1}{d_1} \right|_p &= \left| p^\alpha \frac{a_1 d_1 + p^{\beta-\alpha} c_1 b_1}{b_1 d_1} \right| \\ &= |p^\alpha|_p \cdot |a_1 d_1 + p^{\beta-\alpha} c_1 b_1|_p \\ &= p^{-\alpha} |a_1 d_1 + p^{\beta-\alpha} c_1 b_1|_p \\ &\leq p^{-\alpha} = |a/b|_p \\ &\leq \max(|a/b|_p, |c/d|_p). \end{aligned}$$

Definition. Let $|\ |$ and $|\ |_1$ be vaulations on a field K. We say that they are equivalent if there is a positive number γ such that $|a| = |a|_1^\gamma$ for all $a \in K$.

We can define a metric and hence a topology on K by defining $d: K\,X\,K \to \mathbb{R}$ by

$$d(a, b) = |a - b|.$$

Equivalent valuations induce the same topology. The trivial valuation induces the discrete topology.

Ostrowski proved that every non-trivial valuation on $\mathbb{Q}$ is equivalent to either the ordinary absolute value or to a p-adic valuation for some prime p. Let us denote $|\ |$ on $\mathbb{Q}$ by $|\ |_{p_\infty}$ and we put $S(\mathbb{Q}) = \{p_\infty, p \text{ a prime in } \mathbb{Z}\}$. By the Unique Factorization Thrm for $\mathbb{Z}$, hence for $\mathbb{Q}$, we have for $\alpha \in \mathbb{Q}$,

$$\prod_{v \in S(\mathbb{Q})} |\alpha|_v = \begin{cases} 1 & \text{if } \alpha \neq 0 \\ 0 & \text{if } \alpha = 0 \end{cases} \tag{1}$$

(1) is known as the product formula. Note that (1) holds because we chose $|\ |_p$ from its equivalence class appropriately.

Let K be a finite extension of $\mathbb{Q}$. Then $K = \mathbb{Q}(\alpha)$ for some algebraic number α. Let $\alpha = \alpha_1, \dots, \alpha_d$ be the conjugates of α.

Let K be a finite extension of $\mathbb{Q}$ and let $\mathcal{O}_\mathcal{K}$ denote the ring of algebraic integers of K. Each prime p in $\mathbb{Z}$ is such that the ideal (p) in $\mathcal{O}_\mathcal{K}$ splits into a product of prime ideals $(p) = p_1^{e_1} \cdots p_t^{e_t}$ here $p_1, \dots, p_t$ are distinct prime ideals. e_i is known as the ramification of p_i for $i = 1, \dots, t$. Further if we put $[\mathcal{O}_K/p_i : \mathbb{Z}/p] = f_i$, the residue class degree of p_i, then $e_1 f_1 + \cdots + e_t f_t = [K : \mathbb{Q}]$. If K is a Galois extension of $\mathbb{Q}$ then $e_1 = \cdots = e_t$ and $f_1 = \cdots = f_t$.

Since K is a finite extension of $\mathbb{Q}$ there is an $\beta \in K$ such that $K = \mathbb{Q}(\beta)$. Suppose that $\beta_1, \dots, \beta_d$ be the conjugates of β over $\mathbb{Q}$. Then the $\mathbb{Q}$-isomorphisms σ of K into $\mathbb{C}$ are determined once we knwon $\sigma(\beta)$. Let $\sigma_i(\beta) = \beta_i$ for $i = 1, \dots, d$. We may suppose that $\beta_1, \dots, \beta_{r_1}$ are real and that $\beta_{r_1+1}, \dots, \beta_{r_1+2r_2}$ are not real and that $\beta_{r_1+i} = \overline{\beta_{r_1+r_2+i}}$ for $i = 1, \dots, r_2$. Note that $\sigma_{r_1+r_2+i} = \overline{\sigma_{r_1+i}}$ for $i = 1, \dots, r_2$. We say that $\sigma_1, \dots, \sigma_{r_1+r_2}$ are the infinite primes of K and that the prime ideals of $\mathcal{O}_\mathcal{K}$ are the finite primes. Let $S(K)$ denote the union of the finite and the inifinite primes.

We now define a valuation v on K for each prime in $S(K)$. Let $\alpha \in K$. If $v = p \in S(K)$ then we put

$$|\alpha|_v = \begin{cases} N_{K/\mathbb{Q}}(p)^{-\omega_p(\alpha)/d} & \text{if } \alpha \neq 0 \\ 0 & \text{if } \alpha = 0 \end{cases}$$

where $\omega_p(\alpha)$ is the order of p in the canonical decomposition of the fractional ideal (α) as a product of prime ideals.

Further if $v = \sigma \in S(K)$ we put

$$|\alpha|_v = \begin{cases} |\sigma(\alpha)|^{g/d} & \text{if } \alpha \neq 0 \\ 0 & \text{if } \alpha = 0 \end{cases}$$

where

$$g = \begin{cases} 1 & \text{if } \sigma \in \{\sigma_1, \dots, \sigma_{r_1}\} \\ 2 & \text{if } \sigma \in \{\sigma_{r_1+1}, \dots, \sigma_{r_1+r_2}\} \end{cases}$$

Once again it is possible to check that the product formula holds:

$$\prod_{v \in S(K)} = \begin{cases} 1 & \text{if } \alpha \neq 0 \\ 0 & \text{if } \alpha = 0 \end{cases}.$$

Note that $v \in S(K)$ means the v for each prime in $S(K)$.

We now introduce a new height function $h(\alpha)$ on K. For $\alpha \in K$ we put

$$h(\alpha) = \prod_{v \in S(K)} \max(1, |\alpha|_v).$$

If L is a finite extension of K then for $\alpha \in K$

$$\prod_{v\in S(L)} \max(1, |\alpha|_v) = \prod_{v\in S(K)} \max(1, |\alpha|_v),$$

so h is defined on the set of all algebraic numbers. The definition of h does not depend on the field containing α.

What is the link with $M(\alpha)$? Suppose that α is an algebraic number of degree d over $\mathbb{Q}$ and minimal polynomial $f(x) = a_d x^d + \cdots a_0 = d_d \prod_{i=1}^{d}(x - \alpha_i)$. Put $K = \mathbb{Q}(\alpha)$ then

$$h(\alpha) = M(\alpha)^{1/d}.$$

As before suppose that $\alpha_1, \dots, \alpha_{r_1}$ are the real conjugates of α and that $\alpha_{r_{1+1}}, \dots, \alpha_{r_1+2r_2}$ are the conjugates which are not real and that $\alpha_{r_1+i} = \overline{\alpha_{r_1+r_2+i}}$. Then

$$\prod_{v=\sigma\in S(K)} (\max(1, |\alpha|_v) = \prod_{i=1}^{r_1} \max(1, |\sigma_i(\alpha)|^{1/d}) \prod_{i=r_{1+1}}^{r_1+r_2} \max(1, |\sigma_i(\alpha)|^{2/d})$$

so

$$\prod_{v=\sigma\in S(K)} \max(1, |\alpha|_v)^d = \prod_{i=1}^{d} \max(1, |\alpha_i|).$$

It can be shown that

$$\begin{aligned}\prod_{p\in S(K)} \max(1, |\alpha|_v^d) &= \prod_{p\in S(k)} \max(1, N_{K/\mathbb{Q}}(p)^{-\omega_p(\alpha)}) \\ &= \prod_{p\in S(K)} \max(1, p^{-f\omega_p(\alpha)}) = |a_d|.\end{aligned}$$

Thus

$$\begin{aligned} h(\alpha)^d &= \prod_{v=p\in S(k)} \max(1, |\alpha|_v^d) \cdot \prod_{v=\sigma\in S(k)} \max(1, |\alpha|_v^d) \\ &= M(\alpha).\end{aligned}$$

We can now verify the following : For α an algebraic number and $k \in \mathbb{Z}^+$,

$$h(\alpha^k) = \prod_{v\in S(\mathbb{Q}(\alpha))} \max(1, |\alpha^k|_v) = \prod_{v\in S(\mathbb{Q}(\alpha))} \max(1, |\alpha|_v)^k = h(\alpha)^k.$$

Recall that if α is a non-zero algebraic number with minimal polynomial $f(x)$ of degree d then $g(x) = x^d f(\frac{1}{x})$ is the minimal polynomial of α^{-1}. But

$$M(\alpha) = \exp\left(\int_0^{2\pi} \log|f(e^{i\theta})| d\theta\right) = \exp\left(\int_0^{2\pi} \log|g(e^{i\theta})| d\theta\right) = M(\alpha^{-1}).$$

Thus $h(\alpha) = h(\alpha^{-1})$ for $\alpha \neq 0$. Therefore $h(\alpha^k) = h(\alpha)^{|k|}$ for all $k \in \mathbb{Z}$.

Further if α, β are algebraic numbers then

$$\begin{aligned} h(\alpha\beta) &= \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} \max(1, |\alpha\beta|_v) \\ &\leq \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} (\max(1, |\alpha|_v))(\max(1, |\beta|_v)) \\ &= \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} (\max(1, |\alpha|_v)) \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} (\max(1, |\beta|_v)) \\ &= h(\alpha)h(\beta)\end{aligned}$$

Theorem 6'.
Let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers which are multiplicatively dependent. Suppose that $A_i = \max(h(\alpha_i), e)$ for $i = 1, \ldots, n$ and that $A_1 \leq A_2 \leq \cdots \leq A_n$. Let $d = [\mathbb{Q}(\alpha_1, \ldots, \alpha_n) : \mathbb{Q}]$. There exists integers $t_1, \ldots, t_n$, not all zero, with $\alpha_1^{t_1} \cdots \alpha_n^{t_n} = 1$ and for $k = 1, \ldots, n$

$$|t_k| \leq (11(n-1)d^3)^n \log A_2 \cdots \log A_n$$

Proof:
For each $m \in \mathbb{Z}^+$ let $\phi(m)$ denote Euler's phi function so $\phi(m) = m \prod_{p \,|\, m}(1 - 1/p)$. Notice that

$$\phi(m)^2 = m^2 \prod_{p \,|\, m} (1 - 1/p)^2 = m \left(\frac{m}{\prod_{p \,|\, m} (\frac{p}{p-1})^2} \right).$$

Further

$$\prod_{p \,|\, m} \left(\frac{p}{p-1} \right)^2 \leq \prod_{p \,|\, m} p \prod_{p \,|\, m} \frac{p}{(p-1)^2} \leq 2m$$

Therefore $\phi(m)^2 \geq m/2$. Thus if $\phi(m) = d$ we wee that $m \leq 2d^2$.

First suppose that $n = 1$. Then α_1 is a root of unity. Since $d = [\mathbb{Q}(\alpha_1) : \mathbb{Q}]$ we see that α_1 is a m-th root of unity with $m \leq 2d^2$. In particular $\alpha_1^{t_1} = 1$ with $|t_1| \leq 2d^2$ and the result follows.

Next suppose that $n > 1$ and we may suppose wlog that no subset of $n-1$ elements from $\{\alpha_1, \ldots, \alpha_n\}$ is multiplicatively dependent. Therefore there is a unique n-tuple of relatively prime integers $(k_1, \ldots, k_n)$ with $k_1 > 0$ for which $\alpha_1^{k_1} \cdots \alpha_n^{k_n} = 1$. Let $j \in \mathbb{Z}$ with $1 \leq j \leq n$ for which $|k_j| \geq |k_i|$ for $i = 1, \ldots, n$.

Put $c_i = (11(n-1)d^3 \log A_i)^{-1}$ for $i \neq j$
and $c_j = (11(n-1)d^3)^{n-1} \log A_1 \cdots \log A_n / \log A_j$.

Observe that $c_1 \cdots c_n = 1$.

Consider the system of inequalities:

$$\left| x_i - \frac{k_i}{k_j} x_j \right| \leq c_i \quad \text{for } i = 1, \ldots, n, \ i \neq j \tag{1}$$

$$|x_j| \leq c_j \tag{2}$$

Notice that the associated matrix

$$B = \begin{pmatrix} 1 & & & \frac{-k_1}{k_j} & & \\ & 1 & & \frac{-k_2}{k_j} & & \\ & & \ddots & \vdots & & \\ & & & 1 & & \\ & & & \vdots & \ddots & \\ & & & \frac{-k_n}{k_j} & & 1 \end{pmatrix}$$

has determinant 1.

By Minkowski's Linear Forms Theorem (7'), there exists a non-zero integer point $(b_1, \ldots, b_n)$ which satisfies (1) and (2).

Put $\alpha = \alpha_1^{b_1} \cdots \alpha_n^{b_n}$. We claim that α is a root of unity. We have

$$\alpha^{k_j} = \alpha_1^{k_j b_1} \cdots \alpha_n^{k_j b_n}$$

Since $\alpha_1^{k_1}\cdots\alpha_n^{k_n}=1$ we have

$$\alpha^{k_j}=\alpha_1^{k_jb_1-k_1b_j}\cdots\alpha_n^{k_jb_n-k_nb_j}.$$

Thus

$$h(\alpha^{k_j})=h(\alpha_1^{k_jb_1-k_1b_j}\cdots\alpha_n^{k_jb_n-k_nb_j})$$

so

$$\begin{aligned}h(\alpha)^{|k_j|}&\le h(\alpha_1)^{|k_jb_1-k_1b_j|}\cdots h(\alpha_j)^0\cdots h(\alpha_n)^{|k_jb_n-k_nb_j|}\\&\le h(\alpha_1)^{c_1|k_j|}\cdots h(\alpha_j)^0\cdots h(\alpha_n)^{c_n|k_j|}\end{aligned}$$

Therefore

$$h(\alpha)\le h(\alpha_1)^{c_1}\cdots h(\alpha_{j-1})^{c_{j-1}}h(\alpha_{j+1})^{c_{j+1}}\cdots h(\alpha_n)^{c_n}.$$

Thus

$$\begin{aligned}\log(M(\alpha)^{1/d})\le\log(h(\alpha))&\le\sum_{i=1,i\neq j}^{n}c_i\log h(\alpha_i)\\&\le\sum_{i=1,i\neq j}^{n}c_i\log A_i.\end{aligned}$$

Accordingly,

$$\frac{1}{d}\log M(\alpha)\le\sum_{i=1,i\neq j}^{n}\frac{1}{11(n-1)d^3}\le\frac{1}{11d^3}.$$

Hence $\log M(\alpha)\le\frac{1}{11d^2}$. Therefore, by Theorem 8, α is a root of unity. Since the degree of α is at most d we see that α is an m-th root of unity with $m\le 2d^2$.

Therefore $\alpha^m=\alpha_1^{b_1m}\cdots\alpha_n^{b_nm}=1$. Further by (2),

$$\begin{aligned}|b_jm|&\le c_j\cdot 2d^2=2d^2(11(n-1)d^3)^{n-1}\log A_1\cdots\log A_n/\log A_j\\&\le(11(n-1)d^3)^n\log A_2\cdots\log A_n\end{aligned}$$

Furthermore, since $|k_j|\ge|k_i|$ for $i=1,\dots,n$, and

$$|b_i-\frac{k_i}{k_j}b_j|\le c_i$$

we see that, for $i=1,\dots,n$, $i\neq j$,

$$|b_i|\le c_i+|b_j|\le 1+|b_j|\le 1+c_j\le\frac{3}{2}c_j.$$

Since $m\le 2d^2$ we see that

$$|b_im|\le 3d^2c_j\le(11(n-1)d^3)^n\log A_2\cdots\log A_n.$$

□

Note if $\alpha_1^{k_1}\cdots\alpha_n^{k_n}=1$ then $k_1\log\alpha_1+\cdots+k_n\log\alpha_n=0$ for some choice of branches of the logarithms.

In 1970, Senge and Strauss proved that if $a,b\in\mathbb{Z}$ larger than 1 with $(\log a)/\log b$ irrational then the number of integers n for which the sum of the digits of n in base a plus the sum of the digits of n in base b lies below a fixed bound is finite. The proof was not effective in the sense that given a bound one could not determine the finite set of integers n. We can overcome this difficulty using Theorems 3 and 6'.

Let $\alpha, \beta \in \mathbb{Z}$ with $0 \leq \alpha < a$ and $0 \leq \beta < b$. Denote the number of digits in the base a expansion of n which are different from α by $L_{\alpha,a}(n)$. Similarly denote the number of digits in the base b expansion of n which are different from β by $L_{\beta,b}(n)$. Put

$$L_{\alpha,a,\beta,b}(n) = L_{\alpha,a}(n) + L_{\beta,b}(n).$$

Note that the sum of the digits of n in base a and base b is bounded from below by $L_{0,a,0,b}(n)$. Note that for all $n \in \mathbb{Z}$ with $n > 1$ we have

$$L_{\alpha,a,\beta,b}(n) < c_1 \log n,$$

where c_1 is a positive constant. Further on average,

$$L_{\alpha,a,\beta,b}(n) < c_c \log n, \text{ for } c_2 > 0.$$

To fix ideas $33 = 2^5 + 1 = 3^3 + 2 \cdot 3$. So $L_{0,2,0,3}(33) = 4$. The sum of the digits is this case is 5.

Similarily $63 = 2^5 + 2^4 + 2^3 + 2^2 + 2 + 1 = 2 \cdot 5^2 + 2 \cdot 5 + 3$. So $L_{1,2,2,5}(63) = 1$.

Note the condition that $\frac{\log a}{\log b}$ is irrational is necessary. For suppose that $\frac{\log a}{\log b} = \frac{r}{s}$ so that $a^s = b^r$. Then for each integer $k \in \mathbb{Z}^+$ $L_{0,a,0,b}(a^{sk}) = 2$ since $n = a^{sk}$ and $n = b^{rk}$.

Theorem 12. *(Stewart) Let $a, b \in \mathbb{Z}$ $a, b > 1$ with $\log a/\log b$ irrational and let α and β be integers with $0 \leq \alpha < a$ and $0 \leq \beta < b$. Then there is a positive number C, which is effectively computable in terms of a and b, such that if $n \in \mathbb{Z}$ with $n > 25$ then*

$$L_{\alpha,a,\beta,b}(n) > \frac{\log\log n}{\log\log\log n + C} - 1.$$

Proof:

Suppose that $n > a + b$ and consider the expansions

$$n = a_1 a^{m_1} + \alpha \frac{a_{m_1} - 1}{a - 1} + a_2 a^{m_2} + \cdots + a_r a^{m_r},$$

where $0 \leq a_1 < a$ and $-\alpha \leq a_i < a - \alpha$ for $i = 2, \ldots, r$ and

$$n = b_1 b^{l_1} + \beta \frac{b_{l_1} - 1}{b - 1} + b_2 b^{l_2} + \cdots + b_t b^{l_t},$$

where $0 \leq b_1 < b$ and $-\beta \leq b_i < b - \beta$ for $i = 2, \ldots, t$. Further $m_1 > m_2 > \cdots > m_r \geq 0$ and $l_1 > l_2 > \ldots > l_t \geq 0$.

We put

$$\theta = c_1 \log\log n, \tag{1}$$

where c_1 is a positive number which is effectively computable in terms of a and b and $c_1 > 4$. We shall assume that $n > c_2 \geq 25$, where $c_2, c_3, \ldots$ denote positive number which are effectively computable in terms of a and b and may be determined independently of c_1.

Define k to be the integer satisfying

$$\theta^k \leq \frac{\log n}{4 \log a} < \theta^{k+1} \tag{2}$$

and put

$$\Theta_1 = (1, \theta],\ \Theta_2 = (\theta, \theta^2],\ \ldots\ ,\ \Theta_k = (\theta_{k-1}, \theta^k].$$

If each of the intervals $\Theta_1, \ldots, \Theta_k$ contains at least one term either of the form $m_1 - m_s$ with $2 \leq s \leq r$ or of the form $l_1 - l_j$ with $2 \leq j \leq t$ then

$$L_{\alpha,a,\beta,b}(n) \geq r + t - 2 \geq k. \tag{3}$$

By (2),

$$(k + 1) \log\theta > \log\log n - \log(4 \log a).$$

so

$$k > \frac{\log\log n}{\log\theta} - \frac{\log(4\log a)}{\log\theta} - 1.$$

Thus by (1)

$$k > \frac{\log\log n}{\log\log\log n + \log c_1} - \log(4\log a) - 1. \tag{4}$$

Since $L_{\alpha,a,\beta,b} \geq 0$ our result now follows from (3) and (4) in this case.

Therefore we may assume that there is an integer s with $1 \leq s \leq k$ for which the interval Θ_s contains no term of the form $m_1 - m_i$ or of the form $l_1 - l_j$. Define integers p and q by the inequalities

$$m_1 - m_p \leq \theta^{s-1} \text{ and } m_1 - m_{p+1} > \theta^s, \tag{5}$$

and

$$l_1 - l_q \leq \theta^{s-1} \text{ and } l_1 - l_{q+1} > \theta^s, \tag{6}$$

with the convention that $m_{r+1} = 0$ and $l_{t+1} = 0$. Then

$$\begin{aligned}(b-1)(a-1)n &= (b-1)(a-1)a_1a^{m_1} + (b-1)\alpha a^{m_1} + (b-1)(a-1)a_2a^{m_2}\\ &+ \cdots + (b-1)(a-1)a_ra^{m_r} - (b-1)\alpha\\ &= A_1a^{m_p} + A_2,\end{aligned}$$

where

$$A_1 = (b-1)(a-1)a_1a^{m_1-m_p} + (b-1)\alpha a^{m_1-m_p} + \cdots + (b-1)(a-1)a_p$$

and

$$A_2 = (b-1)(a-1)a_{m_{p+1}}a^{m_{p+1}} + \cdots + (b-1)(a-1)a_ra^{m_r} - (b-1)\alpha.$$

Note that

$$0 < A_1 < (b-1)(a-1)a^{m_1-m_p+1} + (b-1)\alpha a_{m_1-m_p}$$

so

$$0 < A_1 < 2(b-1)(a-1)a^{m_1-m_p+1}. \tag{7}$$

Further

$$0 \leq |A_2| < (b-1)(a-1)a^{m_{p+1}+1} + |(b-1)\alpha|$$

so

$$0 \leq |A_2| < 2(b-1)(a-1)a^{m_{p+1}+1} \tag{8}$$

Similarly

$$(b-1)(a-1)n = B_1b^{l_q} + B_2,$$

where B_1 and B_2 are integers with

$$0 < B_1 < 2(b-1)(a-1)b^{l_1-l_q+1}. \tag{9}$$

$$0 \leq |B_2| < 2(b-1)(a-1)b^{l_{q+1}+1} \tag{10}$$

We have

$$1 = \frac{A_1a^{m_p} + A_2}{B_1b^{l_q} + B_2} = \frac{A_1a^{m_p}}{B_1b^{l_q}}\frac{(1 + \frac{A_2}{A_1a^{m_p}})}{(1 + \frac{B_2}{B_1b^{l_q}})}.$$

If x and y are real numbers with absolute values at most 1/2 then

$$\max\left(\frac{1+x}{1+y}, \frac{1+y}{1+x}\right) \leq 1 + 4\max(|x|, |y|). \tag{11}$$

Notice that

$$\frac{|A_2|}{A_1 a^{m_p}} \leq \frac{2(b-1)(a-1)a^{m_{p+1}+1}}{(b-1)(a-1)a^{m_1}} \leq 2a^{-m_1+m_{p+1}+1}.$$

By (6),

$$m_1 - m_{p+1} \geq \theta^s \geq \theta = c_1 \log\log n$$

and thus for n sufficiently large

$$\frac{|A_2|}{A_1 a^{m_p}} \leq \frac{1}{2}.$$

Similarly

$$\frac{|B_2|}{B_1 b^{l_q}} \leq \frac{1}{2}.$$

We put

$$R = \frac{A_1 a^{m_p}}{B_1 b^{l_q}}.$$

¿From (11) we conclude that

$$1 \leq \max(R, R^{-1}) \leq 1 + 4\max(\frac{|A_2|}{A_1 a^{m_p}}, \frac{|B_2|}{B_1 b^{l_q}})$$

so

$$\max(R, R^{-1}) \leq 1 + 8\max(a^{-m_1+m_{p+1}+1}, b^{-l_1} - l_{q+1} + 1).$$

Since $\log(1+x) < x$ for $x > 0$ we have

$$|\log R| < 8ab\max(a^{-m_1+m_{p+1}}, b^{-l_1} - l_{q+1}).$$

Thus if $\log R \neq 0$ then by (5) and (6),

$$\log|\log R| < c_3 - c_4\theta^s. \tag{12}$$

On the other hand

$$|\log R| = |\log\frac{A_1}{B_1} + m_p \log a - l_q \log b|,$$

and we can apply Prop 3 to give a lower bound for $|\log R|$. We put $n = 3$, $d = 1$ and $\alpha_1, \alpha_2, \alpha_3$ to be $\frac{A_1}{B_1}$, a and b respectively. Note that m_p and l_q are at most $\log n/\log 2$ and that the height of $\frac{A_1}{B_1}$ is at most the maximum of $|A_1|$ and $|B_1|$. Thus, by Prop 3, if $\log R \neq 0$ then

$$|\log R| \geq \exp(-c_5 \log(4\max(|A_1|, |B_1|)) \log\log n)$$

so, by (7) and 9)

$$\log|\log R| \geq -c_6 \max(1, m_1 - m_p, l_1 - l_1) \log\log n.$$

Suppose that $\log R = 0$, hence that

$$\log\frac{A_1}{B_1} + m_p \log a - l_q \log b = 0. \tag{13}$$

By Theorem 6', there exists $x_1, x_2, x_3 \in \mathbb{Z}$ not all zero with

$$x_1 \log\frac{A_1}{B_1} + x_2 \log a + x_3 \log b = 0 \tag{14}$$

and with

$$\max(|x_1|, |x_2|, |x_3|) \leq c_9 \log(\max(|A_1|, |B_1|, e).$$

By (5) and (7),

$$\log A_1 \leq c_{10}\theta^{s-1}$$

and by (6) and (9),

$$\log B_1 \leq c_{11}\theta^{s-1}$$

Therefore

$$|x_2| \leq c_{12}\theta^{s-1} \leq c_{12}\theta^{k-1}$$

and so, by (2),

$$|x_2| < \frac{\log n}{4\log a},$$

for n sufficiently large.

By (5) $m_1 - m_p \leq \theta^{s-1}$ so $m_p \geq m_1 - \theta^{s-1}$. Since $m_1 \geq \frac{\log n}{2\log a}$ and since $\theta^{s-1} \leq \theta^{k-1} < \frac{\log n}{4\log a}$ for n sufficiently large we see that

$$m_p > \frac{\log n}{2\log a} - \frac{\log n}{4\log a} = \frac{\log n}{4\log a}.$$

In particular $m_p > |x_2|$.

Recall that (14) and note that if $x_1 = 0$ then $\log a/\log b$ is rational. Thus we may suppose that $x_1 \neq 0$. Eliminating $\log\frac{A_1}{B_1}$ from (13) and (14) we find that $(x_1 m_p - x_2)\log a + (x_3 - x_1 l_q)\log b) = 0$. Since $m_p > |x_2|$ we see that $x_1 m_p - x_2 \neq 0$ hence $\log a/\log b$ is rational. □

Let p be a prime with $p \geq 3$ and let $1 = n_1 < n_2 < \cdots$ be the sequence of positive integers which are composed of primes size at most p.

In 1898, Stormer proved that

$$\liminf_{i\to\infty}(n_{i+1} - n_i) > 2.$$

In 1908, Thue proved that

$$\lim_{i\to\infty}(n_{i+1} - n_i) = \infty.$$

In 1965, Erdos used a result of Mahler to prove the following. Let $\epsilon > 0$ then there exists a positive number $N(p,\epsilon)$, such that if $n_i > N(p,\epsilon)$, then

$$n_{i+1} - n_i > n_i^{1-\epsilon}.$$

In 1973, Tijdeman used estimates for linear forms in logarithms to prove:

Theorem 13. *Let p be a prime and let $1 = n_1 < n_2 < \cdots$ be the sequence of positive integers all of whose prime factors are at most p. There exists a positive number C, which is effectively computable in terms of p, such that if $n_i \geq 3$ then*

$$n_{i+1} - n_i > n_i/(\log n_i)^C.$$

Proof:

Let $p_1, \ldots, p_k$ be the primes of size at most p. Consider the prime decompositions of n_i and n_{i+1} for some i, say

$$n_{i+1} = p_1^{a_1}\cdots p_k^{a_k} \text{ with } a_j \in \mathbb{N}$$

and

$$n_i = p_1^{b_1}\cdots p_k^{b_k} \text{ with } b_j \in \mathbb{N}$$

Then

$$0 \neq \log\frac{n_{i+1}}{n_i} = (a_1 - b_1)\log p_1 + \cdots + (a_k - b_k)\log p_k$$

and by Thrm 2 with $k = n, d = 1$, and $\alpha_1, \ldots, \alpha_k$ given by $p_1, \ldots, p_k$ respectively. Note that

$$\max_{1\leq j\leq k}(|(a_j - b_j)|) \leq \log(n_{i+1})/\log 2 \leq 1 + \frac{\log n_i}{\log 2}$$

since $n_{i+1} \leq 2n_i$. Further observe that $k \geq 2$ since if $k = 1$ the result is immediate. Therefore, by Theorem 2,

$$\log \frac{n_{i+1}}{n_i} > \exp(-k^{c_1 k} \log p_1 \cdots \log p_k \log\log n_i),$$

where c_1 is a positive number. Note that

$$\begin{aligned} k^{c_1 k} \log p_1 \cdots \log p_k &< e^{c_1 k \log k} (\log p)^k < e^{c_1 k \log k + k \log\log p} \\ &< e^{c_2 k \log k} \text{ since } p < k^2 \end{aligned}$$

and by the prime number theorem

$$e^{c_2 k \log k} < e^{c_3 p}.$$

Therefore

$$\log \frac{n_{i+1}}{n_i} > \exp(-3^{c_3 p} \log\log n_i) = \frac{1}{(\log n_i)^{e^{c_3 p}}}.$$

Since $\log(1+x) < x$ for $x > 0$,

$$\log \frac{n_{i+1}}{n_i} = \log(1 + n_{i+1} - n_i n_i) < \frac{n_{i+1} - n_i}{n_i}.$$

Therefore

$$n_{i+1} - n_i > \frac{n_i}{(\log n_i)^{e^{c_3 p}}}.$$

□

Theorem 14. *(Tijdeman) Let p and q be distinct primes and let $1 = n_1 < n_2 < \cdots$ be the sequence of positive integers whose only prime factors are p and q. There exists positive numbers c and N which are effectively computable in terms of p and q, such that if $n_i \geq N$, then*

$$n_{i+1} - n_i < \frac{n_i}{(\log n_i)^c}.$$

Before we prove Thrm 14, we need some basic facts from Diophantine approximation.

For any real number α we define an associated sequence $(\alpha_0.\alpha_1, \cdots)$ where we put $\alpha_0 = \alpha$ and

$$\alpha_k = \frac{1}{\alpha_{k-1} - [\alpha_{k-1}]}, \text{ for } k = 1, 2, \ldots,$$

provided that $\alpha_{k-1} - [\alpha_{k-1}] \neq 0$; here $[x]$ denotes the greatest integer $\leq x$ for $x \in \mathbb{R}$.

Next put $a_k = [\alpha_k]$ for $k = 0, 1, 2, \ldots$. Then

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{\cdots \frac{1}{\alpha_k}}}$$

We put $\frac{p_k}{q_k} = a_0 + \frac{1}{a_1 + \cdots + \frac{1}{a_k}}$ for $k = 0, 1, 2, \ldots$ with $(p_k, q_k) = 1$. $\frac{p_k}{q_k}$ is said to be a convergent to α. The finite continued fraction is denote by $[a_0, a_1, \ldots, a_k]$.

We'll now show that for $n = 2, 3, \ldots$

$$p_n = a_n p_{n-1} + p_{n-2} \text{ and } q_n = a_n q_{n-1} + q_{n-2} \tag{1}$$

Note that $p_0 = a_0$ and $q_0 = 1$ and $p_1 = a_0 a_1 + 1$ and $q_1 = a_1$. We have

$$\frac{p_2}{q_2} = \frac{a_0(a_1 a_2 + 1) + a_2}{a_1 a_2 + 1}$$

Observe that $p_2 = a_0(a_1 a_2 + 1) + a_2$ and $q_2 = a_1 a_2 + 1$. Further $p_2 = a_2(a_1 a_0 + 1) + a_0 = a_2(p_1 + p_0)$ and $q_2 = a_2 q_1 + q_0$. Thus (1) holds for $n = 2$. Suppose (1) holds for $n \leq k - 1 \geq 2$ and we'll prove by induction that (1) holds for $n = k$. Define a sequence of rationals $\frac{p'_j}{q'_j}$ for $j = 0, 1, 2, \dots$ with p'_j and q'_j coprime integers with $q'_j > 0$ given by

$$\frac{p'_j}{q'_j} = [a_1, \dots, a_{j+1}].$$

By the recuurence relations given by (1) for $n \le k - 1$ we find that

$$p'_{k-1} = a_k p'_{k-2} + p'_{k-3}$$

and

$$q'_{k-1} = a_k q'_{k-2} + q'_{k-3}.$$

But

$$\frac{p_j}{q_j} = a_0 + \frac{1}{\frac{p'_{j-1}}{q'_{j-1}}} = a_0 + \frac{q'_{j-1}}{p'_{j-1}}$$

so

$$p_j = a_0 p'_{j-1} + q'_{j-1} \text{ and } q_j = p'_{j-1}. \tag{2}$$

Thus, on taking $j = k$, we find that

$$p_k = a_0 p'_{k-1} + q'_{k-1} \text{ and } q_k = p'_{k-1}.$$

In particular,

$$\begin{aligned} p_k &= a_0(a_k p'_{k-2} + p'_{k-3}) + (a_k q'_{k-2} + q'_{k-3}) \\ &= a_k(a_0 + p'_{k-2} + q'_{k-2}) + (a_0 p'_{k-3} + q'_{k-3}) \end{aligned}$$

and so by (2)

$$p_k = a_k p_{k-1} + p_{k-2}.$$

Similarly

$$q_k = p'_{k-1} = a_k p'_{k-2} + p'_{k-3} = a_k q_{k-1} + q_{k-2}.$$

The result now follows by induction.

Recall, $\alpha = [a_0, a_1, \dots, a_k, \alpha_{k+1}]$ and $1/\alpha_{k+1} \le 1/a_{k+1}$ and so

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \alpha < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

In particular α lies between p_n/q_n and p_{n+1}/q_{n+1} for $n = 0, 1, 2, \dots$.

Proposition 15. *$p_n q_{n+1} - q_n p_{n+1} = (-1)^{n+1}$ for $n = 0, 1, 2, \dots$.*

Proof:
Notice that the result holds for $n = 0$ since $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, $q_1 = a_1$ hence $p_0 q_1 - q_0 p_1 = a_0 a_1 - a_0 a_1 + 1 = -1$.

Suppose its true for $0, \dots, n - 1$. Then

$$\begin{aligned} p_n q_{n+1} - q_n p_{n+1} = (-1)^{n+1} &= p_n(a_{n+1} q_n + q_{n-1}) - q_n(a_{n+1} p_n + p_{n-1}) \\ &= p_n q_{n-1} - q_n p_{n-1} \\ &= (-1)(p_{n-1} q_n - q_{n-1} p_n) = (-1)^{n+1} \end{aligned}$$

and the result follows by induction. □

By Prop 15,

$$\frac{p_k}{q_k} - \frac{p_{k+1}}{q_{k+1}} = \frac{p_k q_{k+1} - q_k p_{k+1}}{q_k} q_{k+1} = \frac{(-1)^{k+1}}{q_k q_{k+1}}.$$

Since α lies between p_n/q_n and p_{n+1}/q_{n+1} for $n = 0, 1, 2, \ldots$ we conclude that

$$|\alpha - \frac{p_n}{q_n}| \leq \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}.$$

It is not difficult to show that if $|\alpha - p/q| < \frac{1}{2q^2}$ then $p/q = p_n/q_n$ for some n with $n \geq 0$.

Lemma 16. *(Tijdeman) Let p and q be distinct primes let $\frac{h_0}{k_0}, \frac{h_1}{k_1}, \cdots$ be the sequence of convergents to $\frac{\log p}{\log q}$. There exists a positive number c which is effectively computable in terms of of p and q such that*

$$k_{j+1} < k_j^c \log q, \textit{ for } j = 0, 1, 2, \ldots .$$

Proof:
We may suppose that $j \geq 2$ and so $k_j \geq 2$. Since $|\frac{\log p}{\log q} - \frac{h_j}{k_j}| < \frac{1}{k_j k_{j+1}}$. Hence

$$|k_j \log p - h_j \log q| < \frac{\log q}{k_{j+1}}.$$

By Thrm 2 with $\alpha_1 = p$, $\alpha_2 = q$, $d = 1$ we find that

$$|k_j \log p - h_j \log q| > \exp(-c_1 \log(\max(|h_j|, |k_j|))$$

where $c_1, c_2, \ldots$ denote positive numbers which are effectively computable in terms of p and q.

Since $|\frac{\log p}{\log q} - \frac{h_j}{k_j}| < \frac{1}{k_j k_{j+1}}$ we see that $|h_j| < c_2 |k_j|$ hence

$$|k_j \log p - h_j \log q| > \exp(-c_3 \log k_j) = \frac{1}{k^{j^{c_3}}}.$$

Thus

$$\frac{1}{k_j^{c_3}} < \frac{\log q}{k_j + 1}$$

hence

$$k_{j+1} < k_j^{c_3} \log q.$$

□

Proof:
Theorem 14

Put $n_i = n = p^u q^v$. we may suppose that

$$p^u \geq \sqrt{n},$$

hence

$$u \geq \frac{\log n}{2 \log p}. \qquad (1)$$

Let $\frac{h_0}{k_0}, \frac{h_1}{k_1}, \cdots$ be the sequence of convergents to $\frac{\log p}{\log q}$. Choose j so that

$$k_j \leq u < k_{j+1}.$$

We may suppose that N is sufficiently large so that $n \geq 3$ and $j \geq 2$. We shall distinguish two cases according to whether $\frac{h_j}{k_j}$ is bigger or smaller than $\frac{\log p}{\log q}$.

Case 1 $\frac{h_j}{k_j} > \frac{\log p}{\log q}$.

Put $n' = p^{u-k_j}q^{v+h_j}$. Note that $n' \in \mathbb{Z}$ and $n' > n$ since $\frac{h_j}{k_j} > \frac{\log p}{\log q}$. Thus $n' \geq n_{i+1}$. We have

$$\frac{h_j}{k_j} - \frac{\log p}{\log q} < \frac{1}{k_j k_{j+1}},$$

and so

$$h_j \log q - k_j \log p < \frac{\log q}{k_j + 1}.$$

Therefore

$$\log\left(\frac{n'}{n}\right) = \log\left(\frac{q^{h_j}}{p^{k_j}}\right) < \frac{\log q}{k_{j+1}}.$$

we have $k_{j+1} > u \geq \frac{\log n}{2\log p}$ and so

$$0 < \log\left(\frac{n'}{n}\right) < \frac{2\log p \log q}{\log n}. \tag{2}$$

For n sufficiently large in terms of p and q

$$0 < \frac{n'}{n} - 1 < \exp\left(\frac{2\log p \log q}{\log n}\right) - 1 < \frac{1}{2}.$$

Since, for $|x| < 1$,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots .$$

Further, for $0 < x < \frac{1}{2}$, we have

$$\log(1+x) > x - \frac{x^2}{2} = x(1 - x/2) > x/2.$$

Thus

$$\log\left(\frac{n'}{n}\right) = \log\left(1 + (\frac{n'}{n} - 1)\right) > \frac{1}{2}\left(\frac{n'}{n} - 1\right),$$

for n sufficiently large. By (2)

$$\frac{n'}{n} - 1 < \frac{4\log p \log q}{\log n} \text{ so } n' < n + \frac{4\log p \log qn}{\log n}$$

Let $c_1, c_2, \cdots$ denote positive constancts which are effectively compuatable in terms of p and q. Thus

$$n' < n + \frac{c_1 n}{\log n}, \text{ for } n > c_2.$$

Since $n_{i+1} \leq n'$,

$$n_{i+1} < n_i + \frac{c_1 n_i}{\log n_i}, \text{ for } n_i > c_2.$$

<u>Case 2</u> $\frac{h_j}{k_j} < \frac{\log p}{\log q}$.

Hence $\frac{h_{j-1}}{k_{j-1}} > \frac{\log p}{\log q}$. Put $n' = p^{u-k_{j-1}}q^{v+h_{j-1}}$. Again note $n' \in \mathbb{Z}$ and $n' \geq n_{i+1}$. We have

$$\begin{aligned}\frac{h_{j-1}\log q - k_{j-1}\log p}{k_{j-1}\log q} &= \frac{h_{j-1}}{k_{j-1}} - \frac{\log p}{\log q}\\ &< \frac{h_{j-1}}{k_{j-1}} - \frac{h_j}{k_j} = \frac{1}{k_{j-1}k_j}.\end{aligned}$$

Therefore,

$$\begin{aligned}\log\left(\frac{n'}{n}\right) &= \log\left(\frac{\log q^{h_{j-1}}}{p^{k_{j-1}}}\right)\\ &= h_{j-1}\log q - k_{j-1}\log p\\ &< \frac{k_{j-1}\log q}{k_{j-1}k_j}.\end{aligned}$$

Accordingly

$$\log\left(\frac{n'}{n}\right) < \frac{\log q}{k_j}.$$

We find from Lemma 16, that

$$k_j > \left(\frac{k_{j+1}}{\log q}\right)^{\frac{1}{c}}.$$

By (1) we have $u \geq \frac{\log n}{\log p}$ and, since $k_{j+1} > u$, we see that

$$\begin{aligned}\log\left(\frac{n'}{n}\right) &< \frac{\log q}{\left(\frac{\log n}{2\log p\log q}\right)^{\frac{1}{c}}}\\ &= \frac{(\log q)^{1+1/c}(2\log p)^{1/c}}{(\log n)^{1/c}} \qquad (3)\end{aligned}$$

Thus, for n sufficiently large,

$$\log\left(\frac{n'}{n}\right) > \frac{1}{2}\left(\frac{n'}{n} - 1\right).$$

Comparing this estimate with the upper bound (3), we find that

$$n' < n + (2(\log q)^{1+1/c}(2\log p)^{1/c}\frac{n}{(\log n)^{1/c}}.$$

Since $n' \geq n_{i+1}$ we see that

$$n_{i+1} \leq n_i + \frac{c_3 n_i}{(\log n_i)^{1/c}},\ \text{for } n_i > c_4$$

Thus in cases 1 and 2,

$$n_{i+1} \geq n_i + \frac{c_5 n_i}{(\log n_i)^{c_6}},\ \text{for } n_i > c_7.$$

Therefore

$$n_{i+1} \geq n_i + \frac{n_i}{(\log n_i)^{c_8}},\ \text{for } n_i > c_9.$$

□

Background from Algebraic Number Theory

Let K be a finite extension of $\mathbb{Q}$. Thus there exists some irreducible polynomial $f(\in \mathbb{Q}[x])$ such that K is $\mathbb{Q}[x]/f\mathbb{Q}[x]$. Suppose that $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ and that f factors over $\mathbb{C}$ as $f(x) = (x-\alpha_1)\cdots(x-\alpha_n)$. By the Primitive Element Theorem there exists a θ such that $K = \mathbb{Q}(\theta)$. There are n field embeddings of K into $\mathbb{C}$, that is, n isomorphisms of K into $\mathbb{C}$ which are the identity on $\mathbb{Q}$. They are given by $\theta \to \alpha_i$ for $i = 1, \ldots, n$ and so determine n conjugate fields $\mathbb{Q}(\alpha_1), \ldots, \mathbb{Q}(\alpha_n)$ in $\mathbb{C}$, non necessarily distinct). Let $\alpha_1, \ldots, \alpha_{r_1}$ be real numbers and let $\alpha_{r_1+1}, \ldots, \alpha_n$ be complex numbers which are not real. We may suppose that

$$\alpha_{r_i+i} = \overline{\alpha_{r_i+r_2+i}} \text{ for } i = 1, \ldots, r_2.$$

Note that $n = r_1 + 2r_2$.

The ring of algebriac integers of K, denoted $\mathcal{O}_{\mathcal{K}}$, consists of the elements of K which are roots of monic polynomials with integer coefficients. $\mathcal{O}_{\mathcal{K}}$ is a Dedekind domain.

Definition. An integral domain $\mathcal{O}$ is a Dedekind domain if:
1) $\mathcal{O}$ is a Noetherian ring.
2) $\mathcal{O}$ is integrally closed in its field of fractions.
3) All non-zero prime ideals are maximal ideals.

If $\mathcal{O}$ is a Dedekind domain then there is unique factorization of ideals into prime ideals up to reordering. Note that we do not always have unique factorization of elements of $\mathcal{O}_{\mathcal{K}}$ into prime elements in $\mathcal{O}_{\mathcal{K}}$.

Let K be an extension of $\mathbb{Q}$ of degree n with ring of algebraic integers $\mathcal{O}_{\mathcal{K}}$. Then there exists a set $\{\omega_1, \dots, \omega_n\}$ of elements from $\mathcal{O}_{\mathcal{K}}$ such that every element of $\mathcal{O}_{\mathcal{K}}$ has a unique representation as an integral linear combination of $\omega_1, \dots, \omega_n$. We call $\{\omega_1, \dots, \omega_n\}$ an integral basis for $\mathcal{O}_{\mathcal{K}}$. Any two integral bases for $\mathcal{O}_{\mathcal{K}}$ (for K) are related by a matrix with determinant ± 1.

Definition. The discriminant D of K is

$$D = (\det(\sigma_i(\omega_j)))^2,$$

where $\sigma_1, \dots, \sigma_n$ are the embeddings of K in $\mathbb{C}$ and $\{\omega_1, \dots, \omega_n\}$ is an integral basis. Note that D is a non-zero integer.

Consider the set S of non-zero ideals in $\mathcal{O}_{\mathcal{K}}$. We define a relation $\sim$ on S by saying that $a \sim b$, for ideals $a, b \in S$, if there exists non-zero elements α, β in $\mathcal{O}_{\mathcal{K}}$ such that $[\alpha]a = [\beta]b$; here $[\alpha]$ denote the principal ideal generated by α in $\mathcal{O}_{\mathcal{K}}$ and similarly for $[\beta]$. $\sim$ is an equivalence relation on S and if $a \sim [1]$ then a is principal.

We can define a multiplication on the equivalence classes of S by taking multiplication of representatives. This determines a finite abelian group called the ideal class group of K. The order of the group is called the class number and is denoted by h. Thus if a is an ideal of $\mathcal{O}_{\mathcal{K}}$ then $a^h \sim [1]$.

Let K be a finite extension of $\mathbb{Q}$. The group of units $U(K)$ of $\mathcal{O}_{\mathcal{K}}$ is the set of invertable elements in $\mathcal{O}_{\mathcal{K}}$. It forms an abelian group under multiplication. Plainly the roots of unity of K are in $U(K)$.

In 1846, Dirichlet proved that $U(K)$ is isomorphic to

$$\mu(K) \text{ x } \mathbb{Z}^r,$$

where $\mu(K)$ is a finite torsion group, where $r = r_1 + r_2 - 1$ and where K has r_1 real embeddings in $\mathbb{C}$ and $2R_2$ non-real embeddings. Let $\sigma_1, \dots, \sigma_{r_1}$ be the real embeddings of K in $\mathbb{C}$ and suppose that $\sigma_{r_1+1}, \dots, \sigma_{r_1+2r_2}$ are the complex emebbedings and that $\sigma_{r_1+i} = \overline{\sigma_{r_1+r_2+i}}$ for $i = 1, \dots, r_2$.

By Dirichlet's result there exists units $u_1, \dots, u_r$ such that every element x in $U(K)$ can be written in the form

$$X = \zeta \cdot u_1^{b_1} \cdots u_r^{b_r} \text{ where } b_1, \dots, b_r \in \mathbb{Z}, \zeta \text{ a root of unity.}$$

The set $\{u_1, \dots, u_r\}$ is said to be a fundamental system of units. While the fundamental system of units need not be unique it is possible to attach a volume to the system which is independent of the choice of system. We make use of the logarithmic map $L : K^* \to \mathbb{R}^{r+1}$ given by

$$L(\alpha) = (\log|\sigma_1(\alpha)|, \dots, \log|\sigma_{r_1}(\alpha)|, 2\log|\sigma_{r_1+1}(\alpha)|, \dots, 2\log|\sigma_{r_1+r_2}(\alpha)|).$$

L is an abelian group homomorphism.

For any $\alpha \in K$ we let $N_{K/\mathbb{Q}}(\alpha)$ denote the norm of α and we put

$$N_{K/\mathbb{Q}}(\alpha) = \prod_{i=1}^{r_1+2r_2} \sigma_i(\alpha).$$

Note that the norm is multiplicative, in other words, $N_{K/\mathbb{Q}}(\alpha\beta) = N_{K/\mathbb{Q}}(\alpha)N_{K/\mathbb{Q}}(\beta)$. We see that if $x \in U(K)$ then $N_{K/\mathbb{Q}}(x) = \pm 1$. In particular, if $\alpha \in U(K)$ then $L(\alpha)$ lies in the r dimensional subspace of $\mathbb{R}^{r+1}$ given by

$$H = \{(x_1, \dots, x_r) \mid x_1 + \cdots + x_r = 0\}.$$

Thus $L : U(K) \to H$. The image of $U(K)$ under L is a lattice in H and the kernel of L is just the set of roots of unity of K. The volume of a fundamental region for the lattice is called the regulator R_K of K.

Let $\{u_1, \dots, u_r\}$ be a fundamental system of units. We have

$$R_K = \det |\delta_i \log |\sigma_i(u_j)||_{i=1,\dots,r//j=1,\dots,r},$$

where

$$\delta_i = \begin{cases} 1 & \text{if } 1 \le i \le r_1 \\ 2 & \text{if } r_1 < i \end{cases}.$$

In 1981, Zimmert, sharpening work of Remak, proved that $R_K > .056$. In 1918, Landau proved that there is positive number C, which depends on $d = [K : \mathbb{Q}]$, such that

$$\log(hR) < C|D|^{1/2}(\log|D|)^{d-1}.$$

Since the class number h of K is always at least 1 this yields an upper bound for R in terms of the discriminant D and the degree d of K over $\mathbb{Q}$.

Let I be an ideal in $\mathcal{O}_\mathcal{K}$. The norm of I, denoted $N(I)$, is the cardinality of $\mathcal{O}_\mathcal{K}/\mathcal{I}$. It can be shown that if α in $\mathcal{O}_\mathcal{K}$ and $[\alpha]$ denotes the principal ideal of $\mathcal{O}_\mathcal{K}$ generated by α then

$$N([\alpha]) = |N_{K/\mathbb{Q}}(\alpha)|.$$

Dedekind introduced a generalization of the Riemann zeta function $\zeta(s)$. Let $[K : \mathbb{Q}] < \infty$. He defined $\zeta_K(s)$ for $Re(s) > 1$ by

$$\zeta_K(s) = \sum_I \frac{1}{N(I)^s},$$

where the sum is taken over all non-zero ideals of $\mathcal{O}_\mathcal{K}$. This converges uniformly on compact subsets to an analytic function for $Re(s) > 1$. Further $\zeta_\mathbb{Q}(s) = \zeta(s)$. It can be shown that $\zeta_K(s)$ can be analytically continued to all of $\mathbb{C}$ with the exception of a simple pole at $s = 1$. Further there is a functional equation which relates $\zeta_K(s)$ with $\zeta_K(1-s)$. The Generalized Riemann Hypothesis (GRH) is that the only zeros of $\zeta_K(s)$ with $0 \le Re(s) \le 1$ have $Re(s) = 1/2$. There is an Euler product representation for $Re(s) > 1$ given by

$$\zeta_K(s) = \prod_p \left(1 - \frac{1}{N(p)^s}\right)^{-1}$$

where the product is taken over all prime ideals p of $\mathcal{O}_\mathcal{K}$.

Let $\omega(K)$ denote the number of roots of unity in K. Then

$$\lim_{s\to 1}(s-1)\zeta_K(s) = \frac{2^{r_1}(2\pi)^{r_2}hR_K}{\omega(K)\sqrt{|D|}}.$$

Thue Equations

Let $F(x,y) = a_n x^n + a_{n-1}x^{n-1}y + \cdots + a_0 y^n$ be a binary form with integer coefficients, and suppose $n \geq 3$. Let m be a non-zero integer. Suppose that F has non-zero discriminant. The equation

$$F(x,y) = m \tag{4}$$

is known as a Thue equation. Thue proved in 1909 that if $a_n \neq 0$ and $F(x,1)$ is irreducible then (4) has only finitely many solutions in integers x and y.

Eg $x^3 - 2y^3 = 6$. → Thue - only finitely many. In fact (2,1) is the only solution.

By contrast $x^2 - 2y^2 = 1$ has infinitely many solutions in integers x and y.

Let K be a finite extension of $\mathbb{Q}$. It is possible to choose a fundamental system of units $\{u_1, \ldots, u_r\}$ such that

$$\max_{1 \leq i \leq r, 1 \leq \mathrm{J} \leq r} |\log |u_i^{(j)}|| < CR,$$

where R is the regulator of K and C is a number which depends on $[K : \mathbb{Q}]$.

Theorem 17. *Let F be an irreducible binary form with integer coefficients and degree $n \geq 3$. Let m be a non-zero integer. There exists a positive number C, which is effectively computable in terms of F, such that all solutions in integers x and y of the Diophantine equation $F(x,y) = m$ satisfy*

$$\max\{|x|, |y|\} < |2m|^{C \log\log |4m|}.$$

Proof:

Let $c_1, c_2, \ldots$ denote positive numbers which are effectively computable in terms of F. Let $F(x,y) = a_n x^n + \cdots + a_1 x y^{n-1} + a_0 y^n$. By considering $a_n^{n-1}F(x,y) = a_n^{n-1}m$ and letting $a_n^{n-1}F(x,y) = f(X,y)$ with $X = a_n x$ we see that we may suppose without loss of generality $a_n = 1$.

Next let $F(x,y) = (x - \alpha^{(1)}y)\cdots(x - \alpha^{(n)}y)$ be the factorization of F over $\mathbb{C}$. Suppose, as usual, that $\alpha^{(1)}, \ldots, \alpha^{(r_1)}$ are real and that $\alpha^{(r_1+i)} = \overline{\alpha^{(r_1+r_2+i)}}$ for $i = 1, \ldots, r_2$, hence that $n = r_1 + 2r_2$.

Suppose that x and y are integers for which $F(x,y) = m$, and put

$$\beta^{(i)} = x - \alpha^{(i)}y, \text{ for } i = 1, \ldots, n.$$

Then

$$|\beta^{(1)} \cdots \beta^{(n)}| = |m|.$$

Put $K = \mathbb{Q}(\alpha^{(1)})$ and let $\eta_1^{(1)}, \ldots, \eta_r^{(1)}$ be a fundamental system of units for which

$$\max_{\substack{1 \leq i \leq r \\ 1 \leq j \leq r}} |\log |\eta_j^{(i)}|| < c_1 R_K.$$

Every point p in $\mathbb{R}^r$ is within c_2 of the lattice generated by the vectors

$$(\log |\eta_1^{(i)}|, \ldots, \log |\eta_r^{(i)}|) \text{ for } i = 1, \ldots, r.$$

Take

$$p = (\log(|m|^{-1/n}|\beta^{(1)}|), \ldots, \log(|m|^{-1/n}|\beta^{(r)}|)).$$

Then we can find integers $b_1, \ldots, b_r$ such that

$$|b_1 \log |\eta_i^{(1)} + b_2 \log |\eta_i^{(2)}| + \cdots + b_r \log |\eta_i^{(r)}| + \log(|m|^{-1/n}|\beta^{(i)}|)| < c_2,$$

for $i = 1, \ldots, r$. Now put, for $j = 1, \ldots, n$,

$$\gamma^{(j)} = \beta^{(j)} \eta_1^{(}j)^{b_1} \cdots (\eta_r^{(j)})^{b_r}. \tag{1}$$

Observe that, for $j = 1, \dots, r$,

$$|\log |m|^{-1/n}|\gamma^{(j)}|\,| < c_2.$$

Since $|\gamma^{(r_1+i)}| = |\gamma^{(r_1+r_2+i)}|$ for $i = 1, \dots, r$ we see that

$$|\log |m|^{-1/n}|\gamma^{(j)}|\,| < c_2$$

holds for $j = 1, \dots, n$ except for $j = n$ when $r_1 = n$ or for $j = r_1 + r_2$ and $j = r_1 + 2r_2$ when $r_1 < n$. But $|\gamma^{(1)} \cdots \gamma^{(n)}| = |\beta^{(1)} \cdots \beta^{(n)}| = m$ and so

$$\sum_{j=1}^{n} \log |m|^{-1/n}|\gamma^{(j)}| = 0.$$

Therefore

$$|\log(|m|^{-1/n}|\gamma^{(j)}|)| < c_3 \text{ for } j = 1, \dots n. \tag{2}$$

Note that since $\beta^{(j)}$ is an algebraic integer so is $\gamma^{(j)}$ for $j = 1, \dots, n$. Further the height $H(\gamma^{(j)})$ is at most $c_4|m|$ since the coefficients in the minimal polynomial are elementary symmetric polynomials in the $\gamma^{(i)}$'s, hence are at most $c_4|m|$ in absolute value.

Consider now the equations obtained from (1) by taking logarithms:

$$b_1 \log |\eta_1^{(j)}| + \cdots + b_r \log |\eta_r^{(j)}| = \log \left|\frac{\gamma^{(j)}}{\beta^{(j)}}\right|,$$

for $j = 1, \dots, r$. By Cramer's Rule, for $i = 1, \dots, r$:

$$b_i = \frac{\det \begin{pmatrix} \log |\eta_1^{(1)}| & \cdots & \log |\frac{\gamma^{(1)}}{\beta^{(1)}}| & \cdots & \log |\eta_r^{(1)}| \\ \vdots & & \vdots & & \vdots \\ \log |\eta_1^{(r)}| & \cdots & \log |\frac{\gamma^{(r)}}{\beta^{(r)}}| & \cdots & \log |\eta_r^{(r)}| \end{pmatrix}}{\det \begin{pmatrix} \log |\eta_1^{(1)}| & \cdots & \log |\eta_r^{(1)}| \\ \vdots & & \vdots \\ \log |\eta_1^{(r)}| & \cdots & \log |\eta_r^{(r)}| \end{pmatrix}} \tag{3}$$

Our aim is to bound $|b_i|$ in terms of m hence to bound $|x|, |y|$ in terms of m. Let $B = \max(|b_1|, \dots, |b_r|)$ and suppose that $|b_i| = B$. Notice that Δ, the determinant in the denominator, satisfies

$$\Delta = R \cdot 2^{-r_2}.$$

Expanding the determinant in the numerator of (3) along the i-th column we see that

$$\max_{1 \le j \le r} \left| \log \left| \frac{\gamma^{(j)}}{\beta^{(j)}} \right| \right| > c_5 B.$$

Let the maximum occur for $j = J$. Then

$$\begin{aligned} |\log |m|^{-1/n} \log |\beta^{(J)}|\,| &= \left| \log \frac{|\beta^{(J)}|}{|\gamma^{(J)}|} + \log |m|^{-1/n}|\gamma^{(J)}|) \right| \\ &> c_5 B - c_3. \end{aligned}$$

Since $\sum_{j=1}^{n} \log |m|^{-1/n}|\beta^{(j)}| = 0$ it follows that for some l with $1 \le l \le n$

$$\log(|m|^{-1/n}|\beta^{(l)}|) < \frac{c_3 - c_5 B}{n-1}.$$

In particular,

$$|\beta^{(l)}| < |m|^{1/n} c_7 e^{-c_6 B}. \tag{4}$$

Since $|\beta^{(1)} \cdots \beta^{(n)}| = |m|$, there exists an integer k with $k \neq l$ such that

$$|\beta^{(k)}| > |m|^{1/n} c_7^{-1/(n-1)} e^{c_6 B/(n-1)}. \tag{5}$$

Let j be an integer with $1 \leq j \leq n$ and with $j \neq k$ and $j \neq l$. Note that j exists since $n \geq 3$. We have the identity

$$(\alpha^{(k)} - \alpha^{(l)})\beta^{(j)} - (\alpha^{(j)} - \alpha^{(l)})\beta^{(k)} = (\alpha^{(k)} - \alpha^{(j)})\beta^{(l)}.$$

Note

$$(\alpha^{(k)} - \alpha^{(l)})(x - \alpha^{(j)}y) - (\alpha^{(j)} - \alpha^{(l)})(x - (\alpha^{(k)}y) = (\alpha^{(k)} - \alpha^{(j)})(x - \alpha^{(l)}y).$$

Coef of y:

$$\begin{aligned} -(\alpha^{(k)} - \alpha^{(l)})\alpha^{(j)} + (\alpha^{(j)} - \alpha^{(l)})\alpha^{(k)} &= -(\alpha^{(l)}\alpha^{(k)} - \alpha^{(l)}\alpha^{(j)}) \\ &= -(\alpha^{(k)} - \alpha^{(j)})\alpha^{(l)} \end{aligned}$$

Divide by $(\alpha^{(k)} - \alpha^{(l)})\beta^{(k)}\gamma^{(j)}/\gamma^{(k)}$ to get

$$\frac{\beta^{(j)}\gamma^{(k)}}{\gamma^{(j)}\beta^{(k)}} - \frac{\alpha^{(j)} - \alpha^{(l)}}{\alpha^{(k)} - \alpha^{(l)}}\frac{\gamma^{(k)}}{\gamma^{(j)}} = \frac{\alpha^{(k)} - \alpha^{(j)}}{\alpha^{(k)} - \alpha^{(l)}})\frac{\beta^{(l)}\gamma^{(k)}}{\beta^{(k)}\gamma^{(j)}}$$

$$\Rightarrow \left(\frac{\eta_1^{(k)}}{\eta_1^{(j)}}\right)^{b_1} \cdots \left(\frac{\eta_r^{(k)}}{\eta_r^{(j)}}\right)^{b_r} - \alpha_{r+1} = \lambda$$

where

$$\alpha_{r+1} = \frac{\alpha^{(j)} - \alpha^{(l)}}{\alpha^{(k)} - \alpha^{(l)}}\frac{\gamma^{(k)}}{\gamma^{(j)}}$$

and

$$\lambda = \left(\frac{\alpha^{(k)} - \alpha^{(j)}}{\alpha^{(k)} - \alpha^{(l)}}\right)\frac{\beta^{(l)}\gamma^{(k)}}{\beta^{(k)}\gamma^{(j)}}$$

Put $\alpha_i = \frac{\eta_i^{(k)}}{\eta_i^{(j)}}$ Thus

$$\alpha_1^{b_1} \cdots \alpha_r^{b_r}\alpha_{r+1}^{-1} = 1 + \lambda$$

Thus taking the principal branch of the logarithm

$$\log(\alpha_1^{b_1} \cdots \alpha_r^{b_r}\alpha_{r+1}^{-1}) = \log\left(1 + \frac{\lambda}{\alpha_{r+1}}\right)$$

so

$$b_1 \log \alpha_1 + \cdots + b_r\alpha_r - \log \alpha_{r+1} - b_{r+2}\log(-1) = \log\left(1 + \frac{\lambda}{\alpha_{r+1}}\right)$$

and here $|b_{r+2}| \leq |b_1| + \cdots + |b_{r+1}| + 1 \leq (r+1)\max_{i=1,\ldots,r}|b_i| \leq (r+1)B$ and we have introduced this factor since we have taken the principal branch of the logarithm. Put $\Lambda = b_1 \log \alpha_1 + \cdots + b_r \log \alpha_r - \log \alpha_{r+1} - b_{r+2}\log(-1)$. Since $\lambda \neq 0$ we see that $\Lambda \neq 0$. Put

$$A_{r+1} = \max(H(\alpha_{r+1}), e).$$

Put $K = \mathbb{Q}(\alpha^{(k)}, \alpha^{(j)}, \alpha^{(l)})$. Notice that $d = [K : \mathbb{Q}] \leq n^3$. By Thrm 2,

$$|\Lambda| > \exp(-c_6 \log A_{r+1} \log(r+1)B). \tag{6}$$

Put $\Theta_1 = \frac{\alpha^{(j)}-\alpha^{(l)}}{\alpha^{(k)}-\alpha^{(l)}}$, $\Theta_2 = \gamma^{(k)}$, $\Theta_3 = (\gamma^{(j)})^{-1}$, so that $\alpha_{r+1} = \Theta_1\Theta_2\Theta_3$. Then

$$\begin{aligned} H(\alpha_{r+1}) &= H(\Theta_1\Theta_2\Theta_3) \leq 2^d M(\Theta_1\Theta_2\Theta_3) \\ &\leq (2h\Theta_1\Theta_2\Theta_3))^d \leq (2(h(\Theta_1)(\Theta_2)(\Theta_3))^d \\ &\leq (2M(\Theta_1)M(\Theta_2)M(\Theta_3)^{-1})^d \\ &\leq (2d^{3/2}H(\Theta_1)H(\Theta_2)H(\Theta_3^{-1}))^d \qquad \text{by prop 5} \\ &\leq (c_7 m^2)^d. \end{aligned}$$

Therefore

$$\log A_{r+1} \leq c_8 \log 2|m|.$$

¿From (6)

$$\log|\Lambda| > -c_9(\log 2|m|)\log B \tag{7}$$

But

$$|\Lambda| = \left|\log\left(1 + \frac{\lambda}{\alpha_{r+1}}\right)\right| = \left|\log\left(1 + \left(\frac{\alpha^{(k)}-\alpha^{(j)}}{\alpha^{(j)}-\alpha^{(l)}}\right)\frac{\beta^{(l)}}{\beta^{(k)}}\right)\right|.$$

Notice that, by (4) and (5),

$$\left|\left(\frac{\alpha^{(k)}-\alpha^{(j)}}{\alpha^{(j)}-\alpha^{(l)}}\right)\frac{\beta^{(l)}}{\beta^{(k)}})\right| \leq c_{10}e^{-c_{11}B}.$$

Further we have

$$|\log(1+z)| \leq 2|z|$$

for $|z| < 1/2$. Thus there exists a c_{12} such that $B < c_{12}$ or $c_{10}e^{-c_{11}B} < 1/2$ and so $B < c_{12}$ or

$$|\Lambda| < 2c_{10}e^{-c_{11}B}$$

hence

$$\log|\Lambda| < \log(2_c 10) - c_{11}B.$$

Therefore $B < c_{13}$ or

$$\log|\Lambda| < -c_{14}B.$$

Thus, by (7),

$$\frac{B}{\log B} < c_{15}\log 2|m|$$

hence

$$B < c_{16}\log 2|m|\log\log(4|m|).$$

But now

$$x = \frac{\alpha^{(2)}\beta^{(1)} - \alpha^{(1)}\beta^{(2)}}{\alpha^{(2)}-\alpha^{(1)}}$$

and

$$y = \frac{\beta^{(1)}-\beta^{(2)}}{\alpha^{(2)}-\alpha^{(1)}}.$$

whence

$$\max(|x|,|y|) < c_{17}\max(|\beta^{(1)}|,|\beta^{(2)}|).$$

We have

$$\begin{aligned}|\beta^{(i)}| &= |\gamma^{(i)}(\eta_1^{(i)})^{-b_1}\cdots(\eta_r^{(i)})^{-b_r}|\\ &\leq c_{18}|m|^{1/n}e^{c_{19}B}\\ &\leq c_{18}|m|^{1/n}e^{c_{20}\log(2|m|)\log\log(4|m|)}\\ &\leq (2|m|)^{c_{21}\log\log(4|m|)}\end{aligned}$$

□

Comments:

(1) A sharper version of Theorem 2 allows us to conclude that

$$\max(|x|,|y|) < (2|m|)^C;$$

where C is effectively computable in terms of F.

(2) It is possible to extend the theorem to treat the following situation. Let $[K:\mathbb{Q}] < \infty$. For any $\theta \in K$ let

$$\overline{|\theta|} = \max_{\sigma}|\sigma(\theta)|$$

where the max is taken over all embeddings σ of K into $\mathbb{C}$ which fix $\mathbb{Q}$. Let $F \in \mathcal{O}_{\mathcal{K}}[x,y]$ be a binary form with non-zero discriminant and degree at least 3, let $\mu \in \mathcal{O}_{\mathcal{K}}$ with $\mu \neq 0$. Then there exists a positive number c which is effectively computable in terms of F and μ such that if x and y are in $\mathcal{O}_{\mathcal{K}}$ and $F(x,y) = \mu$ then $\max(\overline{|x|},\overline{|y|}) < c$. The proof is essentially the same as that of Thrm 17.

What other Diophantine equations can we treat?

Let $m \in \mathbb{Z}$ with $m \geq 2$ and $f \in \mathbb{Z}[x]$. Subject to some hypotheses we can study the equations

$$y^m = f(x). \tag{8}$$

If $m = 2$, (8) is known as a hyperelliptic equation and for $m > 2$ we have a superelliptic equation. We will proceed with the case $m \geq 3$, $f \in \mathbb{Z}[x]$, monic, with 2 simple roots.

Let $d = [K:\mathbb{Q}] < \infty$. There is a fundamental system of units $\eta_1,\dots,\eta_r$ in $\mathcal{O}_{\mathcal{K}}$ such that

$$(I) \qquad \max_{1\leq i,j\leq r}|\log|\eta_i^{(j)}|| < cR,$$

where R is the regulator and c is a positive number which is effectively computable in terms of d.

Given such a system, every unit η in $U(K)$ can be expressed as

$$(II) \qquad \eta = \eta'\eta_1^{b_1}\cdots\eta_r^{b_r} \text{ with } b_1,\dots,b_r \in \mathbb{Z},$$

and $\overline{|\eta'|} < c_2$ where c_2 depends on d and R.

Let $\alpha \in \mathcal{O}_{\mathcal{K}}$ with

$$|N_{K/\mathbb{Q}}(\alpha)| \leq M.$$

Then there exists a positive number c_3, which is effectively computable in terms of d, R and M, such that there is a unit ϵ in $\mathcal{O}_{\mathcal{K}}$ for which

$$(III) \qquad \overline{|\epsilon\alpha|} < c_3.$$

Let a be an ideal of $\mathcal{O}_{\mathcal{K}}$. There exists an ideal b in $\mathcal{O}_{\mathcal{K}}$ with

$$(IV) \qquad N(b) \leq \sqrt{|D|}$$

such that ab is a principal ideal.

Theorem 18. *Let $f \in \mathbb{Z}[x]$ be a monic polynomial with at least 2 simple roots. Let $m \in \mathbb{Z}$ with $m \geq 3$. Then there exists a positive number c, which is effectively computable in terms of f and m, such that if $x, y \in \mathbb{Z}$ for which $y^m = f(x)$ then*

$$\max(|x|, |y|) < c.$$

For the proof we need the following result.

Lemma 19. *Let $m \in \mathbb{Z}$ with $m \geq 2$, $f \in \mathbb{Z}[x]$ be a monic polynomial with at least 1 simple root α, and let K be the splitting field of f over $\mathbb{Q}$. Suppose that $x, y \in \mathbb{Z}$ for which $y^m = f(x)$. There exists a positive number c, which is effectively computable in terms of m, f and $\gamma, \phi, \delta \in \mathcal{O}_{\mathcal{K}}$ with $\gamma\phi \neq 0$, such that*

$$(x - \alpha) = \left(\frac{\gamma}{\phi}\right) \delta^m,$$

with

$$\max(\overline{|\gamma|}, \overline{|\phi|}) < c.$$

Proof:
Let $f(x) = (x - \alpha)(x - \alpha_2) \cdots (x - \alpha_n)$; here $\alpha_2, \ldots, \alpha_n$ need not be distinct, but $\alpha \neq \alpha_i$ for $i = 2, \ldots, n$. Let $c_1, c_2, \ldots$ denote positive numbers which are effectively computable in terms of f and m.

Let $\eta_1, \ldots, \eta_r$ be a fundamental system of units satisfying I. Put

$$\Delta = \prod_{i=2}^{n} [\alpha - \alpha_i],$$

so Δ is a non-zero ideal of $\mathcal{O}_{\mathcal{K}}$.

Let $x, y \in \mathbb{Z}$ for which $y^m = f(x)$. If $x = \alpha$ we may take $\gamma = \phi = 1$, $\delta = 0$ and the result holds, so we may assume that $x \neq \alpha$. We have

$$[y]^m = [x - \alpha][x - \alpha_2] \cdots [x - \alpha_n] \tag{1}$$

as an equation of ideals. Let $\wp$ be a prime ideal which divides $[x - \alpha]$. Let $l_1, \ldots, l_n$ denote the exact power of $\wp$ which divides $[x - \alpha], [x - \alpha_2], \ldots, [x - \alpha_n]$ respectively. Let $l_j = \max(l_1, \ldots, l_n)$.

First suppose that $j = 1$ so that $l_1 = l_j$. Then

$$\wp^{l_i} \mid [x - \alpha] - [x - \alpha_i] = [\alpha - \alpha_i] \text{ for } i = 2, \ldots, n.$$

Therefore

$$\wp^{l_2 + \cdots + l_n} \mid \Delta.$$

By (1) $l_1 + \cdots + l_n \equiv 0 \pmod{m}$. Put $l_1 \equiv a \pmod{m}$ with $0 \leq a \leq m - 1$. Then either $a = 0$ or $a \leq (m - 1)(l_2 + \cdots + l_n)$.

Next suppose $j \geq 2$. Then

$$\wp^{l_1} \mid [x - \alpha] - [x - \alpha_j] \text{ so } \wp^{l_1} \mid [\alpha - \alpha_j],$$

hence $\wp^{l_1} \mid \Delta$.

In both cases $l_1 \equiv a \pmod{m}$ where a is bounded from above by the power of $\wp$ which divides Δ^{m-1}. In particular, there exists ideals $a, b \in \mathcal{O}_{\mathcal{K}}$ with

$$[x - \alpha] = ab^m \tag{2}$$

where $a \mid \Delta^{m-1}$. By IV there exist ideals $a_1, b_1 \in \mathcal{O}_{\mathcal{K}}$ with aa_1 and bb_1 principal and with

$$\max(N(a_1), N(b_1)) < c_1.$$

Put $aa_1 = [\gamma_1]$ and $bb_1 = [\delta_1]$, with $\gamma_1, \delta_1 \in \mathcal{O}_{\mathcal{K}}$. Then, from (2),

$$a_1 b_1^m [x-\alpha] = aa_1(bb_1)^m$$

so

$$a_1 b_1^m [x-\alpha] = [\gamma_1][\delta_1]^m. \tag{3}$$

Note that $a_1 b_1^m$ is principal, say equal to $[\phi_1]$ with $\phi_1 \in \mathcal{O}_{\mathcal{K}}$. Observe that

$$N([\phi_1]) = N(a_1)N(b_1)^m < c_2.$$

Further, since $a \mid \Delta^{m-1}$,

$$N([\gamma_1]) = N(a)N(a_1) \leq N(\Delta)^{m-1} N(a_1) < c_3.$$

By III we can find associates (equivalent up to multiplication by a unit) γ_2 and ϕ_2 of γ_1 and ϕ_1 such that

$$\max(\overline{|\gamma_2|}, \overline{|\phi_2|}) < c_4.$$

Therefore, by (3)

$$x - \alpha = \epsilon \left(\frac{\gamma_2}{\phi_2}\right) \delta_1^m,$$

for some unit $\epsilon \in \mathcal{O}_{\mathcal{K}}$. By I and II we can find units $\epsilon_1, \epsilon_2 \in \mathcal{O}_{\mathcal{K}}$ such that

$$\epsilon = \epsilon_1 \epsilon_2^m$$

with $\overline{|e_1|} < c_5$. Therefore

$$x - \alpha = \left(\frac{\epsilon_1 \gamma_2}{\phi_2}\right) (\epsilon_2 \delta_1)^m,$$

and we put $\gamma = \epsilon\gamma_2$, $\phi = \phi_2$ and $\delta = \epsilon_2 \delta_1$. We have $\overline{|\epsilon\delta_2|} < c_6$ and $\overline{|\phi_2|} < c_4$ so the result follows. □

Proof of Theorem 18 Suppose that x and y are integers for which $y^m = f(x)$. Let $c_1, c_2, \ldots$ denote positive constants that are effectively computable in terms of of m and f. Let $f(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$ with α_1 and α_2 simple roots of f. Put $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ so K is the splitting field of f. By Lemma 19, there exist $\gamma_1, \gamma_2, \phi_1, \phi_2, \delta_1, \delta_2$ in $\mathcal{O}_{\mathcal{K}}$ with $\gamma_1\phi_1 \neq 0$ and $\gamma_2\phi_2 \neq 0$ such that

$$x - \alpha_i = \left(\frac{\gamma_i}{\phi_i}\right) \delta_i^m \text{ for } i = 1, 2 \tag{4}$$

with

$$\max(\overline{|\gamma_1|}, \overline{|\gamma_2|}, \overline{|\phi_1|}, \overline{|\phi_2|}) < c_1 \tag{5}$$

Therefore

$$(x-\alpha_1) - (x-\alpha_2) = \left(\frac{\gamma_1}{\phi_1}\right)\delta_1^m - \left(\frac{\gamma_2}{\phi_2}\right)\delta_2^m$$

Accordingly

$$\gamma_1\phi_2\delta_1^m - \gamma_2\phi_1\delta_2^m = \phi_1\phi_2(\alpha_2 - \alpha_1).$$

We view this as a Thue equation

$$g(x,y) = \mu$$

in $\mathcal{O}_{\mathcal{K}}$ with $g(x,y) = \gamma_1\phi_2 x^m - \gamma_2\phi_1 y^m$ and $\mu = \phi_1\phi_2(\alpha_2 - \alpha_1)$. Since $\gamma_1\phi_2 \neq 0$ and $\gamma_2\phi_1 \neq 0$ and $m \geq 3$ we have by a generalization of Thrm 17 to algebraic number fields that

$$\max(\overline{|\delta_1|}, \overline{|\delta_2|}) < c_2 \tag{6}$$

By (4), (5), and (6) we see that $|x| < c_4$ and so $|y| < c_5$. □

The hyperelliptic equation $y^2 = f(x)$ in integers x and y with $f \in \mathbb{Z}[x]$ a monic polynomial and at least 3 simple zeros has the property

$$\max(|x|, |y|) < c,$$

where c is effectively computable in terms of f. The proof depends upon a reduction of the problem to a Thue equation in $\mathcal{O}_{\mathcal{K}}$ for a finite extension K of $\mathbb{Q}$.

Bounds for linear forms in logarithms.

Lemma 20. *Let L and $K_1, \dots, K_n$ be integers with $0 \le K_1 < K_2 < \cdots < K_N < L$. Let $\mathcal{E}$ be a set of at least L non-zero complex numbers. There exist $a_1, \dots, a_n \in \mathcal{E}$ such that*

$$\det\left(\left(a_i^{K_j}\right)_{\substack{i=1,\dots,n\\ j=1,\dots,n}}\right) \neq 0.$$

Proof:
By induction on n. True for $n = 1$. Suppose true for $n-1$. Then there exists $a_1, \dots, a_{n-1} \in \mathcal{E}$ such that

$$A = \det\left(\left(a_i^{K_j}\right)_{\substack{i=1,\dots n-1\\ j=1,\dots,n-1}}\right) \neq 0.$$

Consider the polynomial $P(z)$ defined by

$$\begin{aligned} P(z) &= \det\begin{pmatrix} a_1^{K_1} & \cdots & a_{n-1}^{K_1} & z^{K_1} \\ a_1^{K_2} & \cdots & a_{n-1}^{K_2} & z^{K_2} \\ \vdots & \vdots & \vdots & \vdots \\ a_1^{K_n} & \cdots & a_{n-1}^{K_n} & z^{K_n} \end{pmatrix} \\ &= Az + \cdots \end{aligned}$$

Notice that $a_1, \dots, a_{n-1}$ are roots of $P(z)$ so

$$P(z) = (z - a_1) \cdots (z - a_{n-1}) Q(z)$$

for $Q(z) \in \mathbb{C}[z]$ with $\deg Q = K_n - (n-1)$. Since Card $(\mathcal{E} - \{a_1, \dots, a_{n-1}\}) \ge L - (n-1) > K_n - (n-1)$. Therefore there is an element $a_n \in \mathcal{E} \setminus \{a_1, \dots, a_{n-1}\}$ which is not a root of Q. Therefore $P(a_n) \neq 0$ and the result follows by induction. □

Proposition 21. *Let $\alpha_1, \alpha_2, \beta \in \mathbb{C}$ with $\alpha_1\alpha_2 \neq 0$. Let $K, L, R_1, R_2, S_1, S_2 \in \mathbb{Z}^+$ and $P \in \mathbb{C}[x, y]$ be a non-zero polynomial with degree at most $K-1$ in x and degree at most $L-1$ in y. Put $R = R_1 + R_2 - 1$ and $S = S_1 + S_2 - 1$. Suppose that*

$$\mathit{Card}\ \{\alpha_1^r \alpha_2^s \mid 0 \le r < R_1, 0 \le s < S_1\} > L$$

and that

$$\mathit{Card}\ \{r + s\beta \mid 0 \le r < R_2, 0 \le s < S_2\} > (K-1)L.$$

Then for some r and s with $0 \le r < R$ and $0 \le s < S$ we have

$$P(r + s\beta, \alpha_1^r \alpha_2^s) \neq 0.$$

Proof:
Note that we may assume WLOG that $P(x, 0) \neq 0$ since otherwise $P(x, y) = P_1(x, y) y^t$ for $t \in \mathbb{Z}^+$ and since $\alpha_1\alpha_2 \neq 0$ we could replace P by P_1.

Suppose that

$$P(r+s\beta, \alpha_1^r\alpha_2^s) = 0 \tag{1}$$

for $0 \le r < R$ and $0 \le s < S$. Let us write

$$P(x,y) = \sum_{i=1}^{n} Q_i(x)y^{K_i}$$

with $0 = K_1 < K_2 < \cdots < K_n < L$. Put $\mathcal{E} = \{\alpha_1^r\alpha_2^s \mid 0 \le r < R_1, 0 \le s < S_1\}$. By assumption Card $(\mathcal{E}) > L$ and so by Lemma 20 there exists a subset $\mathcal{L}$ of

$$\{(r,s) \mid 0 \le r < R_1, 0 \le s < S_1\}$$

of cardinality n such that

$$B = \det\left(\left((\alpha_1^r\alpha_2^s)^{k_i}\right)_{\substack{i=1,\dots,n\\(r,s)\in\mathcal{L}}}\right) \neq 0.$$

We now consider the polynomial $P_{r,s}(x,y)$ for each pair (r,s) in $\mathcal{L}$ given by

$$P_{r,s}(x,y) = P(x+r+s\beta, \alpha_1^r\alpha_2^s y)$$

so

$$P_{r,s}(x,y) = \sum_{i=1}^{n} Q_i(x+r+s\beta)(\alpha_1^r\alpha_2^s)^{K_i}y^{K_i} \tag{2}$$

We now define the polynomial $\Delta(x)$ by

$$\Delta(x) = \det\left(\left(Q_i(x+r+s\beta)(\alpha_1^r\alpha_2^s)^{K_i}\right)_{\substack{i=1,\dots,n\\(r,s)\in\mathcal{L}}}\right).$$

If $Q_i(x) = b_ix^{m_i} + \cdots$ with $b_i \neq 0$ for $i = 1,\dots,n$ then

$$\Delta(x) = b_1\cdots b_nBx^{m_1+\cdots+m_n} + \cdots$$

with $b_1\cdots b_nB \neq 0$. In particular,

$$\deg\Delta(x) = m_1+\cdots+m_n \le n(K-1) \le L(K-1). \tag{3}$$

We may view (2) as a system of n linear equations in the variables $Z_1,\dots,Z_n$ where $Z_i = y^{K_i}$ for $i = 1,\dots,n$. Then, by Cramer's Rule there exist polynomials $S_{r,s}(x) \in \mathbb{C}[x]$ for each $(r,s) \in \mathcal{L}$ such that

$$\Delta(x) = Z_1\Delta(x) = \sum_{(r,s)\in\mathcal{L}} P_{r,s}(x,y)S_{r,s}(x). \tag{4}$$

Note that by (1) for each $(r,s) \in \mathcal{L}$,

$$\begin{aligned} P_{r,s}(r_0+s_0\beta, \alpha_1^{r_0}\alpha_2^{s_0}) &= P(r_0+s_0\beta+r+s\beta, \alpha_1^{r+r_0}\alpha_2^{s+s_0}) \\ &= P((r_0+r)+(s_0+s)\beta, \alpha_1^{r+r_0}\alpha_2^{s+s_0}) = 0 \end{aligned}$$

for $0 \le r_0 < R_2$ and $0 \le s_0 < S_2$. Therefore by (4)

$$\Delta(r_0+s_0\beta) = 0$$

for $0 \le r_0 < R_2$ and $0 \le s_0 < S_2$. By assumption

$$\text{Card}\,\{r_0+s_0\beta \mid 0 \le r_0 < R_2, 0 \le s_0 < S_2\} > (K-1)L.$$

Therefore $\deg\Delta(x) > (K-1)L$ contradicting (3). □

In 1844, Catalan conjectured that the only two consecutive positive integers which are pure powers are 8 and 9. In particular, he conjectured that the only solution of

$$x^m - y^n = 1 \tag{5}$$

in integers x, y, m, n with $x, y, m, n > 1$ is given by $x = n = 3$, $y = m = 2$.

Recently Mihailescu proved that Catalan's conjecture is correct. In 1976, Tijdeman had reduced the problem to a finite computation.

Theorem 22. *(Tijdeman) There exists a positive effectively computable number c such that if $x, y, m, n > 1$ are integers satisfying (5) then* $\max(x, y, m, n) < c$.

Proof:
Let $c_1, c_2, \dots$ be positive effectively computable numbers. We may assume WLOG that m and n are primes p and q and we can consider the equivallent equation to (5) given by

$$x^p - y^q = \epsilon, \tag{1}$$

with $p > q$ and $x, y > 1$ and ϵ from $\{-1, 1\}$.

Initial Assumption: $\min(p, q, x, y) > c_1$.

Therefore we may suppose that p and q are odd. Note by (1) $\gcd(x, y) = 1$. Further since $p > q$, we see that $x < y$. Notice from (1) that

$$x^p = y^q + \epsilon = (y + \epsilon)(y^{q-1} - \epsilon y^{q-2} + e^2 y^{q-3} - \cdots + \epsilon^{q-1}).$$

Let $d = \gcd(y + \epsilon, y^{q-1} - \epsilon y^{q-2} + e^2 y^{q-3} - \cdots + \epsilon^{q-1})$. Then

$$y^q = (-\epsilon + (y + \epsilon))^q = (-\epsilon)^q + \binom{q}{1}(-\epsilon)^{q-1}(y + \epsilon) + \cdots + (y + \epsilon)^q$$

so

$$\frac{y^q - (-\epsilon)^q}{y + \epsilon} = \binom{q}{1}(-\epsilon)^{q-1} + \binom{q}{2}(-\epsilon)^{q-2}(y + \epsilon) + \cdots + (y + \epsilon)^{q-1}. \tag{2}$$

Thus $\frac{y^q + \epsilon}{y + \epsilon} = q + (y + e) \cdot t$ for some integer t. Therefore $d | q$ and so either $d = 1$ or $d = q$. If $d = q$ then by (2) and the fact that q is odd we see that q divides $\frac{y^q + \epsilon}{y + \epsilon}$ and $q^2 \not| \frac{y^q + \epsilon}{y + \epsilon}$. Therefore there is an integer δ with $\delta \in \{-1, 0\}$ such that, for some $s \in \mathbb{Z}^+$,

$$y + \epsilon = q^\delta s^p \tag{3}$$

Similarly

$$y^q = x^p - \epsilon = (x - \epsilon)\left(\frac{x^p - \epsilon}{x - \epsilon}\right)$$

and so there exists a $\gamma \in \{0, -1\}$ and a positive integer r such that

$$x - \epsilon = p^\gamma r^q. \tag{4}$$

Note that $r, s > 1$ and if $\gamma = -1$ then $p \mid r$ and similarly if $\delta = -1$ then $q \mid s$. Thus

$$p^\gamma r^q \geq 2^{q-1} \text{ and } q^\delta s^p \geq 2^{p-1}.$$

Now by (1),(3) and (4) we have

$$(p^\gamma r^q + \epsilon)^p - (q^\delta s^p - \epsilon)^q = \epsilon. \tag{5}$$

¿From (1), (3) and (4),

$$2^p r^{pq} \geq (r^q + 1)^p + 1 \geq x^p + 1 > y^q \geq (q^\delta s^p - 1)^q \geq \frac{s^{pq}}{(2q)^q}$$

and similarly

$$2^q s^{pq} \geq (s^p+1)^q + 1 \geq y^q + 1 > x^p \geq (p^\gamma r^q - 1)^p \geq \frac{r^{pq}}{(2p)^p}.$$

Thus, since $p > q > c_1$,

$$s^{pq} \leq 2^{p+q} q^q r^{pq} \leq 4^p q^q r^{pq}$$

hence

$$s \leq 4^{\frac{1}{q}} q^{\frac{1}{p}} r \leq 2r \tag{6}$$

Further we have

$$r^{pq} \leq 2^{p+q} p^p s^{pq} \leq (4p)^p s^{pq}$$

hence

$$r \leq (4p)^{\frac{1}{q}} s \tag{7}$$

We shall first prove that q is much smaller than p. It follows from (3) and (4) that

$$\max((x-1)^P, (y-1)^q) < x^p = y^q + \epsilon < \min((x+1)^p, (y+1)^q). \tag{8}$$

Therefore

$$p^{\gamma p} r^{qp} - q^{\delta q} s^{pq} = (x-\epsilon)^p - (y+\epsilon)^q \neq 0.$$

Plainly $p^\gamma r^q \geq 2^{q-1}$. Either $2^{q-1} > 12p^3$ or $12p^3 \geq 2^{q-1}$ in which case

$$\log 12 + 3\log p \geq (q-1)\log 2$$

hence

$$q < c_2 \log p. \tag{9}$$

Assume now that $2^{q-1} > 12p^3$ so in particular $p^\gamma r^q > 12p^3$.

Now $x - \epsilon = p^\gamma r^q$ so

$$\left|\frac{x}{p^\gamma r^q} - 1\right| = \frac{1}{p^\gamma r^q}, \tag{10}$$

$x^p - y^q = \epsilon$ so

$$\left|\frac{y^q}{x^p} - 1\right| = \frac{1}{x^p} \tag{11}$$

and $y + \epsilon = q^\delta s^p$ so

$$\left|\frac{y}{q^\delta s^p} - 1\right| = \frac{1}{q^\delta s^p} \tag{12}$$

Since $|\log(1+z)| \leq 2|z|$ for $|z| < \frac{1}{2}$ we have from (10), (11) and (12) and the fact that $p > q$ and $-1 \leq \gamma, \delta \leq 0$, that

$$\left|\log\frac{x}{p^\gamma r^q}\right| = \left|\log\left(1 + \left(\frac{x}{p^\gamma r^q} - 1\right)\right)\right| < \frac{2}{p^\gamma r^q}$$

hence

$$|p\log x - p\log(p^\gamma r^q)| < 2p^2 r^{-q} \tag{13}$$

and that

$$|p\log x - q\log y| = \left|\log\left(1 + \left(\frac{x^p}{y^q} - 1\right)\right)\right| \leq \frac{2}{x^p} \leq 2pr^{-q} \tag{14}$$

and finally that

$$|q\log y - q\log(q^\delta s^p)| \leq 2q^{1-\delta} s^{-p} \leq 2q^2 s^{-q} < 2p^2 s^{-q}$$

and so by (7)

$$|q\log y - q\log(q^{\delta}s^{p})| < 8p^{3}r^{-q}. \tag{15}$$

Therefore by (13), (14), and (15)

$$|p\log(p^{\gamma}r^{q}) - q\log(q^{\delta}s^{p})| < 12p^{3}r^{-q}$$

Put $\Lambda_1 = p\log(p^{\gamma}r^{q}) - q^{\log(q^{\delta}s^{p})}$ so

$$\Lambda_1 = p\log p^{\gamma} - q\log q^{\delta} + pq\log\left(\frac{r}{s}\right).$$

We have

$$|\Lambda_1| < 12p^{3}r^{-q}. \tag{16}$$

We may employ Thrm 2 with $\alpha_1 = p$, $\alpha_2 = q$, $\alpha_3 = \frac{r}{s}$, $d = 1$ and $n = 3$ to conclude that since $A_1 = p$, $A_2 = q < p$ and $A_3 = 2r$ (since $s \leq 2r$),

$$|\Lambda_1| > \exp(-c_3(\log p)^{3}\log r). \tag{17}$$

Comparing (16) with (17) we find that

$$r^{q} \leq 12p^{3}r^{c_3\log^{3}p} < r^{c_4\log^{3}p} \tag{18}$$

Hence

$$q < c_4\log^{3}p. \tag{19}$$

It follows from (3), (1) and (8) that

$$(p^{\gamma}r^{q}+\epsilon)^{p} - q^{q\delta}s^{qp} = x^{p} - (y+\epsilon)^{q} \neq 0 \tag{20}$$

we have by (14) and (15) that

$$|p\log x - q\log(q^{\delta}s^{p})| \leq 2x^{-p} + 2q^{2}s^{-p}.$$

Further since $x^{p} = y^{q} + \epsilon$ and $y + \epsilon = q^{\delta}s^{p}$,

$$x^{p} \geq y^{q} - 1 > 2^{\frac{q}{2}}y > 2qy > s^{p}.$$

Thus

$$|p\log(p^{\gamma}r^{q}+\epsilon) - q\log(q^{\delta}s^{p})| < 4q^{2}s^{-p}$$

so

$$\left|-q\delta\log q + p\log\left(\frac{p^{\gamma}r^{q}+\epsilon}{s^{q}}\right)\right| < 4q^{2}s^{-p}.$$

Put $\Lambda_2 = -q\delta\log q + p\log\left(\frac{p^{\gamma}r^{q}+\epsilon}{s^{q}}\right)$ and then

$$|\Lambda_2| < 4q^{s}s^{-p}. \tag{21}$$

We now apply Thrm 2 with $A_1 = q < p$, $A_2 \leq 5ps^{q}$ and $B \leq p$ to give a lower bound for $|\Lambda_2|$. Since (20), $\Lambda_2 \neq 0$ we have

$$|\Lambda_2| > \exp(-c_6\log^{2}p\log(5ps^{q})). \tag{22}$$

On comparing (21) and (22) we find that

$$s^{p} \leq 4q^{2}(5ps^{q})^{c_6\log^{2}p}$$

and since by (19), $q < c_5\log^{3}p$,

$$s^{p} \leq s^{c_7\log^{5}p},$$

hence

$$p < c_7\log^{5}p$$

and so $p < c_8$. Therefore, by (19), $q < c_9$.

We may now suppose that p and q are fixed and so by Thrm 18 x and y are bounded. The result now follows subject to our initial assumption that x, y, m and n exceed c_1.

To be Continued □

The initial assumption can be dealt with by appleaing to two results due to Shorey + Tijdeman. Let $P \geq 2$ and let S be the set of integers composed of primes $\leq P$.

Theorem 23. *Let $\tau \geq 1$ and let $P \geq 2$. The exists a positive number C with is effectively computabale in terms of τ and P.such that all solutions of the equation $ax^m - by^n = b^k$ in integers a, b, x and k from S and m, n and y with $x, y, m, n > 1$ and with* $\gcd(ax^m, k) \leq \tau$ *satisfy*

$$\max(|a|, |b|, |k|, x, y, m, n) < C.$$

Theorem 24. *Let $\tau \geq 1$ and let $m \in \mathbb{Z}$ with $m > 1$ and let $p \geq 2$. There exists a positive number c_0 which is effectively computable in terms of τ, p and m, such that all solutions of the equation*

$$ax^m - by^n = k$$

in a, b and k from S and integers x, y and n with $n, |x|, |y| > 1$, $mn > 4$ and $\gcd(ax^m, k) \leq \tau$ *satisfy*

$$\max(|a|, |b|, |k|, |x|, |y|, n) < C_0.$$

Returning to the proof of Thrm 22. note that if $x^p - y^q = \epsilon$ with x fixed then we may take P to be the greatest prime factor of x and apply Thrm 23 to conclude that $(x, y, p, q) < c_{11}$ and similarly if y is fixed. If p is fixed apply Thrm 24 and similarly if q is fixed.

One of the key tools in the proofs of Thrm 23 and Thrm 24 is a p-adic verstion of Thrm 2.

Let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers of naive heights $A_1, \ldots, A_n$ respectively. Put $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ and $d = [K : \mathbb{Q}]$. Let $\wp$ be a prime ideal of $\mathcal{O}_{\mathcal{K}}$ lying above the rational prime p. For any fractional ideal $\mathcal{A}$ of K, let $\mathrm{ord}_\wp \mathcal{A}$ denote the power of $\wp$ dividing $\mathcal{A}$.

In 1977, van der Poorten proved:

Theorem 25. *Let $\wp$ be a prime ideal of $\mathcal{O}_{\mathcal{K}}$ lying above the rational prime p. There exist effectively computable positive numbers c and c_0 such that*

$$\mathrm{ord}_\wp \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1 < (cnd)^{c_0 n} \frac{p^d}{\log p} \log A_1 \cdots \log A_n \log^2 B,$$

for all rational integers $b_1, \ldots, b_n$ of absolute value at most $B(\geq 2)$ for which $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$.

There were some mistakes in van der Poorten's argument but they were fixed by Yu in 1989.

It is possible to use Thrm 25 to treat generalizations of the Thue equation. For any $n \in \mathbb{Z}$ let $P(n)$ denote the greatest prime factor of n with the convention that $P(0) = P(1) = P(-1) = 1$. In 1977 Shorey, van der Poorten, Tijdeman and Schinzel proved the following result.

Theorem 26. *Let f be a binary form with integer coefficients, non-zero discriminant and degree at least 3. There exists a positive number C which is effectively computable in terms of f such that if x and y are coprime integers with $f(x, y) \neq 0$ then*

$$P(f(x, y)) > C \log\log z,$$

where $z = \max(|x|, |y|, 3)$.

Notice that this implies the finiteness of Thue equation $f(x,y) = m$. Let $p_1, \dots, p_s$ be primes. The equation

$$f(x,y) = p_1^{z_1} \cdots p_s^{z_s}$$

in integers x and y and non-negative integers $z_1, \dots, z_s$ is known as a Thue-Mahler equation. By Thrm 26 it has only finitely many solutions.

Let α_1, α_2 be non-zero algebraic numbers. Let $d = [\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}]$. Let $\log \alpha_1$ be a branch of the logarithm evaluated at α_1 and let $\log \alpha_2$ be a branch of the logarithm evaluated at α_2. Let b_1, b_2 be non-zero integers and put

$$\Lambda = b_1 \log \alpha_1 + b_2 \log \alpha_2.$$

Put $A_i = \max(h(\alpha_i)^d, e^{|\log \alpha_i|}, e)$ for $i = 1, 2$. We now prove a result due to Mignotte and Waldschmidt, later established by Laurent by a different method.

Theorem 27. *There exists a positive number C, which is effectively computable in terms of d, such that if $\Lambda \neq 0$ then*

$$|\Lambda| > \exp(-C \log A_1 \log A_2 \log^2 B')$$

where

$$B' = \max\left(3, \frac{|b_1|}{\log A_2} + \frac{|b_2|}{\log A_1}\right).$$

Proof:
We may assume WLOG that $b_1 > 0$, $b_2 < 0$ and that α_1 and α_2 have absolute value at least 1. Replacing b_2 by $-b_2$ we can write

$$\Lambda = b_2 \log \alpha_2 - b_1 \log \alpha_1$$

with b_2 and b_1 positive.

Let $K \geq 3$, $L \geq 2$, R_1, R_2, S_1, S_2 be positive integers. Put $N = KL$, $R = R_1 + R_2 - 1$ and $S = S_1 + S_2 - 1$. If the conditions

$$\text{Card}\,\{\alpha_1^r \alpha_2^s \mid 0 \leq r < R_1, 0 \leq s < S_1\} > L \tag{1}$$

$$\text{Card}\,\{b_2 r + b_1 s \mid 0 \leq r < R_2, 0 \leq s < S_2\} > (K-1)L \tag{2}$$

hold then the KL x RS matrix (make a choice for the indexing of the rows and columns)

$$\left(\binom{rb_2 + sb_1}{k} \alpha_1^{lr} \alpha_2^{ls}\right)$$

is of maximal rank N. For if not there exist $c_{k,l} \in \mathbb{C}$, not all zero, for $k = 0, \dots, K-1$, $l = 0, \dots, L-1$ such that the sum of $c_{k,l}$ times the k,l-th column vector is the zero vector. Equivalently the polynomial $P(x,y)$ defined by

$$P(x,y) = \sum_{\substack{0 \leq k < K \\ 0 \leq l < L}} c_{k,l} \binom{x}{k} y^l$$

of degree at most $K-1$ in x and degree at most $L-1$ in y is zero for $x = rb_2 + sb_1$ with $0 \leq r < R_2$, $0 \leq s \leq S_2$ and $y = \alpha_1^r \alpha_2^s$ with $0 \leq r < R_1$, $0 \leq s < S_2$ which contradicts Prop 21.

Suppose then that (1) and (2) hold. Then we can extract an N x N minor, from the N x RS matrix, with non-zero determinant Δ. Thus

$$\Delta = \det\left(\binom{r_jb_2 + s_jb_1}{k_i}\alpha_1^{l_ir_j}\alpha_2^{l_is_j}\right)_{\substack{1\le i\le N\\1\le j\le N}} \tag{3}$$

where $(k_1, l_1), \ldots, (k_N, l_N)$ is some ordering of the set $\{(0,0), (0,1), \ldots, (0, L-1), (1,0), (1,1), \ldots, (K-1, L-1)\}$ and $(r_1, s_1), \ldots, (r_n, s_n)$ is an ordering of the (r, s)'s chosen to make the minor of maximal rank.

Strategy: we now compare estimates for $|\Delta|$.

Notice that

$$\sum_{i=1}^{N} k_i = L\sum_{k=1}^{K} k = \frac{LK(K-1)}{2} = \frac{N(K-1)}{2}.$$

Put

$$b = ((R-1)b_2 + (S-1)b_1)\left(\prod_{k=1}^{K-1} k!\right)^{\frac{-2}{K^2-K}}.$$

Expand the determinant in (3). We see that there are $N!$ terms of the form $E\alpha_1^{E_1}\alpha_2^{E_2}$ where E is a product of binomail coefficients of the form $\binom{rb_2+sb_1}{k}$, E_1 is a sum of l_ir_j's and E_2 is a sum of l_is_j's. Plainly $0 \le E_1 \le NLR$ and $0 \le E_2 \le NLS$. Further since $\binom{x}{k} = \frac{x!}{k!(x-k)!} \le \frac{x^k}{k!}$ for $x \in \mathbb{Z}^+, 0 \le k \le x$,

$$E \le \frac{(Rb_2 + Sb_1)^{\sum_{i=1}^{n} k_i}}{\prod\limits_{i=1}^{N} k_i!}$$

hence

$$E \le E^* = \frac{(Rb_2 + Sb_1)^{N(K-1)/2}}{\prod\limits_{k=1}^{K-1}(k!)^L}.$$

Thus

$$\begin{aligned} h(\Delta) &\le N!E^*h(\alpha_1)^{NLR}h(\alpha_2)^{NLS}\\ &\le N^Nb^{KN/2}h(\alpha_1)^{NLR}h(\alpha_2)^{NLS}. \end{aligned}$$

Observe that $\Delta \in \mathbb{Q}(\alpha_1, \alpha_2)$. Then

$$\prod(\sigma(\Delta)) = \frac{u}{v} \text{ with } u, v \in \mathbb{Z}, \quad (u,v) = 1$$

where the product is take in over all embeddings σ of $\mathbb{Q}(\alpha_1, \alpha_2)$ in $\mathbb{C}$ which fix $\mathbb{Q}$. We have $|v| \le h(\Delta)$ and $0 \ne \prod_\sigma |\sigma(\Delta)| = \left|\frac{u}{v}\right| \ge \frac{1}{|v|}$.

Further $|\sigma(\Delta)| \le h(\Delta)^{\frac{1}{d'}}$ where $d' = [\mathbb{Q}(\Delta) : \mathbb{Q}]$. Therefore

$$|\Delta| \ge \frac{1}{|V|\prod_{\sigma\ne id}|\sigma(\Delta)|}$$

hence

$$\log|\Delta| > -DN\log N - \frac{DKN\log b}{2} - DLN(R\log h(\alpha_1) + S\log h(\alpha_2)). \tag{4}$$

We now remark that

$$\begin{aligned} b &\leq (Rb_2 + Sb_1)\left(\prod_{k=1}^{K-1} k^{k-K}\right)^{\frac{2}{K^2-K}} \\ &\leq (Rb_2 + Sb_1)\exp\left(\frac{2}{K(K-1)}\sum_{k=1}^{K-1}(k-K)\log k\right) \end{aligned}$$

Note that

$$\begin{aligned} \sum_{k=1}^{K-1} k\log k &\leq 2\log 2 + 3\log 3 + \int_3^{K-1} x\log x dx \\ &< 2\log 2 + 3\log 3 + \int_3^{K-1}\left(x\log x + \frac{1}{2}x\right)dx \\ &< 2\log 2 + 3\log 3 + \frac{1}{2}x^2\log x]_3^{K-1} \\ &< 2\log 2 + 3\log 3 + \frac{1}{2}(K-1)^2\log(K-1) - \frac{9}{2}\log 3 \\ &< \frac{1}{2}(K-1)^2\log(K-1) \end{aligned}$$

and

$$\sum_{k=1}^{K-1}\log k = \log(K-1)! > \log\left(\frac{K-1}{e}\right)^{K-1}.$$

Thus

$$\begin{aligned} b &\leq (Rb_2 + Sb_1)\exp\left(\frac{2}{K^2-K}\left(\frac{1}{2}(K-1)^2\log(K-1) - K(K-1)\log\left(\frac{K-1}{e}\right)\right)\right) \\ &\leq (Rb_2 + Sb_1)\exp\left(\left(1-\frac{1}{K}\right)\log(K-1) - 2\log(K-1) + 2\right) \\ &\leq (Rb_2 + Sb_1)\exp\left(-\log(K-1) - \frac{\log(K-1)}{K} + 2\right) \end{aligned}$$

so

$$b \leq \frac{Rb_2 + Sb_1}{K-1}\cdot e^2 \leq \frac{Rb_2 + Sb_1}{K}e^3. \tag{5}$$

We now introduce Λ'. We put

$$\Lambda' = \Lambda\cdot\max\left(e^{|\Lambda|\frac{LS}{b_2}}(\frac{LS}{b_2}), e^{|\Lambda|\frac{LR}{b_1}}(\frac{LR}{b_1})\right).$$

Let ρ denote a real number larger than 1.

Lemma 28. *If* $|\Lambda'| \leq \rho^{-N+1/2}$ *then*

$$\log|\Delta| \leq -N^2\log\frac{\rho}{2} + N\log N + \frac{KN\log\rho}{2} + \frac{KN\log b}{2} + \rho LRN|\log\alpha_1| + \rho LSN|\log\alpha_2|.$$

Proof:
We may suppose, WLOG, that $b_1|\log\alpha_1| \leq b_2|\log\alpha_2|$. We have

$$\Lambda = b_2\log\alpha_2 - b_1\log\alpha_1.$$

Thus

$$\log\alpha_2 = \frac{\Lambda}{b_2} + \frac{b_1}{b_2}\log\alpha_1.$$

Put $\beta = \frac{b_1}{b_2}$. Thus $\log\alpha_2 = \frac{\Lambda}{b_2} + \beta\log\alpha_1$, hence $\alpha_2 = \alpha_1^{\beta}e^{\frac{\Lambda}{b_2}}$. By the multilinearity of the determinant,

$$\Delta = \det\left(\frac{b_2^{k_i}}{k_i!}(r_j+s_j\beta)^{k_i}\alpha_1^{l_ir_j}\alpha_2^{l_is_j}\right)_{\substack{1\le i\le N\\ 1\le j\le N}}$$

Notice that

$$\alpha_1^{l_ir_j}\alpha_2^{l_is_j} = \alpha_1^{l_i(r_j+s_j\beta)}e^{\frac{\Lambda l_is_j}{b_2}}.$$

Now we put

$$e^{\frac{\Lambda l_is_j}{b_2}} = 1+\Lambda'\Theta_{i,j}$$

where $\Theta_{i,j} = \frac{e^{\frac{\Lambda l_is_j}{b_2}}-1}{\Lambda'}$. Thus

$$|\Theta_{i,j}| \le \frac{b_2(e^{l_is_j|\Lambda|/b_2}-1)}{LS|\Lambda|e^{|\Lambda|LS/b_2}}.$$

For $x\in\mathbb{R}^+$, $e^x-1\le xe^x$, hence

$$|\Theta_{i,j}| \le \frac{\frac{b_2LS|\Lambda|}{b_2}e^{LS|\Lambda|/b_2}}{LS|\Lambda|e^{|\Lambda|LS/b_2}} \le 1.$$

Therefore

$$\Delta = \det\left(\frac{b_2^{k_i}}{k_i!}(r_j+s_j\beta)^{k_i}\alpha_1^{l_ir_j+\beta s_j}(1+\Theta_{i,j}\Lambda')\right)_{\substack{1\le i\le N\\ 1\le j\le N}}.$$

Accordingly the determinant Δ can be expressed as

$$\Delta = \sum_{I\subseteq\{1,\dots,N\}}(\Lambda')^{N-|I|}\Delta_I \tag{1}$$

and

$$\Delta_I = \det\begin{pmatrix}\phi_i(z_1) & \cdots & \phi_i(z_n)\\ \Theta_{i,1}\phi_i(z_1) & \cdots & \Theta_{i,N}\phi_i(z_n)\end{pmatrix}\begin{cases}I\in I\\ I\notin I\end{cases}$$

where $\phi_i(z) = \frac{b_2^{k_i}}{k_i!}z^{k_i}\alpha_1^{l_iz}$, $z_j = r_j+s_j\beta$ for $1\le i,j\le N$. We now define for each $I\subseteq\{1,\dots N\}$ the function $\Phi_I(x)$ given by

$$\Phi_I(x) = \det\begin{pmatrix}\phi_i(xz_1) & \cdots & \phi_i(xz_n)\\ \Theta_{i,1}\phi_i(xz_1) & \cdots & \Theta_{i,N}\phi_i(xz_n)\end{pmatrix}\begin{cases}I\in I\\ I\notin I\end{cases}.$$

Notice that $\Phi_I(1) = \Delta_I$. We claim that $\Phi_I(x)$ has a zero multiplicity $\frac{\nu^2-\nu}{2}$ where $\nu = |I|$. To see this we expand each ϕ_i with $i\in I$ as a Taylor series at the origin, say

$$\phi_i(z) = \sum_{n\ge0}p_{i,n}z^n.$$

We now plug the above Taylor series into the expression for the determinant and expand to get

$$\Phi_I(x) = \sum_{n_i,i\in I}\left(\prod_{i\in I}p_{i,n_i}x_i^{n_i}\right)\det\begin{pmatrix}z_i^{n_i} & \cdots & z_n^{n_i}\\ \Theta_{i,1}\phi_i(x,z_1) & \cdots & \Theta_{i,N}\phi_i(xz_n)\end{pmatrix}\begin{cases}I\in I\\ I\notin I\end{cases}$$

where the summation indicies run independendently from 0 to ∞. Observe that if $n_\gamma = n_\psi$ for $\gamma,\psi\in I$, $\gamma\ne\psi$ then the corresponding term in the sum is zero since two rows in the associated determinant are the same. Therefore we can restrict our attention to terms in the sum for which $n_1,\dots,n_\nu$ are all distinct. Hence for which $n_1+\cdots+n_\nu\le 0+1+\cdots+\nu-1 = \frac{\nu^2-\nu}{2}$.

Thus $\Phi_i(x)$ has a zero of order $\frac{\nu^2-\nu}{2}$ at 0. By the maximum modulus principle applied to the analytic function $\Phi_I(x)/x^{\frac{\nu^2-\nu}{2}}$ we find that

$$|\Delta_I| = |\Phi_I(1)| = \left|\frac{\Phi_I(1)}{1^{\frac{\nu^2-\nu}{2}}}\right| \leq \rho^{-\frac{\nu^2-\nu}{2}} \max_{|x|=\rho} |\Phi_I(x)|.$$

We deduce from (1) that

$$|\Delta| \leq 2^N \max_{0\leq\rho\leq N} \rho^{-(N-\frac{1}{2})(N-\rho)-\frac{\rho^2-\rho}{2}} \max_I \max_{|x|=\rho} |\Phi_I(x)|.$$

Observe that

$$\min_{0\leq\nu\leq N}\left((N-\frac{1}{2})(N-\rho)+\frac{\rho^2-\rho}{2}\right) = \min_{0\leq\nu\leq N} N^2 - N\nu - \frac{1}{2}N + \frac{\nu^2}{2}$$

The min occurs at $\nu = N$ so is

$$N^2 - N^2 - \frac{N}{2} + \frac{N^2}{2} = \frac{N(N-1)}{2}.$$

Therefore

$$|\Delta| \leq 2^N \rho^{-N(N-1)/2} \max_I \max_{|x|=\rho} |\Phi_I(x)| \tag{2}$$

Futher

$$|\Phi_I(x)| \leq N!\left(\prod_{i=1}^N \frac{(b_2x(R+S\beta))^{k_i}}{k_i!}\right)\exp\left(\sum_{n=1}^N l_i(R+S\beta)|x||\log\alpha_i|\right) \tag{3}$$

and since $\beta|\log\alpha_i| \leq |\log\alpha_2|$,

$$|\Phi_I(x)| \leq N!(|x|b)^{(K-1)N/2}\exp(|x|LRN\log|\alpha_1| + |x|LSN|\log\alpha_2|).$$

For $N \geq 6$, $2^N N! \leq N^N$, so from 2) and (3),

$$\begin{aligned}\log|\Delta| \leq &-\frac{N^2\log\rho}{2} + \frac{N\log\rho}{2} + N\log N + \frac{(K-1)N\log\rho}{2} + \frac{(K-1)N\log b}{2}\\ &+\rho LRN|\log\alpha_i| + \rho LSN|\log\alpha_2|,\end{aligned}$$

as required. Lemma 28 Follows. □

We now compare our upper bound for $\log|\Delta|$, obtained under the assumption that $|\Delta'| \leq \rho^{-N+\frac{1}{2}}$, with the lower bound (4). We find that

$$-2d\log N - 2dK\log b - 2dLR\log h(\alpha_1) - 2dLS\log h(\alpha_2) \leq$$

$$-N\log\rho + \log\rho + 2\log N + (K-1)\log\rho + (K-1)\log b + 2\rho LR|\log\alpha_1| + 2\rho LS|\log\alpha_2|.$$

So

$$\begin{aligned}N\log\rho \leq &(2d+1)\log N + 3dK\log b + K\log\rho + 2LR(\rho|\log\alpha_1| + d\log h(\alpha_1))\\ &+ 2LS(\rho|\log\alpha_2| + d\log h(\alpha_2)).\end{aligned}$$

Therefore, from the definition of $\log A_1$ and $\log A_2$,

$$N\log\rho \leq (2d+1)\log N + 3dK\log b + K\log\rho + 2L(\rho+1)(R\log A_1 + S\log A_2). \tag{4}$$

Put $B = \max(\log\rho, d(7+\log B'))$ and take

$$K = [c^2B\log A_1\log A_2], \quad L = [B], \quad R_1 = [B\log A_2]$$

$$S_1 = \lceil B\log A_1\rceil, \quad R_2 = \lceil cB\log A_2\rceil, \quad S_2 = \lceil cB\log A_1\rceil$$

where c is a real number with $c > 1$ to be choosen later.

Notice that $R_1 \geq L$ and $S_1 \geq L$ hence if α_1 and α_2 are not both roots of unity then condition (1) holds. The result follows easily if $alpha_1$ and α_2 are both roots of unity so we may suppose (1) holds. Our argument now splits into two cases.

Case 1. The set $\{rb_2 + sb_1 \mid 0 \leq r < R_2, 0 \leq s < S_2\}$ has R_2S_2 elements. Then

$$R_2S_2 \geq c^2B^2 \log A_1 \log A_2 > (K-1)L,$$

and so condition (1) and (2) apply. Now by (5), $d \log b \leq B$ and so the right hand side of (4) is at most

$$\begin{aligned} &2(d+1)\log(c^2B^2\log A_1 \log A_2) + 3c^2B^2 \log A_1 \log A_2 + c^2 \log \rho B \log A_1 \log A_2 \\ &+2(\rho + B(2(1+2c)B \log A_1 \log A_2) \leq 4c^2 12c(\rho+1) + 2c^2)B^2 \log A_1 \log A_2 \\ &\qquad < 6c(c+4\rho)B^2 \log A_1 \log A_2 \qquad (5)\end{aligned}$$

On the other hand

$$N \log \rho = KL \log \rho \geq \frac{1}{2}c^2 \log \rho B^2 \log A_1 \log A_2.$$

Take $\rho = c$. Then comparing the lower bound with (5), using (4), we find that $\frac{1}{2}c^2 \log c < 30c^2$ which is false for $c > e^{60}$. Thus our assumption that $|\Lambda'| \leq \rho^{-N+\frac{1}{2}}$ is false hence $|\Lambda'| > \exp((-N+\frac{1}{2})\log c)$, and so

$$|\Lambda| > \exp((-N+\frac{1}{2})\log c) \max\left(e^{|\Lambda|LS/b_2}(LS/b_2), e^{|\Lambda|LR/b_1}(LR/b_1)\right).$$

We may assume that $|\Lambda| < (LS)^{-1}$ and $|\Lambda| < (LR)^{-1}$ since otherwise the results holds. Then $1 + \log L + \vee R + \log S < B^2 \log A_1 \log A_2$ hence

$$|\Lambda| > \exp(-N2\log c).$$

Since $N = KL < 2c^2B \log A_1 \log A_2$ and our results follows.

Case 2. The set $\{rb_2 + sb_1 \mid 0 \leq r < R_2, 0 \leq s < S_2\}$ has fewer than R_2S_2 elements. Then there exist integers r and s with $|r| \leq R-1$ and $|s| \leq S-1$ for which $rb_2 + sb_1 = 0$. Accordingly,

$$\begin{aligned}|\Lambda| = |b_1 \log \alpha_1 + b_2 \log \alpha_2| &= \frac{b_1}{r}|r \log \alpha_1 + \frac{b_2}{b_1} r \log \alpha_2| \\ &= \frac{b_1}{r}|r \log \alpha_1 - s \log \alpha_2|.\end{aligned}$$

Now by Prop 4 and the fact that for α algebraic of degree d,

$$H(\alpha) \geq 2^d h(\alpha)^d,$$

our result follows. □

Lemma 29. *Let $N, Q, b_1, \dots, b_n$ be integers with $N \geq Q > 0$. There exists a $r \in \mathbb{Z}^+$ with $\left[\frac{N}{Q}\right] \leq r \leq N$ and integers $p_1, \dots, p_n$ such that*

$$|b_i - rp_i| \leq rQ^{-1/n} + \frac{|b_i|}{(2r-1)} \textit{ for } i = 1, \dots, n.$$

Proof:

Consider the system of linear inequalities

$$|x_0| \leq Q$$

and

$$|x_0 \frac{b_i}{N} - x_i| < Q^{-1/n} \text{ for } i = 1, \dots, n.$$

By Minkowski's Linear Forms Theorem, Thrm 7', there exists a non-zero vector $(p_0, p_1, \ldots, p_n) \in \mathbb{Z}^{n+1}$ such that $|P_0| \leq Q$ and

$$\left|p_0 \frac{b_i}{N} - p_i\right| < Q^{-1/n} \text{ for } i = 1, \ldots, n.$$

Multiplying $(p_0, \ldots, p_n)$ by -1 if necessary, we may suppose that $0 \leq p_0 \leq Q$. Note that $p_0 \neq 0$ hence $0 < p_0 \leq Q$.

Let r be the nearest integer to N/p_0. Since $N \geq Q \geq p_0 > 0$ we have $N \geq r \geq [N/Q]$. Further for $1 \leq i \leq n$,

$$\begin{aligned} |b_i - rp_i| &= \left| r\left(\frac{p_0 b_i}{N} - p_i\right) + \left(\frac{N}{p_0} - r\right)\frac{b_i}{N/P_0}\right| \\ &\leq rQ^{-1/n} + \frac{1}{2}\frac{|b_i|}{(r - \frac{1}{2})}. \end{aligned}$$

□

Estimate for linear forms in n logarithms from Thrm 27 and Lemma 29 :

Let $\alpha_1, \ldots, \alpha_n, \alpha_{n+1}$ be non-zero algebraic numbers. We put $K = Q(\alpha_1, \ldots, \alpha_{n+1})$ and $d = [K : \mathbb{Q}]$. We let $b_1, \ldots, b_n$ be non-zero integers. Let $\log \alpha_1, \ldots, \log \alpha_{n+1}$ be determinations of the logarithm of $\alpha_1, \ldots, \alpha_{n+1}$ respectively. Let $A_1, \ldots, A_{n+1}$ be numbers for which

$$\log A_i \geq \max(d \log h(\alpha_i) |\log \alpha_i|, 1) \text{ for } i = 1, \ldots, n+1$$

and put

$$B = \max(|b_1|, \ldots, |b_n|).$$

Put

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n + \log \alpha_{n+1}.$$

Using an idea of Bombieri, two mathematicians Bilu and Bugeaud, proved from Lemma 29 and Thrm 27.

Theorem 30. *Let $\epsilon > 0$. Suppose that $0 < |\Lambda| < e^{-\epsilon B}$. There exists a positive number C, which is effectively computable in terms of ϵ and $A_1, \ldots, A_n$, such that $B < C \log A_{n+1}$.*

Note: This gives enough to prove Thrm 17 on Thue equations.

Idea of proof.

Replace $|\Lambda|$ by $r \log \alpha + \log \gamma$ where

$$\log \alpha = p_1 \log \alpha_1 + \cdots + p_n \log \alpha_n$$

and

$$\log \gamma = (b_1 - rp_1) \log \alpha_1 + \cdots + (b_n - rp_n) \log \alpha_n + \log \alpha_{n+1}.$$

We apply Thrm 27 and choose p_i as in Lemma 29. In particular take

$$N = Q[\max(B^{1/2}Q^{1/2}, \log A_{n+1} Q^{1+\frac{1}{n}}]$$

and then choose Q so that $B \geq N \geq Q \geq 1$ appropriately to get a contradiction.