Diophantine Analysis

Daniel Naylor

December 4, 2024

Contents

Lecture 1

Fact: If $n \neq m \in \mathbb{Z}$, then $|n - m| \geq 1$.

Although this fact sounds very obvious, in this course it will be one of our most used tools.

1 Diophantine Approximation

Theorem 1.1 (Dirichlet)**.** Assuming that:

- α is an irrational real number

Then there exist infinitely many $\frac{p}{q} \in \mathbb{Q}$ such that

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

Proof. Consider the numbers $0, \alpha, 2\alpha, \ldots, N\alpha$ for some fixed $N \in \mathbb{Z}_{>0}$. Consider them in $\mathbb{R}/\mathbb{Z} \equiv [0,1]$. Note

$$\left[0, \frac{1}{N}\right) \sqcup \left[\frac{1}{N}, \frac{2}{N}\right) \sqcup \cdots \sqcup \left[\frac{N-1}{N}, 1\right).$$

By the box principle (pigeonhole principle), there exists $N \geq n_2 > n_1 \geq 0$ such that $n_2\alpha$ and $n_1\alpha$ belong to the same interval. Then:

$$|n_2\alpha - n_1\alpha - p| \leq \frac{1}{N}$$

for some $p \in \mathbb{Z}$. Take $q = n_2 - n_1$. Then

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{Nq} \leq \frac{1}{q^2}.$$

Take $N \to \infty$, then you get an infinite sequence of rationals. If α is not raitonal, then this sequence cannot stabilise, so we get infinitely many $\frac{p}{q}$ as desired. □

Can we do better?

In particular for $\alpha \in \overline{\mathbb{Q}}$.

Theorem (Liouville)**.** Assuming:

- α is algebraic of degree d

Then there exists $c > 0$ such that for all $\frac{p}{q} \in \mathbb{Q}$ with $\alpha \neq \frac{p}{q}$, we have

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^d}.$$

Proof. Let $P \in \mathbb{Z}[x]$ be the minimal polynomial of α, so $P(\alpha) = 0$. Now note that $P\left(\frac{p}{q}\right) \neq 0$ (by

irreducibility when $d \geq 2$, and for $d = 1$ using the hypothesis that $\alpha \neq \frac{p}{q}$). Then

$$\left| P\left(\frac{p}{q}\right) \right| \geq \frac{1}{q^d}.$$

Note that $P\left(\frac{p}{q}\right)$ is rational with denominator q^d. On the other hand,

$$\left| P\left(\frac{p}{q}\right) \right| \leq \left(\max_{x \in [\alpha-1, \alpha+1]} |P'(x)| \right) \cdot \left| \alpha - \frac{p}{q} \right|.$$

provided $\left|\alpha - \frac{p}{q}\right| \leq 1$, which we may assume. Hence

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^d}.$$

□

Improvements of the exponent d in Liouville:

- Thue: $\frac{d}{2} + 1 + \varepsilon$
- Siegel: little better than $2\sqrt{d} + \varepsilon$
- Dyson: $\sqrt{2d} + \varepsilon$

Theorem 1.2 (Roth)**.** Assuming that:

- α is an irrational real algebraic number

Then there exists $c = c(\alpha, \varepsilon) > 0$ such that for all $\frac{p}{q} \in \mathbb{Q}$ we have

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{c}{q^{2+\varepsilon}}.$$

Theorem 1.3 (Thue)**.** Assuming that:

- $P(X, Y) \in \mathbb{Z}[X, Y]$ homogeneous of degree $d \geq 3$
- without repeated factors
- $m \in \mathbb{Z}$

Then the equation

$$P(X, Y) = m$$

has only finitely many solutions in $\mathbb{Z}^2$ with $\gcd(X, Y) = 1$.

Liouville's theorem $\leftrightarrow |P(p, q)| \geq 1$.

Lemma 1.4. Assuming that:

- $P \in \mathbb{R}[X, Y]$ be homogeneous of degree d
- without repeated factors

Then for all $p, q \in \mathbb{Z}$, there exists α root of $P(X, 1)$ such that

$$cq^{-d}P(p, q) \leq \left|\alpha - \frac{p}{q}\right| \leq Cq^{-d}P(p, q).$$

Here c, C depend on P, and a fixed compact set that contains $\frac{p}{q}$.

Proof. Let

$$P(X, 1) = a(X - \alpha_1) \cdots (X - \alpha_d),$$

with $\alpha_1, \ldots, \alpha_d$ distinct (since we assumed no repeated factors, and characteristic 0 fields are always separable). Without loss of generality assume that α_1 is the closest to $\frac{p}{q}$.

Then $c_0 < \left|\frac{p}{q} - \alpha_j\right| < C_0$ for some constants depending on P and the compact set for $j \neq 1$. So we get lower and upper bounds on $P\left(\frac{p}{q}, 1\right) = P(p, q) \cdot \frac{1}{q^d}$. □

Proof of Thue. Suppose $P(p, q) = m$. The lemma tells us that there exists α a root of P such that

$$\left|\frac{p}{q} - \alpha\right| < C \cdot q^{-d} |\underbrace{P(p, q)}_{m}| = C \cdot m \cdot q^{-d}.$$

If the degree of $\alpha \geq 2$, then Roth or already Thue implies that q must be bounded, hence only finitely many solutions.

For $\alpha \in \mathbb{Q}$, we use Liouville. □

Lecture 2

Let $(x_1, \ldots, x_n) \in \mathbb{Z}^n$. The height of it is

$$H(x_1, \ldots, x_n) = \max(|x_1|, \ldots, |x_n|).$$

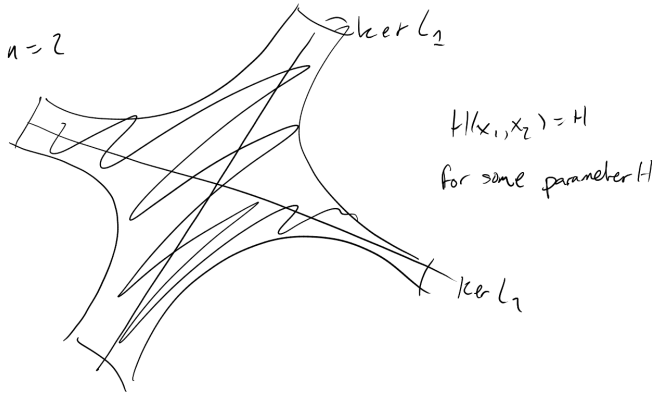
Theorem 1.5 (Subspace theorem, Archimedean version, Schmidt)**.** Assuming that:

- $n \in \mathbb{Z}_{\geq 2}$
- $L_1, \ldots, L_n$ linearly independent linear forms with algebraic coefficients in n-variables

Then for all $\varepsilon > 0$ the solutions of

$$\prod_{j=1}^{n} |L_J(x_1, \ldots, x_n)| < H(x_1, \ldots, x_n)^{-\varepsilon}, \qquad (*)$$

for $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ are contained in a finite collection of proper linear subspaces of $\mathbb{Q}^n$, which depend only on $L_1, \ldots, L_n, \varepsilon$.



The volume of the region is

$$H(x_1, \ldots, x_n) \leq H \qquad \text{and} \qquad \prod_{j=1}^{n} |L_j(x_1, \ldots, x_n)| < H^{-\varepsilon}$$

is $\sim (\log H)^{n-1} H^{-\varepsilon}$. Consider the paralellepipeds:

$$|L_j(x_1, \ldots, x_n)| < H^{\kappa_k}$$

for some $\kappa_j \in \mathbb{R}$ with $\sum \kappa_j = -\varepsilon$.

This implies Roth's theorem:

Let $\alpha \in \overline{\mathbb{Q}} \cap \mathbb{R}$ irrational. Consider the linear forms

$$L_1(X_1, X_2) = X_1 - \alpha X_2$$
$$L_2(X_1, X_2) = X_2$$

Let $p, q \in \mathbb{Z}$. Then $(*)$ is equivalent to $|p - \alpha q||q| < \max(p, q)^{-\varepsilon}$. If $\left|\frac{p}{q} - \alpha\right| < \frac{|\alpha|}{2}$, then this is equivalent to $\left|\frac{p}{q} - \alpha\right| < Cq^{-2-\varepsilon}$. Roth's theorem is true apart from p, q contained in a finite collection of subspaces. A subspace is of the form $p + \beta q = 0$ for some $\beta \in \mathbb{Q}$ or maybe $q = 0$.

Obvious subspaces:

- $\ker(L_j)$
- Example $n = 3$: $L_1 = X_1 - \sqrt{2}X_2$, $L_2 = X_1 - \sqrt{2}X_2 + X_3$, $L_2 = X_2$. Consider the subspace $V = \{(p, q, 0) : p, q \in \mathbb{Q}\}$. Now $(*)$ becomes:

  $$|p - \sqrt{2}q|^2|q| < \max(p, q)^{-\varepsilon},$$

  or alternatively

  $$\left|\frac{p}{q} - \sqrt{2}\right|^2 < q^{-3} \max(p, q)^{-\varepsilon/2}.$$

  This has plenty of solutions by Dirichlet if $\varepsilon < 1$.
- A line, that is a 1-dimensional subspace may contain only finitely many solution.

The places of $\mathbb{Q}$ is $M_{\mathbb{Q}}$ and it consists of all prime numbers and ∞. For each $v \in M_{\mathbb{Q}}$, we define an absolute value on $\mathbb{Q}$. $|\bullet|_\infty$ is the ordinary absolute value. If $v \in M_{\mathbb{Q}}$ is a prime number, this is the v-adic absolute value, that is, for $a \in \mathbb{Z}$, $|a|_v = v^{-b}$ where $b \in \mathbb{Z}$ is maximal such that $v^b \mid a$. For $\frac{a}{b} \in \mathbb{Q}$, we define $\left|\frac{a}{b}\right|_v = \frac{|a|_v}{|b|_v}$. If $x, y \in \mathbb{Q}$, then:

- $|x|_v|y|_v = |xy|_v$
- $|x + y|_v \leq |x|_v + |y|_v$

When $v \neq \infty$,

$$|x + y| \leq \max(|x|_v, |y|_v).$$

This is called the ultrametric inequality.

Theorem 1.6 (Subspace theorem, p-adic version with Q coeffs)**.** Assuming that:

- $n \in \mathbb{Z}_{\geq 2}$
- $S \subset M_{\mathbb{Q}}$ with $\infty \in S$
- for each $v \in S$, let $L_1^{(v)}, \ldots, L_n^{(v)}$ be linearly independent forms with rational coefficients in n variables

Then the solutions of

$$\prod_{v \in S} \prod_{j=1}^{n} |L_j^{(v)}(x_1, \ldots, x_n)|_v < H(x_1 7 \ldots, x_n)^{-\varepsilon},$$

with $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ are contained in a finite collection of proper subspaces of $\mathbb{Q}^n$.

$n = 2$, $S = \{2, 3, \infty\}$, $L_j^{(v)} = X_j$, $v \in S$, $j = 1, 2$. Consider $a \in \mathbb{Z}$. Let $a = 2^k 3^l b$ with b not divisible by 2 or 3.

$$|a|_2 |a|_3 |a|_\infty = 2^{-k} 3^{-l} |a| = |b|.$$

Consider $X_1 = 2^k$, $X_2 = 3^l$, then

$$\prod_{v \in S} \prod_{j=1}^{2} |L_j^{(v)}(2^k, 3^l)|_v = 1.$$

Lecture 3 What happens if you replace $L_2^{(\infty)}$ with $X_1 - X_2$?

Proposition 1.7. Assuming that:

- $\varepsilon > 0$

Then there exists $c = c(\varepsilon) > 0$ such that for $p, q, k, m \in \mathbb{Z}_{>0}$, we have

$$|p2^k - q3^m| > c \frac{\max(2^k, 3^m)^{1-\varepsilon}}{\max(p, q)}$$

or $p2^k = q3^m$.

Proof. Take $n = 2$, $S = \{2, 3, \infty\}$. Let $L_j^{(v)} = X_j$ for all j, v, except: $L_2^{(\infty)} = X_2 - X_1$. Then the solutions of

$$\prod_{v \in S} \prod_{j=1}^{2} |L_j^{(v)}(x_1, x_2)|_v < H(x_1, x_2)^{-\varepsilon/2}$$

with $x_1, x_2 \in \mathbb{Z}$ are contained in the lines: $X_1 = 0$, $X_2 = 0$, $X_1 = X_2$ plus finitely many points.

Plug in $X_1 = p2^k$, $X_2 = q3^m$. Then

$$|L_1^{(\infty)}(x_1, x_2)|_\infty = p2^k \qquad |L_2^{(\infty)}(x_1, x_2)|_\infty \leq \frac{\max(2^k, 3^m)^{1-\varepsilon}}{\max(p, q)}$$

provided p, k, q, m does not satisfy the claim with $c = 1$. Also,

$$|L_1^{(2)}(x_1, x_2)|_2 \leq 2^{-k} \qquad |L_2^{(2)}(x_1, x_2)|_2 \leq 1$$

$$|L_1^{(3)}(x_1, x_2)|_3 \leq 1 \qquad |L_2^{(3)}(x_1, x_2)|_3 \leq 3^{-m}$$

so

$$(*) \leq \frac{p}{3^m} \cdot \frac{\max(2^k, 3^m)^{1-\varepsilon}}{\max(p, q)} \leq \frac{\max(2^k, 3^m)^{1-\varepsilon}}{3^m}.$$

Assume $3^m \geq 2^k$ by symmetry. Then

$$(*) \leq \max(2^k, 3^m)^{-/eps}.$$

We can assumme that $p, q \leq 3^m$, for otherwise the claim is trivial. Then $H(p2^k, q3^m) \leq 3^{2m}$. Then

$$(*) < H(p2^k, q3^m)^{-\varepsilon/2}.$$

Then either $p2^k = q3^m$ or p, q, k, m is one of finitely many exceptions.

Make c small enough to rule out the exceptions. □

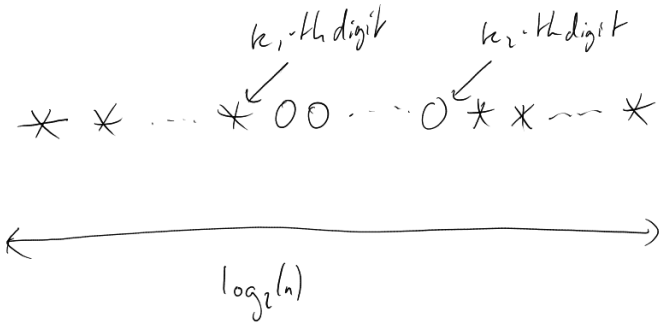
For $a, b \in \mathbb{Z}_{>0}$, let $N(a, b)$ denote the number of non-zero digits in the base b expansion of a.

Theorem 1.8 (Senge, Strauss)**.** We have $N(a, 2) + N(a, 3) \to \infty$ as $a \to \infty$.

Despite the fact that this statement looks quite modest, the proof is not so simple.

Proof. Take $a \in \mathbb{Z}$: we assume that $N(a, 2) + N(a, 3) < N$ for some fixed N. Consider its base 2 expansion.

First we will explore the consequences of having a large string of 0s in the base 2 expansion.



Then $a = p \cdot 2^{k_1} + e_1$. We know:

$$|p| < 2^{\log_2(a) - k_1 + 1}, \qquad |e_1| < 2^{k_2}.$$

Similarly: $a = q \cdot 3^{m_1} + e_2$ with $|q| < 3^{\log_3(a) - m_1 + 1}$ and $|e_2| < 3^{m_2}$.

We will make sure that $\frac{2^{k_1}}{3^{m_1}}, \frac{2^{k_2}}{k^{m_2}} \in \left[\frac{1}{3}, 3\right]$.

$$|p2^{k_1} - q3^{m_1}| = |e_1 - e_2| < 3 \cdot 2^{k_2}.$$

Want to use the proposition. So we need:

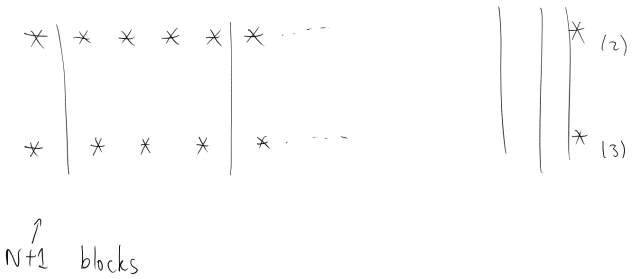
$$|p2^k - q3^m| \cdot \max(p, q) < c \max(2^k, 3^m)^{1-\varepsilon}.$$

So we want

$$C \cdot 2^{k_2} \cdot 2^{\log_2(a) - k_1} < c \cdot 2^{k_1(1-\varepsilon)}.$$

We want

$$\log_2(a) - k_1 < k_1 - k_j - \varepsilon \log_2(a).$$



Since at most N blocks have a non-zero number, one of the blocks only has zeroes, which can be used with the above to show that a cannot be too large. □

The constants in all results so far (except Liouville) are *ineffective*!

Are there any improvements of

$$\left|2^{1/3} - \frac{p}{q}\right| < \frac{100}{q^3}$$

(suppose 100 is the best you can get with Liouville) for $q < 10^{10^{10^{10^{10^{10^{10}}}}}}$. No!

To demonstrate what it means that the above results are ineffective:

Suppose that we want to find all the solutions of $x^3 - 2y^3 = 11$. Thue says that we have finitely many. But because it is ineffective, we have no idea how to bound the largest of these is, so would struggle
Lecture 4 to find all solutions, even with an arbitrarily powerful computer (or an army of postdocs).

1.1 Transcendence

Liouville proved $\alpha = \sum_{n=0}^{\infty} \frac{1}{10^{n!}}$ is transcendental.

What about e, π, $2^{\sqrt{2}}$?

Hermite: e is transcendental.

Lindemann: If $\alpha \neq 0$, then at least one of α or e^{α} is transcendental.

Theorem 1.9 (Lindemann-Weierstrass)**.** Assuming that:

- $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ distinct

Then $e^{\alpha_1}, \ldots, e^{\alpha_n}$ are linearly independent over $\overline{\mathbb{Q}}$ (algebraic closure of $\mathbb{Q}$).

Hilbert's 7th problem: Let $\alpha \neq 0, 1$, algebraic, β irrational algebraic. Then α^β is transcendental.

Note (for this problem): $\alpha^\beta = \exp(\beta \cdot \log \alpha)$ where $\log \alpha$ is any complex number with $e^{\log \alpha} = \alpha$. So in the above problem we can think of "α^β is transcendental" as meaning "any choice for α^β is transcendental".

Convention: If $\alpha \in \mathbb{R}_{>0}$, then $\log \alpha \in \mathbb{R}$.

Theorem. Let α_1, α_2 be non-zero algebraic numbers. Then $\log \alpha_1, \log \alpha_2$ are linearly independent over $\overline{\mathbb{Q}}$ if and only if they are linearly independent over $\mathbb{Q}$.

Proof of Hilbert's 7th $\iff$ above Theorem is true.

$\Rightarrow$ Suppose $\log \alpha_1, \log \alpha_2$ are dependent over $\mathbb{Q}$. Then $\exists \beta \in \overline{\mathbb{Q}}$ such that $\beta \log \alpha_1 = \log \alpha_2$. Then $\alpha_1^\beta = \alpha_2$ either $\beta \in \mathbb{Q}$ or $\alpha_1 = 1$.

$\Leftarrow$ Suppose there exists α_1, α_2 non-zero algebraic such that $\alpha_1^\beta = \alpha_2$ for some $\beta \in \overline{\mathbb{Q}}$. Then $\beta \log \alpha_1 = \log \alpha_2$ for some choice of the logarithms. If the logarithms are 0, then we deduce $\alpha_1 = 1$, a contradiction. Otherwise, we deduce that $\beta \in \mathbb{Q}$ (by the above theorem), which is also a contradiction. □

Theorem (Baber)**.** Let $\log \alpha_1, \ldots, \log \alpha_n$ be $\mathbb{Q}$-linearly independent logarithms of algebraic numbers.
Then $1, \log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over $\overline{\mathbb{Q}}$.

Conjecture 1.10 (Schanuel)**.** Let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be linearly independent over $\mathbb{Q}$. Then the transcendence degree of $\mathbb{Q}(\alpha_1, \ldots, \alpha_n, e^{\alpha_1}, \ldots, e^{\alpha_n})$ is at least n.

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}_{>0}$, and $b_1, \ldots, b_n \in \mathbb{Z}$. Let A_j be the max of the numerator and the denominator of a_j.

Let $B = \max(|b_1|, \ldots, |b_n|)$. Then

$$b_1 \log a_1 + \cdots b_n \log a_n \text{ close to } 0 \iff a_1^{b_1} \cdots a_n^{b_n} \text{ close to } 1.$$

$$|a_1^{b_1} \cdots a_n^{b_n} - 1| \geq A_1^{-b} \cdots A_n^{-B} = \exp(-(\log A_1 + \log A_n)B).$$

$$|b_1 \log a_1 + \cdots + b_n \log a_n| \geq \frac{1}{2} \exp(-(\log A_1 + \cdots + \log A_n)B).$$

Notation. Let $\alpha \in \overline{\mathbb{Q}}$, denote its minimal polynomial in $\mathbb{Z}[X]$ by f_α.
If $f \in \mathbb{C}[X]$, then $H(f)$ (the height of f) is the maximal absolute value of its coefficients.

Theorem. Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}_{\neq 0}$, $\beta_0, \dots, \beta_n \in \overline{\mathbb{Q}}$. Fix some choices of $\log \alpha_j$. Let $A_j = \max(H(f_{\alpha_j}) \exp(|\log \alpha_j|), 10)$.
Let $\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n$. Then there exists an *effective* constant C depending on n and the degree of $\mathbb{Q}(\alpha_1, \dots, \alpha_n, \beta_0, \beta_1, \dots, \beta_n)$ such that either $\Lambda = 0$ or

$$|\Lambda| > \exp(-C(\log A_1) \cdots (\log A_n)(\log B)).$$

Conjecturially: this should be

$$|\Lambda| > \exp(-C \max(\log A_1, \dots, \log A_n, \log B)).$$

Lecture 5

Theorem. Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}_{\neq 0}$. Let $\log \alpha_j$ be a choice of their logarithms. Let $b_1, \dots, b_n \in \mathbb{Z}$. Let

$$\begin{aligned} A_j &= \max(H(f(\alpha_1)), \dots, H(f(\alpha_n)), \exp(|log\alpha_1|), \dots, \exp(|\log \alpha_n|), 10) \\ B^* &= \max\left(\frac{|b_1|}{\log A_n}, \dots, \frac{|b_{n-1}|}{\log A_n}, |b_n|, 10\right) \\ \Lambda &= b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n \qquad \leftarrow \text{homogeneous} \end{aligned}$$

Then there is an effective constant C that depends only on n and the degree of $\mathbb{Q}(\alpha_1, \dots, \alpha_n)$ such that

$$|\Lambda| > \exp(-C(\log A_1), \dots, (\log A_n)(\log B^*)) \qquad \text{or} \qquad \Lambda = 0.$$

Observe

$$\exp(\operatorname{Re} \log \alpha_j) = |\alpha_j| \leq H(f(\alpha_j)).$$

Recall

$$B = \max(|b_1|, \dots, |b_n|, \log A_1, \dots, \log A_n, 10).$$

Typical scenario: $\alpha_1, \dots, \alpha_{n-1}$ fixed numbers, $b_n = 1$, $b_j \sim \log A_n$.

In the setting of Diophantine approximations, it is possible to show

$$\left|\alpha - \frac{p}{q}\right| > c(\alpha) \cdot \frac{1}{q^{d-\varepsilon(\alpha)}},$$

with $c(\alpha)$ and $\varepsilon(\alpha)$ being effective constants.

Proposition. There is an effective absolute constant C such that for all p, q, k, m:

$$|p2^k - q3^m| > \frac{\max(2^k, 3^m)}{\max(p, q, 10)^{-C \log(\max(k,m)/\log \max(p,q,10)+10)}},$$

or $p2^k = q3^m$.

Proof. Suppose $3^m > 2^k$.

$$\Lambda = k \log 2 - m \log 3 + 1 \cdot \log(p/q).$$

$A_2 = A_1 = 10$, $A_3 = \max(p, q, 10)$

$$B^* = \frac{\max(k, m)}{\log A_3} + 1.$$

Then:

$$|\Lambda| > \exp(-C \log A_3 \log B^*) = A_3^{-C \log B^*}.$$

$|\exp(\Lambda) - 1| > \frac{1}{10}|\Lambda|$

$$\exp(\Lambda) - 1 = \left| 2^k \cdot^{-m} \cdot \frac{p}{q} - 1 \right| \geq A_3^{-\tilde{C} \log B^*}$$

Multiply by $q \cdot 3^m$. □

Before:

$$|p2^k - q3^m| > C \frac{\max(2^k, 3^m)^{1-\varepsilon}}{\max(p, q)}.$$

The new bound wins when $\max(p, q) < \max(2^k, 3^m)^{o(1)}$.

In particular, when $p = q = 1$:

$$|2^k - 3^m| > \frac{\max(2^k, 3^m)}{\max(k, m)^C} \qquad \text{vs} \qquad |2^k - 3^m| > C 2^{1-\varepsilon k}.$$

$p_1 2^{k_1} + p_2 3^{k_2} + p_3 5^{k_3}$ for $k_1, k_2, k_3 \in \mathbb{Z}_{>0}$, $p_1, p_2, p_3 \in \mathbb{Z}$.

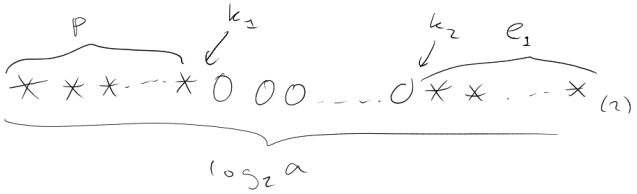
Recall: $N(a, b)$ is the number of non-zero digits in the base b expansion of a.

Theorem (Stewart)**.** There is an effective absolute constant C such that

$$N(a, 2) + N(a, 3) \geq \frac{\log \log a}{\log \log \log a + C} - 1,$$

for $a \in \mathbb{Z}_{\geq 0}$.

Digit expansion of a



$a = p2^{k_1} + e_1$.

We need $p^K e_1 < 2^{k_1}$ where $K = C \log\log_2 a$ (this is an upper bound for the exponent of $\max(p, q, 10)$ in the proposition). Previously we have $pe_1 < 2^{k_1(1-\varepsilon)}$.

Alternative to heights of minimal polynomials (is better behaved under operations like addition):

Definition 1.11 (Mahler measure)**.** Let $P \in \mathbb{C}[X]$

$$\begin{aligned} P(X) &= a_d X^d + a_{d-1} X^{d-1} + \cdots + A_0 \\ &= a_d (X - \alpha_1) \cdots (X - \alpha_d) \end{aligned}$$

Then we define

$$M(P) = |a_d| \cdot \prod_{j=1}^{d} \max(1, |\alpha_j|).$$

We could define the height of an algebraic number α as

$$H(\alpha) = M(f_\alpha)^{\frac{1}{\deg f_\alpha}},$$

Lecture 6 but instead we will define it in a different (but equivalent) way.

Consider two algebraic integers α, β, and assume

$$[\mathbb{Q}[\alpha + \beta] : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}] \times [\mathbb{Q}(\beta) : \mathbb{Q}].$$

This means that the Galois-conjugates are $\alpha_i + \beta_j$ where α_i runs through the conjugates of α and β_j runs through the conjugates of β.

Then

$$\begin{aligned} M(f_{\alpha+\beta}) &= \prod_{i,j} \max(1, |\alpha_i + \beta|) \\ &\leq \prod_{i,j} 2 \max(1, |\alpha_i|) \max(1, |\beta_j|) \\ &= 2^{d_1 d_2} \left(\prod_i \max(1, |\alpha_i|) \right)^{d_2} \left(\prod_j \max(1, |\beta_i|) \right)^{d_1} \\ &= 2^{d_1 d_2} M(f_\alpha)^{d_2} M(f_\beta)^{d_1} \end{aligned}$$

Recall that we mentioned that we could define

$$H(\alpha) = M(f_\alpha)^{\frac{1}{\deg f_\alpha}}.$$

Then would have

$$H(\alpha + \beta) \leq 2H(\alpha)H(\beta).$$

Similarly,

$$H(\alpha\beta) \leq H(\alpha)H(\beta).$$

Proposition. Let $P \in \mathbb{C}[X]$ of degree d. Then

$$2^{-d}H(P) \leq M(P) \leq (d+1)H(P).$$

Proof. For the upper bound:

$$\log M(P) = \int_0^1 \log |P(e^{-2\pi it})| \mathrm{d}t.$$

Known as Jensen's formula (enough to prove for P of degree 1).

Note that

$$|P(X)| \leq (d+1)H(P)$$

for all $|X| = 1$. This with Jensen's formula gives the upper bound. For the lower bound:

$$P(X) = a_d X^d + \cdots + a_1 X + a_0.$$

Then

$$\left|\frac{a_j}{a_d}\right| = \sum_{\{k_1,\ldots,k_j\} \subset \{1,\ldots,d\}} \underbrace{|a_{k_1}| \cdots |a_{k_j}|}_{\leq M(f)/|a_d|}.$$

The number of terms is $\leq 2^d$. Hence $|a_j| \leq 2^d M(P)$. □

Absolute Values

Let K be a number field. Then a function $|\bullet| : K \to \mathbb{R}_{\geq 0}$ is an absolute value if:

- $|\alpha\beta| = |\alpha||\beta|$
- $|\alpha + \beta| \leq |\alpha| + |\beta|$ for all $\alpha, \beta \in K$

Example.

- Trivial absolute value: $|\alpha| = 0$ for all $\alpha \in K$.
- Let $\sigma : K \to \mathbb{C}$ be an embedding. Then $|\alpha|_\sigma = |\sigma(\alpha)|$.
- Let $P \subset \mathcal{O}_K$ be a non-zero prime ideal lying above $p \in \mathbb{Z}$. (This means $p \in P$).

  Then we define ord_P on K as follows: for $\alpha \in \mathcal{O}_K$, $\mathrm{ord}_P(\alpha)$ is the largest m such that $P^m \mid \alpha\mathcal{O}_K$. For $\alpha, \beta \in \mathcal{O}_K$, $\mathrm{ord}_P(\alpha/\beta) = \mathrm{ord}_P(\alpha) - \mathrm{ord}_p(\beta)$.

  Let $e_P = \mathrm{ord}_P(p)$ (ramification index). Then we define

$$|\alpha|_P = p^{-\mathrm{ord}_P(\alpha)/e_p}.$$

Comment on the normalisation: for $\alpha \in \mathbb{Q}$, we have $|\alpha|_\sigma = |\alpha|_\infty$, and $|\alpha|_P = |\alpha|_p$.

The places of K are M_K comprises:

- all embeddings $\sigma : K \to \mathbb{C}$ such that $\sigma(K) \in \mathbb{R}$
- one from each complex conjugate pairs from the rest
- all (non-zero) prime ideals

For $v \in M_K$, $|\bullet|_v$ denotes the absolute value given above.

Infinite places: $M_{K,\infty}$: embeddings.

Finite places: $M_{K,f}$: prime ideals.

For $v \in M_K$, we define d_v as follows:

- if v is a real embedding, then $d_v = 1$.
- if v is complex, then $d_v = 2$.
- if v is a prime ideal, then $d_v = e_v \cdot f_v$, where: $[\mathcal{O}_K/v : \mathbb{Z}/p\mathbb{Z}] = f_v$ (where p is the rational prime below v).

Comment:

$$d_v = [K : \mathbb{Q}_p]$$

Lecture 7 where p is the place of $\mathbb{Q}$ below v.

L/K extension of number fields, then $w \in M_L$ lies above $v \in M_K$; in notation $w \mid v$.

If both are embeddings and $w|_K = v$ or $w|_K = \overline{v}$ or both are finite and w lies over v as prime ideals, i.e. $w \mid v\mathcal{O}_L$.

Remark. $\sum_{v|\infty} d_v = [K : \mathbb{Q}]$, $\sum_{v|p} d_V = [K : \mathbb{Q}]$.

Proposition (Product formula)**.** Let K be a number field. Then for all $\alpha \in K \neq 0$, we have

$$\prod_v |\alpha|_v^{d_v} = 1.$$

Proof. We compute $N(\alpha\mathcal{O}_K)$ in two ways.

$$N(\alpha\mathcal{O}_K) = \prod_{v \in M_{K,f}} N(v)^{\mathrm{ord}_v(\alpha)} = \prod_{v \in M_{K,f}} p^{f_v \cdot \mathrm{ord}_v(\alpha)},$$

where p is the rational prime lying below v.

Recall $|\alpha|_v = p^{-\operatorname{ord}_v(\alpha)/e_v} = p^{-\operatorname{ord}_v(\alpha)\cdot f_v/d_v}$. So

$$N(\alpha\mathcal{O}_K) = \prod_{v\in M_{\alpha,f}} |\alpha|_v^{-d_v}.$$

Also,

$$N(\alpha\mathcal{O}_K) = |N(\alpha)| = \prod_{v\in M_{K,\infty}} |\alpha|_v^{d_v}.$$

Dividing the equations gives the desired result. □

Now we define

$$H(\alpha) = \left(\prod_{v\in M_K} \max(1, |\alpha|_v)\right)^{\frac{1}{[K:\mathbb{Q}]}}.$$

We will also use $h(\alpha) = \log H(\alpha)$. We won't be using that much, but we mention it mostly because it is used in the literature.

H is known as "multiplicative height", while h is known as "logarithmic / absolute / Weil height".

Proposition 1.12. Let L/K be an extension of number fields. Let $\alpha \in K$. Then $H(\alpha)$ as defined above is the same for K and L.

Proof. **Claim 1:** If $w \in M_L$, $v \in M_K$ such that $w \mid v$ then $|\alpha|_w = |\alpha|_v$ for all $\alpha \in K$.

Claim 2: $\sum_{w|v} d_w = [L : K]d_v$.

Assuming these claims are true, then for $\alpha \in K$

$$\prod_{w|v} \max(1, |\alpha|_w)^{d_w} = \max(1, |\alpha_v|)^{[L:K]d_v}$$

Then

$$\left(\prod_{w|v} \max(1, |\alpha|_w)^{d_w}\right)^{\frac{1}{[L:\mathbb{Q}]}} = \max(1, |\alpha_v|)^{\frac{d_v}{[K:\mathbb{Q}]}}$$

Which implies the desired result.

Proof of Claim 1: Will show if v, w are embeddings then

$$|\alpha|_w = |w(\alpha)| = |v(\alpha)| = |\alpha|_v.$$

If w, v are prime ideals, then we need

$$\frac{\operatorname{ord}_w(\alpha)}{e_w} = \frac{\operatorname{ord}_v(\alpha)}{e_v}.$$

For this, note that for all ideals $I \subset \mathcal{O}_K$, we have

$$\mathrm{ord}_w(I \cdot \mathcal{O}_L) = \mathrm{ord}_w(v \cdot \mathcal{O}_L) \cdot \mathrm{ord}_v(I).$$

Use this for $p\mathcal{O}_K$ and $\alpha\mathcal{O}_K$ in the role of I:

$$\begin{aligned} e_w &= \mathrm{ord}_w(v \cdot \mathcal{O}_L) e_v \\ \mathrm{ord}_w(\alpha) &= \mathrm{ord}_w(v \cdot \mathcal{O}_L) \cdot \mathrm{ord}_v(\alpha) \end{aligned}$$

Proof of Claim 2: Omitted. □

Proposition. Let $\alpha \in \overline{\mathbb{Q}}_{\neq 0}$. Then

$$H(\alpha) = M(f_\alpha)^{\frac{1}{\deg(f_\alpha)}}.$$

Remark. Recall $2^{-d} H(f_\alpha) \leq H(\alpha)^d \leq (1+d) H(f_\alpha)$.

Proof. Enough to prove

$$|a_d|^{[K:\mathbb{Q}]} = \prod_{v \in M_{K,f}} \max(1, |x|_v)^{d_v},$$

where K is a number field with $\alpha \in K$.

Lecture 8 If $K = \mathbb{Q}(\alpha)$, then this is immediate from the definitions.

For a polynomial $P \in K[X]$, we write $|P|_v$ for the maximum $|\bullet|_v$ of all the coefficients of P.

A variant of Gauss's lemma can be stated as follows: Let $Q_1, Q_2 \in K[X]$. Then $|Q_1 Q_2|_v = |Q_1|_v |Q_2|_v$ for $v \in M_{K,f}$.

Observe that $|f_\alpha|_v = 1$ (for all $v \in M_{K,f}$) because the coefficients are coprime rational integers. We write $f_\alpha = a_d(X - \alpha_1) \cdots (X - \alpha_d)$ (we take K to be the splitting field of f_α). Gauss's lemma gives

$$\prod_{v \in M_{K,f}} |a_d|_v^{d_v} \cdot \prod_{v \in K_{K,f}} \prod_{j=1}^{d} \max(1, |\alpha_j|_v))^{d_v} = 1.$$

Let σ be an automorphism of K such that $\sigma\alpha_j = \alpha$ for some fixed j. This permutes $M_{K,f}$. That is, $\forall v \in M_{K,f}$, there exists $\sigma v \in M_{K,f}$ such that $|\sigma\beta|_{\sigma v} = |\beta|_v$. So

$$\begin{aligned} \prod_{v \in M_{K,f}} \max(1, |\alpha_j|_v)^{d_v} &= \prod_{v \in M_{K,f}} \max(1, |\underbrace{\sigma\alpha_j}_{=\alpha}|_{\sigma v})^{d_{\sigma v}} \\ &= \prod_{v \in M_{K,f}} \max(1, |\alpha|_v)^{d_v} \end{aligned}$$

By the product formula:

$$\prod_{v\in M_{K,f}} |a_d|_v^{d_v} = \prod_{v\in M_{K,\infty}} |a_d|_v^{-d_v} = |a_d|^{-[K:\mathbb{Q}]}.$$

So

$$\left[\prod_{v\in M_{K,f}} \max(1,|\alpha|_v)^{d_v}\right]^{\overbrace{d}^{=[\mathbb{Q}(\alpha):\mathbb{Q}]}} = |a_d|^{[K:\mathbb{Q}]} \qquad \square.$$

Lemma. Let $\alpha \in \overline{\mathbb{Q}}$, and $k \in \mathbb{Z}$. Then

$$H(a^k) = H(\alpha)^{|k|}.$$

Proof. If $k > 0$, then this is immediate from the definition. So just need to consider $k = -1$:

$$H(\alpha^{-1})^d = \prod_{v\in M_K} \max(1,|\alpha|_v^{-1})^{d_v}$$

($d = \deg \alpha$). We multiply this by

$$\prod_{v\in M_v} |\alpha|_v^{d_v} = 1.$$

So

$$H(\alpha^{-1})^d = \prod_{v\in M_k} \max(|\alpha|_v,1)^{d_v} = H(\alpha)^d.$$

$\square$

Let P be a polynomial in possibly several variables, with complex coefficients. Then $\mathcal{L}(P)$ is defined to be the sum of the absolute values of all the coefficients. This is sometimes called the length of P.

Proposition. Let $k \in \mathbb{Z}_{>1}$, $n_1,\dots,n_k \in \mathbb{Z}_{>0}$. Let $P, Q \in \mathbb{Z}[X_1,\dots,X_k]$ of degree $\le n_j$ in X_j. Let $\alpha_1,\dots,\alpha_k \in \overline{\mathbb{Q}}_{\neq 0}$. Then:

$$H\left(\frac{P(\alpha_1,\dots,\alpha_k)}{Q(\alpha_1,\dots,\alpha_k)}\right) \le \max(\mathcal{L}(P),\mathcal{L}(Q)) \cdot \prod_{j=1}^{k} H(\alpha_j)^{k_j}.$$

In particular: $H(\alpha\beta) \le H(\alpha)H(\beta)$ and $H(\alpha+\beta) \le 2H(\alpha)H(\beta)$.

Proof. Let K be a number field containing all α_i.

$$\begin{aligned} H\left(\frac{P(\dots)}{Q(\dots)}\right)^{[K:\mathbb{Q}]} &= \prod_{v\in M_K} \max\left(1, \left|\frac{P(\dots)}{Q(\dots)}\right|_v\right)^{d_v} \\ &= \prod_{v\in M_K} \max(|Q(\dots)|_v, |P(\dots)|_v)^{d_v} \qquad \text{from product formula for } Q(\dots) \end{aligned}$$

Let first $v \in M_{K,f}$. Then

$$\begin{aligned} |P(\alpha_1,\dots,\alpha_k)|_v &\le \max_{\substack{j_1=0,\dots,n_1 \\ \vdots \\ j_k=0,\dots,n_k}} |\alpha_1|_v^{j_1}\cdots|\alpha_k|^{j_k} \\ &= \prod_{i=1}^{k} \max(1, |\alpha_i|_v^{n_i}) \end{aligned}$$

For $v \in M_{K,\infty}$:

$$|P(\alpha_1,\dots,\alpha_k)|_v \le \mathcal{L}(P)\cdot\prod_{i=1}^{k}\max(1,|\alpha_i|_v^{n_i}).$$

So

$$H\left(\frac{P(\dots)}{Q(\dots)}\right)^{[K:\mathbb{Q}]} \le \max(\mathcal{L}(P),\mathcal{L}(Q))^{[K:\mathbb{Q}]}\prod_{i=1}^{k}\prod_{v\in M_K}\max(1,|\alpha_i|_v)^{n_i d_v}.$$

Then taking a $[K:\mathbb{Q}]$ root of both sides gives the desired inequality. □

Lemma. Let $\alpha \in \overline{\mathbb{Q}} \subset \mathbb{C}$. Then:

$$H(\alpha)^{-\deg\alpha} \le |\alpha| \le H(\alpha)^{\deg\alpha}.$$

This is sometimes known as "trivial bound" or "Liouville's bound".

Proof.

$$H(\alpha)^{\deg\alpha} = \prod_{v\in M_K}\max(1,|\alpha|_v)^{d_v} \ge |\alpha|$$

Apply this for α^{-1}:

$$\begin{aligned} |\alpha^{-1}| &\le H(\alpha^{-1})^d = H(\alpha)^d \\ |\alpha| &\ge H(\alpha)^{-d} \end{aligned}$$

□

Theorem (Siegel)**.** Let α be a real algebraic irrational number. Then for all $\varepsilon > 0$, there exists $c = c(\alpha, \varepsilon) > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| \geq cq^{-\sqrt{2d}-\varepsilon}$$

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

Lecture 9 We will spend the next 3-5 lectures proving this.

We will spend today's lecture discussing an outline of the proof, discussing why certain parts are necessary and also some intuition as to why one would expect this method to work.

(1) Suppose to the contrary that there are infinitely many $\frac{p_1}{q_1}, \frac{p_2}{q_2}, \ldots$ such that $\left|\alpha - \frac{p_j}{q_j}\right| > \frac{1}{q^{\sqrt{2d}+\varepsilon}}$.

(2) Choose two among these appropriately, which I will denote $\frac{p_1}{q_1}, \frac{p_2}{q_2}$.

(3) Construct a polynomial $P \in \mathbb{Z}[X_1, X_2]$ that vanishes at (α, α) to high order.

(4) Give a lower bound on $P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)$.

(5) Give an upper bound on $P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)$.

(6) Realise that they give a contradiction.

1 variable is not enough: let $P(X)$ be of degree n. Then P may vanihs at α to order n/d. Then we have a lower bound of

$$\left|P\left(\frac{p}{q}\right)\right| \geq \frac{1}{q^n},$$

and we might hope for an upper bound like

$$\left|P\left(\frac{p}{q}\right)\right| \lesssim \left|\alpha - \frac{p}{q}\right|^{n/d}.$$

To get a contradiction, we need $\left|\alpha - \frac{p}{q}\right|^{n/d} < \frac{1}{q^n}$, i.e. $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^d}$.

Lower bound

$$\left|P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \geq \frac{1}{q_1^{n_1} q_2^{n_2}}$$

where n_1 is the degree in X_1 and n_2 is the degree in n_2.

Upper bound:

$$P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) = \sum_{j_1, j_2} P_{j_1, j_2}(\alpha, \alpha) \left(\alpha - \frac{p_1}{q_1}\right)^{j_1} \left(\alpha - \frac{p_2}{q_2}\right)^{j_2}$$

where $P_{j_1,j_2}(X_1,X_2) = \frac{1}{j_1!j_2!}\frac{\partial^{j_1+j_2}}{\partial X_1^{j_1}\partial X_2^{j_2}}P(X_1,X_2)$. Note

$$\left(\alpha - \frac{p_1}{q_1}\right)^{j_1}\left(\alpha - \frac{p_2}{q_2}\right)^{j_2} \leq \frac{1}{q_1^{j_1(\sqrt{2d}+\varepsilon)}} \cdot \frac{1}{q_2^{j_2(\sqrt{2d}+\varepsilon)}} = \exp(-(\sqrt{2d}+\varepsilon)(j_1\log q_1 + j_2 \log q_j 2)).$$

Index of P at (β_1,β_2) with respect to the weights w_1, w_2.

$$I_P(\beta_1,\beta_2;w_1,w_2) = \min(j_1w_1 + j_2w_2, P_{j_1,j_2}(\beta_1,\beta_2) \neq 0).$$

Use $w_1 = \log q_1$, $w_2 = \log q_2$. With this, we get the upper bound

$$\left|P\left(\frac{p_1}{q_1},\frac{p_2}{q_2}\right)\right| \lesssim \exp(-(\sqrt{2d}+\varepsilon)\cdot I_P(\alpha,\alpha)).$$

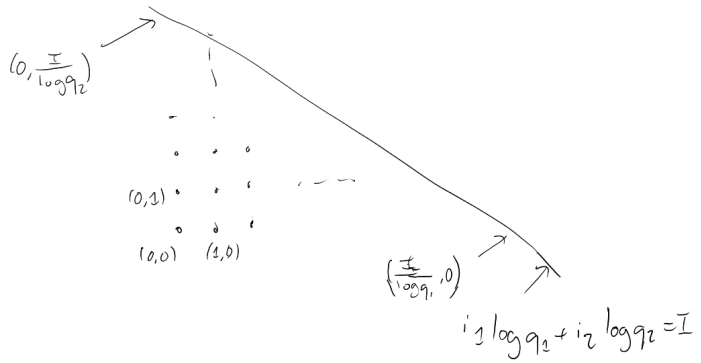
How big can $I_P(\alpha,\alpha)$ be made? We look for P in the form

$$P(X_1,X_2) = \sum_{i_1=0}^{n_1}\sum_{i_2=0}^{n_2} a_{i_1,i_2}X_1^{i_1}X_2^{i_2}.$$

The condition that $P_{j_1,j_2}(\alpha,\alpha) = 0$ is a linear equation for a_{i_1,i_2} over $\mathbb{Q}[\alpha]$.

By picking a basis of $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$, this becomes a system of d linear equations. To find P such that $I_P(\alpha,\alpha) \geq I$ we need to solve:

$$d\cdot|\{(j_1,j_2) : j_1\log q_1 + j_2 \log q_2 \leq i\}| \sim \frac{I^2}{2\log q_1 \cdot \log q_2}$$



I can choose n, n_2, I, and I want to do the following:

$$\frac{dI^2}{2\log q_1 \log q_2} \lesssim n_1 n_2.$$

$$\exp(-(\sqrt{2d}+\varepsilon)I) \lesssim \frac{1}{q_1^{n_1}q_2^{n_2}}$$

$$(\sqrt{2d}+\varepsilon)I \gtrsim n_1 \log q_1 + n_2 \log q_2$$

Take $n_k \sim \frac{\sqrt{2d}+\varepsilon}{2} \cdot \frac{I}{\log q_k}$ for some large I.

Subtleties that still need to be considered:

- Siegel's Lemma will be needed to make sure that the P_{j_1,j_2} are not too large.
- $P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0$.

$P_{j_1,j_2} \to$ coefficient of $x_1^{i_2} x_2^{i_2}$ is $a_{i_1+j_1,i_2+j_2} \cdot \binom{i_1+j_1}{i_1}\binom{i_2+j_2}{j_2}$, where $a_{i_1+i_2,j_1+j_2}$ is the coefficient of $X_1^{i_1+j_1} X_2^{i_2+j_2}$ in P.

Lecture 10 $H(P_{j_1,j_2}) \leq 2^{n_1+n_2} H(P)$.

Thue: $P(X,Y) = R_1(X) + Y R_2(X)$.

Let L be a linear form in $K[X_1, \ldots, X_N]$ where K is a number field.

For $v \in M_K$: $|L|_v = \max(|a_j|_v)$ where $L = a_1 X_1 + \cdots + a_N X_N$. Then define

$$H(L) = \left(\prod_{v \in M_K} |L|_v^{d_v} \right)^{\frac{1}{[K:\mathbb{Q}]}}.$$

By the product formula, this is invariant under multiplication by an element $\alpha \in K^\times$:

$$|\alpha L|_v = |\alpha|_v |L|_v,$$

so

$$H(\alpha L) = \prod_{v \in M_K} |\alpha L|_v^{d_v} = H(L) \prod_{v \in M_K} |\alpha|_v^{d_v} = H(L).$$

Lemma (Siegel's lemma)**.** Let K be a number field of degree D. Let $M, N \in \mathbb{Z}_{>0}$ such that $N > MD$ and let $\mathcal{H} \in \mathbb{R}_{\geq 1}$. Let $L_1, \ldots, L_M \in K[X_1, \ldots, X_N]$ be linear forms such that $H(L_j) \leq \mathcal{H}$. Then there exist $x_1, \ldots, x_N \in \mathbb{Z}$ (not all 0) such that $L_h(x_1, \ldots, x_N) = 0$ for $j = 1, \ldots, M$ and

$$|x_i| \leq (N\mathcal{H})^{\frac{MD}{N-MD}}.$$

In particular, if $N \geq MD$, then the bound is $N\mathcal{H}$.

There is a refinement of this lemma which is due to Bombieri and Vaaler.

Corollary. Let α be an algebraic number of degree D. Let $w_1, w_2, \delta \in \mathbb{R}_{>0}$, and let $I \in \mathbb{R}_{>0}$.

Let $n_1, n_2 \in \mathbb{Z}_{>0}$. Suppose that

$$|\{(i_1, i_2) \in \mathbb{Z}^2_{\geq 0} : i_1 w_1 + i_2 w_2 < I\}| \leq \frac{(n_1+1)(n_2+1)}{(1+\delta)D}.$$

Then there exists $P \neq 0 \in \mathbb{Z}[X_1, X_2]$ of degree n_j in X_j such that $I_P(\alpha, \alpha, w_1, w_2) \geq I$ and

$$H(P) \leq (4H(\alpha))^{(n_1+n_2)\delta^{-1}}$$

where $H(P)$ is the maximal absolute value of hte coefficients.

Proof. For (i_1, i_2) consider:

$$L_{i_1,i_2} = \sum_{j_1=0}^{n_1} \sum_{j_2=0}^{n_2} \binom{j_1}{i_1} \binom{j_2}{i_2} a_{j_1,j_2} \cdot \alpha^{j_1-i_1+j_2-i_2}$$

where a_{j_1,j_2} are variables of L_{i_1,i_2}. Then

$$L_{i_1,i_2}((a_{j_1,j_2})_{j_1,j_2}) = 0 \iff P_{i_1,i_2}(\alpha, \alpha) = 0$$

where

$$P = \sum_{j_1=0}^{n_1} \sum_{j_2=0}^{n_2} a_{j_1,j_2} X_1^{j_1} X_2^{j_2}.$$

Need to find $(a_{j_1,j_2})_{j_1,j_2}$ such that $L_{i_1,i_2}((a_{j_1,j_2})) = 0$ for all i_1, i_2 with $i_1 w_1 + i_2 w_2 \leq I$.

Apply Siegel's lemma:

$$N = (n_1+1)(n_2+1), \qquad M \leq \frac{N}{(1+\delta)D}.$$

Then

$$\frac{MD}{N - MD} \leq \frac{MD}{(1+\delta)MD - MD} = \delta^{-1}.$$

We need to estimate $H(L_{i_1,i_2})$. For finite places v,

$$|L_{i_1,i_2}|_v \leq \max(1, |\alpha|_v)^{n_1+n_2}.$$

For infinite places:

$$|L_{i_1,i_2}|_v \leq 2^{n_1} \cdot 2^{n_2} \max(1, |\alpha|_v)^{n_1+n_2}$$

Then

$$H(L_{i_1,i_2}) \leq 2^{n_1+n_2} \cdot H(\alpha)^{n_1+n_2} =: \mathcal{H}.$$

Then Siegel's lemma gives the bound

$$[2^{n_1+n_2} H(\alpha)^{n_1+n_2} \underbrace{(n_1+1)(n_2+1)}_{\leq 2^{n_1+n_2}}]^{\delta^{-1}}.$$

□

Proof of Siegel's lemma for $K = \mathbb{Q}$. We can assume that the coefficients of each L_j are integers, and that they are relatively prime. Then each coefficient is bounded by $\mathcal{H}$. Take $Y = \left\lfloor (N\mathcal{H})^{\frac{MD}{N-MD}} \right\rfloor$.

Consider $(y_1, \ldots, y_N) \in \{0, 1, \ldots, Y\}^N$. Evaluating L_j at all such $(y_1, \ldots, y_N)$ we have

$$\max L_j(y_1, \ldots, y_N) - \min L_j(y_1, \ldots, y_N) \leq Y \cdot \mathcal{H}N.$$

The number of possible values of $L_j(y_1, \ldots, y_N)$ is $\leq Y \cdot H \cdot N + 1$.

Claim: $(Y\mathcal{H} \cdot N + 1)^M < (Y + 1)^N$.

Indeed:

$$Y = \left\lfloor (N \cdot \mathcal{H})^{\frac{M}{N-M}} \right\rfloor$$
$$Y + 1 > (N \cdot \mathcal{H})^{\frac{M}{N-M}}$$
$$(Y + 1)^N > (N \cdot \mathcal{H})^M \cdot (Y + 1)^M$$

The claim follows by

$$N\mathcal{H}Y + 1 < N\mathcal{H}(Y + 1).$$

Note that the above line uses the fact that $\mathcal{H} \geq 1$!

By the box principle, there exist $(y_1, \ldots, y_N) \neq (z_1, \ldots, z_N)$, with entries bounded by Y, such that

$$L_j(y_1, \ldots, y_N) = L_j(z_1, \ldots, z_N) \qquad \forall j = 1, \ldots, M. \qquad \square$$

Lecture 11

In the $K = \mathbb{Q}$ case, a key step is that for $L \in \mathbb{Z}[X_1, \ldots, X_N]$ and $H(L) \leq \mathcal{H}$, the points $L(y_1, \ldots, y_N)$ are integers confined in an interval of length $N\mathcal{H}Y$ (where $y_1, \ldots, y_N = 0, \ldots, N$).

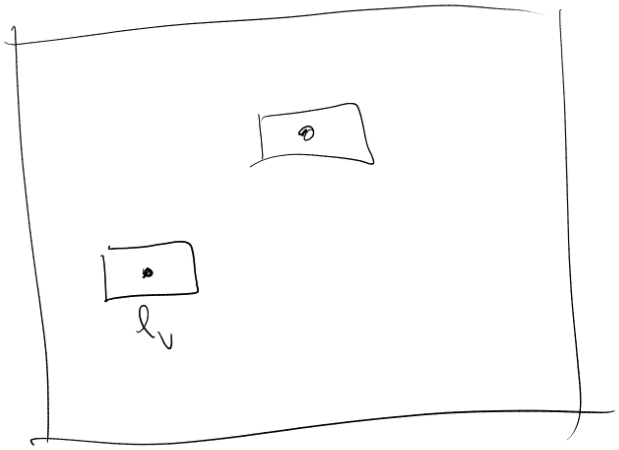
In the general case, consider the map:

$$\Phi : K \to \mathbb{R}^n \cdot \mathbb{C}^s \cong \mathbb{R}^D$$
$$\alpha \mapsto (v(\alpha))_{v \in M_{K,\infty}}$$

The v-component of $\Phi(L(y_1, \ldots, y_N))$ is confined in an interval (or box) of size $NY \cdot |L|_v$.

Let $\alpha = L(y_1, \ldots, y_N) - (z_1, \ldots, z_N) \neq 0$. By the product formula,

$$\prod_{v \in M_{K,\infty}} |\alpha|_v^{d_v} = \prod_{v \in M_{K,f}} |\alpha|_v^{-d_v} \geq \prod_{v \in M_{K,f}} |L|_v^{d_v}.$$



Make sure $\prod_v l_v \leq$ RHS of above.

Non-vanishing:

Proposition. For every $\varepsilon > 0$, there exists $C = C(\varepsilon)$ such that the following holds. Let $n_1, n_2 \in \mathbb{Z}_{>0}$, and let $\frac{p_1}{q_1}, \frac{p_2}{q_2} \in \mathbb{Q}$. Suppose that

$$\exp(n_1 + n_2) < q_j^{n_j/C}$$

for $j = 1, 2$, and that $\log q_2 > C \log q_1$.
Let $P \neq 0 \in \mathbb{Z}[X_1, X_2]$ of degree in n_j in X_j for $j = 1, 2$ such that

$$H(P) < q_j^{n_j/C}$$

for $j = 1, 2$. Then

$$I_P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}, \log q_1, \log q_2\right) \leq \varepsilon(n_1 \log q_1 + n_2 \log q_2).$$

Note: from now on, whenever we say $\frac{p}{q} \in \mathbb{Q}$, we also mean $\gcd(p, q) = 1$.

When we apply this we will have $n_1 \log q_1 \sim n_2 \log q_2$.

Without the asymmetry assumption ($\log q_2 > C \log q_1$), we have the counterexample: $P = (X_1 - X_2)^n$, with $\frac{p_1}{q_1} = \frac{p_2}{q_2}$.

Alternatively: $P = (R(X_1) - X_2 Q(X_1))^n$ (for R, Q some small degree polynomials) for any $\frac{p_1}{q_1}, \frac{p_2}{q_2}$ such

that

$$\frac{p_2}{q_2} = \frac{R\left(\frac{p_1}{q_1}\right)}{Q\left(\frac{p_2}{q_2}\right)}$$

Lemma. Let $F, F^{(1)}, F^{(2)} \in \mathbb{Z}[X_1, X_2]$, and let $i_1, i_2 \in \mathbb{Z}_{\geq 0}$. Let $\alpha_1, \alpha_2 \in \mathbb{R}$ and $w_1, w_2 \in \mathbb{R}_{>0}$. Then the following holds:

$$\begin{aligned} I_{F_{i_1,i_2}}(\alpha_1, \alpha_2) &\geq I_F(\alpha_1, \alpha_2) - i_1 w_1 - i_2 w_2 \\ I_{F^{(1)}+F^{(2)}}(\alpha_1, \alpha_2) &\geq \min_{j=1,2} I_{F^{(j)}}(\alpha_1, \alpha_2) \\ I_{F^{(1)}F^{(2)}} &= I_{F^{(1)}}(\alpha_1, \alpha_2) + I_{F^{(2)}}(\alpha_1, \alpha_2) \end{aligned}$$

Baby case: $P(X_1, X_2) = F(X_1)G(X_2)$ for some F, G polynomials.

In this case if $I_P \geq \varepsilon(n_1 \log q_1 + n_2 \log q_2)$ then either $I_F \geq \varepsilon n_1 \log q_1$ or $I_G \geq \varepsilon n_2 \log q_2$.

If F vanishes at $\frac{p_1}{q_1}$ to order m for some m, then

$$(q_1 X_1 - p_1)^m \mid F.$$

The leading coefficient of $F$is divisible by q_1^m. In particular, $H(F) > q_1^m$. Then $H(F) > q_1^{\varepsilon n_1}$ or $H(G) > q_2^{\varepsilon n_2}$.

Hence $H(P) > \min(q_1^{\varepsilon n_1}, q_2^{\varepsilon n_2})$, which contradicts the assumptions.

In general, we can always write

$$P(X_1, X_2) = F^{(1)}(X_1)G^{(1)}(X_2) + \cdots + F^{(h)}(X_1)G^{(h)}(X_2)$$

with $h \leq n_2$.

Consider $h = 2$.

$$\begin{aligned} P(X_1, X_2) &= F^{(1)}(X_1) \cdot G^{(1)}(X_2) + F^{(2)}(X_1)G^{(2)}(X_2) \\ \frac{\partial}{\partial X_2} P &= F^{(1)} \cdot \frac{\partial}{\partial x_2} G^{(1)} + F^{(2)} \cdot \frac{\partial}{\partial X_2} G_2 \end{aligned}$$

Lecture 12

$$\frac{\partial}{\partial X_2} G^{(2)} P - G^{(2)} \frac{\partial}{\partial X_2} P = F^{(1)} \left(G^{(1)} \frac{\partial}{\partial X_2} G^{(2)} - \frac{\partial}{\partial X_2} G^{(1)} \cdot G^{(2)} \right)$$

We will later have to worry about whether the resulting polynomial is 0.

For any h:

$$\begin{vmatrix} P & G^{(2)} & \cdots & G^{(h)} \\ \frac{\partial}{\partial X_2}P & \frac{\partial}{\partial X_2}G^{(2)} & \cdots & \frac{\partial}{\partial X_2}G^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}P & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(2)} & \cdots & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(h)} \end{vmatrix}$$
$$= \begin{vmatrix} F_1G^{(1)} & G^{(2)} & \cdots & G^{(h)} \\ F_1\frac{\partial}{\partial X_2}G^{(1)} & \frac{\partial}{\partial X_2}G^{(2)} & \cdots & \frac{\partial}{\partial X_2}G^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ F_1\frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(1)} & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(2)} & \cdots & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(h)} \end{vmatrix}$$
$$= F_1 \begin{vmatrix} G^{(1)} & G^{(2)} & \cdots & G^{(h)} \\ \frac{\partial}{\partial X_2}G^{(1)} & \frac{\partial}{\partial X_2}G^{(2)} & \cdots & \frac{\partial}{\partial X_2}G^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(1)} & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(2)} & \cdots & \frac{\partial^{h-1}}{\partial X_2^{(h-1)}}G^{(h)} \end{vmatrix}$$

The degree increases h-fold, but not the index.

$$\begin{vmatrix} P_{0,0} & P_{0,1} & \cdots & P_{0,h-1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{h-1,0} & P_{h-1,1} & \cdots & P_{h-1,h-1} \end{vmatrix}$$
$$= \begin{vmatrix} F^{(1)} & F^{(2)} & \cdots & F^{(h)} \\ F_1^{(1)} & F_1^{(2)} & \cdots & F_1^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ F_{h-1}^{(1)} & F_{h-1}^{(2)} & \cdots & F_{h-1}^{(h)} \end{vmatrix} \cdot \begin{vmatrix} G^{(1)} & G_1^{(1)} & \cdots & G_{h-1}^{(1)} \\ G^{(2)} & G_1^{(2)} & \cdots & G_{h-1}^{(2)} \\ G^{(h)} & G_1^{(h)} & \cdots & G_{h-1}^{(h)} \end{vmatrix}$$

where $P_{ij} = \frac{1}{i!j!}\frac{\partial^{i+j}}{\partial X_1^i \partial X_2^j}P$, $F_i = \frac{1}{i!}\frac{\partial^i}{\partial X_1^i}F$.

Lemma. Let $F^{(1)}, F^{(2)}, \ldots, F^{(h)}$ be $\mathbb{Q}$-linearly independent polynomials in $\mathbb{Z}[X]$. Then

$$\begin{vmatrix} F^{(1)} & F^{(2)} & \cdots F^{(h)} & \\ F_1^{(1)} & F_1^{(2)} & \cdots & F_1^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ F_{h-1}^{(1)} & F_{h-1}^{(2)} & \cdots & F_{h-1}^{(h)} \end{vmatrix} \neq 0.$$

(Wronskian)

Proof of Proposition assuming the lemma. Suppose to the contrary that the proposition does not hold for some $P, \frac{p_1}{q_1}, \frac{p_2}{q_2}$. Write $P = F^{(1)}G^{(1)} + \cdots + F(h)G^{(h)}$ such that h is minimal. Then $h \leq n_2 + 1$ and

the $F^{(1)}, \cdots, F^{(k)}$ and $G^{(1)}, \ldots, G^{(h)}$ are $\mathbb{Q}$-linearly independent. Then consider

$$\mathcal{P} = \begin{vmatrix} P_{0,0} & \cdots & P_{0,h-1} \\ \vdots & \ddots & \vdots \\ P_{h-1,0} & \cdots & P_{h-1,h-1} \end{vmatrix}$$

and

$$\mathcal{F} = \begin{vmatrix} F_{0,0} & \cdots & F_{0,h-1} \\ \vdots & \ddots & \vdots \\ F_{h-1,0} & \cdots & F_{h-1,h-1} \end{vmatrix} \mathcal{G} = \begin{vmatrix} G_{0,0} & \cdots & G_{0,h-1} \\ \vdots & \ddots & \vdots \\ G_{h-1,0} & \cdots & G_{h-1,h-1} \end{vmatrix}$$

Then $\mathcal{P}(X_1, X_2) = \mathcal{F}(X_1)\mathcal{G}(X_2) \neq 0$ by the above Lemma.

Note $\deg_{X_j} \mathcal{P} \leq hn_j$, $\deg \mathcal{F} \leq n_1$, $\deg \mathcal{G} \leq n_2$. Also

$$\begin{aligned} H(\mathcal{P}) &\leq \underbrace{h!}_{\text{ways to multiply entries}} \Big(\underbrace{(n_1+1)(n_2+1)}_{\text{monomials in the entries}} \Big)^h \underbrace{(2^{n_1+n_2} HP)^h}_{\text{coefficients of entries}} \\ &\leq 2^{(n_1+n_2)h} 2^{(n_1+n_2)h} q_j^{hn_j/C} \end{aligned}$$

for $j = 1, 2$.

$H(\mathcal{P}) = H(\mathcal{F})H(\mathcal{G})$. Then

$$\begin{aligned} H(\mathcal{F}) &\leq (8^{n_1+n_2} q_j^{n_j/C}) \\ &\leq (q_j^{hn_1/C})^h \end{aligned}$$

$$\begin{aligned} H(\mathcal{G}) &\leq (8^{n_1+n_2} q_2^{n_2/C}) \\ &\leq (q_j^{hn_2/C})^h \end{aligned}$$

$I_{P_{i,j}} \geq I_P - i \log q_1 - j \log q_2$. If $j \leq \frac{\varepsilon h}{10} + 1$, $\log q_1 < \frac{\varepsilon}{10} \log q_2$. By the indirect assumption

$$I_P \geq \varepsilon(n_1 \log q_1 + n_2 \log q_2),$$

$$I_{P_{i,j}} \geq \frac{\varepsilon}{2} n_2 \log q_2 + \frac{\varepsilon}{2} n_1 \log q_1.$$

Lecture 13

$$I_{\mathcal{P}} \left(\frac{p_1}{q_1}, \frac{p_2}{q_2} \right) \geq \frac{\varepsilon^2}{20} h(n_1 \log q_1 + n_2 \log q_2).$$

If F vanishes to order m at $\frac{p_1}{q_1}$, then q_1^m divides the leading coefficient of F. In particular, $q_1^m \leq H(F)$.

Then

$$I_F \left(\frac{p_1}{q_1}; \log q_1 \right) \leq \log H(\mathcal{F}) \leq \frac{10hn_1 \log q_1}{C}$$

$$I_G\left(\frac{p_2}{q_2}; \log q_2\right) \leq \log H(\mathcal{G}) \leq \frac{10hn_2 \log q_2}{C}$$

If C is sufficiently large in terms of ε, then

$$I_{\mathcal{P}}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) < I_{\mathcal{F}}\left(\frac{p_1}{q_1}\right) + I_{\mathcal{G}}\left(\frac{p_2}{q_2}\right).$$

A contradiction. □

Now we prove the lemma from earlier:

Lemma. Let $F^{(1)}, F^{(2)}, \ldots, F^{(h)}$ be $\mathbb{Q}$-linearly independent polynomials in $\mathbb{Z}[X]$. Then

$$\begin{vmatrix} F^{(1)} & F^{(2)} & \cdots F^{(h)} & \\ F_1^{(1)} & F_1^{(2)} & \cdots & F_1^{(h)} \\ \vdots & \vdots & \ddots & \vdots \\ F_{h-1}^{(1)} & F_{h-1}^{(2)} & \cdots & F_{h-1}^{(h)} \end{vmatrix} \neq 0.$$

(Wronskian)

Proof. The statement does not change if we replace $F^{(j)}$ by $aF^{(i)} + bF^{(j)}$ for some $a, b \in \mathbb{Q}$ and $i \in \{1, \ldots, h\}$ provided $b \neq 0$.

Then we may assume: $F^{(i)} = X^{mi} +$ lower order terms and the m_i are distinct.

We will prove that:

$$\begin{vmatrix} X^{m_1} & \cdots & X^{m_h} \\ \binom{m_1}{1} X^{m_1-1} & \cdots & \binom{m_h}{1} X^{mh-1} \\ \vdots & \ddots & \vdots \\ \binom{m}{h-1} X^{m_1-h+1} & \cdots & \binom{mh}{h-1} X^{m_h-h+1} \end{vmatrix} \neq 0.$$

Then this is the leading term of th Wronskian, so this will prove the claim. The determinant is equal to:

$$\begin{vmatrix} \binom{m_1}{0} & \cdots & \binom{m_h}{0} \\ \vdots & \ddots & \vdots \\ \binom{m_1}{h-1} & \cdots & \binom{m_h}{h-1} \end{vmatrix} \cdot X^M$$

Supose to the contrary that a non-trivial linear combination of the rows is $(0, 0, \ldots, 0)$. Now the i-th row is a polynomial of degree $i - 1$evaluated at $m_1, \ldots, m_h$. Then the linear combination of the rows is a non-zero polynomial of degree $\leq h - 1$ evaluated at $m_1, \ldots, m_h$. □

Theorem. Let α be an irrational, real algebraic number of degree $d \geq 2$. Then for all $\varepsilon > 0$,

there exists $C = C(\alpha, \varepsilon)$ such that

$$\left|\alpha - \frac{p}{q}\right| > Cq^{-\sqrt{2d}-\varepsilon},$$

for all $\frac{p}{q} \in \mathbb{Q}$.

Proof. Suppose to the contrary that there are infinitely many $\frac{p}{q}$ with

$$\left|\alpha - \frac{p}{q}\right| < q^{-\sqrt{2d}-\varepsilon}.$$

Then fix $\varepsilon_0 > 0$ sufficiently small in terms of α, ε and let C be the constant when the proposition is applied with ε_0 in place of ε.

Now let $\frac{p_1}{q_1}, \frac{p_2}{q_2}$ be such that

$$\left|\alpha - \frac{p_1}{q_1}\right|, \left|\alpha - \frac{p_2}{q_2}\right| < q^{-\sqrt{2d}-\varepsilon}$$

and

$$\log q_1 > C \cdot \varepsilon_0^{-1} \qquad \log q_2 > C \log q_1.$$

We use Siegel's lemma to construct $P(X_1, X_2)$ that vanishes at (α, α) to high order.

We choose $n_1, n_2 \in \mathbb{Z}$ such that

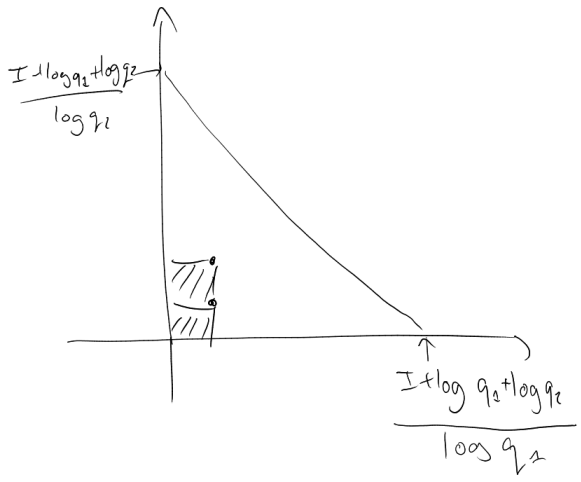
$$n_1 \log q_1 \leq n_2 \log q_2 \leq n_1 \log q_1 + \log q_1.$$

We want a polynomial P such that

$$I_P(\alpha, \alpha) \geq \frac{n_1 \log q_1 + n_2 \log q_2}{\sqrt{2d} + \frac{\varepsilon}{10}}.$$

For this we need to estimate

$$\begin{aligned} |\{(i_1, i_2) \in \mathbb{Z}^2_{\geq 0} : i_1 \log q_1 + i_2 \log q_2 \leq I\}| &\leq \frac{(I + \log q_1 + \log q_2)^2}{2 \log q_1 \log q_2} \\ &\leq \frac{(n_1 + 1)(n_2 + 1)}{(1 + \delta)d} \end{aligned}$$



This is because

$$I \sim \frac{2n_1 \log q_1}{\sqrt{2d}} \sim \frac{2n_2 \log q_2}{\sqrt{2d}}$$

so

$$\frac{I^2}{2 \log q_1 \log q_2} \sim \frac{2n_1 \cdot 2n_2}{2 \cdot 2d} = \frac{n_1 n_2}{d}.$$

Lecture 14

So we find $P \in \mathbb{Z}[X_1, X_2]$ such that $I_P(\alpha, \alpha; \log q_1, \log q_2) \geq I$ and $H)(P) \leq (4H(\alpha))^{\delta^{-1}(n_1+n_2)}$. We need:

$$H(P), \exp(n_1 + n_2) \leq q_j^{n_j/C} \sim q_1^{n_1/C}$$

for $j = 1, 2$ and $\log q_2 > C \log q_1$. This will be fine if $(4H(\alpha))^{\delta^{-1}} < q_1^C$. This is fine if ε_0 is sufficiently small with respect to α and δ.

Then $I_P\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \leq \varepsilon_0(n_1 \log q_1 + n_2 \log q_2)$. Then there exists $\tilde{P}$ a partial derivative of P such that

$$H(\tilde{P}) \leq (8H(\alpha))^{\delta^{-1}(n_1+n_2)},$$

$$I_{\tilde{D}}(\alpha, \alpha) \geq I - \varepsilon_0(n_1 \log q_1 + n_2 \log q_2) \geq \frac{n_1 \log q_1 + n_1 \log q_2}{\sqrt{2d} + \frac{\varepsilon}{5}},$$

if ε_0 is sufficiently small.

$\tilde{P}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0$. Then

$$\left|\tilde{P}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| > \frac{1}{q_1^{n_1} q_2^{n_2}}.$$

Taylor's formula:

$$\tilde{P}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) = \sum_{i_1, i_2} \tilde{P}_{i_1,i_2}(\alpha, \alpha) \left(\alpha - \frac{p}{q}\right)^{i_1} \left(\alpha - \frac{p_2}{q_2}\right)^{i_2}$$

If i_1, i_2 are such that $P_{i_1,i_2}(\alpha, \alpha) \neq 0$, then

$$i_1 \log q_1 + i_2 \log q_2 > \frac{n_1 \log q_1 + n_2 \log q_2}{\sqrt{2d} + \frac{\varepsilon}{5}}$$

hence

$$\begin{aligned}\left|\alpha - \frac{p_1}{q_1}\right|^{i_1} \left|\alpha - \frac{p_2}{q_2}\right|^{i_2} &< \exp\left(-(\sqrt{2d} + \varepsilon) \cdot \frac{n_1 \log q_1 + n_2 \log q_2}{\sqrt{2d} + \frac{\varepsilon}{5}}\right) \\ &< (q_1^{n_1} q_2^{n_2})^{-\frac{\sqrt{2d}+\varepsilon}{\sqrt{2d}+\frac{\varepsilon}{5}}}\end{aligned}$$

The exponent is smaller than -1!

Now estimate the coefficients:

$$\begin{aligned}\tilde{P}_{i_1,i_2}(\alpha, \alpha) &\leq (n_1 + 1)(n_2 + 1)(8H(\alpha))^{\delta^{-1}(n_1+n_2)} \cdot \max(1, |\alpha|)^{n_1+n_2} \\ &< C_1(\alpha, \varepsilon)^{n_1+n_2}\end{aligned}$$

and

$$\begin{aligned}\tilde{P}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) &\leq (n_1 + 1)(n_2 + 1) C_1(\alpha, \varepsilon)^{n_1+n_2} \cdot (q_1^{n_1} q_2^{n_2})^{-\frac{\sqrt{2d}+\varepsilon}{\sqrt{2d}+\frac{\varepsilon}{5}}} \\ &\leq (2C_1(\alpha, \varepsilon))^{n_1+n_2} \cdot (q_1^{n_1} q_2^{n_2})^{-\frac{\sqrt{2d}+\varepsilon}{\sqrt{2d}+\frac{\varepsilon}{5}}} \\ &< (q_1^{n_1} q_2^{n_2})^{-1}\end{aligned}$$

Contradiction. □

Theorem (Gelfond-Schneider)**.** Let λ_1, λ_2 be logarithms of non-zero algebraic numbers. Then λ_1, λ_2 are linearly independent over $\overline{\mathbb{Q}}$ if and only if they are linearly independent over $\mathbb{Q}$.

We will prove this by assuming $\frac{\lambda_1}{\lambda_2} \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$, and then showing that a particular determinant is both equal to zero and not equal to zero, hence getting a contradiction.

Before doing this, we will discuss how the previous proof could have been instead been phrased using determinants.

We considered some functions $\varphi_1, \ldots, \varphi_L$ which were some enumeration of $X_1^{j_1} X_2^{j_2}$. Then we used Siegel's lemma to find $a_1, \ldots, a_L$ such that $D = a_1\varphi_1 + \cdots + a_L\varphi_L$ vanishes at $u_1 = (\alpha, \alpha)$ to some order. (Note that P also vanishes at all Galois-conjugates of (α, α): $u_2, \ldots, u_d$). Then we find an argument to show that P also vanishes at $u_{d+1} = \left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)$ to some order.

This means that for $i = 1, \dots, L$ there exists $k(i) \in \{1, \dots, d+1\}$ and some partial differentiation operator ∂_i such that $\partial_i P(u_{k(i)}) = 0$. We also showed that P with so much vanishing cannot exist.

Let:

$$M = \begin{pmatrix} \partial_1\varphi_1(u_{k(1)}) & \cdots & \partial_L\varphi_1(u_{k(l)}) \\ \vdots & \ddots & \vdots \\ \partial_1\varphi_L(u_{k(1)}) & \cdots & \partial_L\varphi_L(u_{k(L)}) \end{pmatrix}$$

Then P having all that vanishing is equivalent to

$$(a_1, \dots, a_L)M = (0, \dots, 0).$$

Lecture 15 Now the existence of P is equivalent to $\det M = 0$.

Let $\lambda_1, \lambda_2 \in \mathbb{R}_{\neq 0}$, and $\alpha_1 = e^{\lambda_1}$, $\alpha_2 = e^{\lambda_2} \in \overline{\mathbb{Q}}$. Let $\beta = \frac{\lambda_2}{\lambda_1} \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$. So we assumed that Gelfond-Schneider is false. We aim for a contradiction.

Let $T_0, T_1, S \in \mathbb{Z}_{>0}$ with

$$L := (T_0 + 1)(2T_1 + 1) = (2S + 1)^2.$$

Consider the "monomials"

$$X^\tau \exp(t\lambda_1 X)$$

for $\tau = 0, \dots, T_0$, $t = -T_1, \dots, T_1$ and the points $s_1 + \beta s_2$ for $s_1, s_2 = -s, \dots, s$.

Notation. $[-]_{\substack{\tau,t \\ s_1,s_2}}$ means a matrix with rows indexed by τ, t and columns indexed by s_1, s_2.

Let

$$\begin{aligned} \Delta &= \det[(s_1 + \beta s_2)^\tau \cdot \exp(t\lambda_1(s_1 + \beta s_2))]_{\substack{\tau,t \\ s_1,s_2}} \\ &= \det[(s_1 + \beta s_2)^\tau \alpha_1^{ts_1} \alpha_2^{ts_2}]_{\substack{\tau,t \\ s_1,s_2}} \end{aligned}$$

Steps:

(1) Give an analytic upper bound on Δ

(2) Give an arithmetic lower bound on Δ

(3) "zero estimate" $\implies \Delta \neq 0$.

Steps (1) and (2) will be done in such a way that together they will give $\Delta = 0$. Then this will contradict (3).

We will alternate between viewing $(s_1 + \beta s_2)^\tau \cdot \exp(t\lambda_1(s_1 + \beta s_2))$ as a function of a single variable (function of $s_1 + \beta s_2$) and thinking of it as a function of two variables (function of s_1 and s_2).

Upper bound

Proposition. For $n \in \mathbb{Z}_{>0}$, there exists $c = c(n) > 0$ such that the following holds:
Let $L \in \mathbb{Z}_{>0}$, $E \in \mathbb{R}_{>1}$. Let $f_1, \dots, f_L : \mathbb{C}^n \to \mathbb{C}$ be *analytic* functions (here, analytic means convergent power series on $\mathbb{C}^n$). Let $\xi_1, \dots, \xi_L \in \mathbb{C}^n$. Let $r = \max_{\substack{s=1,\dots,L \\ j=1,\dots,n}} |\xi_{s,j}|$. Then

$$\det[f_t(\xi_s)]_{\substack{t=1,\dots,L \\ s=1,\dots,L}} \leq E^{-cL^{1+\frac{1}{n}}} \cdot L! \cdot \prod_{t=1}^{L} |f|_{Er}.$$

Notation. $|f|_R = \max_{|x_1|,\dots,|x_n| \leq R} |f(x_1, \dots, x_n)|$.

Corollary. With Δ, T_0, T_1, S, L as above, there exists $c, C > 0$ depending only on β, z, such that for all $E \in \mathbb{R}_{\geq e}$:

$$|\Delta| \leq \exp(-cL^2 \log E + CL \cdot T_0 \log(ES) + CLT_1 ES).$$

Proof. We take $n = 1$ and some $E \geq e$. We have $|s_1 + \beta s_2| < C_0 \cdot S$ with $C_0 = C_0(\beta)$.

$$|z^\tau \exp(t\lambda_1 z)| < \exp(C_1 T_0 \cdot \log ES + C_1 T_1 ES)$$

for $|z| < E \cdot C_0 \cdot S$, with $C_1 = C_1(\beta, \lambda_1)$. □

One possible choice of the parameters: $E = e$. $S \sim L^{\frac{1}{2}}$, $T_0 \sim L^{1-\varepsilon}$, $T_1 \sim L^{\varepsilon}$. In this case:

$$|\Delta| = \exp(-cL^2).$$

(for large L).

Lemma (Schwart's Lemma)**.** Let f be a holomorphic function on D_R the disc of radius R with a zero of order k at 0. Then: for all $z \in D_R$:

$$|f(z)| \leq \frac{|z|^K \cdot |f|_R}{R^K}.$$

Proof. The maximum modulus principle for $\frac{f(z)}{z^K}$. □

[Proof of Proposition.] We apply Schwart's Lemma for

$$f(z) = \det[f_t(z \cdot \xi_s)]_s^t$$

and $R = E$. Note: $|F|_E \leq L! \cdot \prod_{t=1}^{T} |f_t|_{Er}$.

So the proposition follows if we show that F vanishes to order $cL^{1+\frac{1}{n}}$ at 0. We prove this. Enough to do it when each f_t is of the form $z_1^{a_1} \cdots z_n^{a_n}$ for some $a_1, \ldots, a_n \in \mathbb{Z}$ depending on t.

This is because all f_ts are infinite linear combinations of such f_ts, and hence the determinant can be written as an infinite combination of special determinants. Furthermore we may assume that the $(a_1, \ldots, a_n)$ are distinct for different ts.

Observe: $\det[f_t(z \cdot \xi_s)]_{\substack{t \\ s}} = z^{\sum \deg f_t} \cdot \det[f_t(\xi_s)]_{\substack{t \\ s}}$ if each f_t is of the special form.

The number of monomials with degree $\leq d$ is at most d^n. We take $d = \left\lfloor \left(\frac{L}{2}\right)^{\frac{1}{n}} \right\rfloor$. Then at least half of the f_ts have degree $\geq d$. So $\sum \deg f_t \geq \left(\frac{L}{2}\right) \cdot d \geq c \cdot L^{1+\frac{1}{n}}$. □

Lecture 16

Proposition (1)**.** Let $S = (T_0 + 1)T_1$ be non-negative integers. Let $w_1, \ldots, w_{T_1}$ and $\xi_1, \ldots, \xi_S$ be two sets of distinct real numbers.
Then

$$\det[\xi_S^\tau \exp(w_t \xi_S)]_{\substack{\tau,t \\ s}} \neq 0,$$

with: $\tau = 0, \ldots, T_0$, $t = 1, \ldots, T_1$, $s = 0, \ldots, S$.

alternant / interpolation determinant

Proposition (2)**.** Let $T \in \mathbb{Z}_{\geq 1}$, let $w_1, \ldots, w_T$ be distinct real numbers. Let $P_1, \ldots, P_T \in \mathbb{R}[X]$ be non-zero. Then the function

$$F(x) = P_1(x)e^{w_1 x} + \cdots + P_T(x)e^{w_T x}$$

has at most $\deg P_1 + \cdots + \deg P_T + T - 1$ real zeroes counting multiplicities.

Proposition (2) $\implies$ *Proposition (1).* Suppose to the contrary that det $= 0$. Then there exists $a_{\tau,t} \in \mathbb{R}$ not all 0 such that

$$\sum a_{\tau,t} x^\tau \exp(w_t x)$$

vanishes for all $x = \xi_1, \ldots, \xi_S$. This is a function of the type in Proposition (2). Each polynomil is of degree $\leq T_0$, and there are T_1 many of them, so there can be no more than $T_0 \cdot T_1 + T_1 - 1 < S$ zeroes. □

Lemma 1.13. Let f be a C^∞ function on $\mathbb{R}$ with N real zeroes. Then f' has at least $N - 1$ zeroes.

Corollary of Rolle's Theorem.

Proof of Proposition (2). By induction on $N := \deg P_1 + \cdots + \deg P_T + T - 1$. If $N = 0$, then $T = 1$ and $\deg P_1 = 0$. So $F(x) = a \cdot \exp(w_1 x)$ for some $a \neq 0$. This indeed has no zeroes.

Suppose $N > 0$ and the claim holds for $N - 1$.

We assume as we may that $w_1 = 0$ (if not, then replace w_j by $w_j - w_1$, which has the effect of replacing F by $F \cdot e^{-w_1 \cdot x}$).

Then by the lemma, F has at most one more zero than

$$F' = \underbrace{P_1(x)'}_{\deg P_1 - 1} + \underbrace{(P_2'(x) + P_2(x) w_w) e^{w_2 x}}_{\deg P_2} + \cdots$$

By the induction hypothesis, F' has at most $N - 1$ zeroes, so F has at most N zeroes. □

Now we return to proving Gelfond-Schneider.

Let $z_1, z_2 \in \mathbb{R}_{\neq 0}$ such that $\alpha_j = e^{\lambda_j} \in \overline{\mathbb{Q}}$ for $j = 1, 2$.

We aim for a contradiction. We have integers L, T_0, T_1, S such that

$$L = (T_0 + 1)(2T_1 + 1) = (2S + 1)^2.$$

Let

$$\Delta = \det[(s_1 + \beta s_2)^\tau \exp(\lambda_1 t (s_1 + \beta s_2))]_{\substack{\tau, t \\ s_1, s_2}}.$$

Last time:

$$\log |\Delta| \leq -cL^2 \log E + CLT_0 \log ES + CLT_1 ES$$

where $E \in \mathbb{R}_{>1}$ arbitrary.

Apply Proposition (1) with $\xi_S = (s_1 + \beta s_2)$ with some enumeration of s_1, s_2 and $w_t = \lambda_1 t$. Then $\Delta \neq 0$.

Recall:

$$\Delta = \det[(s_1 + \beta s_2)^\tau \alpha_1^{t s_1} \alpha_2^{t s_2}]_{\substack{\tau, t \\ s_1, s_2}}.$$

Then

$$\Delta = P(\beta, \alpha_1, \alpha_2)$$

for some $P \in \mathbb{Z}[X, Y, Z]$. So:

$$H(\Delta) \leq \mathcal{L}(P) \cdot H(\beta)^{T_0 \cdot L} \cdot H(\alpha_1)^{T_1 S \cdot L} H(\alpha_2)^{T_1 S L}$$

using

$$\mathcal{L}(P_1, P_2) \leq \mathcal{L}(P_1) \mathcal{L}(P_2)$$

and

$$\mathcal{L}\left(\sum P_j\right) \leq \sum \mathcal{L}(P_j)$$

we get

$$\mathcal{L}(P) \leq L! \cdot (2S)^{T_0 L}.$$

Liouville bound:

$$\log|\Delta| > -C(\log L! + T_0 L \log S).$$

Take: $E = 10$.

Then we have a contradiction if

$$-cL^2 + CLT_0 \log S + CLT_1 S < -C(L \cdot \log L + T_0 L \log S + T_1 LS).$$

I want:

$$L^2 > C(T_0 L \log S + LT_1 S).$$

Lecture 17 Take: $S \approx L^{\frac{1}{2}}$, $T_0 \approx L^{1-\varepsilon}$, $T_1 \approx L^{\varepsilon}$.

Theorem (Nesterenko)**.** Let $T_0, T_1, N, M \in \mathbb{Z}_{>0}$. Let $\Sigma_1, \Sigma_2 \subset \mathbb{C}^2$ such that $|\Sigma_1| = N$, $|\Sigma_2| = M$, and the exponentials of the second coordinates of Σ_1 and the first coordinates of Σ_2 are distinct. Let $P \in \mathbb{C}[X, Y]$ of degree $\leq T_0$ in X, and $\leq T_1$ in Y. Suppose that $P(X, \exp(y))$ vanishes on $\Sigma_1 + \Sigma_2$. Then

$$N \leq T_1 \qquad \text{or} \qquad M \leq T_0(T_1 + 1).$$

Proof. If $P(X, Y) = \tilde{P}(X, Y) \cdot Y$, then $P(X, \exp(y))$ vanishes at exactly the same places as $\tilde{P}(X, \exp(y))$. So we may assume $Y \nmid P(X, Y)$. Suppose that $N > T_1$, and write $\Sigma_1 = \{(\xi_1, \eta_1), \ldots, (\xi_N, \eta_N)\}$. Then $P(\xi_j + X, \exp(\eta_j + y))$ vanishes on Σ_2 for all $j = 1, \ldots, N$. We write $P(X, Y) = R_1(X)Y^{k_1} + \cdots + R_K Y^{k_K}$ with $0 = k_1 < k_2 < \cdots < k_K \leq T_1$.

Then

$$P(\xi_j + X, \exp(\eta_j + y)) = R_1(\xi_j + X) \cdot \exp(\eta_j)^{k_1} \cdot \exp(y)^{k_1} + \cdots .$$

Write

$$Q_{i,j}(X) = R_i(\xi_j + X) \exp(\eta_j)^{k_i}.$$

Then

$$P(\xi_j + X, \exp(\eta_j + y)) = \sum_{i=1}^{n} Q_{i,j}(X)(\exp(y))^{k_i}.$$

I look for polynomials $A_1, \ldots, A_k \in \mathbb{C}[X]$ such that

$$\sum_{j=1}^{K} A_j(X) P(\xi_j + X, \exp(\eta_j + y)) = B(X) \in \mathbb{C}[X] \tag{$*$}$$

such that $\deg B \leq T_0(T_1 + 1)$, and then since B vanishes at the first coordinates of Σ_2, $M \leq T_0(T_1 + 1)$ will follow. □

Lemma. Let $Q_{ij} \in \mathbb{C}[X]$ for $i, j = 1, \dots, K$ for some $K \in \mathbb{Z}_{>0}$. Then there exists $A_1, \dots, A_k \in \mathbb{C}[X]$ such that

$$\sum_i A_i Q_{ij} = \begin{cases} \det[Q_{ij}] & \text{if } j = 0 \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let $[\tilde{Q}_{ij}]$ be the adjugate of $[Q_{ij}]$. Then

$$[\tilde{Q}_{ij}] \cdot [Q_{jk}] = \det[Q_{jk}] \cdot \mathrm{id}\,.$$

Let $A_1, \dots, A_k$ be the first row of $[\tilde{Q}_{ij}]$. □

$$\begin{pmatrix} Q_{11}(X) & \cdots & Q_{1k}(X) \\ \vdots & \ddots & \vdots \\ Q_{k1}(X) & \cdots & Q_{kk}(X) \end{pmatrix} \begin{pmatrix} \exp(y)^{e_1} \\ \vdots \\ \exp(y)^{e_k} \end{pmatrix} = \begin{pmatrix} P(\xi_1 + X, \exp(\eta_i + y)) \\ \vdots \\ P(\xi_k + X, \exp(\eta_k + y)) \end{pmatrix}$$

Premultiply this by the row vector $(A_1(X), \dots, A_k(X))$. We get $(*)$ with $B = \det[Q_{ij}$.

$\deg B \le T_0 K \le T_0(T_1 + 1)$.

We need to make sure that $B \neq 0$

The leading term of Q_{ij} is $a_i \cdot \exp(\eta_j)^{k_i} \cdot X^{\deg R_i}$, where a_i is the leading coefficient of R_i.

To show $B \neq 0$, we will consider the leading term of B:

$$\det[a_i \exp(\eta_j)^{k_i} X^{\deg R_i}]_{ij} = \det[\exp(\eta_j)^{k_i}]_{ij} X^{\sum \deg R_i} \prod a_i.$$

Lemma. Let $K \in \mathbb{Z}_{\ge 1}$, wne let $0 = k_1 < \cdots < k_K \in \mathbb{Z}$. Let $A \subset \mathbb{C}$ such that $|\{\exp(\eta) : \eta \in A\}| > k_K$. Then there exists a choice of $\eta_1, \dots, \eta_K \in A$ such that

$$\det[\exp(\eta_i)^{k_j}] \neq 0.$$

Proof. By induction on K. $K = 1$ is true.

Suppose $K > 1$, and the claim holds for $K - 1$. Consider the determinant:

$$\begin{vmatrix} \exp(\eta_1)^{k_1} & \cdots & \exp(\eta_K)^{k_K} \\ \vdots & \ddots & \vdots \\ \exp(\eta_{K-1})^{k_1} & \cdots & \exp(\eta_{K-1}^{k_K} \\ z^{k_1} & \cdots & z^{k_K} \end{vmatrix} = D(z)$$

which has the property that the upper left $(K-1) \times (K-1)$ minor is $\neq 0$.

Now D is a polynomial which is $\neq 0$ of degree k_K, so it has at most k_K many 0s. Choose η_K such that $\exp(\eta_K)$ is not one of them. $\square$

Lecture 18

Theorem. Let $d \geq 3$. Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree d without repeated factors. Let $G(X,Y) \in \mathbb{Z}[X,Y]$ be of degree $\leq d-1$. Assume $F-G$ is irreducible. Then

$$F(X,Y) = G(X,Y) \qquad X,Y \in \mathbb{Z}$$

has at most finitely many solutions.

Schinzel proved this only assuming that $F \neq aQ^n$ for some irreducible Q of degree ≤ 2. He used Siegel's theorem on integral points. If an algebraic curve has infinitely many points, then it has genus D and at most 2 points at infinity. Our proof is based on an argument of Corvaja and Zannier for proving Siegel's theorem.

Subspace theorem: Let V be a vector space of dimension n over $\overline{\mathbb{Q}}$. Let $e_1^{(0)}, \ldots, e_n^{(0)}$ and $e_1, \ldots, e_n$ be two bases of V. Then for all $\varepsilon > 0$, there exists a finite number of elements $f_1, \ldots, f_m \in V$ such that all $\varphi \in V^*$ that solves:

$$\prod_{i=1}^{n} |\varphi(e_i)| \leq H(\varphi(e_1^{(0)}), \ldots, \varphi(e_n^{(0)}))^{-\varepsilon} \tag{$*$}$$

with $\varphi(e_i^{(0)}) \in \mathbb{Z}$ for all $i = 1, \ldots, n$, φ satisfies $\varphi(f_j) = 0$ for some $j \in \{1, \ldots, n\}$.

$\exists \alpha_{i,j} \in \overline{\mathbb{Q}}$ such that

$$e_i = \sum_j \alpha_{ij} e_j^{(0)}$$

and $L_i = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n$. φ satisfies $(*)$ if and only if

$$(x_1, \ldots, x_n) = (\varphi(e_1^{(0)}), \ldots, \varphi(e_n^{(0)})) \in \mathbb{Z}^n$$

satisfies

$$\prod_{i=1}^{\infty} |L_i(x_1, \ldots, x_n)| < H(x_1, \ldots, x_n)^{-\varepsilon}.$$

Let F, G be as in the theorem, and write $P = F - G$.

We assume that $Y \nmid F$.

Then there exists $\alpha_1, \ldots, \alpha_d \in \overline{\mathbb{Q}}$ distinct such that

$$F(X,Y) = (X - \alpha_1 Y) \cdots (X - \alpha_d Y).$$

Write Γ for the set of $(x,y) \in \mathbb{C}^2$ with $P(x,y) = 0$. Then for $(x,y) \in \Gamma$ we have

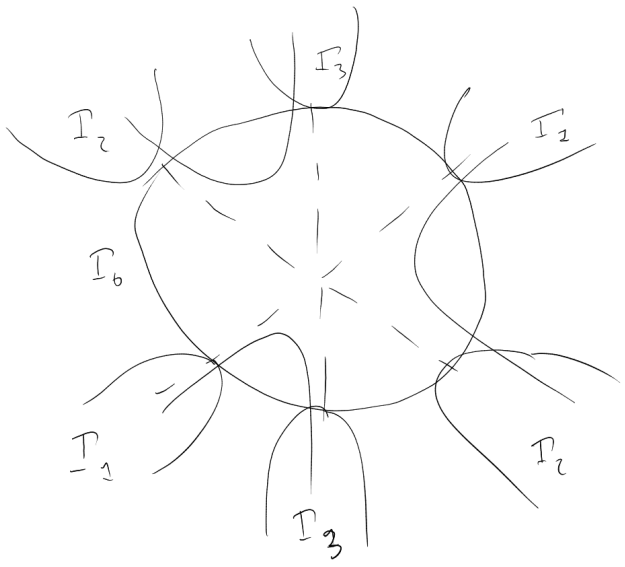
$$F(x,y) \leq C(|x| + |y|)^{d-1}.$$

By a similar argument to the lemma for Thue's equation, for all $\varepsilon > 0$ there exists $R = R(P, \varepsilon)$ such that $(x, y) \in \Gamma$ with $|x| + |y| > R$, then $\left|\frac{x}{y} - \alpha_j\right| < \varepsilon$ for some j.

We pick a small $\varepsilon > 0$, in particular $|\alpha_i - \alpha_j| > 2\varepsilon$ for $i \neq j$. We define

$$\Gamma_0 = \{(x, y) \in \Gamma : |x| + |y| < R\}$$
$$\Gamma_j = \left\{(x, y) \in \Gamma : |x| + |y| \geq R, \left|\frac{x}{y} - \alpha_j\right| < \varepsilon\right\}$$

for $j = 1, \ldots, d$.



Γ_0 is bounded so only has finitely many integer points. We want to show this also for $\Gamma_1, \ldots, \Gamma_j$. Write $I = P\mathbb{Q}[X, Y]$ for the ideal generated by P. Take some $D \in \mathbb{Z}_{\geq 1}$ and large enough. Write $\overline{\mathbb{Q}}[X, Y]^{(D)}$ for polynomials of degree $\leq D$. We will apply the subspace theorem in the vector space

$$V = \overline{\mathbb{Q}}[X, Y]^{(D)}/(I \cap \overline{\mathbb{Q}}[I, Y]^{(D)}).$$

Elements $f \in V$ can be evaluated on Γ.

In particular, for $(x, y) \in \Gamma$, the map $f \mapsto f(x, y)$ is an element of V^*. Reference basis: the monomials $X^k Y^m$ for $k + m \leq D$ span V. Pick a linearly independent family for $e_1^{(0)}, \ldots, e_n^{(0)}$, where $n = \dim V$.

If $(x, y) \in \Gamma \cap \mathbb{Z}^2$, then $e_i^{(0)}(x, y) \in \mathbb{Z}$. Also,

$$H(e_1^{(0)}(x, y), \ldots, e_n^{(0)}(x, y)) < C|y|^D.$$

We need to find some l_j's that decay on a fixed Γ_i.

For $j = 1, \ldots, d$ we introduce a symbol p_j and call these the "points of Γ at infinity". We define for $f \in V$:

$$\operatorname{ord}_{p_j}(f) = \sup\{m \in \mathbb{Z} : f(x,y) \cdot y^m \text{ is bounded on } \Gamma_j\}.$$

Note $\operatorname{ord}_{p_j}(f) \geq -D$.

Lemma. Let $f \in V$ and let $j \in \{1, \ldots, d\}$. If $\operatorname{ord}_{p_j}(f) < \infty$, then the limit

$$\lim_{(x,y)\in\Gamma, |y|\to\infty} f(x,y) y^{\operatorname{ord}_{p_j}(f)}$$

exists and $\neq 0$. In addition, we have

$$\lim_{(x,y)\in\Gamma_j, |y|\to\infty} (X - \alpha Y) Y^{-1} = \alpha_j - \alpha$$

for all $\alpha \in \overline{\mathbb{Q}}$.

Can be proved that $\operatorname{ord}_{p_j}(f) = \infty$ if and only if $f = 0$.

Lecture 19 $\frac{Z}{Y}$ is a local uniformiser at $p_j = (\alpha_i, 1, 0)$.

Proof. Let $j = 1$, and by taking the substitution $X - \alpha_1 Y \mapsto X$, we may assume $\alpha_1 = 0$.

First, we show X is bounded on Γ_1. To this end:

$$X = \frac{G(X,Y)}{a(X - \alpha_2 Y) \cdots (X - \alpha_d Y)}.$$

Note

$$a(X - \alpha_2 Y) \cdots (X - \alpha_d Y) \geq c Y^{d-1}$$

on Γ, with some $c = c(P) > 0$. We may write $P = 0$ as:

$$aXY^{d-1} + bY^{d-1} + \tilde{P}(X,Y)$$

($\tilde{P}$ of degree $\leq d - 2$ in Y). a is not the same as in the factorisation of F and $a \neq 0$, but b may be 0.

This gives:

$$X = \frac{-b}{a} + Y^{-1} \cdot \underbrace{Q(X, Y^{-1})}_{\text{bounded}}. \tag{$**$}$$

For some polynomial Q. Then $\lim X = \frac{-b}{a}$ on Γ.

Proving the first claim, suppose we can write

$$f(X,Y) = R_1(X) Y^k + R_2(X) Y^{k-1} + \cdots. \tag{$**$}$$

Here, negative exponents of Y are allowed, but the sum must be finite. You can always do this with $k = D$ if $R_1\left(-\frac{b}{a}\right) \neq 0$. Then $f(X,Y) \cdot Y^{-k} \to R\left(-\frac{b}{a}\right) \neq 0$ and $\operatorname{ord}_{p_1}(f) = -k$ and the claim holds.

If $R_1\left(-\frac{b}{a}\right) = 0$, then use $(**)$ to write $(**)$ with k replaced by $k-1$.

Iterate this. □

Lemma. For each $j = 1, \dots, d$, there is a basis $l_1, \dots, l_n$ $(n = \dim V)$ of V such that

$$\operatorname{ord}_{p_j}(l_i) \leq -D + i - 1.$$

Proof. By induction, we show that there $l_1, \dots, l_{i-1}$ and $V_i \subset V$ such that

$$\begin{aligned} V &= \overline{Qbb}l_1 \oplus \cdots \oplus \overline{\mathbb{Q}}l_{i-1} \oplus V_i \\ \operatorname{ord}_j(l_k) &\leq -D + k - 1 && \text{for } k = 1, \dots, i-1 \\ \operatorname{ord}_{p_j}(f) &\leq -D + i && \text{for } f \in V_i \end{aligned}$$

$i = 1$ is trivial: $V = V_1$.

So suppose $i > 1$ and the claim holds for $i-1$. We define: l_{i-1} to be an element in V_{i-1} of minimal order at p_j. Let $V_i = \{f \in V_{i-1} : \operatorname{ord}_{p_j}(f) > \operatorname{ord}_{p_j}(l_{i-1})\}$.

Just need to show: $V_{i-1} = l_{i-1}\overline{\mathbb{Q}} \oplus V_i$. To this end, let $g \in V_{i-1}$. Write $m = \operatorname{ord}_{p_j}(l_{i-1})$. Then

$$\lim_{\Gamma_j} g \cdot V_1^m =: b < \infty.$$

$$f = g - \frac{b}{\lim_{\Gamma_j} l_{i-1} \cdot Y^m} l_{i-1}.$$

Then

$$\lim_{\Gamma_j} fY^m = 0$$

so by the previous lemma, $\operatorname{ord}_{p_j} f > m$. So $f \in V_i$. □

For this to be useful, we need n to be large. (We need $n \geq 2D + 2$).

Lemma.

$$\dim V \geq dD - d(d-1).$$

Remark. Thinking about Γ as a projective curve, V is the space of rational functions with poles of order at most D at each point at ∞. By Riemann-Roch: $\dim V = dD - g + 1$, provided D is large enough.

Proof. Let $R(X, Y) = \prod(X - \alpha_j Y)$ $(= \frac{F(X,Y)}{a})$. The point is that the polynomials

$$Q_{jl}(X, Y) = \frac{R(X, Y)}{X - \alpha_j Y} \cdot Y^l \in V$$

are linearly independent in V. $j = 1, \dots, d$, $l = 1, \dots, D - d + 1$. Suppose $Q = \sum_{j,l} \beta_{jl} Q_{jl}$ for some $\beta_{jl} \in \overline{\mathbb{Q}}$ not all 0.

Want to show $Q \neq 0$. To that end, let $\beta_{j',l'} \neq 0$ such that l' is maximal with this property.

We can show that:

$$\lim_{\Gamma_j} Q(X,Y) \cdot Y^{-l'-d+1} = \beta_{j',l'} \prod_{i=\{1,\dots,d\}\setminus\{j'\}} (\alpha'_j - \alpha_i).$$

Uses the first lemma today. □

Lecture 20

Lemma. Let $f, P \in \mathbb{Z}[X,Y]$ without common factors in $\mathbb{Z}[X,Y]$. Then the system of equations $f(X,Y) = P(X,Y) = 0$ has only finitely many solutions.

Proof. $\mathbb{Z}[X,Y] \cong \mathbb{Z}[X][Y]$ (poynomials in Y with coefficients in $\mathbb{Z}[X]$). f, P have no common factors in $\mathbb{Z}[X][Y]$. Then Gauss's lemma gives us that they have no common factors in $\mathbb{Q}(X)[Y]$. This is because $\mathbb{Z}[X]$ is a UFD and $\mathbb{Q}(X)$ is its quotient field.

Since $\mathbb{Q}(X)[Y]$ is a Euclidean domain, there exists $F, G \in \mathbb{Q}(X)$ such that

$$F \cdot P + G \cdot f = 1.$$

Multiply by the common denominator D of F, G, and we get

$$\tilde{F} \cdot P + \tilde{G} \cdot f = D(X)$$

for some $\tilde{F}, \tilde{G} \in \mathbb{Z}[X]$. Hence the common solutions of $f = P = 0$ has finitely many X-coordinates. Then swap X and Y. □

Theorem. Let $F \in \mathbb{Z}[X,Y]$ homogeneous of degree d, without repeated factors. Let $G \in \mathbb{Z}[X,Y]$ of degree $< d$. Assume $F - G$ is irreducible in $\mathbb{Z}[X,Y]$. Then there are at most finitely many solutions of $F(X,Y) = G(X,Y)$ with $X, Y \in \mathbb{Z}$.

$F(X,Y) = (X - \alpha_1 Y) \cdots (X - \alpha_d Y)$. $\Gamma = \Gamma_0 \cup \Gamma_1 \cup \cdots \cup \Gamma_d$. $P = F - G$, $I = P \cdot \overline{\mathbb{Q}}[X,Y]$. $V = \overline{\mathbb{Q}}[X,Y]^{(D)} / I \cap \overline{\mathbb{Q}}[X,Y]^{(D)}$. $\operatorname{ord}_{p_j}(f) = \sup(t \in \mathbb{Z} : f(X,Y) \cdot Y^t$ bounded on $\Gamma_j)$. $n = \dim V$.

$\forall j \exists l_1, \dots, l_n \in V$ a basis such that $\operatorname{ord}_{p_j}(l_i) \geq -D + i - 1$. $n = \dim V > dD - d(d-1)$.

Subspace Theorem: Let V be a vector space of dimension n over $\overline{\mathbb{Q}}$. Let $l_1, \dots, l_n$, $l_1^{(0)}, \dots, l_n^{(0)} \in V$ be two bases. $\forall \varepsilon > 0$ there exists $f_1, \dots, f_m \in V_{\neq 0}$ such that $\forall \varphi \in V^*$ that satisfies

$$\prod_{i=1}^{n} |\varphi(l_j)| \leq H(\underbrace{\varphi(l_1^{(0)})}_{\in \mathbb{Z}}, \dots, \underbrace{\varphi(l_n^{(0)})}_{\in \mathbb{Z}})^{-\varepsilon}$$

then $\varphi(f_j) = 0$ for some $j = 1, \dots, m$.

Proof of Schinzel's Theorem. We show that $\mathbb{Z}^2 \cap \Gamma_j$ is finite for any $j = 1, \dots, d$. Let $l_1, \dots, l_n \in V$ be a basis with $\operatorname{ord}(l_i) \geq -D + i - 1$. Then

$$\begin{aligned} \prod_{i=1}^{n} |l_i(X,Y)| &\leq C \cdot Y^{\sum - \operatorname{ord}_{p_j}(l_i)} && \text{on } \Gamma_j \\ &\leq C \cdot Y^{-D-1} \end{aligned}$$

if $n \geq 2D + 2$. We set D to be large enough so that this holds.

Recall the reference basis $l_j^{(0)}$ are suitable monomials of degree $\leq D$, so

$$|l_i^{(0)}(X,Y)| < CY^D.$$

Then for $x, y \in \mathbb{Z}^2 \cap \Gamma_j$, we have:

$$H(l_i^{(0)}(x,y), \dots, l_n^{(0)}(x,y)) \leq C \cdot |Y|^D.$$

Hence

$$\prod_{i=1}^{n} |l_i(x,y)| < H(l_1^{(0)}(x,y), \dots)^{-1}$$

provided y is still large.

By the subspace theorem, $f_i(x,y) = 0$ for some $i = 1, \dots, m$.

To apply the lemma, we need $f_i \in \mathbb{Z}[X,Y]$. This can be assumed: indeed, multiplying f_i by an element of $\overline{\mathbb{Q}}$, we can make the leading coefficient to be in $\mathbb{Z}$, and all other coefficients will be algebraic integers. Then replace f_i by the sum of its Galois conjugates. □

Theorem. For $q \in \mathbb{Z}_{>0}$ with $\gcd(q, 6) = 1$, we write $\operatorname{ord}(q)$ for the order of the multiplicative group generated by $2, 3$ in $\mathbb{Z}/q\mathbb{Z}$.
Then:

$$\lim_{q \to \infty} \frac{\operatorname{ord}(q)}{(\log q)^2} = \infty.$$

Remark. $2^n 3^m$ for $n < \frac{1}{2} \log_2 q$, $m < \frac{1}{2} \log_3 q$. Hence

$$\operatorname{ord}(q) \geq \left(\frac{1}{2} \log_2 q\right) \left(\frac{1}{2} \log_3 q\right).$$

Theorem (Corvaja, Zannier; Hernández, Luca)**.** Write $\mathcal{S} = \{2^n 3^m : n, m \in \mathbb{Z}_{\geq 0}\}$. Then for all $\varepsilon > 0$, there are only finitely many pairs of multiplicatively independent $a, b \in \mathcal{S}$ such that

$$\gcd(a - 1, b - 1) \geq \max(a, b)^{\varepsilon}.$$

a, b are multiplicatively independent if there does not exist $n, m \in \mathbb{Z}$ such that $a^n = b^m$.

Fact: there exist infinitely many n such that

$$\gcd(2^n - 1, 3^n - 1) \geq 3^{n^{c/\log\log n}}.$$

Lecture 21

Theorem (1)**.** If $2, 3 \nmid q$, then

$$\frac{\operatorname{ord}(q)}{(\log q)^2} \to \infty.$$

Theorem (2)**.** For all $\varepsilon > 0$, there are only finitely many pairs of multiplicatively independent $a, b \in \mathcal{S}$ such that

$$\gcd(a - 1, b - 1) > \max(a, b)^{\varepsilon}.$$

Proof of Theorem 1 using Theorem 2. Let

$$\Lambda = \{(n, k) \in \mathbb{Z}^2 : 2^n \cdot 2^k \equiv \pmod q\}.$$

This is a subgroup of $\mathbb{Z}^2$, and $|\mathbb{Z}^2/\Lambda = \operatorname{ord}(q)$ The volume $\mathbb{R}^2/\Lambda$ is $\operatorname{ord}(q)$.

Our aim is to find $(n_1, k_1), (n_2, k_2) \in \Lambda \cap \mathbb{Z}^2_{\geq 0}$ linearly idnependent and $n_1, k_1, n_2, k_2 \leq C \operatorname{ord}(q)/\log q$, where C is absolute.

If we can do this, then: $q \mid \gcd(2^{n_1}2^{k_1} - 2^{n_2}3^{k_2} - 1)$. By Theorem (2), since $2^{n_1}3^{k_1} - 1$ and $2^{n_2}3^{k_2} - 1$ are multiplicatively independent, we would get

$$\begin{aligned} q &< \max(2^{n_1}3^{k_1}, 2^{n_2}3^{k_2})^{\varepsilon} \\ &< \exp(C \operatorname{ord}(q)/\log q)^{\varepsilon} \end{aligned}$$

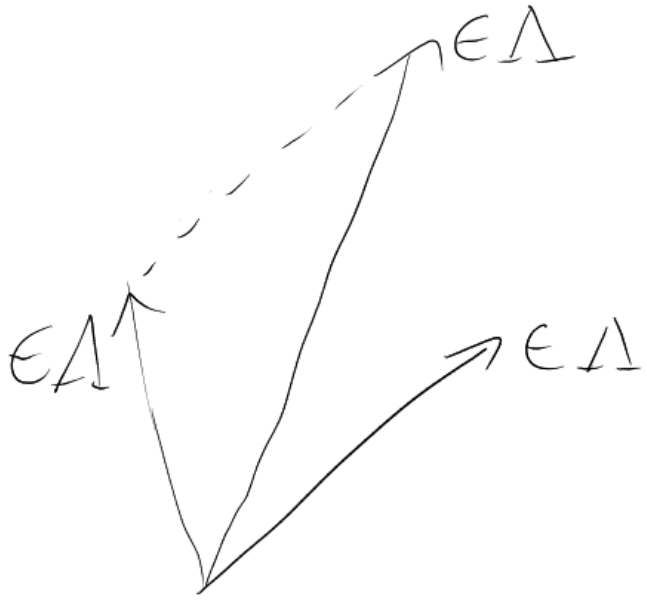
Taking log:

$$\begin{aligned} \log q &< C \cdot \varepsilon \operatorname{ord}(q)/\log q \\ \operatorname{ord}(q) &> C^{-1} \cdot \varepsilon^{-1} \cdot (\log q)^2 \end{aligned}$$

provided q is sufficiently large in terms of ε.

Now to the proof of the above stated aim: Let $(\tilde{n}_1, \tilde{k}_1), (\tilde{n}_2, \tilde{k}_3 \in \Lambda$ that generate Λ and such that their angle is as close to $\frac{\pi}{2}$ as possible.

Then this angle is between $\frac{\pi}{3}$ and $\frac{2\pi}{3}$:



The area of the parallelogram spanned by $(\tilde{n}_1, \tilde{k}_1)$ and $(\tilde{n}_2, \tilde{k}_2)$ is at least

$$\frac{2}{\sqrt{3}} \|(\tilde{n}_1, \tilde{k}_1)\|_2 \|(\tilde{n}_2, \tilde{k}_2)\|_2 \leq \mathrm{ord}(q).$$

Minkowski's second theorem in the geometry of numbers.

We know that $q \mid 2^{|\tilde{n}_1|} \cdot 3^{|\tilde{k}_1|} - 1$ or $q \mid 2^{|\tilde{n}_1|} - 3^{|\tilde{k}_1|}$. Then: either $|\tilde{n}_1|$ or $|\tilde{k}_1|$ has to be $\geq \frac{1}{2}\log_3(q)$. In particular: $\|(\tilde{n}_1, \tilde{k}_1)\|_2 \geq c \log q$ (for some absolute constant c).

Then $\|(\tilde{n}_1, \tilde{k}_1)\|_2, \|(\tilde{n}_2, \tilde{k}_2)\|_2 \leq c\frac{\mathrm{ord}(q)}{\log q}$. □

Proposition 1.14. Let $L \in \mathbb{Q}[X_1, \ldots, X_n]$ be a linear form. Then there exists $C = C(L)$ such that any solution $x_1, \ldots, x_n \in \mathcal{S}$ of $L(x_1, \ldots, x_n) = 0$ satisfies

$$|x_i - x_i|_\infty |x_i - x_j|_2 |x_i - x_j|_3 < C \qquad (*)$$

for some $i \neq j \in \{1, \ldots, n\}$.

Remark. For $x \in \mathbb{Z}$ such that $x = 2^n 3^k y$ with $n, k \in \mathbb{Z}_{\geq 0}$, $2, 3 \nmid y$, then

$$|x|_\infty |x|_2 |x|_3 = |y|.$$

Note that $(*)$ is invariant under multiplication by elements of $\mathcal{S}$.

Theorem. Let V be a vector space of dimension n over $\overline{\mathbb{Q}}$. Let $S \subset M_{\mathbb{Q}}$ be finite with $\infty \in S$. For each $v \in S$, let $\Lambda_1^{(v)}, \dots, \Lambda_n^{(v)}$ be a basis of V^*. Furthermore, let $\Lambda_1^{(0)}, \dots, \Lambda_n^{(0)}$ be another basis. Fix an extension of each $|\bullet|_v$ from $\mathbb{Q}$ to $\overline{\mathbb{Q}}$.
Then for all $\varepsilon > 0$, there are finitely many $\varphi_1, \dots, \varphi_n \in V^*$ such that all solutions $x \in V$ of

$$\prod_{v \in S} \prod_{j=1}^{n} |\Lambda_j^{(v)}(x)|_v \leq H(\Lambda_1^{(0)}(x), \dots, \Lambda_n^{(0)}(x))^{-\varepsilon}$$

with $\Lambda_1^{(0)}(x), \dots, \Lambda_n^{(0)}(x) \in \mathbb{Z}$ satisfy $\varphi_i(x) = 0$ for some $i = 1, \dots, n$.

Proof of Proposition. By induction on n. Suppose $n = 2$. As we observed the conclusion, is invariant under dividing x_1, x_2 by the same element of $\mathcal{S}$. Now $\gcd(x_1, x_2) \in \mathcal{S}$. So it is enough to prove for solutions with $\gcd(x_1, x_2) = 1$.

Let $L(X_1, X_2) = aX_1 + bX_2$. Then $ax_1 + bx_2 = 0$ with $\gcd(x_1, x_2) = 1$ implies $x_1 \mid b$ and $x_2 \mid a$.

So there are finitely many possibilities for x_1, x_2 in terms of L. Pick C that works for all.

(to be continued). □

Lecture 22

"generalised S-unit equations".

Let K be a number field: $\mathcal{O}_K = \{x \in K : |x|_v \leq 1 \text{ for all } v \in M_{K,f}\}$. Let $S \in M_K$ be a finite set containing $M_{K,\infty}$: $\mathcal{O}_{K,s} = \{x \in K : |x|_v \leq 1 \text{ for all } v \notin S\}$ ("S-integers"). $\mathcal{O}_{K,s}^{\times}$ units in $\mathcal{O}_{K,s}$ ("S-units").

Unit eqution $x + y = 1$ with x, y units.

Proof (continued). Induction on d. $d = 2$ was checked before.

Suppose $d > 2$, and the claims hold for $d - 1$. We make some simplifying assumptions to be specified later. We apply the subspace theorem on $\mathbb{Q}^{d-1} = V$. The reference basis is $\Lambda_j^{(0)} = X_j$, $j = 1, \dots, d-1$. As a first approximation, we try $\Lambda_j^{(v)} = X_j$ for all j, v. Let $S = \{\infty, 2, 3\}$. Let $x = (x_1, \dots, x_d)$ be a solution of $L(x_1, \dots, x_d) = 0$. Then

$$\prod_{v \in S} \prod_{j=1}^{d-1} |\Lambda_j^{(v)}(x)|_v = 1.$$

We can replace $\Lambda_1^{(w)}$ by

$$\frac{a_1}{a_n} X_1 + \cdots + \frac{a_{d-1}}{a_d} X_{d-1},$$

where $L = a_1X_1 + \cdots + a_dX_d$. Then we replace $|x_1|_w$ by $|x_n|_w$. We do this for some choice w.

Now back to the simplifying assumptions: We assume that $|x|_\infty$ is maximal for $j = n$. Then $|x_n|_2|x_n|_3 = |x_n|_\infty^{-1}$. So let $w \in \{2, 3\}$ such that $|x_n|_w \leq |x_n|_\infty^{-\frac{1}{2}}$. We may also assume $|x_1|_w = 1$. For this, we may need to divide x by the common divisor, and rearrange the indices.

For these augmented $\Lambda_j^{(v)}$'s, we get

$$\prod_{v\in S}\prod_{j=1}^{d-1} |\Lambda_j^{(v)}(x)|_v \leq |x_n|_\infty^{-\frac{1}{2}} \leq H(\Lambda_1^{(0)}(x), \ldots, \Lambda_{d-1}^{(0)}(x))^{-\frac{1}{2}}.$$

So the subspace theorem applies with $\varepsilon = \frac{1}{2}$. So $x_1, \ldots, x_{d-1}$ satisfies one of finitely many linear equations. Apply the induction hypothesis for each of them. □

Theorem 1.15. For all $\varepsilon > 0$, there exist finitely many multiplicatively independent pairs $a, b \in \mathcal{S}$ such that

$$\gcd(a-1, b-1) > \max(a, b)^\varepsilon.$$

Proof. Fix some $\varepsilon > 0$. Let $a, b \in \mathcal{S}$ multiplicatively independent and such that

$$d = \gcd(a-1, b-1) > \max(a, b)^\varepsilon.$$

Our goal is to show $d < C$ for some $C = C(\varepsilon)$. Note: $2, 3 \nmid d$, because otherwise $2 \nmid a, b$ or $3 \nmid a, b$. Then a and b would be a power of the same prime. Not possible due to multiplicative independence.

Fix some $n \in \mathbb{Z}_{>0}$ sufficiently large depending on ε. We apply the subspace theorem on $V = \mathbb{Q}^{n^2}/\{(x, \ldots, x) : x \in \mathbb{Q}\}$.

We will evaluate our functionals at the point $e/d = (e_1/d, \ldots, e_{n^2}/d)$ where $e_1, \ldots, e_{n^2}$ is an enumeration of $a^k b^l$ for $k = 0, \ldots, n-1$, $l = 0, \ldots, n-1$ such that $e_1 = 1$, $e_{n^2} = a^{n-1}b^{n-1}$.

Note: $\frac{e_i}{d} - \frac{e_j}{d} \in \mathbb{Z}$. This is because $e_i \equiv 1 \pmod{d}$. Also: $\left|\frac{e_i}{d} - \frac{e_j}{d}\right|_v \leq \min\left(\left|\frac{e_i}{d}\right|_v, \left|\frac{e_j}{d}\right|_v\right)$ for all $v \in S = \{\infty, 2, 3\}$. The coordinates on $\mathbb{Q}^{n^2}$ will be denoted by $Y_1, \ldots, Y_{n^2}$. All our linear forms on V will be of the form $Y_i - Y_j$ for some $i \neq j$. This is indeed well defined on V. Reference basis $\Lambda_j^{(0)} = Y_j - Y_{n^2}$.

$$H\left(\Lambda_1^{(0)}\left(\frac{e}{d}\right), \ldots, \Lambda_{n^2-1}^{(0)}\left(\frac{e}{d}\right)\right) \leq a^n b^n.$$

For $v = \infty$:

$$\begin{aligned}
\Lambda_j^{(\infty)} &= Y_{j+1} - Y_1 \\
|\Lambda_j^{(\infty)}(e/d)| &= |e_{j+1}/d|_\infty \\
\Lambda_j^{(v)} &= Y_j - Y_{n^2} \\
|\Lambda_j^{(v)}(e/d)|_v &= |e_j|_v \\
\prod_{j=1}^{n^2-1} |\Lambda_j^{(\infty)}(e/d)|_\infty &\le \left(\prod_{j=1}^{n^2} |e_j/d|_\infty \right) \cdot d \\
\prod_{j=1}^{n^2-1} |\Lambda_j^{(v)}(e/d)|_v &\le \left(\prod_{j=1}^{n^2} |e_j/d|_v \right) / |a^{n-1}b^{n-1}|_v \\
\prod_{v \in S} \prod_{j=1}^{n^2-1} |\Lambda_j(e/d)| &\le d \cdot (a^{n-1}b^{n-1}) \cdot d^{-n^2} \\
|e_i/d|_\infty |e_j/d|_2 |e_j/d|_3 &= \frac{1}{d}
\end{aligned}$$

Lecture 23

$d = \gcd(a-1, b-1)$ where $a, b \in \mathcal{S}$ are multiplicatively independent. We assume: $d > \max(a,b)^\varepsilon$ for some $\varepsilon > 0$. Our goal is to prove $d < C(\varepsilon)$.

$$\prod_{v \in S = \{\infty, 2, 3\}} \prod_{j=1}^{n^-1} |\Lambda_j^{(v)}(e/d)|_v \le d a^{n-1} b^{n-1} d^{-n^2}. \tag{$*$}$$

$e_1, \ldots, e_{n^2}$ is an enumeration of $a^k b^l$, $k, l = 0, \ldots, n-1$.

$$(*) \le \max(a,b)^{2n-2} \cdot \max(a,b)^{-\varepsilon(n^2-1)}.$$

Let's take $n > 3\varepsilon^{-1}$, $(*) < \max(a,b)^{-n}$.

$$H(\Lambda_1^{(0)}(e/d), \ldots, \Lambda_{n-1}^{(0)}(e/d)) \le a^{n-1} b^{n-1}.$$

$(*) < H(\cdots)^{-\frac{1}{2}}$. Subspace theorem applies hence there exists a linear relation between $e_1, \ldots, e_{n^2} \in \mathcal{S}$ (distinct by multiplicative independence of a, b).

Proposition implies

$$|e_i - e_j|_\infty |e_i - e_j|_2 |e_i - e_j| < C = C(\varepsilon)$$

for some $i \ne j$. Then $e_i \ne e_j$ so $e_i - e_j \ne 0$. However, $d \mid e_i - e_j$.

$$d \le |e_i - e_j|_\infty |e_i - e_j|_2 |e_i - e_j|_3 < C.$$

□

Theorem 1.16 (Feldman)**.** Let $\alpha \in \overline{\mathbb{Q}}$ of degree $d \geq 3$. Then there exists effective $C = C(\alpha) > 0$ and $\varepsilon = \varepsilon(\alpha) > 0$ such that for all $\frac{p}{q} \in \mathbb{Q}$,

$$\left|\alpha - \frac{p}{q}\right| > \frac{C}{q^{d-\varepsilon}}.$$

Remark. This is enough to solve $P(x, y) = m$, where P is a degree d homogeneous polynomial without repeated factors. Thue equation.

Proposition. Let K be a number field. Then there exists $r \in \mathbb{Z}_{\geq 0}$ and $u_1, \ldots, u_r \in \mathcal{O}_K^\times$ and a constant $C = C(K)$ such that $\forall \alpha \in \mathcal{O}_K$, there exists $\tilde{\alpha} \in \mathcal{O}_K$ and $b_1, \ldots, b_r \in \mathbb{Z}$ such that

$$\begin{aligned} H(\tilde{\alpha}) &\leq C \cdot |N_{K/\mathbb{Q}}(\alpha)|^{\frac{1}{[K:\mathbb{Q}]}} \\ |b_1|, \ldots, |b_r| &\leq C \log H(\alpha) \\ \alpha &= \tilde{\alpha} u_1^{b_1} \cdots u_r^{b_r} \end{aligned}$$

Define $\Phi : K^\times \to \mathbb{R}^{M_{K,\infty}} : (\Phi(\alpha))_v = d_v \cdot \log |\alpha|_v$ (logarithmic embedding). Note that here $K^\times$ is the group under multiplication, while $\mathbb{R}^{M_{K,\infty}}$ is the additive group.

$$\begin{aligned} |N_{K/\mathbb{Q}}(\alpha)| &= \exp\left(\sum_{v \in M_{K,\infty}} (\Phi(\alpha))_v\right) \\ H(\alpha)^{[K:\mathbb{Q}]} &= \exp\left(\sum_{v \in M_{K,\infty}} \max(0, (\Phi(\alpha))_v)\right) \end{aligned}$$

For $\alpha \in \mathcal{O}_K$, $\Sigma(\Phi(\alpha))_v \geq 0$. Then:

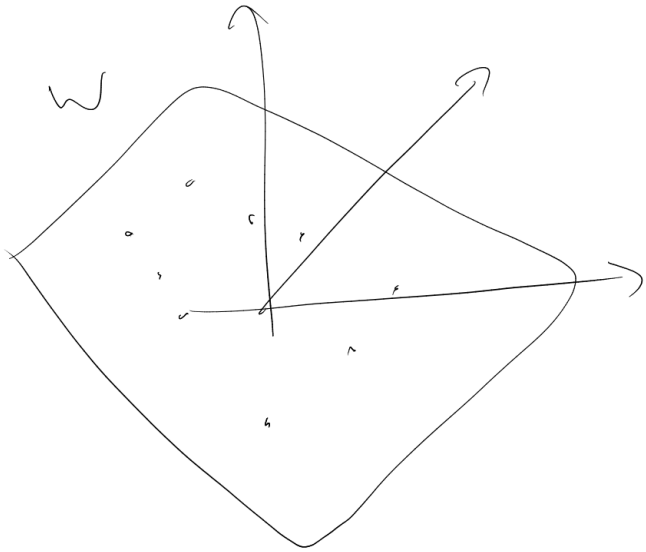
$$\exp(\|\Phi(\alpha)\|_1/2) \leq H(\alpha)^{[K:\mathbb{Q}]} \leq \exp(\|\Phi(\alpha)\|_1).$$

For $\alpha \in \mathcal{O}_K^\times$, $N_{K/\mathbb{Q}}(\alpha) = 1$. So

$$\Phi(\alpha) \in W = \{x \in \mathbb{R}^{M_{K,\infty}} : \sum x_v = 0\}.$$

Kronecker's theorem: $\Phi^{-1}(0) = \ker \Phi$ are the roots of unity.

Dirichlet's unit theorem: $\Phi(\mathcal{O}_K^\times)$ is a lattice in W that is a $\mathbb{Z}$-module of rank $\dim W = r$ which spans W.



Let $u_1, \ldots, u_r$ be a fundamental system of units, that is $\Phi(u_1), \ldots, \Phi(u_r)$ is a basis for the lattice $\Phi(\mathcal{O}_K^\times)$. Fix some $\alpha \in \mathcal{O}_K$. Pick some $x \in \mathbb{R}_{\geq 0}^{M_{K,\infty}}$ such that

$$\sum x_v = \underbrace{\log |N_{K/\mathbb{Q}}(\alpha)|}_{\geq 0}.$$

$x \in \Phi(\alpha) + W$.

Then there exist $y_1, \ldots, y_r \in \mathbb{R}$ such that

$$x = \Phi(\alpha) + y_1\Phi(u_1) + \cdots + y_r\Phi(u_r).$$

There exists $C = C(K)$ such that $|y_j| \leq C \cdot \|\Phi(\alpha)\|_1$.

Let $b_j \in \mathbb{Z}$ with $|y_j - b_j| \leq 1$ and $|b_j| \leq |y_j|$. This gives $|b_j| \leq C \cdot \|\Phi(\alpha)\|_1 \leq C' \log H(\alpha)$.

Take: $\tilde{\alpha} = \alpha u_1^{b_1} \cdots u_r^{b_r}$.

$$\Phi(\tilde{\alpha}) = \Phi(\alpha) + b_1\Phi(u_1) + \cdots + b_r\Phi(u_r) = x + \underbrace{(b_1 - y_1)\Phi(u_1) + \cdots}_{(*)}.$$

$(*)$ is in a fixed, compact region of W.

$\|\Phi(\tilde{\alpha})\|_1 \leq C + \|x\|_1$.

$$H(\tilde{\alpha})^{[K:\mathbb{Q}]} \leq \exp(\|\Phi(\tilde{\alpha})\|_1) \leq \exp(C) \cdot N_{K/\mathbb{Q}}(\alpha).$$

Lecture 24

Theorem. α algebraic of degree $d \leq 3$. Then there exists $C = C(\alpha) > 0$, $\varepsilon = \varepsilon(\alpha) > 0$ such that for all $\frac{p}{q} \in \mathbb{Q}$:

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

Proof. Fix some α and $\varepsilon > 0$ small enough. Suppose that

$$\left|\alpha - \frac{p}{q}\right| < q^{-(d-\varepsilon)}$$

for some $\frac{p}{q} \in \mathbb{Q}$. We aim to show that $q < C = C(\alpha)$. We assume as we may that α is an algebraic integer.

Let

$$P(X) = (X - \alpha_1) \cdots (X - \alpha_d)$$

be the minimal polynomial of $\alpha = \alpha_1$. Then:

$$(p - \alpha_1 q) \cdots (p - \alpha_d q) = Q < Cq^{\varepsilon}.$$

With $Q \in \mathbb{Z}$. Then

$$N_{\mathbb{Q}(\alpha_j)/\mathbb{Q}}(p - \alpha_j q) \mid Q^d.$$

In particular:

$$N(p - \alpha_j q) < Cq^{d\varepsilon}.$$

Therefore: $\exists \tilde{\alpha}_j$, $u_1, \ldots, u_r, b_1, \ldots, b_r \in \mathbb{Z}$ such that $p - \alpha_j q = \tilde{\alpha}_j u_1^{b_1} \cdots u_r^{b_r}$. Then

$$\begin{aligned} H(\tilde{\alpha}_j) &< C \cdot q^{\varepsilon} \\ |b_j| &< C \cdot \log q \end{aligned}$$

Use $|p - \alpha_1 q| < q^{-(d-1-\varepsilon)}$.

Then $p - \alpha_j q$ is very close to $(\alpha_1 - \alpha_j)q$. Consider: $(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)q$, which is similar to both of $(p - \alpha_2 q)(\alpha_1 - \alpha_3)$ and $(\alpha_1 - \alpha_2)(p - \alpha_3 q)$.

Now more formally:

$$\begin{aligned} \left|1 - \frac{(p - \alpha_2 q)(\alpha_1 - \alpha_3)}{(\alpha_1 - \alpha_2)(p - \alpha_3 q)}\right| &= \left|1 - \frac{((p - \alpha_1 q) + (\alpha_1 - \alpha_2)q)(\alpha_1 - \alpha_3)}{(\alpha_1 - \alpha_2)((p - \alpha_1 q) + (\alpha_1 - \alpha_3)q)}\right| \\ &< Cq^{-(d-\varepsilon)} \end{aligned}$$

$$\left|\frac{A - \kappa_1}{B - \kappa_2} - \frac{A}{B}\right| < \frac{\max(\kappa_1, \kappa_2)}{q}.$$

$A \sim B \sim q$. Now use the proposition:

$$\begin{aligned} p - \alpha_2 q &= \tilde{\alpha}_2 u_1^{b_1} \cdots u_r^{b_r} \\ p - \alpha_3 q &= \tilde{\alpha}_3 w_1^{e_1} \cdots w_r^{e_r} \\ H(\tilde{\alpha}_2), H(\tilde{\alpha}_3) &\leq C \cdot q^{\varepsilon} \\ |b_1|, \ldots, |b_r|, |e_1|, \ldots, |e_r| &< C \cdot \log q \end{aligned}$$

Writing $\alpha^* = \frac{\tilde{\alpha}_2(\alpha_1 - \alpha_3)}{\tilde{\alpha}_3(\alpha_1 - \alpha_2)}$ we have:

$$|1 - \alpha^* u_1^{b_1} \cdots u_r^{b_r} w_1^{-e_1} \cdots w_r^{-e_r}| < C q^{-(d-\varepsilon)}.$$

$H(\alpha^*) < C \cdot q^{2\varepsilon}$. Take log to be the principal branch, that is $|\operatorname{Im} \log(\bullet)| \leq \pi$. Warning: $\log(xy) \neq \log(x) + \log(y)$ in general. This is Lipschitz around 1, so we get

$$|\log(\alpha^*) + b_1 \log(u_1) + \cdots + b_r \log(u_r) - e_1 \log(w_1) - \cdots - e_r \log(w_r) + 2k \cdot \underbrace{\log(-1)}_{=\pi i}| < C q^{-(d-\varepsilon)}$$

for a suitable $k \in \mathbb{Z}$, and $|k| < C \log q$.

Reminder:

Theorem. Let $n \in \mathbb{Z}_{\geq 1}$. Let $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}_{\neq 0}$, and let $\log \alpha_j$ be any choice of the log of α_j. Let $b_1, \ldots, b_n \in \mathbb{Z}$ and let $\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n$. Let

$$\begin{aligned} A_j &= \max(H(\alpha_j), \exp(|\log \alpha_j|), 10) \\ B^* &= \max\left(\frac{|b_1|}{\log A_n}, \ldots, \frac{|b_{n-1}|}{\log A_n}, |b_n|, 10\right) \end{aligned}$$

Then there exists an effective constant C (a function of n and the degree of $\mathbb{Q}(\alpha_1, \ldots, \alpha_n)$) such that $\Lambda \neq 0$ implies

$$|\Lambda| > \exp(-C \log(A_1) \cdots \log(A_n) \log(B^*)).$$

So the lower bound gives:

We apply the theorem with $\alpha_n = \alpha^*$. $A_1, \ldots, A_{n-1} < C$, $A_n < C \cdot q^{2\varepsilon}$. $B^* \leq \frac{C \log q}{\log A_n} \leq \frac{C}{\varepsilon}$. $|k| < C \log q$. So

$$|\Lambda| > \exp(-C \cdot \varepsilon \log q \cdot \log \varepsilon^{-1}) > q^{-C\varepsilon \log \varepsilon^{-1}}.$$

We still need to consider $\Lambda = 0$. This is equivalent to:

$$1 = \frac{(p - \alpha_2 q)(\alpha_1 - \alpha_3)}{(\alpha_1 - \alpha_2)(p - \alpha_3 q)}.$$

Solving this equation gives $\alpha_2 = \alpha_3$ or $p = \alpha_1 q$. Neither is the case.

If we use the weaker bound for $|\Lambda|$, then we would prove:

$$\left|\alpha - \frac{p}{q}\right| > C \cdot q^{-(d - \frac{\varepsilon}{\log \log q})}.$$ □

Index