
Analytic Number Theory III

Lecture notes

Prof. Dr. Damaris Schindler

L^AT_EX version by Alex Dalist Howl Sennewald

Mathematical Institute
Georg-August-University Göttingen
Winter term 2023/24

Contents

1	Number fields	1
1.1	Number fields and number rings	1
1.2	Embeddings, Norm and Trace	3
1.3	Discriminant	7
1.4	Cyclotomic fields	9
2	Prime ideal factorisation	11
2.1	Unique prime ideal factorisation	11
2.2	Splitting of primes	15
3	Dirichlet's unit theorem, class groups and lattices	23
3.1	Finiteness of the ideal class group	23
3.2	Geometry of numbers	24
3.3	Bounds for class numbers	32
3.4	Dirichlet's unit theorem	34
4	Diophantine Approximation	39
4.1	Introduction	39
4.2	Transcendence	42
4.3	More on transcendence results	47
4.4	Siegel's lemma	50
4.5	Approaches towards the Gelfond-Schneider theorem	54
5	Applications to Diophantine problems	57
5.1	Linear forms in logarithms	57
5.2	Roth's theorem	60
5.3	Schmidt's subspace theorem	62
	Definitions	69
	Important theorems	69

List of lectures

Lecture 1 from 24.10.2023	1
Lecture 2 from 27.10.2023	3
Lecture 3 from 03.11.2023	6
Lecture 4 from 07.11.2023	8
Lecture 5 from 10.11.2023	10
Lecture 6 from 17.11.2023	12
Lecture 7 from 21.11.2023	14
Lecture 8 from 24.11.2023	17
Lecture 9 from 28.11.2023	20
Lecture 10 from 01.12.2023	21
Lecture 11 from 05.12.2023	24
Lecture 12 from 08.12.2023	28
Lecture 13 from 12.12.2023	30
Lecture 14 from 15.12.2023	32
Lecture 15 from 19.12.2023	34
Lecture 16 from 22.12.2023	36
Lecture 17 from 09.01.2024	40
Lecture 18 from 12.01.2024	44
Lecture 19 from 16.01.2024	46
Lecture 20 from 19.01.2024	49
Lecture 21 from 23.01.2024	53
Lecture 22 from 26.01.2024	57
Lecture 23 from 30.01.2024	59
Lecture 24 from 02.02.2024	60
Lecture 25 from 06.02.2024	63
Lecture 26 from 09.02.2024	66

This script is not a substitute for Prof. Schindler's lecture notes and will not be reviewed by her again. Basically, these are just my personal notes, so I do not guarantee correctness or completeness and I might add further examples and notes if necessary. In general I will not include proofs (because this is no fun in \LaTeX).

If you have any corrections, you can write to me at [Stud.IP](#) or make a pull request directly at the [GitHub repository](#) (which is much more convenient for me than the way via Stud.IP).

glhf,
Alex

1 Number fields

Example (Pell¹ equation): Let $d > 1$ be an integer, which is not a square, and find all integer solutions to

Lecture 1,
24.10.2023

$$x^2 - dy^2 = 1. \quad (1.1)$$

Write $\mathbb{Z}[\sqrt{d}] = \{a + \sqrt{d}b \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{Q}[\sqrt{d}]$ with its natural ring structure. If $(x, y) \in \mathbb{Z}^2$ is a solution to (1.1), then

$$(x + \sqrt{d}y)(x - \sqrt{d}y) = x^2 - dy^2 = 1$$

and for every $k \in \mathbb{N}$

$$(x + \sqrt{d}y)^k(x - \sqrt{d}y)^k = x_k^2 - dy_k^2 = 1,$$

with $x_k, y_k \in \mathbb{Z}$. I.e. if $(x, y) \neq (\pm 1, 0)$ we can generate new solutions as above. Define the norm map $N : \mathbb{Z}[\sqrt{d}] \rightarrow \mathbb{Z}$, $a + \sqrt{d}b \mapsto a^2 - db^2$. Then solutions to (1.1) can be described as units $x + \sqrt{d}y \in \mathbb{Z}[\sqrt{d}]^*$ in the ring $\mathbb{Z}[\sqrt{d}]$ with $N(x + \sqrt{d}y) = 1$.

Example (Gaussian integers): The question is to find all primes p which can be written as a sum of two integer squares

$$p = a^2 + b^2.$$

I.e. we ask for primes p which factor as $p = (a + ib)(a - ib)$ in the ring $\mathbb{Z}[i]$.

1.1 Number fields and number rings, first definitions and examples

Definition (Number field)

A *number field* is a finite field extension of \mathbb{Q} .

Example: a) For $d \in \mathbb{Z}$, where d is not a square, the fields $\mathbb{Q}[\sqrt{d}] = \mathbb{Q}[x]/(x^2 - d)$

¹after John Pell (1611 - 1685), an English mathematician

are number fields (with degree 2 over \mathbb{Q}). We call $\mathbb{Q}[\sqrt{d}]$ a *real quadratic field* if $d > 0$ and an *imaginary quadratic field* if $d < 0$.

b) $\mathbb{Q}[\sqrt{d_1}, \sqrt{d_2}]$ are number fields for $d_1, d_2 \in \mathbb{Z}$, usually called *biquadratic fields*.

c) Let $m \in \mathbb{N}$ and $\omega = e^{\frac{2\pi i}{m}}$. Then $\mathbb{Q}[\omega]$ is a number field, called the *m-th cyclotomic field*.

?) What could be an analogue of the integers in a general number field?

$$\mathbb{Z} \subset \mathbb{Q} \quad ? \subset \mathbb{Q}[\sqrt{d}] \quad ? \subset \mathbb{F}$$

Definition (Algebraic integer)

A complex number $\alpha \in \mathbb{C}$ is called an *algebraic integer*, if there is a monic polynomial $P(x) \in \mathbb{Z}[x]$ with $P(\alpha) = 0$.

Example: • Every $n \in \mathbb{Z}$ is an algebraic integer.

- \sqrt{d} for $d \in \mathbb{Z}$ is an algebraic integer (take $P(x) = x^2 - d$).
- $e^{\frac{2\pi i}{m}}$ is an algebraic integer for every $m \in \mathbb{N}$ (take $P(x) = x^m - 1$).

Theorem 1.1

Let α be an algebraic integer and $f(x) \in \mathbb{Z}[x]$ a monic polynomial with $f(\alpha) = 0$. If $f(x)$ is of minimal degree with these properties, then f is irreducible.

Remark: Theorem 1.1 shows, that the minimal polynomial of an algebraic integer over \mathbb{Q} has coefficients in \mathbb{Z} .

Lemma 1.2

Let $f \in \mathbb{Z}[x]$ be a monic polynomial and $g, k \in \mathbb{Q}[x]$ monic polynomials with $f = gh$. Then, $g, k \in \mathbb{Z}[x]$.

Corollary 1.3

If $\alpha \in \mathbb{C}$ is an algebraic integer, then $\alpha \in \mathbb{Z}$.

Theorem 1.4 (Characterization of algebraic integers)

Let $\alpha \in \mathbb{C}$. Then the following statements are equivalent:

- (i) α is an algebraic integer.
- (ii) $\mathbb{Z}[\alpha]$ is a finitely generated group (under addition).
- (iii) There exists a subring $R \subset \mathbb{C}$ with $\alpha \in R$ and such that $(R, +)$ is a finitely generated group.
- (iv) There is a non-trivial finitely generated subgroup $(A, +)$ of \mathbb{C} , such that $\alpha A \subseteq A$.

Corollary 1.5

The set of algebraic integers in \mathbb{C} is a ring.

Lecture 2,
27.10.2023

Definition (Ring of algebraic integers)

Let K be a number field. Then we write \mathcal{O}_K for the set of algebraic integers contained in K and we call \mathcal{O}_K the ring of integers of K .

Example: $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$

Proposition 1.6

Let $d \in \mathbb{Z}$ be a squarefree integer.

- If $d \equiv 2, 3 \pmod{4}$ then $\mathcal{O}_{\mathbb{Q}[\sqrt{d}]} = \{a + \sqrt{d}b \mid a, b \in \mathbb{Z}\}$.
- If $d \equiv 1 \pmod{4}$, then $\mathcal{O}_{\mathbb{Q}[\sqrt{d}]} = \left\{ \frac{a + \sqrt{d}b}{2} \mid a \equiv b \pmod{2} \right\}$.

1.2 Embeddings, Norm and Trace

Recall: Let L/K be a finite field extension. If $\text{char } K = 0$, then L/K is separable. Let \bar{K} be an algebraic closure of K . If L/K is separable, then $[L : K] = \# \text{Hom}_K(L, \bar{K})$.

Theorem

Let L/K be a finite separable field extension. Then there exists an element $\alpha \in L$ such that $L = K(\alpha)$. In particular, for number fields $\mathbb{Q} \subseteq K \subseteq L$ we obtain the following:

- There exists $\alpha \in L$ such that $L = K(\alpha)$
- If there is an embedding $\hat{\iota} : K \hookrightarrow \mathbb{C}$, then there exist $[L : K]$ embeddings $L \hookrightarrow \mathbb{C}$, which extend $\hat{\iota}$. If $g(x)$ is a minimal polynomial of α over K then

the embeddings are given by $\sigma_i : \alpha \mapsto \beta_i$, where $\beta_1, \dots, \beta_{[L:K]}$ are the $[L:K]$ distinct conjugates of α .

Example: 1. Let $d \in \mathbb{Z}$ be not a square. Then there are exactly two embeddings of $\mathbb{Q}[\sqrt{d}]$ into \mathbb{C} , namely $\sigma_1 : a + \sqrt{d}b \mapsto a + \sqrt{d}b$ and $\sigma_2 : a + \sqrt{d}b \mapsto a - \sqrt{d}b$.

2. We have $[\mathbb{Q}[\sqrt[3]{2} : \mathbb{Q}]] = 3$ and the three embeddings are given by

$$\sigma_1(\sqrt[3]{2}) = \sqrt[3]{2}, \quad \sigma_2(\sqrt[3]{2}) = e^{\frac{2\pi i}{3}} \sqrt[3]{2}, \quad \sigma_3(\sqrt[3]{2}) = e^{\frac{4\pi i}{3}} \sqrt[3]{2}.$$

Note that $\sigma_1(\mathbb{Q}[\sqrt[3]{2}]) \subseteq \mathbb{R}$, whereas σ_2 and σ_3 are "complex embeddings". $\mathbb{Q}[\sqrt[3]{2}]/\mathbb{Q}$ is not a normal extension.

Definition (Trace and norm)

Let K be a field and V an n -dimensional K -vector space. For $\varphi : V \rightarrow V$ a K -endomorphism, we define the characteristic polynomial

$$\chi_\varphi(x) = \det(xI_n - \varphi) = \sum_{i=0}^n c_i x^{n-i}$$

for some $c_0, \dots, c_n \in K$. We define the determinant and trace of φ by $\det \varphi = (-1)^n c_n$ and $\text{trace } \varphi = -c_1$

Note that if $\varphi, \psi : V \rightarrow V$ are both K -endomorphisms of V , then $\det(\varphi \circ \psi) = \det(\varphi) \det(\psi)$ and $\text{trace}(a\varphi + b\psi) = a \text{trace}(\varphi) + b \text{trace}(\psi) \quad \forall a, b \in K$.

Definition (Trace and norm of a number field)

Let $\mathbb{Q} \subseteq K \subseteq L$ be number fields and $\alpha \in L$. We write $\varphi_\alpha : L \rightarrow L, x \mapsto \alpha x$ and define the (relative) norm and trace of α by

$$N_{L/K}(\alpha) = \det \varphi_\alpha, \quad \text{Tr}_{L/K}(\alpha) = \text{trace}(\varphi_\alpha).$$

Remark: The map $N_{L/K} : L^* \rightarrow K^*$ is a group homomorphism as

$$N_{L/K}(\alpha\beta) = N_{L/K}(\alpha)N_{L/K}(\beta) \quad \forall \alpha, \beta \in L \setminus \{0\}.$$

Similarly, $\text{Tr}_{L/K} : L \rightarrow K$ is a K -linear map, as

$$\text{Tr}_{L/K}(u\alpha + v\beta) = u \text{Tr}_{L/K}(\alpha) + v \text{Tr}_{L/K}(\beta) \quad \forall u, v \in K, \alpha, \beta \in L.$$

Example: Let $K = \mathbb{Q}$, $L = \mathbb{Q}(i)$ and $\alpha = a + ib \in \mathbb{Q}(i)$. Then φ_α can be represented with respect to the basis $1, i$ by

$$\varphi_\alpha = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

and hence

$$N_{L/\mathbb{Q}}(a + ib) = a^2 + b^2, \quad \text{Tr}_{L/\mathbb{Q}}(a + ib) = 2a.$$

Lemma 1.7

Let L/K be an extension of number fields with $[L : K] = n$. For $a \in K$ we have

$$N_{L/K}(a) = a^n, \quad \text{Tr}_{L/K}(a) = na.$$

Lemma 1.8

Let L/K be an extension of number fields with $L = K(\alpha)$ and $[L : K] = n$. Let $f(x) = x^n + c_1x^{n-1} + \cdots + c_n$ be the minimal polynomial of α over K . Then

$$N_{L/K}(\alpha) = (-1)^n c_n, \quad \text{Tr}_{L/K}(\alpha) = -c_1.$$

Lemma 1.9

Let L/K be a number field extension, $\alpha \in L$, $[L : K(\alpha)] = r$. Then we have

$$N_{L/K}(\alpha) = \left(N_{K(\alpha)/K}(\alpha) \right)^r, \quad \text{Tr}_{L/K}(\alpha) = r \text{Tr}_{K(\alpha)/K}(\alpha).$$

Corollary 1.10

Let L/K be number fields and $\alpha \in \mathcal{O}_L$. Then $N_{L/K}(\alpha), \text{Tr}_{L/K}(\alpha) \in \mathcal{O}_K$. In particular $N_{L/\mathbb{Q}}(\alpha), \text{Tr}_{L/\mathbb{Q}}(\alpha) \in \mathbb{Z}$.

Theorem 1.11

Let L/K be number fields, $[L : K] = n$ and $\sigma_1, \dots, \sigma_n : L \hookrightarrow \mathbb{C}$ be the n distinct K -linear embeddings of L into \mathbb{C} . Then, for $\alpha \in L$, we have

$$N_{L/K}(\alpha) = \prod_{i=1}^n \sigma_i(\alpha), \quad \text{Tr}_{L/K}(\alpha) = \sum_{i=1}^n \sigma_i(\alpha).$$

Corollary 1.12

Let L/K be a Galois extension of number fields. Then, for $\alpha \in L$ and $\sigma \in \text{Gal}(L/K)$, we have

$$N_{L/K}(\sigma(\alpha)) = N_{L/K}(\alpha), \quad \text{Tr}_{L/K}(\sigma(\alpha)) = \text{Tr}_{L/K}(\alpha).$$

Theorem 1.13

Let $K \subseteq L \subseteq M$ be a tower of number fields and $\alpha \in M$. Then

$$N_{M/K} = N_{L/K}(N_{M/L}(\alpha)), \quad \text{Tr}_{M/K}(\alpha) = \text{Tr}_{L/K}(\text{Tr}_{M/L}(\alpha)).$$

Lecture 3,
03.11.2023

An application of the norm map

Given a number field K with ring of integers \mathcal{O}_K , how can we find \mathcal{O}_K^* , i.e. the units in \mathcal{O}_K ?

- If $\alpha \in \mathcal{O}_K^*$, $\alpha^{-1} \in \mathcal{O}_K$ and $1 = N_{K/\mathbb{Q}}(\alpha\alpha^{-1}) = N_{K/\mathbb{Q}}(\alpha) \cdot N_{K/\mathbb{Q}}(\alpha^{-1})$. By Corollary 1.10, $N_{K/\mathbb{Q}}(\alpha), N_{K/\mathbb{Q}}(\alpha^{-1}) \in \mathbb{Z} \implies N_{K/\mathbb{Q}}(\alpha) = \pm 1$.
- If $\alpha \in \mathcal{O}_K$ with $N_{K/\mathbb{Q}}(\alpha) = \pm 1$, then $\alpha \in \mathcal{O}_K^*$.

Example: Let $d \in \mathbb{Z}$, d squarefree. Then, for $a, b \in \mathbb{Q}$, $N_{\mathbb{Q}[\sqrt{d}]/\mathbb{Q}}(a + \sqrt{d}b) = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2$. For $d \equiv 2, 3 \pmod{4}$, we find that

$$\mathcal{O}_{\mathbb{Q}[\sqrt{d}]} = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}, a^2 - db^2 = \pm 1\}.$$

The trace as a bilinear form

Let L/K be number fields. Then $\text{Tr}_{L/K}$ induces a bilinear form

$$\text{Tr}_{L/K} : L \times L \rightarrow K, (x, y) \mapsto \text{Tr}_{L/K}(x \cdot y). \quad (1.2)$$

Write L^* for the dual vector space of L , i.e. the set of all K -linear vector space homomorphisms.

Theorem 1.14

The bilinear form (1.2) induces an isomorphism of K -vector spaces

$$\psi : L \rightarrow L^*, x \mapsto \text{Tr}_{L/K}(x, \cdot).$$

Corollary 1.15

Let L/K be number fields and (v_1, \dots, v_n) a K -basis with $n = [L : K]$. Then there exists a unique K -basis (w_1, \dots, w_n) of L , such that $\text{Tr}_{L/K}(v_i w_j) = \delta_{ij}$, $1 \leq i, j \leq n$.

1.3 Discriminant

Let K/\mathbb{Q} be a number field of degree $n = [K : \mathbb{Q}]$ and $\sigma_1, \dots, \sigma_n : K \rightarrow \mathbb{C}$ its embeddings.

Definition (Discriminant)

For $\alpha_1, \dots, \alpha_n \in K$, we define the *discriminant* as

$$\text{disc}(\alpha_1, \dots, \alpha_n) = \det \left((\sigma_i(\alpha_j))_{1 \leq i, j \leq n} \right)^2.$$

Theorem 1.16

Let $\alpha_1, \dots, \alpha_n \in K$. Then $\alpha_1, \dots, \alpha_n$ are \mathbb{Q} -linearly independent if and only if $\text{disc}(\alpha_1, \dots, \alpha_n) \neq 0$.

Lemma 1.17

Let $\alpha_1, \dots, \alpha_n \in K$. Then

$$\text{disc}(\alpha_1, \dots, \alpha_n) = \det \left(\text{Tr}_{K/\mathbb{Q}}(\alpha_i \alpha_j) \right)_{1 \leq i, j \leq n}.$$

Corollary 1.18

Let $\alpha_1, \dots, \alpha_n \in K$. Then $\text{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Q}$. If moreover $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$, then $\text{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}$.

Theorem 1.19

Let α be algebraic over \mathbb{Q} with $[\mathbb{Q}[\alpha] : \mathbb{Q}] = n$, and $\alpha_1, \dots, \alpha_n$ the n different conjugates of α . Then

$$\text{disc}(1, \alpha, \dots, \alpha^{n-1}) = \prod_{1 \leq i, j \leq n} (a_i - a_j)^2.$$

If moreover $f(x)$ is the minimal polynomial of α over \mathbb{Q} , then

$$\text{disc}(1, \alpha, \dots, \alpha^{n-1}) = (-1)^{\frac{n(n-1)}{2}} N_{\mathbb{Q}[\alpha]/\mathbb{Q}}(f'(\alpha)).$$

Question: Let K be a number field with ring of integers \mathcal{O}_K and of degree $n = [K : \mathbb{Q}]$. Then K is an n -dimensional \mathbb{Q} -vector space. How can we describe the structure of the group $(\mathcal{O}_K, +)$?

Example: For $d \in \mathbb{Z}$ squarefree and $K = \mathbb{Q}[\sqrt{d}]$, the ring of integers \mathcal{O}_K is a free abelian group of rank 2, where a \mathbb{Z} -basis is given by $(1, \omega)$, with

$$\omega = \begin{cases} \sqrt{d} & d \equiv 2, 3 \pmod{4}, \\ \frac{1+\sqrt{d}}{2} & d \equiv 1 \pmod{4}. \end{cases}$$

Theorem 1.20

Let K/\mathbb{Q} be a number field of degree $n = [K : \mathbb{Q}]$. Then \mathcal{O}_K is a free abelian group of rank n , i.e. there exists $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$, such that every $\beta \in \mathcal{O}_K$ can be uniquely written in the form

$$\beta = m_1\alpha_1 + \dots + m_n\alpha_n$$

with $m_1, \dots, m_n \in \mathbb{Z}$.

Remark: In the notation of Theorem 1.20, we call $(\alpha_1, \dots, \alpha_n)$ an integral basis of \mathcal{O}_K (over \mathbb{Z}).

Lecture 4,
07.11.2023

Lemma 1.21

Let K be a number field as above. Then there exists a \mathbb{Q} -basis of the number field, say $(\alpha_1, \dots, \alpha_n)$, with $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$.

Proposition 1.22

Let $(\alpha_1, \dots, \alpha_n)$ be a \mathbb{Q} -basis of a number field K with $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$, $d = \text{disc}(\alpha_1, \dots, \alpha_n)$ and $\beta \in \mathcal{O}_K$. Then there exist $m_1, \dots, m_n \in \mathbb{Z}$, such that

$$\beta = \frac{m_1\alpha_1 + \dots + m_n\alpha_n}{d}$$

and $d \mid m_i^2$ for $1 \leq i \leq n$.

Lemma 1.23

Let K be a number field with integral bases $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$. Then

$$\text{disc}(\alpha_1, \dots, \alpha_n) = \text{disc}(\beta_1, \dots, \beta_n).$$

Definition (Discriminant of K)

Let K be a number field and $(\alpha_1, \dots, \alpha_n)$ a \mathbb{Z} -basis for \mathcal{O}_K . We define the *discriminant* $\text{disc}(K)$ of K as

$$\text{disc}(K) = \text{disc}(\alpha_1, \dots, \alpha_n).$$

Example: Let $d \in \mathbb{Z}$ be squarefree. Then

$$\text{disc}([\sqrt{d}]) = \begin{cases} 4d & d \equiv 2, 3 \pmod{4}, \\ d & d \equiv 1 \pmod{4}. \end{cases}$$

1.4 Cyclotomic fields

Definition

For $m \in \mathbb{N}$ we call $\mathbb{Q}[e^{\frac{2\pi i}{m}}]$ the m -th cyclotomic field.

Example: • The first two cyclotomic fields are equal to \mathbb{Q} .

- Let $m = 6$ and write $\omega = e^{\frac{2\pi i}{6}}$. Then $\omega^5 = -\omega^2$, i.e. $\omega = -\omega^4$ and $\mathbb{Q}[\omega] = \mathbb{Q}[\omega^2]$. This means that the third and sixth cyclotomic fields are equal.

In the following let $m \in \mathbb{N}$ and write $\omega = e^{\frac{2\pi i}{m}}$.

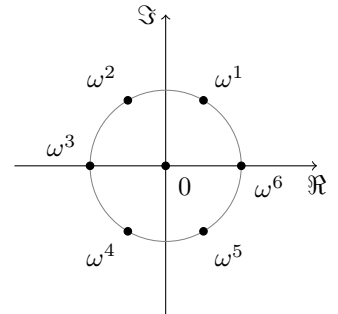
Theorem 1.24

The extension $\mathbb{Q}[\omega]$ over \mathbb{Q} is Galois with degree equal to $\varphi(m)$, where φ is Euler's totient function. Moreover, the Galois group is isomorphic to $(\mathbb{Z}/m\mathbb{Z})^* = \{k \in \mathbb{Z}/m\mathbb{Z} \mid \gcd(k, m) = 1\}$.

For $k \in (\mathbb{Z}/m\mathbb{Z})^*$ the corresponding automorphism is given by $\omega \mapsto \omega^k$.

Proposition 1.25

The conjugates of ω are exactly given by ω^k with $\gcd(m, k) = 1$.



Corollary 1.26

Let $m \in \mathbb{N}$ be even. Then the roots of unity contained in $\mathbb{Q}(e^{\frac{2\pi i}{m}})$ are exactly the m -th roots of unity.

Corollary 1.27

The m -th cyclotomic fields, for m even, are all non-isomorphic.

Theorem 1.28

Let $m = p^r$ for some prime p and $\omega = e^{\frac{2\pi i}{m}}$. Then $\mathcal{O}_{Q[\omega]} = \mathbb{Z}[\omega]$.

Remark: More generally, $\mathbb{Z}[\omega] = \mathcal{O}_{Q[\omega]}$ for every cyclotomic field.

Notation: We write $\text{disc}(\alpha) = \text{disc}(1, \alpha, \dots, \alpha^{n-1})$.

Lemma 1.29

For $m \in \mathbb{N}$ we have $\text{disc}(\omega) \mid m^{\varphi(m)}$.

Lecture 5,
10.11.2023

Lemma 1.30

For $m \geq 3$ we have $\text{disc}(1 - \omega) = \text{disc}(\omega)$.

Lemma 1.31

Let $m = p^r$ be a prime power, $r \in \mathbb{N}$. Then

$$\prod_{\substack{1 \leq k \leq m \\ \gcd(k, m) = 1}} (1 - \omega^k) = p.$$

Remark: In particular for $m = p^r$ we have $\frac{p}{(1 - \omega)^{\varphi(m)}} \in \mathbb{Z}[\omega]$.

2 Prime ideal factorisation

2.1 Unique prime ideal factorisation

Motivation: If K is a number field with ring of integers \mathcal{O}_K , then we may not have a unique factorisation in \mathcal{O}_K into irreducible elements (up to units and ordering).

Example: Let $K = \mathbb{Q}(\sqrt{-5})$, then $\mathcal{O}_K = \mathbb{Z}[\sqrt{-5}]$. In $\mathbb{Z}[\sqrt{-5}]$ we have $2 \cdot 3 = 6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, where 2 and 3 are irreducible elements.

Our next goal is to replace factorisation into irreducible elements by prime *ideal* factorisation. Instead of number fields, we consider more generally *Dedekind domains*.

Definition (Integrally closed ring)

Let R be an integral domain and $K = \left\{ \frac{a}{b} \mid a, b \in R, b \neq 0 \right\}$ its field of fractions. We call R *integrally closed*, if every element $\frac{a}{b} \in K$, which is a zero of a monic polynomial with coefficients in R is contained in R .

Example: Let K be a number field with ring of integers \mathcal{O}_K . Then \mathcal{O}_K is integrally closed. Indeed let $\alpha \in K$ satisfy $\alpha^n + b_1\alpha^{n-1} + \dots + b_n = 0$, with $b_1, \dots, b_n \in \mathcal{O}_K$. Then $\mathbb{Z}[\alpha, b_1, \dots, b_n]$ is finitely generated as an additive group and we have $\alpha \in \mathcal{O}_K$.

Definition (Noetherian¹ ring)

We call a commutative ring R *noetherian* if every ideal is finitely generated.

Remark: The following statements about a commutative ring R are equivalent:

1. R is noetherian.
2. Every increasing sequence of ideals is eventually constant, i.e. if $I_1 \subseteq I_2 \subseteq \dots$, then there is some $n_0 \in \mathbb{N}$, such that $I_n = I_{n_0}$ for every $n > n_0$.
3. Every non-empty set S of ideals has a maximal element, i.e. there is some $M \in S$, such that if $M' \in S$ with $M \subseteq M'$, then $M = M'$.

¹after Emmy Noether (1882 - 1935), a German mathematician

Example: Principal ideal domains and polynomial rings $\mathbb{Z}[x_1, \dots, x_n]$ or $K[x_1, \dots, x_n]$ for any field K are noetherian.

Definition (Dedekind² domain)

A *Dedekind domain* is a noetherian integrally closed domain, in which every non-zero prime ideal is maximal.

Theorem 2.1

Let K be a number field. Then its ring of integers \mathcal{O}_K is a Dedekind domain.

Example: Coordinate rings of irreducible smooth curves over an algebraically closed field, e.g. $\mathbb{C}[T]$ is a Dedekind domain.

First properties of Dedekind domains

Lemma 2.2

Let R be a Dedekind domain, which is not a field, and $0 \neq I \subseteq R$ an ideal. Then I contains a product of non-zero prime ideals $P_1 \cdots P_k \subseteq I$.

Lemma 2.3

Let R be a Dedekind domain with field of fractions K and $0 \neq I \subsetneq R$ an ideal. Then there exists $\alpha \in K \setminus R$ with $\alpha I \subseteq R$.

Theorem 2.4

Let R be a Dedekind domain and $0 \neq I \subseteq R$ an ideal. Then there is an ideal $0 \neq J \subseteq R$, such that IJ is principal.

Example: Let $R = \mathbb{Z}[\sqrt{-5}]$ and $I = (2, 1 + \sqrt{-5})$. Then I is not principal, but $(2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}) = (2)$ is principal.

Observation: Note that $\alpha \in I$ implies that $J \subset A = \frac{1}{\alpha}IJ$. Hence $\gamma JI = \gamma\alpha \left(\frac{1}{\alpha}JI\right) = \alpha\gamma A \subseteq (\alpha)$. As $\gamma J \subseteq \gamma A \subseteq R$, we find that $\gamma J \subseteq J$.

²after Richard Dedekind (1831 - 1916), a German mathematician

The ideal class group

Definition (Equivalence of ideals)

Let R be an integral domain. We say that two non-zero ideals I, J are equivalent if and only if there exist $\alpha, \beta \in R \setminus \{0\}$ with $\alpha I = \beta J$.

Remark: 1. This really is an equivalence relation. We call the equivalence classes under this relation *ideal classes*.

2. We can define a multiplication on the set of ideal classes by multiplication of representatives, $[I][J] = [IJ]$, with the neutral element $[R]$.

3. All principal ideals form one ideal class.

Corollary 2.5

Let R be a Dedekind domain. Then the ideal classes form a group under multiplication.

Definition (Ideal class group)

We call the group given by ideal classes under multiplication in the Dedekind domain R the *ideal class group* of R , denoted by $Cl(R)$.

Example: \mathbb{Z} is a principal ideal domain, hence $|Cl(\mathbb{Z})| = 1$.

Remark: There are only finitely many imaginary quadratic fields K with $|Cl(\mathcal{O}_K)| = 1$.

Question (Gauss): Do there exist as many real quadratic number fields K with $|Cl(\mathcal{O}_K)| = 1$?

Corollary 2.6

Let R be a Dedekind domain and A, B, C ideals with $A \neq 0$.

1. *If $AB = AC$ then $B = C$.*
2. *We have $B \mid A$, i.e. $A = BJ$ for some ideal J , if and only if $A \subseteq B$.*

Theorem 2.7 (Unique prime ideal factorisation)

Every ideal $I \neq 0$ in a Dedekind domain R can be written as a product $I = P_1 \cdots P_r$

with non-zero prime ideals P_1, \dots, P_r and this representation is unique up to ordering of P_1, \dots, P_r .

Example: In $\mathbb{Z}(\sqrt{-5})$ we don't have unique factorisation into reducible elements, e.g. $2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, but in terms of ideals we have $(2) = (2, 1 + \sqrt{-5})^2 = P_1^2$, $(3) = (3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5}) = P_2 \cdot P_3$. Note that P_1, P_2, P_3 are all prime ideals as $|\mathbb{Z}[\sqrt{-5}]/P_i| \in \{2, 3\}$ for $1 \leq i \leq 3$. In the ideal class group we find that

$$\begin{aligned} (2) \cdot (3) &= P_1^2 P_2 P_3 \\ &= P_1 P_2 P_1 P_3 \\ &= (1 + \sqrt{-5})(1 - \sqrt{-5}). \end{aligned}$$

Definition (Greatest common divisor, least common multiple)

Let R be a Dedekind domain and $I, J \neq 0$ ideals with prime factorisation

$$I = \prod_{i=1}^r P_i^{a_i}, \quad J = \prod_{i=1}^r P_i^{b_i},$$

where P_1, \dots, P_r are distinct prime ideals and $a_1, \dots, a_r, b_1, \dots, b_r \in \mathbb{Z}_{\geq 0}$. We define the *greatest common divisor* $\gcd(I, J)$ and *least common multiple* $\text{lcm}(I, J)$ by

$$\gcd(I, J) = \prod_{i=1}^r P_i^{\min(a_i, b_i)}, \quad \text{lcm}(I, J) = \prod_{i=1}^r P_i^{\max(a_i, b_i)}.$$

Exercise

Show that

$$\gcd(I, J) = I + J, \quad \text{lcm}(I, J) = I \cap J.$$

Question: Given the ring of integers \mathcal{O}_K in a number field K , we know that every ideal is finitely generated. Can we say something about the numbers of generators we need? E.g. in $\mathbb{Z}[\sqrt{-5}]$, the prime ideal $(2, 1 + \sqrt{-5})$ is not a principal ideal, but generated by two elements.

Remark: Chinese Remainder Theorem: Let R be a commutative ring with 1 and

a_1, \dots, a_n coprime ideals, i.e. $a_i + a_j = R \forall i \neq j$. Then there is an isomorphism

$$R / \bigcap_{i=1}^n a_i \rightarrow R/a_1 \times \cdots \times R/a_n.$$

Theorem 2.8

Let R be a Dedekind domain, $I \subseteq R$ a non-zero ideal and $\alpha \in I \setminus \{0\}$. Then there exists $\beta \in I$ with $I = (\alpha, \beta)$.

Corollary 2.9

A Dedekind domain is a unique factorisation domain (UFD) if and only if it is a principal ideal domain (PID).

Remark: In general, a PID is a UFD but the reverse implication does not hold. For example $\mathbb{Z}[x]$ is a UFD, but not a PID.

2.2 Splitting of primes

Let p be a (rational) prime number. Then (p) is a prime ideal in \mathbb{Z} , but the ideal $(p) = p\mathcal{O}_K$ need not be a prime ideal in \mathcal{O}_K . For example, let $p \equiv 1 \pmod{4}$, then in $\mathbb{Z}[i]$ we have

$$(p) = (a + ib)(a - ib), \quad (2.1)$$

where $a^2 + b^2 = p$ with $a, b \in \mathbb{Z}$. Note that $N_{\mathbb{Q}[i]/\mathbb{Q}}(a + ib) = p$ and hence $a + ib$ is a prime element in the PID $\mathbb{Z}[i]$, and (2.1) is the prime ideal factorisation of (p) . Moreover, $a + ib$ and $a - ib$ do not differ by multiplication with one of the units $\pm 1, \pm i$, and hence

$$P_1 = (a + ib) \neq (a - ib) = P_2$$

in $\mathbb{Z}[i]$. The ideal (2) splits in $\mathbb{Z}[i]$ as $2 = (1 + i)^2$, where $(1 + i)$ is a prime ideal. If $p \equiv 3 \pmod{4}$ is a rational prime, then (p) remains a prime ideal in $\mathbb{Z}[i]$. (check!)

Question: More generally, let $K \subseteq L$ be number fields with rings of integers $\mathcal{O}_K, \mathcal{O}_L$. Given a non-zero prime ideal P in \mathcal{O}_K , how does $P\mathcal{O}_L$ split into prime ideals in \mathcal{O}_L ?

Notation: In the following, we keep the notation $K \subseteq L$, $\mathcal{O}_K \subseteq \mathcal{O}_L$ as above.

Definition (Primes)

We say that $P \subseteq \mathcal{O}_K$ or $Q \subseteq \mathcal{O}_L$ is a *prime* if P or respectively Q is a non-zero

prime ideal in \mathcal{O}_K or respectively \mathcal{O}_L . Moreover, we say that Q lies above P or P lies under Q if $Q \mid P\mathcal{O}_L$.

Lemma 2.10

Let P resp. Q be primes in \mathcal{O}_K resp. \mathcal{O}_L . Then Q lies above P if and only if one of the following equivalent conditions holds:

1. $P\mathcal{O}_L \subseteq Q$.
2. $P \subseteq Q$.
3. $Q \cap \mathcal{O}_K = P$.
4. $Q \cap K = P$.

Theorem 2.11

Every prime Q in \mathcal{O}_L lies above a unique prime P in \mathcal{O}_K and for every prime P in \mathcal{O}_K there is some prime Q in \mathcal{O}_L , which lies above P .

Lemma 2.12

Let Q be a prime in \mathcal{O}_L lying above P in \mathcal{O}_K . Then \mathcal{O}_L/Q and \mathcal{O}_K/P are finite fields with $\mathcal{O}_K/P \hookrightarrow \mathcal{O}_L/Q$.

Let P be a prime in \mathcal{O}_K and consider in \mathcal{O}_L the prime ideal factorisation

$$P\mathcal{O}_L = \prod_{i=1}^r Q_i^{e_i}$$

with distinct primes Q_1, \dots, Q_r .

Definition (Ramification index, inertia degree)

We call

$$e_i = e(Q_i \mid P)$$

the *ramification index* of Q_i above P and

$$f_i = f(Q_i \mid P) = [\mathcal{O}_L/Q_i : \mathcal{O}_K/P]$$

the *inertia degree* of Q_i over P . Moreover, we call \mathcal{O}_L/Q_i and \mathcal{O}_K/P *residue fields* of Q_i or respectively P .

Remark: Let $K \subseteq L \subseteq M$ be number fields with primes $P \subseteq Q \subseteq R$. Then

$$e(R | P) = e(R | Q)e(Q | P), \quad f(R | P) = f(R | Q)f(Q | P).$$

Example: Let $K = \mathbb{Q}$, $L = \mathbb{Q}(i)$. If p is a rational prime with $p \equiv 1 \pmod{4}$, then $(p) = P_1 \cdot P_2$, $P_1 = (a + ib)$, $P_2 = (a - ib)$ for some $a, b \in \mathbb{Z}$. We have

$$e(P_i | (p)) = 1, \quad f(P_i | (p)) = 1.$$

For a rational prime $p \equiv 3 \pmod{4}$ we obtain

$$e\left(\underbrace{(p)}_{\subseteq \mathbb{Z}[i]} \mid \underbrace{(p)}_{\subseteq \mathbb{Z}}\right) = 1, \quad f\left(\underbrace{(p)}_{\subseteq \mathbb{Z}[i]} \mid \underbrace{(p)}_{\subseteq \mathbb{Z}}\right) = 2.$$

For $p = 2$ note that $(2) = (1 + i)^2$ and $|\mathbb{Z}[i] | (1 + i)| = 2$, hence

$$e((1 + i) | (2)) = 2, \quad f((1 + i) | (2)) = 1.$$

In this example, independent of the rational prime p we find that

$$\sum_{i=1}^r e_i f_i = [\mathbb{Q}(i) : \mathbb{Q}].$$

Our goal now is to show the above statement for number fields $K \subseteq L$.

Lecture 8,
24.11.2023

Norms of ideals

Definition (Norm of an ideal)

Let K be a number field and $I \subseteq \mathcal{O}_K$ a non-zero ideal. Then we define the *norm* $N(I)$ of the ideal I as

$$N(I) := |\mathcal{O}_K / I|.$$

Lemma 2.13

Let $I, J \subseteq \mathcal{O}_K$ be non-zero ideals. Then

$$N(IJ) = N(I)N(J).$$

Proposition 2.14

Let K be a number field of degree $n = [K : \mathbb{Q}]$ and $p \in \mathbb{Z}$ a prime with prime ideal

factorisation

$$(p) = \prod_{i=1}^r P_i^{e_i}$$

in \mathcal{O}_K and $f_i = f(P_i \mid p)$ for $1 \leq i \leq r$. Then

$$\sum_{i=1}^r e_i f_i = n.$$

Next, we will look at general number field extensions $L \subseteq K$. We start with some preparations:

Lemma 2.15

Let $0 \neq B \subseteq A \subsetneq R$ be ideals in a Dedekind domain R . Then there exists $\alpha \in K = \text{Quot}(R)$, such that

$$\alpha B \subseteq R, \text{ but } \alpha B \subsetneq A.$$

Lemma 2.16

Let $I \neq 0$ be an ideal in \mathcal{O}_K and $n = [L : K]$. Then

$$N(I\mathcal{O}_L) = N(I)^n.$$

Example: For $K = \mathbb{Q}$ we have already used this identity above, in which case it reduces to

$$N((p)) = p^n,$$

with $(p) \subseteq \mathcal{O}_L$ and p a rational prime.

Theorem 2.17

Let P be a prime in \mathcal{O}_K and $P\mathcal{O}_L = \prod_{i=1}^r Q_i^{e_i}$ the prime ideal factorisation in \mathcal{O}_L with distinct ideals Q_1, \dots, Q_r and inertia degrees $f_i = f(Q_i \mid P)$. Then

$$[L : K] = \sum_{i=1}^r e_i f_i.$$

Example: (a) Let p be a rational prime and $\omega = e^{\frac{2\pi i}{p^r}}$ for some $r \in \mathbb{N}$. By Lemma 1.31 we have

$$p = \prod_{\substack{1 \leq k \leq m \\ \gcd(k, m) = 1}} (1 - \omega^k).$$

We show on the exercise sheet that for $p \nmid k$

$$(1 - \omega^k) = u_k(1 - \omega)$$

for some $u_k \in \mathbb{Z}[\omega]$. Hence in $\mathbb{Z}[\omega]$ we have

$$(p) = (1 - \omega)^{\varphi(p^r)}.$$

By Theorem 2.17, we deduce that $(1 - \omega)$ is a prime ideal in $\mathbb{Z}[\omega]$ and

$$f((1 - \omega) \mid (p)) = 1$$

- (b) Let α be a root of $\alpha^3 = \alpha + 1$. Then $\mathbb{Q}(\alpha)/\mathbb{Q}$ is an extension of degree 3. One can compute $\text{disc}(1, \alpha, \alpha^2) = -23$. As 23 is square-free, we find that $\mathcal{O}_{\mathbb{Q}(\alpha)} = \mathbb{Z}[\alpha]$ with integral basis $(1, \alpha, \alpha^2)$. Moreover, in $\mathbb{Z}[\alpha]$, we have

$$23 \cdot \mathbb{Z}[\alpha] = (23, \alpha - 10)^2(23, \alpha - 3), \quad (2.2)$$

where $(23, \alpha - 10)$ and $(23, \alpha - 3)$ are coprime. Hence (2.2) is the prime ideal factorisation of (23) in $\mathbb{Z}[\alpha]$ and

$$f((23, \alpha - 10) \mid 23) = f((23, \alpha - 3) \mid 23) = 1.$$

Remark: In these examples we have found ramification indices $e > 1$, which however is not the "typical" case, as we will see below.

Definition (Ramified prime)

Let P be a prime in \mathcal{O}_K . We say that P is *ramified in \mathcal{O}_L* , if there is a prime Q in \mathcal{O}_L , lying above P , with

$$e(Q \mid P) > 1.$$

Theorem 2.18

Let p be a rational prime (i.e. a prime number in \mathbb{Z}), which is ramified in \mathcal{O}_K . Then

$$p \mid \text{disc}(\mathcal{O}_K).$$

Remark: One can even show, that $p \mid \text{disc}(\mathcal{O}_K)$ implies that p is ramified in \mathcal{O}_K .

Corollary 2.19

There are only finitely many primes P in \mathcal{O}_K which are ramified in \mathcal{O}_L .

Lecture 9,
28.11.2023

Galois extensions

In the proof of Theorem 2.18 we noted that if L/\mathbb{Q} is a Galois extension and Q a prime in \mathcal{O}_L above $p \in \mathbb{Z}$, so is the ideal $\sigma(Q)$ for all $\sigma \in \text{Gal}(L/\mathbb{Q})$.

Theorem 2.20

Let L/K be Galois and Q a prime in \mathcal{O}_L lying above the prime P in \mathcal{O}_K . Then $\sigma(Q)$ is a prime above P for every $\sigma \in \text{Gal}(L/K)$. Moreover, if Q' is another prime in \mathcal{O}_L over P , then there exists an automorphism $\sigma \in \text{Gal}(L/K)$ with $\sigma(Q) = Q'$.

Example: $K = \mathbb{Q}$, $L = \mathbb{Q}(i)$, $p \in \mathbb{Z}$ a prime with $p \equiv 1 \pmod{4}$. Write $p = a^2 + b^2$ with $a, b \in \mathbb{Z}$. In $\mathbb{Z}[i]$ we have $(p) = (a + ib)(a - ib)$.

Corollary 2.21

Let L/K be a Galois extension, P a prime in \mathcal{O}_K and Q_1, Q_2 primes in \mathcal{O}_L lying above P . Then

$$e(Q_1 | P) = e(Q_2 | P), \quad f(Q_1 | P) = f(Q_2 | P).$$

Remark: In the notation above, we hence obtain

$$P\mathcal{O}_L = (Q_1 \cdots Q_r)^e \text{ with } f(Q_i | P) = f(Q_j | P).$$

Question: Let L/K be any number fields (not necessarily Galois) and P a prime in \mathcal{O}_K . Find explicitly the factorisation

$$P\mathcal{O}_L = \prod_{i=1}^r Q_i^{e_i}$$

with Q_1, \dots, Q_r prime.

Example: Let $m \in \mathbb{Z} \setminus \{1\}$ be odd and square-free and let $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt{m})$. Consider an odd prime $p \in \mathbb{Z}$ with $p \nmid m$. By Theorem 2.18, p is not ramified in \mathcal{O}_K as $\text{disc}(K) \in \{m, 4m\}$. Hence we either have $p\mathcal{O}_L = Q_i Q_2$ with distinct primes

Q_1, Q_2 and $f(Q_i | p) = 1$ for $i = 1, 2$, or $p\mathcal{O}_L$ is prime with $f(p\mathcal{O}_L | p) = 2$.

Let Q be a prime above p . Consider the polynomial $g(X) = X^2 - m$. Then $g(X)$ has a zero in \mathcal{O}_L and hence a zero in \mathcal{O}_L/Q .

1. If m is not a square modulo p , then $X^2 - m$ has no zero in $\mathbb{Z}/p\mathbb{Z}$ and $\mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathcal{O}_L/Q$ is a non-trivial field extension, i.e. $f(Q | p) = 2$.
2. Let $a \in \mathbb{Z}$ be a solution to $a^2 - m \equiv 0 \pmod{p}$. Then in \mathcal{O}_L we have the factorisation $(a - \sqrt{m})(a + \sqrt{m}) \in p\mathcal{O}_L$ and in fact

$$(p, a - \sqrt{m})(p, a + \sqrt{m}) = p\mathcal{O}_L. \quad (2.3)$$

As neither of the factors $(p, a - \sqrt{m}), (p, a + \sqrt{m})$ contains 1, and $p\mathcal{O}_L$ factors into a product of at most two primes, we have already found in (2.3) the prime ideal factorisation of $p\mathcal{O}_L$ and

$$f((p, a \pm \sqrt{m}) | p) = 1.$$

More generally, let L/K be number fields, say of degree $n = [L : K]$. Fix an element $\alpha \in \mathcal{O}_L$, such that $L = K(\alpha)$. Note, that by Proposition 1.22 the quotient $\mathcal{O}_L/\mathcal{O}_K[\alpha]$ is finite. Let $g(X) \in \mathcal{O}_K[X]$ be the minimal polynomial of α over K .

Theorem 2.22

With notation as above, let P be a prime in \mathcal{O}_K and factor $g(X)$ in $(\mathcal{O}_K/P)[X]$ as

$$g(X) \equiv g_1(X)^{e_1} \cdots g_r(X)^{e_r} \pmod{P[X]},$$

where $g_1(X), \dots, g_r(X) \in \mathcal{O}_K[X]$ are monic polynomials, pairwise distinct and irreducible in $(\mathcal{O}_K/P)[X]$. Let $(p) \in P \cap \mathbb{Z}$ and assume $p \nmid |\mathcal{O}_L/\mathcal{O}_K[\alpha]|$. Then we have the factorisation

$$P\mathcal{O}_L = \prod_{i=1}^r Q_i^{e_i},$$

where $Q_i = (P, g_i(\alpha))$ is a prime and $f(Q_i | P) = \deg g_i$ for $1 \leq i \leq r$.

Example: Let α be a root of $\alpha^3 - \alpha - 1 = 0$. We have from earlier that $\mathcal{O}_{\mathbb{Q}[\alpha]} = \mathbb{Z}[\alpha]$ and $\text{disc}(\mathbb{Q}[\alpha]) = -23$. Modulo 23 we find that

$$X^3 - X - 1 \equiv (X - 10)^2(X - 3) \pmod{23}$$

and hence by Theorem [2.22](#)

$$23\mathbb{Z}[\alpha] = (23, \alpha - 10)^2(23, \alpha - 3).$$

3 Number fields - Dirichlet's unit theorem, class groups and lattices

3.1 Finiteness of the ideal class group

Let K be a number field with ring of integers \mathcal{O}_K . We will keep this notation throughout this chapter.

Recall: We call two non-zero ideals $I, J \subseteq \mathcal{O}_K$ equivalent, if $\exists \alpha, \beta \in \mathcal{O}_K \setminus \{0\}$, such that $\alpha I = \beta J$, and we write $Cl(\mathcal{O}_K)$ for the group of equivalence classes under multiplication.

Question: Is $Cl(\mathcal{O}_K)$ finite?

Theorem 3.1

For every number field K there is a constant C_K , such that every non-zero ideal I contains an element $\alpha \in I \setminus \{0\}$ with

$$|N_{K/\mathbb{Q}}(\alpha)| \leq C_K N(I).$$

Corollary 3.2

Let K and C_K be as in Theorem 3.1. Then every ideal class $C \in Cl(\mathcal{O}_K)$ contains an ideal I with $N(I) \leq C_K$.

Corollary 3.3

For every number field K we have $|Cl(\mathcal{O}_K)| < \infty$.

Example: Let $K = \mathbb{Q}[\sqrt{2}]$, i.e. $\mathcal{O}_K = \mathbb{Z}[\sqrt{2}]$. As in the proof of Theorem 3.1, we can take $C_K = (1 + \sqrt{2})^2$ (using the integral basis $(1, \sqrt{2})$). Note that $(1 + \sqrt{2})^2 < 6$. We consider the prime ideals in $\mathbb{Z}[\sqrt{2}]$, which lie above 2, 3, 5. Note that $2\mathbb{Z}[\sqrt{2}] = (\sqrt{2})^2$ and that $(3), (5)$ are prime ideals (see Theorem 2.22, noting that $X^2 - 2$ remains

irreducible modulo 3, 5). Hence $|Cl(\mathbb{Z}[\sqrt{2}])| = 1$.

Remark: In the example above and other examples, we would like to take C_K as small as possible.

Our next goal will be to find improvements for the value of C_K using results from the geometry of numbers.

Idea: Let K be a number field of degree n , $\sigma_1, \dots, \sigma_r : K \hookrightarrow \mathbb{R}$ its real embeddings and $\tau_1, \bar{\tau}_1, \tau_2, \bar{\tau}_2, \dots, \tau_s, \bar{\tau}_s : K \hookrightarrow \mathbb{C}$ its different complex embeddings, where we sort them into pairs $\tau_i, \bar{\tau}_i$, which differ by complex conjugations. Then $n = r + 2s$ and we can define an injective map

$$\varphi : K \rightarrow \mathbb{R}^n, \quad \alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \Re \tau_1(\alpha), \Im \tau_1(\alpha), \dots, \Re \tau_s(\alpha), \Im \tau_s(\alpha)).$$

Let $(\alpha_1, \dots, \alpha_n)$ be an integral basis of \mathcal{O}_K . Then we can view $\varphi(\mathcal{O}_K) = \mathbb{Z}\varphi(\alpha_1) + \dots + \mathbb{Z}\varphi(\alpha_n) \subseteq \mathbb{R}^n$ as an additive group. Also, if $I \subseteq \mathcal{O}_K$ is a non-zero ideal, then I is a free \mathbb{Z} -module of rank n , say with basis $(\beta_1, \dots, \beta_n)$. Then

$$\varphi(I) = \mathbb{Z}\varphi(\beta_1) + \dots + \mathbb{Z}\varphi(\beta_n) \subseteq \mathbb{R}^n$$

and we can interpret $\varphi(I)$ as a *lattice* in \mathbb{R}^n . In order to improve upon C_K in Theorem 3.1, we would like to find a "small" non-zero element in this lattice.

Lecture 11,
05.12.2023

3.2 Geometry of numbers

Motivation: Consider a lattice L , e.g. $\mathbb{Z}^n \subseteq \mathbb{R}^n$, and a "nice" subset $C \subseteq \mathbb{R}^n$, e.g. a ball of radius r . When does C contain a point in $L \setminus \{0\}$?

Definition (Lattice)

Let $v_1, \dots, v_n \in \mathbb{R}^n$ be linearly independent vectors (over \mathbb{R}). Then we call the group

$$L = \{z_1 v_1 + \dots + z_n v_n \mid z_1, \dots, z_n \in \mathbb{Z}\} \subseteq \mathbb{R}^n$$

a (full) *lattice* in \mathbb{R}^n and v_1, \dots, v_n a basis of L . We define the determinant $d(L)$ of the lattice L as

$$d(L) = |\det(v_1, \dots, v_n)|.$$

Remark: As additive groups we have $L \cong \mathbb{Z}^n$. If $x \in L$ and v_1, \dots, v_n as above, then there is exactly one way to write x as $\sum_{i=1}^n x_i v_i$ with $x_1, \dots, x_n \in \mathbb{Z}$.

Notation: We write $M_{n \times n}(\mathbb{Z})$ for the set of $n \times n$ matrices with coefficients in \mathbb{Z} . and $GL(n, \mathbb{Z}) = \{A \in M_{n \times n}(\mathbb{Z}) \mid \det(A) = \pm 1\}$ for the group of invertible matrices in $M_{n \times n}(\mathbb{Z})$.

Lemma 3.4

Let $L \subseteq \mathbb{R}^n$ be a lattice and $\{v_1, \dots, v_n\}, \{w_1, \dots, w_n\}$ bases of L . Then there exists a matrix $A \in GL(n, \mathbb{Z})$, say $A = (a_{i,j})_{1 \leq i,j \leq n}$, such that

$$w_i = \sum_{j=1}^n a_{i,j} v_j, \quad 1 \leq i \leq n.$$

Moreover,

$$|\det(v_1, \dots, v_n)| = |\det(w_1, \dots, w_n)|.$$

Remark: In particular, the determinant $d(L)$ of a lattice $L \subseteq \mathbb{R}^n$ is well-defined.

Next, we want to compare the relative "size" of two lattices $M \subseteq L \subseteq \mathbb{R}^n$. Let $L = \{\sum_{i=1}^n z_i v_i \mid z_1, \dots, z_n \in \mathbb{Z}\}$ and $M = \{\sum_{i=1}^n t_i w_i \mid t_1, \dots, t_n \in \mathbb{Z}\}$ with $M \subseteq L$. Then $w_i \in L \forall 1 \leq i \leq n$ and hence there exists an $a_{i,j} \in \mathbb{Z}$ with $w_i = \sum_{j=1}^n a_{i,j} v_j \forall 1 \leq i \leq n$. Let $A = (a_{i,j})_{1 \leq i,j \leq n} \in M_{n \times n}(\mathbb{Z})$.

Definition (Index of a sublattice)

In the notation above, we define the *index* $[L : M]$ of M in L as

$$[L : M] = |\det(A)|.$$

Remark: 1. The index $[L : M]$ does not depend on the choice of bases of L, M .

By $w_i = \sum_{j=1}^n a_{i,j} v_j$, we have

$$\underbrace{|\det(w_1, \dots, w_n)|}_{d(M)} = |\det(A)| \underbrace{|\det(v_1, \dots, v_n)|}_{d(L)},$$

and hence $[L : M] = \frac{d(M)}{d(L)}$.

2. One can show that $[L : M] = |L/M|$, where L/M is the quotient group.

Example: Let e_1, \dots, e_n be the unit vectors in \mathbb{R}^n , i.e. $e_i = (0, \dots, 0, 1, 0, \dots, 0)$.

1. $\mathbb{Z}^n = \{\sum_{i=1}^n e_i z_i \mid z_1, \dots, z_n \in \mathbb{Z}\}$ is a lattice with $d(\mathbb{Z}^n) = 1$. Let $d_1, \dots, d_n \in \mathbb{N}$ and set $w_i = d_i e_i$ for all $1 \leq i \leq n$. Then $M = \{\sum_{i=1}^n z_i w_i \mid z_1, \dots, z_n \in \mathbb{Z}\} \subseteq \mathbb{Z}^n$ is a sublattice with $d(M) = |\det(d_1 e_1, \dots, d_n e_n)| = d_1 \cdots d_n$ and $[\mathbb{Z}^n : M] = d_1 \cdots d_n$. Hence, as abelian groups, $\mathbb{Z}^n / M \cong \mathbb{Z}/d_1 \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_n \mathbb{Z}$.
2. $L = \left\{ \frac{a_1}{2} e_1 + \cdots + \frac{a_n}{2} e_n \mid a_1, \dots, a_n \in \mathbb{Z}, a_1 \equiv \cdots \equiv a_n \pmod{2} \right\}$ is a lattice in \mathbb{R}^n with basis $e_1, \dots, e_{n-1}, \frac{e_1 + \cdots + e_n}{2}$.

Convex bodies

Definition (Convex set)

We call a subset $C \subseteq \mathbb{R}^n$ *convex* if for all $x, y \in C$ the line segment

$$\{tx + (1-t)y \mid 0 \leq t \leq 1\}$$

is contained in C as well.

Definition (Central symmetric convex body)

A subset $C \subseteq \mathbb{R}^n$ is called a *central symmetric convex body* if it has the following properties:

- (a) C is compact (i.e. closed and bounded) and convex. (convex body)
- (b) 0 is in the interior of C . (central)
- (c) If $x \in C$, then $-x \in C$. (symmetric)

Example: 1. Let $C \subseteq \mathbb{R}^n$ be a central symmetric convex body and $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ an invertible linear map. Then $A(C)$ is a central symmetric convex body.

2. The norm $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}$ leads to the n -dimensional unit ball

$$B_n = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1\}.$$

$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ induces the n -dimensional unit cube

$$K_n = \left\{ x \in \mathbb{R}^n \mid \max_{1 \leq i \leq n} |x_i| \leq 1 \right\}.$$

$\|x\|_1 = \sum_i^n |x_i|$ give the n -dimensional unit octahedron

$$O_n = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n |x_i| \leq 1 \right\}.$$

Lemma 3.5

Let $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}^n$ be a norm. Then $B_{\|\cdot\|} = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ is a central symmetric convex body.

So far we have found that every norm on \mathbb{R}^n "produces" a central symmetric convex body in \mathbb{R}^n . Is there a one-to-one correspondence, i.e. are these all the different classes of central symmetric convex bodies?

Remark: Let $C \subseteq \mathbb{R}^n$ be a central symmetric convex body. For $\lambda \geq 0$, set $\lambda C = \{\lambda x \mid x \in C\}$. If $\lambda > 0$, then λC is again a central symmetric body. For $x \in \mathbb{R}^n$, we define $\|x\|_C = \min \{\lambda \in \mathbb{R}_{\geq 0} \mid x \in \lambda C\}$.

Lemma 3.6

Using the same notation as above, the following statements hold:

1. $\|\cdot\|_C$ is well-defined.
2. $\|\cdot\|_C$ defines a norm on \mathbb{R}^n .
3. $\lambda C = \{x \in \mathbb{R}^n \mid \|x\|_C \leq \lambda\}$ for $\lambda > 0$.

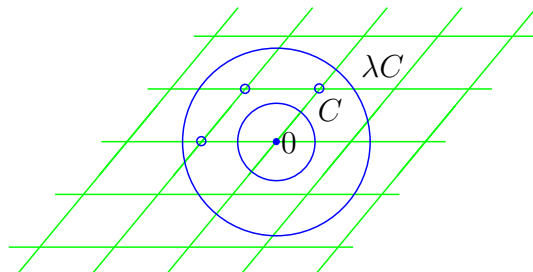
In particular, we recover C via $C = \{x \in \mathbb{R}^n \mid \|x\|_C \leq 1\}$.

Minkowski's¹ first convex body theorem

Let $L \subseteq \mathbb{R}^n$ be a lattice and $C \subseteq \mathbb{R}^n$ a central symmetric convex body. When is $C \cap L \neq \{0\}$, i.e. when does C contain more lattice points than just 0?

Theorem 3.7 (Minkowski's first convex body theorem, 1896)

With the same notation as above, let $\text{vol}(C) \geq 2^n d(L)$. Then $C \cap L \neq \{0\}$, i.e. there exists a $x \in L \setminus \{0\}$ with $x \in C$.



¹after Hermann Minkowski (1864 - 1909), a German mathematician

Lecture 12,

08.12.2023 **Notation:** For a lattice $L \subseteq \mathbb{R}^n$ with basis v_1, \dots, v_n , we define

$$F = \left\{ \sum_{i=1}^n x_i v_i \mid 0 \leq x_i \leq 1 \ \forall 1 \leq i \leq n \right\}$$

as the *fundamental parallelepiped* for L . This is the fundamental domain for \mathbb{R}^n/L . (see below)

Example: $[0, 1]^n$ is the fundamental parallelepiped for \mathbb{Z}^n .

Remark: A fundamental parallelepiped depends on the choice of basis v_1, \dots, v_n , but we have $\text{vol}(F) = |\det(v_1, \dots, v_n)| = d(L)$.

Lemma 3.8

Using the notation as above and for $u \in \mathbb{R}^n$ we write $u + F = \{u + x \mid x \in F\}$. Then

$$\mathbb{R}^n = \bigcup_{u \in L} (u + F)$$

is a disjunction.

Remark: Recall Landau's O -notation: Let $f, g, h : \mathbb{R}_{\geq x_0} \rightarrow \mathbb{R}$ for some $x_0 \in \mathbb{R}$. We write $f(x) = g(x) + O(h(x))$ if there exists $x_1 \geq x_0$ and $C \geq 0$, such that

$$|f(x) - g(x)| \leq Ch(x) \quad \forall x > x_1.$$

Example: $x^{-1} = O(1)$, $[x] = x + O(1)$, $(x+a)^n = x^n + O(x^{n-1})$ for any $a \in \mathbb{R}$, $n \in \mathbb{N}$, $(x+1)^{\frac{1}{2}} = x^{\frac{1}{2}} + O(x^{-\frac{1}{2}})$

Lemma 3.9

Let $L \subseteq \mathbb{R}^n$ be a lattice and $C \subseteq \mathbb{R}^n$ a central symmetric convex body. Then, as $\lambda \rightarrow \infty$, we have

$$|\lambda C \cap L| = \frac{\text{vol}(C)}{d(L)} \lambda^n + O(\lambda^{n-1}).$$

Question: Do we need C to be central symmetric or convex in Minkowski's theorem?

Minkowski's second convex body theorem

Let $L \subseteq \mathbb{R}^n$ be a lattice and $C \subseteq \mathbb{R}^n$ a central symmetric convex body. When is $L \cap C \neq \{0\}$?

Definition (Successive minima)

We let

$$\lambda_1 = \min \{ \lambda > 0 \mid \lambda C \cap L \neq \{0\} \}$$

and for $2 \leq i \leq n$ we define

$$\lambda_i = \min \{ \lambda \in \mathbb{R}_{\geq 0} \mid \lambda C \cap L \text{ contains at least } i \text{ linearly independent points} \}.$$

We call $\lambda_1, \dots, \lambda_n$ the *successive minima* of L with respect to C .

Lemma 3.10

Let $L, C \subseteq \mathbb{R}^n$ be as above. The successive minima $\lambda_1, \dots, \lambda_n$ of L with respect to C are well defined and we have $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n < \infty$. Moreover, there exist linearly independent elements $v_1, \dots, v_n \in L$ with $v_i \in \lambda_i C \forall 1 \leq i \leq n$.

Caveat: The vectors v_1, \dots, v_n from Lemma 3.10 may not be a basis of L . Let

$$L = \left\{ \frac{x_1 e_1 + \dots + x_n e_n}{2} \mid x_i \in \mathbb{Z}, x_1 \equiv \dots \equiv x_n \pmod{2} \right\}.$$

For $n > 4$ and $C = B_n$ the unit ball, we have

$$\left\| \frac{e_1 + \dots + e_n}{2} \right\| = \frac{1}{2} \sqrt{n} > 1,$$

but $\|e_1\|_2 = \dots = \|e_n\|_2 = 1$.

Question: Is there a relation between $d(L)$ and the product $\lambda_1 \dots \lambda_n$?

Example: The lattice $L = \mathbb{Z}d_1 e_1 \oplus \dots \oplus \mathbb{Z}d_n e_n$ with $0 < d_1 \leq \dots \leq d_n$ has with respect to $\|\cdot\|_\infty$ the successive minima $d_1 \leq \dots \leq d_n$ and $d_1 \dots d_n = d(L)$.

Theorem 3.11 (Minkowski's second convex body theorem, 1910)

Let $L \subseteq \mathbb{R}^n$ be a lattice, $C \subseteq \mathbb{R}^n$ a central symmetric convex body and $\lambda_1, \dots, \lambda_n$

successive minima of L with respect to C . Then

$$\frac{1}{n!} \frac{2^n d(L)}{\text{vol}(C)} \leq \lambda_1 \cdots \lambda_n \leq \frac{2^n d(L)}{\text{vol}(C)}$$

Lecture 13,

12.12.2023

Remark: The upper bound is sharp. Take for example $L = \mathbb{Z}^n$ and $C = \{x \in \mathbb{R}^n \mid \|x\|_\infty \leq 1\}$, then $\text{vol}(C) = 2^n$, $d(L) = 1$, $\lambda_1 = \cdots = \lambda_n = 1$. The following example shows that the lower bound is sharp as well.

Example: Let $0 < \lambda_1 \leq \cdots \leq \lambda_n$, $L = \mathbb{Z}^n$, $C = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n \lambda_i |x_i| \leq 1\}$. Then L has successive minima $\lambda_1, \dots, \lambda_n$ with respect to C and $\text{vol}(C) = \frac{2^n}{n!} (\lambda_1 \cdots \lambda_n)^{-1}$.

Minkowski's second convex body theorem implies Minkowski's first convex body theorem. Let L, C be as above and assume that $\text{vol}(C) \geq 2^n d(L)$. Then

$$\lambda_1^n \leq \lambda_1 \cdots \lambda_n \leq \frac{2^n d(L)}{\text{vol}(C)} \leq 1,$$

i.e. $\lambda_1 \leq 1$ and $C \cap L \neq \{0\}$.

Remark: Theorem 3.11 is invariant under linear transformation. Let $L, C, \lambda_1, \dots, \lambda_n$ be as above and $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a linear invertible map. Then $\phi(L)$ is a lattice, $\phi(C)$ is a central symmetric convex body and one can show that $\lambda_1, \dots, \lambda_n$ are the successive minima of $\phi(L)$ with respect to $\phi(C)$ as for $x \in \mathbb{R}^n$ we have $\|x\|_C = \|\phi(x)\|_{\phi(C)}$. We note that

$$\frac{d(\phi(L))}{\text{vol}(\phi(C))} = \frac{|\det \phi| d(L)}{|\det \phi| \text{vol}(C)} = \frac{d(L)}{\text{vol}(C)}.$$

This means it suffices to prove Theorem 3.11 for $L = \mathbb{Z}^n$.

Lemma 3.12

Let $v_1, \dots, v_r \in \mathbb{R}^n$. Then $S = \{\sum_{i=1}^r x_i v_i \mid x_i \in \mathbb{R}, \sum_{i=1}^r |x_i| \leq 1\}$ is the smallest convex subset in \mathbb{R}^n that is symmetric about 0 and contains v_1, \dots, v_r . I.e. S is symmetric about 0 and if $R \subseteq \mathbb{R}^n$ is convex, symmetric about 0 and $v_1, \dots, v_r \in R$, then $S \subseteq R$.

Theorem 3.13

Let $L \subseteq \mathbb{R}^n$ be a lattice. Then there exist $v_1, \dots, v_n \in L$, such that v_1, \dots, v_n are a

basis of L and

$$\|v_1\|_2 \cdots \|v_n\|_2 \leq \left(\frac{4}{3}\right)^{\frac{n(n-1)}{4}} d(L).$$

Remark: This is a weaker version of the upper bound in Theorem 3.11. Our constant $\left(\frac{4}{3}\right)^{\frac{n(n-1)}{4}}$ is in general larger than 2^n (and is for large n actually pretty far off, as the exponent grows in n^2), and each successive minimum λ_i is bounded above by $\|v_i\|_2$, so they might be even smaller.

Corollary 3.14

Let $\lambda_1, \dots, \lambda_n$ be the successive minima of a lattice $L \subseteq \mathbb{R}^n$ with respect to B_n . Then

$$\lambda_1 \cdots \lambda_n \leq \left(\frac{4}{3}\right)^{\frac{n(n-1)}{4}} d(L).$$

Corollary 3.15

Let $E \subseteq \mathbb{R}^n$ be an ellipsoid, symmetric about 0 and $L \subseteq \mathbb{R}^n$ a lattice. Let $\lambda_1, \dots, \lambda_n$ be the successive minima of L with respect to E . Then

$$\lambda_1 \cdots \lambda_n \leq \left(\frac{4}{3}\right)^{\frac{n(n-1)}{4}} V(n) \frac{d(L)}{\text{vol}(E)},$$

where we write $V(n) = \text{vol}(B_n)$.

Theorem (Jordan's² theorem)

Let $C \subseteq \mathbb{R}^n$ be a central symmetric convex body. Then there exists an ellipsoid $E \subseteq \mathbb{R}^n$ with

$$E \subseteq C \subseteq \sqrt{n}E.$$

Corollary 3.16

For all $n \in \mathbb{N}$ there exists a constant $c(n) > 0$ with the following property: Let $L \subseteq \mathbb{R}^n$ be a lattice, $C \subseteq \mathbb{R}^n$ a central symmetric convex body, and $\lambda_1, \dots, \lambda_n$ the successive minima of L with respect to C . Then

$$\lambda_1 \cdots \lambda_n \leq c(n) \frac{d(L)}{\text{vol}(C)}.$$

²after M. E. Camille Jordan (1838 - 1922), a French mathematician

Let $v_1 \in L \setminus \{0\}$ be such that $\|v_1\|_2 = \lambda_1$, where λ_1 is the first successive minimum of L with respect to B_n . Fix an orthonormal basis $\{e_1, \dots, e_n\}$ of \mathbb{R}^n , such that $e_1 = \lambda_1^{-1}v_1$. Consider the projection $\rho : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$, $\sum_{i=1}^n x_i e_i \mapsto (x_2, \dots, x_n)$. Let $L' = \rho(L)$, e.g. if $L = \mathbb{Z}d_1e_1 \oplus \dots \oplus \mathbb{Z}d_ne_n$, then $L' = \mathbb{Z}d_2e_2 \oplus \dots \oplus \mathbb{Z}d_ne_n$.

Lemma 3.17

Using the same notation as above, $L' \subseteq \mathbb{R}^{n-1}$ is a lattice and if v_1, \dots, v_n is a basis of L then $\rho(v_2), \dots, \rho(v_n)$ is a basis of L' .

Lecture 14,
15.12.2023

Lemma 3.18

Let $\{v'_2, \dots, v'_n\}$ be a basis of L' and $v_2, \dots, v_n \in L$ with $\rho(v_i) = v'_i$ for $2 \leq i \leq n$. Then $\{v_1, \dots, v_n\}$ is a basis of L .

Lemma 3.19

$$d(L) = \lambda_1 d(L').$$

Lemma 3.20

Let $v' \in L'$. Then there exists $v \in L$, such that $\rho(v) = v'$ and

$$\|v\|_2^2 \leq \frac{4}{3} \|v'\|_2^2.$$

Remark: We always have $\prod_{i=1}^n \|v_i\|_2 \geq d(L)$.

3.3 Bounds for class numbers

For the rest of this section, let K be a number field with ring of integers \mathcal{O}_K .

Question: Can we improve upon our earlier upper bounds on $|Cl(\mathcal{O}_K)|$?

Idea: We could interpret the non-zero ideal $I \subseteq \mathcal{O}_K$ as a lattice and apply Minkowski's first convex body theorem to find an element $\alpha \in I \setminus \{0\}$ of small norm.

More concretely, let $\sigma_1, \dots, \sigma_r : K \hookrightarrow \mathbb{R}$ be the real embeddings and $\tau_1, \bar{\tau}_1, \dots, \tau_s, \bar{\tau}_s :$

$K \hookrightarrow \mathbb{C}$ be the complex embeddings of K . Note that $r + 2s = n$, where $n = [K : \mathbb{Q}]$. Define the map

$$\varphi : K \rightarrow \mathbb{R}^n, \quad \alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \Re\tau_1(\alpha), \Im\tau_1(\alpha), \dots, \Re\tau_s(\alpha), \Im\tau_s(\alpha)).$$

Lemma 3.21

The image $\varphi(\mathcal{O}_K) =: \Lambda$ is a (full) lattice in \mathbb{R}^n with determinant

$$d(\Lambda) = \frac{1}{2^s} \sqrt{|\text{disc } \mathcal{O}_K|}.$$

Remark: If I is a non-zero ideal, then the same argument shows that $\varphi(I)$ is a sublattice of \mathcal{O}_K . More precisely, $d(\varphi(I)) = d(\varphi(\mathcal{O}_K)) \underbrace{\left| \varphi(\mathcal{O}_K) / \varphi(I) \right|}_{=|\mathcal{O}_K/I|}$, i.e.

$$d(\varphi(I)) = \frac{1}{2^s} \sqrt{|\text{disc } \mathcal{O}_K|} N(I).$$

Corollary 3.22

$\varphi(K)$ is dense in \mathbb{R}^n .

Our next goal is for a non-zero ideal $I \subseteq \mathcal{O}_K$ to find a $\alpha \in I \setminus \{0\}$, such that $|N_{K/\mathbb{Q}}(\alpha)|$ is small. We write $\varphi(\alpha) = (y_1, \dots, y_n) \in \mathbb{R}^n$. Then

$$N_{K/\mathbb{Q}}(\alpha) = y_1 \cdot y_2 \cdots y_r \cdot (y_{r+1}^2 + y_{r+2}^2) \cdots (y_{n-1}^2 + y_n^2).$$

The problem here is that the function $N : \mathbb{R}^n \rightarrow \mathbb{R}$ is not a norm on \mathbb{R}^n .

Idea: Construct a central symmetric convex body $A \subseteq \mathbb{R}^n$, such that $x \in A$ implies that $|N(x)| \leq 1$.

We define

$$A = \left\{ x \in \mathbb{R}^n \mid |x_1| + \cdots + |x_r| + 2 \left(\sqrt{x_{r+1}^2 + x_{r+2}^2} + \cdots + \sqrt{x_{n-1}^2 + x_n^2} \right) \leq n \right\}$$

Lemma 3.23

A is a central symmetric convex body with the property that $x \in A$ implies $|N(x)| \leq 1$. Moreover,

$$\text{vol}(A) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2} \right)^s.$$

Theorem 3.24

Let $I \subseteq \mathcal{O}_K$ be a non-zero ideal. Then there exists an $\alpha \in I \setminus \{0\}$ with

$$|N_{K/\mathbb{Q}}(\alpha)| \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|\text{disc } \mathcal{O}_K|} N(I).$$

Lecture 15,
19.12.2023

Corollary 3.25

Every ideal class $C \in \text{Cl}(\mathcal{O}_K)$ contains a representative I with

$$N(I) \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|\text{disc } \mathcal{O}_K|}.$$

Corollary 3.26

If $K \neq \mathbb{Q}$ (i.e. $n \neq 1$), then

$$|\text{disc } \mathcal{O}_K| > 1.$$

Example: We try to find the class group of $\mathbb{Z}[\sqrt{-5}]$, i.e. we have $K = \mathbb{Q}[\sqrt{-5}]$, $\mathcal{O}_K = \mathbb{Z}[\sqrt{-5}]$, $n = 2$, $s = 1$. By Corollary 3.26 it is sufficient to consider ideals $I \subseteq \mathcal{O}_K$ with

$$N(I) \leq \frac{2!}{4} \frac{4}{\pi} \underbrace{\sqrt{|\text{disc}(\mathbb{Z}[\sqrt{-5}])|}}_{=2\sqrt{5}} = \frac{4\sqrt{5}}{\pi} \leq 3,$$

i.e. ideals lying above 2. Recall that

$$2\mathbb{Z}[\sqrt{-5}] = (2, 1 + \sqrt{-5})^2$$

and $(2, 1 + \sqrt{-5})$ is not principal. Hence

$$|\text{Cl}(\mathbb{Z}[\sqrt{-5}])| = 2.$$

3.4 Dirichlet's unit theorem

Let K be a number field with ring of integers \mathcal{O}_K . What can we say about the group of units \mathcal{O}_K^* ?

Example: • For $K = \mathbb{Q}$ we have $\mathbb{Z}^* = \{\pm 1\}$, for $K = \mathbb{Q}(i)$ we have $\mathbb{Z}[i]^* = \{\pm 1, \pm i\}$. In the exercises we have seen that \mathcal{O}_K^* is finite for all imaginary quadratic number fields K .

- If $K = \mathbb{Q}(\sqrt{d})$ with $d \in \mathbb{N}$ square-free, then the Pell equation $x^2 - dy^2 = 1$ has

a non-trivial solution (x_0, y_0) and $x_0 + \sqrt{d}y_0$ generates infinitely many units in \mathcal{O}_K

Let $n = [K : \mathbb{Q}]$, $\sigma_1, \dots, \sigma_r : K \hookrightarrow \mathbb{R}$ and $\tau_1, \bar{\tau}_1, \dots, \tau_s, \bar{\tau}_s : K \hookrightarrow \mathbb{C}$ be the real and complex embeddings of K . As in Section 3.3, let $\varphi : K \rightarrow \mathbb{R}^n$ be defined by

$$\alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \Re \tau_1(\alpha), \Im \tau_1(\alpha), \dots, \Re \tau_s(\alpha), \Im \tau_s(\alpha)).$$

Definition

In the notation above we define the maps $\log : \varphi(K \setminus \{0\}) \rightarrow \mathbb{R}^{r+s}$ as

$$(x_1, \dots, x_n) \mapsto (\log |x_1|, \dots, \log |x_r|, \log(x_{r+1}^2 + x_{r+2}^2), \dots, \log(x_{n-1}^2 + x_n^2))$$

and $\psi : K \setminus \{0\} \rightarrow \mathbb{R}^{r+s}$ as $\psi = \log \circ \varphi$.

First properties of ψ :

(a) For $\alpha, \beta \in K \setminus \{0\}$ we have

$$\psi(\alpha\beta) = \psi(\alpha) + \psi(\beta).$$

(b) Let $H \subseteq \mathbb{R}^{r+s}$ be the hyperplane given by $y_1 + \dots + y_{r+s} = 0$. Then we have $\psi(\mathcal{O}_K^*) \subseteq H$, because every $\alpha \in \mathcal{O}_K^*$ satisfies

$$1 = |N_{K/\mathbb{Q}}(\alpha)| = |\sigma_1(\alpha)| \cdots |\sigma_r(\alpha)| |\tau_1(\alpha)|^2 \cdots |\tau_s(\alpha)|^2,$$

$$\text{i.e. } 0 = \log |\sigma_1(\alpha)| + \cdots + \log |\tau_s(\alpha)|^2.$$

(c) Let $B \subseteq \mathbb{R}^{r+s}$ be a bounded subset. Then $\log^{-1}(B) \cap \varphi(\mathcal{O}_K \setminus \{0\})$ is finite.

Our next goal is to study the image $\psi(\mathcal{O}_K^*) \subseteq H \subseteq \mathbb{R}^{r+s}$. Note that by (a) above, $\psi(\mathcal{O}_K^*)$ is an (additive) subgroup of H .

Lemma 3.27

Let $G \subseteq \mathbb{R}^m$ be a subgroup, such that every bounded subset of G is finite. Then there exist over \mathbb{R} linearly independent vectors $v_1, \dots, v_d \in \mathbb{R}^m$ for some $d \leq m$ such that

$$G = \left\{ \sum_{i=1}^d x_i v_i \mid x_1, \dots, x_d \in \mathbb{Z} \right\}.$$

Corollary 3.28

$\psi(\mathcal{O}_K^*)$ is a lattice in some linear subspace of H .

Next we will show that $\psi(\mathcal{O}_K^*)$ spans H , i.e. $\psi(\mathcal{O}_K^*)$ is a lattice of full rank in H .

Lemma 3.29

Let $1 \leq k \leq r+s$ and $\alpha \in \mathcal{O}_K \setminus \{0\}$. Write $\psi(\alpha) = (a_1, \dots, a_{r+s})$. Then there exists $\beta \in \mathcal{O}_K \setminus \{0\}$ with

$$|N_{K/\mathbb{Q}}(\beta)| \leq \left(\frac{2}{\pi}\right)^2 \sqrt{|\text{disc } \mathcal{O}_K|}$$

and with the property that if $\psi(\beta) = (b_1, \dots, b_{r+s})$, then $b_j < a_j$ for all $1 \leq j \leq r+s$, $j \neq k$

Lemma 3.30

There exist units $u_1, \dots, u_{r+s} \in \mathcal{O}_K^*$ with the following property: If

$$\psi(u_l) = (u_{l,1}, \dots, u_{l,r+s}),$$

then $u_{l,j} < 0$ for all $j \neq l$.

Remark: If we construct a matrix

$$\begin{pmatrix} \psi(u_1) \\ \vdots \\ \psi(u_l) \\ \vdots \\ \psi(u_{r+s}) \end{pmatrix} = \begin{pmatrix} u_{1,1} & \dots & u_{1,l} & \dots & u_{1,r+s} \\ \vdots & \ddots & \vdots & & \vdots \\ u_{l,1} & \dots & u_{l,l} & \dots & u_{l,r+s} \\ \vdots & & \vdots & \ddots & \vdots \\ u_{r+s,1} & \dots & u_{r+s,l} & \dots & u_{r+s,r+s} \end{pmatrix}$$

Lemma 3.30 tells us that the diagonals are positive while all other entries are negative.

Next we will let u_1, \dots, u_{r+s} be units as in Lemma 3.30. We would like to show that $\psi(u_1), \dots, \psi(u_{r+s})$ span H .

Lecture 16,
22.12.2023

Lemma 3.31

Let $A = (a_{ij})_{1 \leq i,j \leq m} \in \text{Mat}_{m \times m}(\mathbb{R})$ and assume the following properties:

$$(i) \sum_{j=1}^m a_{ij} = 0 \text{ for all } 1 \leq i \leq m$$

$$(ii) a_{ii} > 0 \text{ for all } 1 \leq i \leq m$$

$$(iii) a_{ij} < 0 \text{ for } i \neq j, 1 \leq i, j \leq m$$

Then $\text{rank}(A) = m - 1$.

Corollary 3.32

The image $\psi(\mathcal{O}_K^*) \subseteq H$ is a lattice of rank $r + s - 1$.

Theorem 3.33 (Dirichlet's³ unit theorem)

Let K be a number field with r real and $2s$ complex embeddings and \mathcal{O}_K its ring of integers. Then there exist units $u_1, \dots, u_{r+s-1} \in \mathcal{O}_K^*$, such that every unit $u \in \mathcal{O}_K^*$ can be written uniquely in the form

$$u = \mu \cdot u_1^{e_1} \cdot u_2^{e_2} \cdots u_{r+s-1}^{e_{r+s-1}}$$

with $\mu \in K$ a root of unity and $e_1, \dots, e_{r+s-1} \in \mathbb{Z}$.

Remark: We call u_1, \dots, u_{r+s-1} as in Theorem 3.33 a fundamental system of units.

- Example:**
1. If K is a cubic field with exactly one real embedding, then the only roots of unity in K are ± 1 (as they are the only roots of unity in \mathbb{R}). Hence there exists a fundamental unit $u \in \mathcal{O}_K^*$, such that $\mathcal{O}_K^* = \{\pm u^k \mid k \in \mathbb{Z}\}$.
 2. The only number fields with a finite group of units \mathcal{O}_K^* are \mathbb{Q} and imaginary quadratic number fields.

³after Peter Gustav Lejeune Dirichlet (1805 - 1859), a German mathematician

4 Diophantine Approximation

4.1 Introduction

Motivation: Let $\alpha \in \mathbb{R}$, how well can we approximate α with rational numbers of small denominator? Given $\varepsilon > 0$, what is the "smallest" fraction $\frac{x}{y}$ (i.e. y small), such that $\left| \alpha - \frac{x}{y} \right| < \varepsilon$, $x \in \mathbb{Z}$, $y \in \mathbb{N}$?

Theorem 4.1 (Dirichlet, 1842)

Let $\alpha \in \mathbb{R}$ and $Q \in \mathbb{N}$. Then there exist $x, y \in \mathbb{Z}$, such that $\left| \alpha - \frac{x}{y} \right| \leq \frac{1}{yQ}$, $0 < y \leq Q$ and with $\gcd(x, y) = 1$.

Corollary 4.2

Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then there exist infinitely many pairs $(x, y) \in \mathbb{Z}^2$, such that $y > 0$, $\gcd(x, y) = 1$ and $\left| \alpha - \frac{x}{y} \right| \leq \frac{1}{y^2}$.

Theorem 4.3 (Dirichlet, 1842)

(a) Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ for some $n \in \mathbb{N}$. For all $Q \in \mathbb{N}$ there exists a tuple $(x_1, \dots, x_n, y) \in \mathbb{Z}^{n+1}$ with $0 \leq y \leq Q^n$, such that

$$|\alpha_i y - x_i| \leq \frac{1}{Q} \quad \forall 1 \leq i \leq n.$$

(b) Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, not all in \mathbb{Q} . Then there exist infinitely many tuples $(x_1, \dots, x_n, y) \in \mathbb{Z}^{n+1}$ with $\gcd(x_1, \dots, x_n, y) = 1$, $y > 0$, such that

$$\left| \alpha_i - \frac{x_i}{y} \right| \leq \frac{1}{y^{1+\frac{1}{n}}} \quad \forall 1 \leq i \leq n.$$

Another application of Minkowski's convex body theorem: Rational points close to hyperplanes.

Theorem 4.4

Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, such that $1, \alpha_1, \dots, \alpha_n$ are linearly independent over \mathbb{Q} . Then

there exist infinitely many tuples $(x, y_1, \dots, y_n) \in \mathbb{Z}^{n+1}$ with $y_1, \dots, y_n \neq (0, \dots, 0)$ and

$$|\alpha_1 y_1 + \dots + \alpha_n y_n - x| \leq \left(\max_{1 \leq i \leq n} |y_i| \right)^{-n}.$$

An open problem: Recall the notation $\|y\| = \min_{m \in \mathbb{Z}} |y - m|$ for $y \in \mathbb{R}$. Let $\alpha \in \mathbb{R}$. By Dirichlet's theorem there exist infinitely many $y \in \mathbb{N}$ with $y\|\alpha y\| \leq 1$. Let $\alpha, \beta \in \mathbb{R}$. Then there exist infinitely many $y \in \mathbb{N}$ with $y\|\alpha y\|\|\beta y\| \leq 1$.

Conjecture (Littlewood¹ conjecture)

Let $\alpha, \beta \in \mathbb{R}$. Then

$$\liminf_{y \rightarrow \infty} y\|\alpha y\|\|\beta y\| = 0.$$

Borel² showed in 1909 that the exceptional set has Lebesgue measure 0. Einsiedler³, Katok⁴ and Lindenstrauss⁵ showed in 2006 that the exceptional set also has Hausdorff dimension 0.

Lecture 17,
09.01.2024

Question: Can we do better than Corollary 4.2?

Example: Let $A > \sqrt{5}$ and $\alpha = \frac{1+\sqrt{5}}{2}$. Then the inequality $|\alpha - \frac{x}{y}| \leq \frac{1}{Ay^2}$ has only finitely many solutions $x, y \in \mathbb{N}$.

For $\delta > 0$, consider the inequality

$$\left| \alpha - \frac{x}{y} \right| \leq \frac{1}{y^{2+\delta}} \quad (4.1)$$

in $x, y > 0$, $\gcd(x, y) = 1$. For what α does (4.1) have infinitely many solutions? Khinchin⁶ showed in 1927 that the set of such α has Lebesgue measure 0.

Example: Let $a \in \mathbb{N}_{\geq 3}$ and set $\alpha = \sum_{n=1}^{\infty} 10^{-a^{2n}}$. The claim is that there exist infinitely many $(x, y \in \mathbb{Z}^2)$ with $y > 0$ and $\gcd(x, y) = 1$, such that

$$\left| \alpha - \frac{x}{y} \right| \leq \frac{1}{y^a}.$$

¹after John Edensor Littlewood (1885 - 1977), a British mathematician

²Émile Borel (1871 - 1956), a French mathematician and politician

³Manfred Einsiedler (*1973), an Austrian mathematician

⁴Anatole Katok (1944-2018), an American mathematician

⁵Elon Lindenstrauss (*1970), an Israeli mathematician

⁶Aleksandr Khinchin (1894 - 1959), a Soviet mathematician

Idea: To construct such well-appropriable numbers we pick α in the decimal expansion (or use any other base) with very few digits 1, which get more and more sparse, and set all other digits equal to zero.

Theorem (Roth⁷, 1955)

Let $\alpha \in \mathbb{R}$ be an algebraic number and $\delta > 0$. Then there are only finitely many tuples $(x, y) \in \mathbb{Z}^2$ with $y > 0$, $\gcd(x, y) = 1$ and

$$\left| \alpha - \frac{x}{y} \right| \leq \frac{1}{y^{2+\delta}}.$$

Roth's theorem implies that $\alpha = \sum_{n=1}^{\infty} 10^{-a^{2n}}$ for $a \geq 3$ is transcendental.

Definition (Linearly independent complex numbers)

We call a set $\{\alpha_1, \dots, \alpha_n\} \in \mathbb{C}^n$ linearly independent over \mathbb{Q} if the relation $x_1\alpha_1 + \dots + x_n\alpha_n = 0$ with $x_1, \dots, x_n \in \mathbb{Q}$ implies $x_1 = \dots = x_n = 0$.

Theorem (Schmidt⁸, 1971)

Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ algebraic such that $\{1, \alpha_1, \dots, \alpha_n\}$ is linearly independent over \mathbb{Q} . Let $\delta > 0$. Then there exist only finitely many tuples $(x_1, \dots, x_n, y) \in \mathbb{Z}^{n+1}$ with $y > 0$, $\gcd(x_1, \dots, x_n, y) = 1$ and

$$\left| \alpha_i - \frac{x_i}{y} \right| \leq y^{-1-\frac{1}{n}} \quad \forall 1 \leq i \leq n.$$

Theorem (Subspace Theorem, Schmidt, 1972)

Let $n > 2$ and $L_i = \alpha_{i1}x_1 + \dots + \alpha_{in}x_n$, $1 \leq i \leq n$, be n linearly independent linear forms with coefficients in $\overline{\mathbb{Q}}$. Let $C, \delta > 0$. Then the solution of the inequality

$$|L_1 \cdot L_2 \cdots L_n| \leq C \max\{|x_1|, \dots, |x_n|\}^{-\delta}$$

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

⁷Klaus Roth (1925 - 2015), a British mathematician

⁸Wolfgang M. Schmidt (*1933), an Austrian mathematician

Example: Let α be an algebraic number and consider the linear forms $ax_2 - x_1, x_2$.

$$|ax_2 - x_1||x_2| \leq \max\{|x_1|, |x_2|\}^{-\delta}$$

The application of the Subspace Theorem leads us back to Roth's theorem.

4.2 Transcendence

Definition (Algebraic and transcendental numbers)

We call $\alpha \in \mathbb{C}$ *algebraic* (over \mathbb{Q}) if there exists a non-zero polynomial $P(x) \in \mathbb{Q}[x]$ such that $P(\alpha) = 0$. If $\alpha \in \mathbb{C}$ is not algebraic, then we call it *transcendental*.

Notation: We write $\overline{\mathbb{Q}} = \{\alpha \in \mathbb{C} \mid \alpha \text{ is algebraic}\}$.

Definition (Algebraically independent numbers)

We call $\alpha_1, \dots, \alpha_r \in \mathbb{C}$ *algebraically independent* if there is no non-zero polynomial $P \in \overline{\mathbb{Q}}[x_1, \dots, x_r]$ with $P(\alpha_1, \dots, \alpha_r) = 0$.

Example: 1. $\alpha \in \mathbb{C}$ is transcendental if and only if α is algebraically independent.

2. e is transcendental.

3. $\alpha_1 = e, \alpha_2 = e^2$ are linearly independent over $\overline{\mathbb{Q}}$ but not algebraically independent as $\alpha_1^2 - \alpha_2 = 0$.

Definition (Transcendence degree, transcendence basis)

Let $S \subseteq \mathbb{C}$. We define the *transcendence degree* of S as the maximal number $t \in \mathbb{Z}_{\geq 0}$ (or $t = \infty$), such that S contains t algebraically independent elements. We denote it by $\text{trdeg } S$. If $B \subseteq S$ is an algebraically independent subset with $|B| = \text{trdeg } S$, then we call B a *transcendence basis* of S .

Example: 1. $\text{trdeg } \mathbb{Q}(e) = 1$ and $\{e\}$ and $\{e^2\}$ are examples of a transcendence basis for $\mathbb{Q}(e)$.

2. Let $S \subseteq \mathbb{C}$ with transcendence basis $B = \{\alpha_1, \dots, \alpha_r\}$. Then every $x \in S$ is algebraic over $\overline{\mathbb{Q}}(\alpha_1, \dots, \alpha_r)$.

Lemma 4.5

Let $\alpha \in \mathbb{R}$ and assume that there exists a sequence of tuples of integers $(x_n, y_n) \in \mathbb{Z}^2$, $n \in \mathbb{N}$, with $y_n > 0$, $\frac{x_n}{y_n} \neq \alpha \forall n \in \mathbb{N}$ and

$$|x_n - \alpha y_n| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Then $\alpha \notin \mathbb{Q}$.

Theorem 4.6

$e \notin \mathbb{Q}$.

Proof. Write $e = \sum_{k=0}^{\infty} \frac{1}{k!}$. For $n \in \mathbb{N}$ set $x_n = n! \sum_{k=0}^n \frac{1}{k!}$ and $y_n = n!$. Then

$$\begin{aligned} 0 < |x_n - ey_n| &= n! \sum_{k=n+1}^{\infty} \frac{1}{k!} = \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \dots \\ &< \frac{1}{n+1} + \frac{1}{(n+1)^2} = \frac{1}{n+1} \sum_{q=0}^{\infty} \frac{1}{(n+1)^q} = \frac{1}{n+1} \cdot \frac{1}{1 - \frac{1}{n+1}} \\ &= \frac{1}{n} \rightarrow 0 \text{ for } n \rightarrow \infty \end{aligned}$$

□

Theorem 4.7

The number $\alpha = \sum_{k=1}^{\infty} 10^{-k!}$ is transcendental.

Transcendence of e

For $z \in \mathbb{C}$ we set $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$.

Theorem 4.8 (Hermite⁹, 1873)

e is transcendental.

For a polynomial $f \in \mathbb{C}[x]$ we define the integral transform $F(z) = \int_0^z e^{z-u} f(u) du$, where $z \in \mathbb{C}$, and we integrate over the line segment $\{tz \mid 0 \leq t \leq 1\}$, i.e.

$$F(z) = \int_0^1 e^{z(1-t)} f(tz) z dt.$$

⁹Charles Hermite (1822 - 1901), a French mathematician

Example: If $f(u) = u$, then

$$\begin{aligned} F(z) &= \int_0^1 e^{z(1-t)} z^2 t dt = \left[\frac{1}{z} e^{z(1-t)} z^2 t \right]_0^1 + \int_0^1 \frac{1}{z} e^{z(1-t)} z^2 dt \\ &= -z + \left[-e^{z(1-t)} \right]_0^1 = e^z - z - 1 \end{aligned}$$

Lemma 4.9

Let $f \in \mathbb{C}[x]$ be of degree m . Then

$$F(z) = e^z \left(\sum_{j=0}^m f^{(j)}(0) \right) - \sum_{j=0}^m f^{(j)}(z).$$

Lemma 4.10

Let $f \in \mathbb{C}[x]$ and $z \in \mathbb{C}$. Then

$$|F(z)| \leq |z| e^{|z|} \sup_{\substack{u \in \mathbb{C} \\ |u| \leq |z|}} |f(u)|.$$

Now, assume that e is algebraic. Then there exists $q_0, \dots, q_n \in \mathbb{Z}$, $n \geq 0$, $q_n \neq 0$, such that

$$q_0 + q_1 e + \dots + q_n e^n = 0 \quad (4.2)$$

Lemma 4.11

Let $f \in \mathbb{C}[x]$ be of degree n and q_0, \dots, q_n as in (4.2). Then

$$\sum_{a=0}^n q_a F(a) = - \sum_{a=0}^n \sum_{j=0}^m q_a f^{(j)}(a). \quad (4.3)$$

Lecture 18, 12.01.2024 Our next step will be to construct a polynomial $f(x) \in \mathbb{C}[x]$, such that $|F(0)|, \dots, |F(n)|$ are very small and the right-hand side of (4.3) is a non-zero integer.

Let p be a prime number to be chosen later. Define

$$f(X) = \frac{1}{(p-1)!} X^{p-1} ((X-1)(X-2) \dots (X_n))^p.$$

Lemma 4.12

Let f be as above. Then we have

$$(i) \quad f^{(p-1)}(0) = ((-1)^n n!)^p$$

$$(ii) \quad f^{(j)}(a) \text{ if either } a \in \{1, \dots, n\} \text{ and } 0 \leq j \leq p-1 \text{ or } a=0 \text{ and } 0 \leq j \leq p-2$$

(iii) Let $0 \leq a \leq n$ and $j \geq p$. Then $f^{(j)}(a) \equiv 0 \pmod{p}$.

Lemma 4.13

Let $p > |q_0 n|$. Then

$$M := \sum_{a=0}^n q_a F(a) \in \mathbb{Z} \setminus \{0\}.$$

Lemma 4.14

Let q_0, \dots, q_n and M, p like above. Then $|M| \rightarrow 0$ for $p \rightarrow \infty$.

We summarise: If $q_0 + q_1 e + \dots + q_n e^n = 0$ for $q_0, \dots, q_n \in \mathbb{Z}$, $q_0 \neq 0$, and $f(X) = \frac{1}{(p-1)!} X^{p-1} ((X-1) \dots (X-n))^p$ for a sufficiently large prime p , then $M = \sum_{a=0}^n q_a F(a) \in \mathbb{Z} \setminus \{0\}$ and $|M| < \frac{1}{2}$, which is a contradiction. Hence, e is transcendental.

Remark: In the proof of Theorem 4.8 we showed that for any $n \in \mathbb{N}$, the numbers $1, e, e^2, \dots, e^n$ are linearly independent over \mathbb{Q} (and hence over $\overline{\mathbb{Q}}$).

Question: Let $\alpha_0, \dots, \alpha_n \in \overline{\mathbb{Q}}$. Under which assumptions are the numbers $e^{\alpha_0}, \dots, e^{\alpha_n}$ linearly dependent over \mathbb{Q} or $\overline{\mathbb{Q}}$?

We certainly need the α_i to be distinct, as for example $1 \cdot e^\alpha + (-1) \cdot e^\alpha = 0$ for all $\alpha \in \overline{\mathbb{Q}}$.

Theorem 4.15 (Baker¹⁰, Lindemann¹¹-Weierstraß¹²)

Let $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$ for some $n \in \mathbb{N}$. Assume that $\alpha_1, \dots, \alpha_n$ are pairwise distinct and $\beta_1 \dots \beta_n \neq 0$. Then

$$\beta_1 e^{\alpha_1} \dots \beta_n e^{\alpha_n} \neq 0.$$

Remark: This implies that if $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ are pairwise distinct, then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are linearly independent over $\overline{\mathbb{Q}}$.

¹⁰Alan Baker (1939 - 2018), an English mathematician

¹¹after Ferdinand von Lindemann (1852-1939), a German mathematician,

¹²and Karl Weierstraß (1815-1879), a German mathematician

Corollary 4.16

Let $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$. Then e^α is transcendental.

Corollary 4.17

π is transcendental.

Proof. Assume $\pi \in \overline{\mathbb{Q}}$. Then $i\pi \in \overline{\mathbb{Q}}$, but $e^{i\pi} = -1$ is not transcendental. \square

Corollary 4.18

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ be linearly independent over \mathbb{Q} . Then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent.

Remark: Corollary 4.18 is in fact equivalent to Theorem 4.15.

Example: Imagine we try to show that

$$1 \cdot e^0 + 2 \cdot e^{\sqrt{3}} \neq 0.$$

For $\sigma \in \text{Gal}(\mathbb{Q}(\sqrt{3})/\mathbb{Q})$ and $\alpha \in \mathbb{Q}(\sqrt{3})$, set $\sigma(e^\alpha) = e^{\sigma(\alpha)}$. Then the non-trivial automorphism $\sigma \in \text{Gal}(\mathbb{Q}(\sqrt{3})/\mathbb{Q})$ maps $1 + 2e^{\sqrt{3}}$ to $1 + 2e^{-\sqrt{3}}$. However,

$$(1 + e^{\sqrt{3}})(1 + 2e^{-\sqrt{3}}) = 1 + 4 + 2e^{\sqrt{3}} + 2e^{-\sqrt{3}}$$

is invariant under $\text{Gal}(\mathbb{Q}(\sqrt{3})/\mathbb{Q})$.

We can reduce Theorem 4.15 to the following result:

Theorem 4.19 ("Weak Lindemann-Weierstraß theorem")

Let $\mathbb{Q} \subseteq L \subseteq \mathbb{C}$ be a normal number field. Let $\gamma_1, \dots, \gamma_t, \delta_1, \dots, \delta_t \in L$, such that $\gamma_1, \dots, \gamma_t$ are pairwise distinct and $\delta_1 \cdots \delta_t \neq 0$. Assume that each $\tau \in \text{Gal}(L/\mathbb{Q})$ permutes the pairs $(\gamma_1, \delta_1), \dots, (\gamma_t, \delta_t)$. Then

$$\delta_1 e^{\gamma_1} + \dots + \delta_t e^{\gamma_t} \neq 0.$$

$1 \leq k \leq t$ we define

$$f_K(X) = \frac{1}{(p-1)!} l^{pt} (X - \gamma_k)^{p-1} \prod_{\substack{i=1 \\ i \neq k}}^t (X - \gamma_i)^p.$$

Set $F_k(z) = \int_0^1 e^{z-u} f_k(u) du$ and $M_k = \delta_1 F_k(\gamma_1) + \cdots + \delta_t F_k(\gamma_t)$.

Lemma 4.20

If $\tau \in \text{Gal}(L/\mathbb{Q})$, then $\tau(M_1) \in \{M_1, \dots, M_t\}$.

Notation: For $\alpha, \beta \in \overline{\mathbb{Q}}$, $m \in \mathbb{Z} \setminus \{0\}$, we write $\alpha \equiv \beta \pmod{m}$ if $\frac{\alpha - \beta}{m}$ is an algebraic integer.

Lemma 4.21

Let $1 \leq m \leq t$. Then

$$(i) \quad f_1^{(p-1)}(\gamma_1) = l^{pt} \left(\prod_{i=2}^t (\gamma_i - \gamma_1) \right)^p$$

(ii) If either $2 \leq m \leq t$ and $0 \leq j \leq p-1$ or $m = 1$ and $0 \leq j \leq p-2$, then $f_1^{(j)}(\gamma_m) = 0$.

(iii) $f_1^{(j)}(\gamma_m) \equiv 0 \pmod{p}$ if $1 \leq m \leq t$ and $j \geq p$.

Lemma 4.22

If p is sufficiently large, then $M_1 \neq 0$ is an algebraic integer.

Lemma 4.23

Let $1 \leq k \leq t$. Then $|M_k| \rightarrow 0$ for $p \rightarrow \infty$.

4.3 More on transcendence results

Recall our definition $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$ for $z \in \mathbb{C}$. For $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$, we set $\alpha^\beta = e^{\beta \log \alpha}$, where $\log \alpha$ is some solution of the equation $\alpha = e^z$. I.e. if we fix one solution $\log \alpha$, then all possibilities for $e^{\beta \log \alpha}$ are given by $e^{\beta(\log \alpha + 2\pi i k)}$, $k \in \mathbb{Z}$. In Section 4.2 we have seen that if $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ are pairwise distinct, then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are linearly independent over $\overline{\mathbb{Q}}$. As a corollary: if $\beta \in \overline{\mathbb{Q}} \setminus \{0\}$, then e^β is transcendental.

Theorem 4.24 (Gelfond¹³, Schneider¹⁴, 1934)

Let $\alpha, \beta \in \overline{\mathbb{Q}}$ with $0 \neq \alpha \neq 1$ and $\beta \notin \mathbb{Q}$. Then $\alpha^\beta = e^{\beta \log \alpha}$ is transcendental for any solution $\log \alpha$.

Corollary 4.25

Let $\alpha \in \overline{\mathbb{Q}}$ with $\alpha \notin i\mathbb{Q}$. Then $e^{\pi\alpha}$ is transcendental.

Corollary 4.26

Let $\alpha_1, \alpha_2 \in \overline{\mathbb{Q}} \setminus \{0\}$. Fix a choice of logarithms of $\log \alpha_1, \log \alpha_2$ and assume that $\log \alpha_1, \log \alpha_2$ are linearly independent over \mathbb{Q} . Then if $\beta_1, \beta_2 \in \overline{\mathbb{Q}} \setminus \{0\}$, we have $\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 \neq 0$.

Example: The real logarithms $\log 2$ and $\log 3$ are linearly independent over \mathbb{Q} and $\overline{\mathbb{Q}}$.

Question: How about elements $\log \alpha_1, \dots, \log \alpha_n$ for $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$?

Theorem 4.27 (Baker, 1965)

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$ and fix choices for $\log \alpha_1, \dots, \log \alpha_n$, such that $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over \mathbb{Q} . Let $\beta_1, \dots, \beta_n \in \overline{\mathbb{Q}} \setminus \{0\}$. Then $\beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n$ is transcendental.

Remark: Baker's theorem gives us the stronger conclusion that $1, \log \alpha_1, \dots, \log \alpha_n$ are linearly independent over $\overline{\mathbb{Q}}$.

Definition

Let $\alpha \in \overline{\mathbb{Q}}$ with primitive minimal polynomial $f \in \mathbb{Z}[X]$, i.e. a minimal polynomial $f(X) = a_0 + a_1X + \dots + a_dX^d$ with a_0, \dots, a_d and $\gcd(a_0, \dots, a_d) = 1$. Then we set $H(\alpha) = \max_{0 \leq i \leq d} |a_i|$.

Theorem 4.28

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, $\gamma \in \overline{\mathbb{Q}}$ and $\beta_1, \dots, \beta_n \in \overline{\mathbb{Q}} \setminus \{0\}$. Pick choices of

¹³Alexander Gelfond (1906-1968), a Soviet mathematician, who did his Ph. D with Khinchin

¹⁴Theodor Schneider (1911-1988), a German mathematician, who worked in Göttingen until 1953 and later became the director of the MRI Oberwolfach

$\log \alpha_1, \dots, \log \alpha_n$ and assume that $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over \mathbb{Q} . Then

$$|\gamma + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n| \geq (eB)^{-c}$$

with $B = \max(H(\gamma), H(\beta_1), \dots, H(\beta_n))$ and $c > 0$ an effectively computable constant depending on $n, H(\alpha_1), \dots, H(\alpha_n)$ and the choices for $\log \alpha_1, \dots, \log \alpha_n$.

Question: How can we recognise if $\log \alpha_1, \dots, \log \alpha_n$ are \mathbb{Q} -linearly independent.

Assume that $\log \alpha_1, \dots, \log \alpha_n$ are \mathbb{Q} -linearly dependent. Then there exist $b_1, \dots, b_n \in \mathbb{Z}$, not all zero, such that

$$b_1 \log \alpha_1 + \dots + b_n \log \alpha_n = 0,$$

i.e.

$$\alpha_1^{b_1} \alpha_2^{b_2} \dots \alpha_n^{b_n} = 0.$$

Definition (Multiplicative dependency)

We say that $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ are *multiplicatively dependent* if there exist $b_1, \dots, b_n \in \mathbb{Z}$, not all zero, such that

$$\alpha_1^{b_1} \alpha_2^{b_2} \dots \alpha_n^{b_n} = 0.$$

Remark: If $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0\}$ are multiplicatively independent, then $\log \alpha_1, \dots, \log \alpha_n$ are \mathbb{Q} -linearly independent.

Corollary 4.29

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0\}, \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$, such that $\alpha_1, \dots, \alpha_n$ are multiplicatively independent and $(\beta_1, \dots, \beta_n) \notin \mathbb{Q}^n$, then $\alpha_1^{\beta_1} \dots \alpha_n^{\beta_n}$ is transcendental (for any choice of $\log \alpha_1, \dots, \log \alpha_n$).

Conjecture 4.30 (Schanuel's¹⁵ conjecture)

Let $x_1, \dots, x_n \in \mathbb{C}$ be linearly independent over \mathbb{Q} . Then

$$\text{trdeg}(x_1, \dots, x_n, e^{x_1}, \dots, e^{x_n}) \geq n.$$

Schanuel's conjecture implies the Lindemann-Weierstraß theorem. Let $\alpha_1, \dots, \alpha_n \in$

¹⁵after Stephen Schanuel (1933 - 2014), an American mathematician

$\overline{\mathbb{Q}}$ be linearly independent over \mathbb{Q} . Then $\text{trdeg}(\underbrace{\alpha_1, \dots, \alpha_n}_{\text{trdeg}(e^{\alpha_1}, \dots, e^{\alpha_n})}, e^{\alpha_1}, \dots, e^{\alpha_n}) \geq n$, i.e. $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent. It also implies Baker's theorem. Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, such that $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over \mathbb{Q} . Then $n \leq \text{trdeg}(\log \alpha_1, \dots, \log \alpha_n, \alpha_1, \dots, \alpha_n) = \text{trdeg}(\log \alpha_1, \dots, \log \alpha_n)$. I.e. there is a non-trivial algebraic relation over $\overline{\mathbb{Q}}$ of $\log \alpha_1, \dots, \log \alpha_n$, in particular if $\beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$, not all zero, then $\beta \log \alpha_1 + \dots + \beta_n \log \alpha_n \notin \overline{\mathbb{Q}}$.

What is known towards Schanuel's conjecture?

$n = 1$: If $x \in \mathbb{C}$ is transcendental, then $\text{trdeg}(x, e^x) \geq 1$, if $x \in \overline{\mathbb{Q}} \setminus \{0\}$, then e^x is transcendental by Lindemann-Weierstraß.

$n = 2$: This is still open. For example, we don't know if $(\log 2)(\log 3)$ is transcendental (by Baker's theorem they are $\overline{\mathbb{Q}}$ -linearly independent). Schanuel's conjecture would imply that $\text{trdeg}(\log 2, \log 3, 2, 3) = \text{trdeg}(\log 2, \log 3) \geq 2$, i.e. $\log 2, \log 3$ are algebraically independent.

Conjecture 4.31

e and π are algebraically independent.

Conjecture 4.32

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$ and assume that $\log \alpha_1, \dots, \log \alpha_n$ are \mathbb{Q} -linearly independent for some choices of $\log \alpha_1, \dots, \log \alpha_n$. Then $\log \alpha_1, \dots, \log \alpha_n$ are algebraically independent.

Theorem 4.33

Let $\alpha, \beta \in \overline{\mathbb{Q}} \cap \mathbb{R}$ with $\alpha > 0$, $\alpha \neq 1$ and $\beta \notin \mathbb{Q}$. Let $\log \alpha$ be the real logarithm of α . Then $\alpha^\beta 0e^{\beta \log \alpha}$ is transcendental.

4.4 Siegel's lemma

Consider m linear equations in n variables, say

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned} \tag{4.4}$$

Assume $m \leq n$ and $a_{ij} \in \mathbb{Z}$ for $1 \leq i, j \leq n$. Then there exists a solution $x \in \mathbb{Q}^n \setminus \{0\}$ and hence also a solution $x \in \mathbb{Z}^n \setminus \{0\}$. What can we say about the size of a "smallest" solution?

Example: • Let $m = 1$, $n = 2$, $a \in \mathbb{Z} \setminus \{0\}$. Then every solution $(x_1, x_2) \in \mathbb{Z}^2 \setminus \{0\}$ to the equations $x_1 + ax_2 = 0$ satisfies $x_1 \neq 0$ and $a \mid x_1$, i.e. $\|x\|_\infty \geq a$.

- Let $a \in \mathbb{Z} \setminus \{0\}$, m arbitrary and $n = m - 1$. Consider the system

$$\begin{aligned} x_1 + ax_2 &= 0 \\ x_2 + ax_3 &= 0 \\ &\vdots \\ x_{n-1} + ax_n &= 0 \end{aligned}$$

Every non-trivial solution $x \in \mathbb{Z}^n$ satisfies $x_1 \cdots x_n \neq 0$ and $a^{n-1} \mid x_1$, i.e. $\|x\|_\infty \geq a^{n-1}$

Theorem 4.34 (Siegel's¹⁶ lemma)

Let $n > m > 0$, $A \geq 1$ and $a_{ij} \in \mathbb{Z}$ for $1 \leq i \leq m$, $1 \leq j \leq n$, such that $|a_{ij}| \leq A \forall i, j$. Then there exists a solution $x \in \mathbb{Z}^n \setminus \{0\}$ to the system (4.4), such that

$$\max_{1 \leq i \leq n} |x_i| \leq (nA)^{\frac{m}{n-m}}$$

An alternate point of view: Consider $m = 1$. Let $a_1, \dots, a_n \in \mathbb{Z}$ with $\gcd(a_1, \dots, a_n) = 1$ and define the hyperplane $H = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n a_i x_i = 0\}$ and the lattice $\Lambda = \{x \in \mathbb{Z}^n \mid \sum_{i=1}^n a_i x_i = 0\}$. Then $\Lambda \subset H$ is a lattice of rank $n - 1$.

Lemma 4.35

Assume that $\gcd(a_1, \dots, a_n) = 1$. Then $d(\Lambda) = \|a\|_2$.

If $\lambda_1, \dots, \lambda_n$ are successive minima of Λ with respect to $\|\cdot\|_\infty$, then by Minkowski's second convex body theorem, $\lambda_1 \cdots \lambda_{n-1} \ll d(\Lambda) = \|a\|_2$, i.e. $\lambda_1^{n-1} \leq \|a\|_2, \lambda_1 \ll \|a\|_2^{\frac{1}{n-1}}$.

Next, we want to solve (4.4) in $x \in \mathbb{Z}^n$, where the coefficients a_{ij} are elements in some number field.

Example: Let $d \in \mathbb{Z} \setminus \{0, 1\}$ be square-free and $K = \mathbb{Q}(\sqrt{d})$. Set

$$\omega_d = \begin{cases} \sqrt{d}, & d \equiv 2, 3 \pmod{4} \\ \frac{1+\sqrt{d}}{2}, & d \equiv 1 \pmod{4} \end{cases}$$

¹⁶after Carl Ludwig Siegel (1896-1981), a German mathematician

Then $1, \omega_d$ is an integral basis. For $n > 2m > 0$, consider the system

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{4.5}$$

with $a_{ij} \in \mathcal{O}_K$. Our goal is to find solutions $(x_1, \dots, x_n) \in \mathbb{Z}^n \setminus \{0\}$. Our first idea would be to write $a_{ij} = b_{ij} + c_{ij}\omega_d$. Then rewrite (4.5) as

$$\begin{aligned} (b_{11} + c_{11}\omega_d)x_1 + \cdots + (b_{1n} + c_{1n}\omega_d)x_n &= 0 \\ &\vdots \\ (b_{m1} + c_{m1}\omega_d)x_1 + \cdots + (b_{mn} + c_{mn}\omega_d)x_n &= 0 \end{aligned}$$

This is equivalent to

$$\begin{aligned} b_{11}x_1 + \cdots + b_{1n}x_n &= 0 \\ &\vdots \\ b_{m1}x_1 + \cdots + b_{mn}x_n &= 0 \\ c_{11}x_1 + \cdots + c_{1n}x_n &= 0 \\ &\vdots \\ c_{m1}x_1 + \cdots + c_{mn}x_n &= 0 \end{aligned}$$

If $n > 2m$, then this is a system as in Theorem 4.34 and it has a non-zero solution $x \in \mathbb{Z}^n$ with

$$\max_{1 \leq i \leq n} |x_i| \leq (nA)^{\frac{2m}{n-2m}},$$

where $A = \max_{i,j} \{ |b_{ij}|, |c_{ij}| \}$.

Remark: Note that in this construction A depends on the choice of basis $1, \omega_d$. We will use a basis-independent approach below.

Our general set-up will be to let K be a number field of degree d with ring of integers \mathcal{O}_K . Let $\sigma_1, \dots, \sigma_d : K \hookrightarrow \mathbb{C}$ be the d distinct embeddings, such that $\sigma_1, \dots, \sigma_r : K \hookrightarrow \mathbb{R}$ are the real embeddings and $\sigma_{r+1}, \dots, \sigma_{r+s} : K \hookrightarrow \mathbb{C}$ complex embeddings with $r + 2s = d$ and $\sigma_{r+s+1} = \overline{\sigma_{r+1}}, \dots, \sigma_{r+2s} = \overline{\sigma_{r+s}}$. Define the map $\varphi : K \rightarrow \mathbb{R}^d$,

$$x \mapsto (\sigma_1(x), \dots, \sigma_r(x), \Re \sigma_{r+1}(x), \Im \sigma_{r+1}(x), \dots, \Re \sigma_{r+s}(x), \Im \sigma_{r+s}(x))$$

Definition (House)

Let $\alpha \in K$. We define the *house* of α as

$$|\overline{\alpha}| = \max(|\sigma_1(\alpha)|, \dots, |\sigma_d(\alpha)|).$$

Remark: The definition of $|\overline{\alpha}|$ is independent of the field K with $\alpha \in K$. If $\alpha \in \overline{\mathbb{Q}}$ has a minimal polynomial of degree m over \mathbb{Q} and conjugates $\alpha^{(1)}, \dots, \alpha^{(m)}$, then

$$|\overline{\alpha}| = \max(|\alpha^{(1)}|, \dots, |\alpha^{(m)}|).$$

Question: The rational integers $\mathbb{Z} \subset \mathbb{Q}$ are discrete, in particular if $m \in \mathbb{Z}$, $|m| < 1$, then $m = 0$. Is there a similar statement for $\mathcal{O}_K \subset K$?

Observation: Let $\alpha \neq 0$ be an algebraic integer with conjugates $\alpha^{(1)}, \dots, \alpha^{(m)}$. Then there is at least one index j with

$$|\alpha^{(j)}| \geq 1.$$

Lecture 21,
23.01.2024

Lemma 4.36

Let $\alpha \in \mathcal{O}_K$ with $\|\varphi(\alpha)\|_\infty \leq \frac{2}{3}$, then $\alpha = 0$.

Theorem 4.37

Let K/\mathbb{Q} be as above, i.e. $[K : \mathbb{Q}] = d$. Let $m, n \in \mathbb{N}$ with $n > dm$ and $A \in \mathbb{R}_{\geq 1}$. Suppose $a_{ij} \in \mathcal{O}_K$ with $|\overline{a_{ij}}| \leq A$ for all $1 \leq i \leq m, 1 \leq j \leq n$. Then the system

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned}$$

has a solution $x \in \mathbb{Z}^n \setminus \{0\}$ with

$$\max_{1 \leq i \leq n} |x_i| \leq (3nA)^{\frac{dm}{n-dm}}.$$

4.5 Approaches towards the Gelfond-Schneider theorem

Theorem 4.38 (Weak version of Gelfond-Schneider theorem)

Let $\alpha, \beta \in \overline{\mathbb{Q}} \cap \mathbb{R}$ with $\alpha > 0$, $\alpha \neq 1$, and $\beta \notin \mathbb{Q}$. Let $\log \alpha$ be the real logarithm of α . Then $\alpha^\beta = e^{\beta \log \alpha}$ is transcendental.

We will give an outline to the proof. Assume that $\gamma = \alpha^\beta \in \overline{\mathbb{Q}}$ and let $K = \mathbb{Q}(\alpha, \beta, \gamma)$, $d = [K : \mathbb{Q}]$. Choose $m \in \mathbb{N}$, such that $m\alpha, m\beta, m\gamma$ are all algebraic integers. Let $D_1, D_2, L \in \mathbb{N}$ be parameters to be chosen later.

Step 1: We construct a function

$$F(z) = \sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} a_{ij} z^i \alpha^{jz}$$

with the following properties:

- $a_{ij} \in \mathbb{Z}$ for all $0 \leq i < D_1$, $0 \leq j < D_2$, not all zero
- $F(a + b\beta) = 0$ for all $1 \leq a, b \leq L$
- there is a constant $c_1 > 0$ only depending on α, β, γ, d , such that

$$|a_{ij}| \leq e^{c_1(D_1 \log L + D_2 L)}$$

for all $0 \leq i < D_1$, $0 \leq j < D_2$.

Assume $D_1 D_2 \geq 2dL^2$. Then such a function can be constructed using Siegel's lemma.

Next we choose D_1, D_2 , such that $D_1 = D_2 L$ and $D_1 D_2 = 2dL^2$, i.e. $D_1 = \sqrt{2d} L^{\frac{3}{2}}$, $D_2 = \sqrt{2d} L^{\frac{1}{2}}$. If $L = 2dM^2$ for some $M \in \mathbb{N}$, then D_1, D_2 are both integers, and $|a_{ij}| \leq e^{c_3 L^{\frac{3}{2}} \log L}$ for some c_3 .

Step 2: We show that $F(z)$ has at most $D_1 D_2 = 2dL^2$ zeros in \mathbb{R} . For this we view $F(z)$ as an exponential polynomial, i.e. a function of the form

$$E(z) = \sum_{k=1}^r p_k(z) e^{\gamma_k z} \tag{4.6}$$

with $p_k(z) \in \mathbb{R}[z] \setminus \{0\}$ and $\gamma_k \in \mathbb{R}$ distinct.

Lemma 4.39

Let $E(z)$ be as in (4.6). Then $E(z)$ has at most

$$\left(\sum_{k=1}^r (1 + \deg p_k) \right) - 1$$

zeros in \mathbb{R} .

Step 3: Let $c = 1 + \lfloor \sqrt{2d} \rfloor$. Using tools from complex analysis we show that $|F(a + b\beta)|$ is very small for $1 \leq a, b \leq cL$.

Lemma 4.40

Using the same notation as above, let $1 \leq a, b \leq cL$. Then we have the following estimates for some constants $c_4, c_5 > 0$:

$$(i) \quad |F(a + b\beta)| \leq e^{c_4 L^{\frac{3}{2}} \log L - L^2}$$

(ii) if $\sigma : K \hookrightarrow \mathbb{C}$ is some embedding, then

$$\left| \sigma(F(a + b\beta)) \right| \leq e^{c_5 L^{\frac{3}{2}} \log L}$$

From Lemma 4.40 we deduce the following: Let $1 \leq a, b \leq cL$. Then there is a constant $c_6 > 0$ such that

$$\begin{aligned} \left| N_{K/\mathbb{Q}} \left(m^{D_1+2cLD_2} F(a + b\beta) \right) \right| &= m^{d(D_1+2cLD_2)} \left| \prod_{\sigma: K \hookrightarrow \mathbb{C}} \sigma(F(a + b\beta)) \right| \\ &\leq m^{d(D_1+2cLD_2)} e^{c_4 L^{\frac{3}{2}} \log L - L^2 + (d-1)c_5 L^{\frac{3}{2}} \log L} \\ &\leq e^{c_6 L^{\frac{3}{2}} \log L - L^2} \end{aligned}$$

Hence, for L sufficiently large, we find

$$\left| N_{K/\mathbb{Q}} \left(\underbrace{m^{D_1+2cLD_2} F(a + b\beta)}_{\in \mathcal{O}_K} \right) \right| < 1$$

and $F(a + b\beta) = 0$ for $1 \leq a, b \leq cL$. As $1, \beta$ are \mathbb{Q} -linearly independent, this implies that F has at least $c^2 L^2 > 2dL^2 = D_1 D_2$ real zeros, which is a contradiction to step 2.

5 Diophantine Approximations and applications to Diophantine problems

5.1 Linear forms in logarithms

Lecture 22,
26.01.2024

We recall

Theorem 5.1 (Baker, 1975)

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, $\gamma, \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$ and pick a choice of logarithms $\log \alpha_1, \dots, \log \alpha_n$. If $\gamma + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n \neq 0$, then

$$\left| \gamma + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n \right| \geq (eB)^{-C},$$

where $B = \max(H(\gamma), H(\beta_1), \dots, H(\beta_n))$ and $C = C(\alpha_1, \dots, \alpha_n)$ is an effectively computable constant.

We move towards applications:

Corollary 5.2

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, $b_1, \dots, b_n \in \mathbb{Z}$ and assume that $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$. Then

$$\left| \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1 \right| \geq (eB)^{-\tilde{C}},$$

where $B = \max(|b_1|, \dots, |b_n|)$ and $\tilde{C} > 0$ is an effectively computable constant depending on n and $\alpha_1, \dots, \alpha_n$.

Corollary 5.3

Let $a, b \in \mathbb{Z}$ with $a, b \geq 2$. Then there exists an effectively computable constant $C_1 = C_1(a, b) > 0$, such that for every $m, n \in \mathbb{N}$ with $a^m \neq b^n$, we have

$$|a^m - b^n| \geq \frac{\max(a^m, b^n)}{(e \max(m, n))^{C_1}}.$$

Remark: a) Let $a, b \in \mathbb{Z}_{\geq 2}$ and $k \in \mathbb{Z}_{\neq 0}$. Then there is a constant $C_2 > 0$, such that all solutions $m, n \in \mathbb{N}$ to the equation $a^m - b^n = k$ satisfy

$$\max(|m|, |n|) \leq C_2.$$

b) Using techniques from linear forms in logarithms, Tijdeman¹ (1976) even proved that there exists an effectively computable constant $C_3 > 0$, such that if $a^m - b^n = 1$ with $a, b, m, n \in \mathbb{Z}_{\geq 2}$, then

$$a^m, b^n \leq C_3.$$

Catalan's conjecture (1844), that the equation $a^m - b^n = 1$ with $a, b, m, n \in \mathbb{N}_{\geq 2}$ has only one solution, namely $3^2 - 2^3 = 1$, was proven by Mihăilescu² in 2002.

Corollary 5.4

Let p_1, \dots, p_t be prime numbers and let $(a_n)_{n \in \mathbb{N}}$ be the monotonically increasing sequence of natural numbers, which are composed of the primes p_1, \dots, p_t . Then there exist effectively computable constants C_4, C_5 only depending on t, p_1, \dots, p_n , such that

$$a_n - a_{n-1} \geq \frac{a_n}{C_4(\log a_n)^{C_5}}, \quad n \in \mathbb{N}_{\geq 2}.$$

Example: For $t = 1$, $p_1 = 2$, we obtain the sequence 2^n , $n \geq 0$, and the stronger statement

$$a_n - a_{n-1} = 2^n - 2^{n-1} = \frac{1}{2}2^n = \frac{1}{2}a_n.$$

Unit equations

Let K be a number field with ring of integers \mathcal{O}_K . Let $\alpha, \beta \in K^*$. What can we say about solutions to the equation

$$\alpha x + \beta y = 1 \tag{5.1}$$

in $x, y \in \mathcal{O}_K^*$? For example, are there infinitely many units x, y , such that $x = 1 - y$?

Theorem 5.5 (Baker 1960s, Gyöny 1978)

For every number field K , (5.1) has at most finitely many solutions and they can be effectively determined.

¹Robert Tijdeman (*1943), a Dutch mathematician

²Preda Mihăilescu (*1955), a Romanian mathematician, who currently teaches in Göttingen

Remark: The finiteness statement, though ineffective, has already been proved by Siegel in 1921.

We first make some preparations for the proof of Theorem 5.5: Let $d = [K : \mathbb{Q}]$ and assume that K has r real embeddings $\sigma_1, \dots, \sigma_r : K \hookrightarrow \mathbb{R}$ and $\sigma_{r+1}, \dots, \sigma_{r+s} : K \hookrightarrow \mathbb{C}$ complex embeddings with $r + 2s = d$ and $\sigma_{r+s+1} = \overline{\sigma_{r+1}}, \dots, \sigma_{r+2s} = \overline{\sigma_{r+s}}$. Let u_1, \dots, u_{r+s-1} be a fundamental system of units.

Lemma 5.6

Let $u \in \mathcal{O}_K^*$ and write $u = \xi u_1^{b_1} \cdots u_{r+s-1}^{b_{r+s-1}}$ with $\xi \in K$ a root of unity and $b_1, \dots, b_{r+s-1} \in \mathbb{Z}$. There exists an effectively computable constant $C_6 > 0$ depending only on K, u_1, \dots, u_{r+s-1} , such that

$$\max(|b_1|, \dots, |b_{r+s-1}|) \leq C_6 \log |\overline{u}|.$$

Lecture 23,
30.01.2024

Thue equations

Consider a binary form $F(X, Y) \in \mathbb{Z}[X, Y]$ of degree d , i.e. an expression of the form $F(X, Y) = a_0 X^d + a_1 X^{d-1} Y + \cdots + a_d Y^d$ with $a_0, \dots, a_d \in \mathbb{Z}$. Let $m \in \mathbb{Z}$.

Question: What can we say about solutions to the equation $F(X, Y) = m$ with $x, y \in \mathbb{Z}$? For $d > 3$ we call these Thue³ equations.

Our next goal will be to prove a finiteness result.

Theorem 5.7

Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a binary form of degree $d > 3$ with $F(1, 0) \neq 0$ and such that $F(X, 1)$ has at least three distinct roots in \mathbb{C} . Let $m \in \mathbb{Z} \setminus \{0\}$. Then there are finitely many solutions $x, y \in \mathbb{Z}$ to the equation $F(X, Y) = m$.

Lemma 5.8

Let K be a number field of degree d . Then there exists an effectively computable constant $C_8 > 1$, such that for every $\alpha \in \mathcal{O}_K$ there exists $u \in \mathcal{O}_K^*$ with

$$C_8^{-1} |N_{K/\mathbb{Q}}(\alpha)|^{\frac{1}{d}} \leq |\overline{u\alpha}| \leq C_8 |N_{K/\mathbb{Q}}(\alpha)|^{\frac{1}{d}}.$$

³after Axel Thue (1863 - 1922), a Norwegian mathematician

Remark: Note that $\left| \prod_{j=1}^d \sigma_j(u\alpha) \right| = |N_{K/\mathbb{Q}}(u\alpha)| = |N_{K/\mathbb{Q}}(\alpha)|$, and hence for every $1 \leq j \leq d$ we have $|N_{K/\mathbb{Q}}(\alpha)| \leq |\sigma_j(u\alpha)| \left(C_8 |N_{K/\mathbb{Q}}(\alpha)|^{\frac{1}{d}} \right)^{d-1}$, i.e.

$$|\sigma_j(u\alpha)| \geq C_8^{d-1} |N_{K/\mathbb{Q}}(\alpha)|^{\frac{1}{d}}$$

and we find that all of the conjugates $|\sigma_j(u\alpha)|$ are of computable size.

Lemma 5.9

Let $\alpha \in \mathcal{O}_K \setminus \{0\}$. Then there exist divisors β_1, \dots, β_t of α in \mathcal{O}_K , which can be determined effectively and with the following property: If $\gamma \mid \alpha$ in \mathcal{O}_K , then there exists $1 \leq j \leq t$ and $u \in \mathcal{O}_K^*$, such that $\gamma = u\beta_j$.

An application of Theorem 5.7: Consider the equation

$$y^2 = 2x(x-3) \tag{5.2}$$

in $x, y \in \mathbb{Z}$. For $x \in \mathbb{Z}$ we have $\gcd(2x, x-3) \mid 6$. Hence if $(x, y) \in \mathbb{Z}^2$ is a solution to (5.2), then we can write $2x = au^3$, $x-3 = bv^3$, $u, v \in \mathbb{Z}$ with $a, b \in \{\pm 2^k 3^l \mid 0 \leq k, l \leq 2\}$ and such that ab is a cube. Fix such a tuple (a, b) . Then we find a solution (u, v) to the equation $au^3 - 2bv^3 = 6$, which is a Thue equation. Using the finiteness result for Thue equations we can show that (5.2) has at most finitely many solutions.

Theorem (Baker, 1968)

Let $f(X) \in \mathbb{Z}[X]$ of degree d , $b \in \mathbb{Z} \setminus \{0\}$ and $n \in \mathbb{N}_{\geq 2}$. Assume that f has no multiple zeros and $d \geq 2$ if $n > 3$ and $d \geq 3$ if $n = 2$. Then the equation

$$by^n = f(x)$$

has at most finitely many integer solutions, which can be determined effectively.

Remark: For $n = 2$ we obtain hyperelliptic equations.

5.2 Roth's theorem

Definition (Approximation exponent)

Let $\alpha \in \mathbb{R}$. We define the *approximation exponent* τ_α of α as the infimum of real

numbers $\tau > 0$, such that for all $\delta > 0$ the inequality

$$|\alpha - \xi| \leq H(\xi)^{-\tau-\delta}$$

has only finitely many solutions $\xi \in \mathbb{Q}$.

Remark: a) For $\xi = \frac{p}{q} \in \mathbb{Q}$ with $p, q \in \mathbb{Z}$, $q \neq 0$, and $\gcd(p, q) = 1$, we have

$$H(\xi) = \max(|p|, |q|).$$

b) For $\alpha \in \mathbb{Q}$ we have $\tau_\alpha = 1$.

c) If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then by Dirichlet's theorem we have $\tau_\alpha \geq 2$.

Theorem (Roth, 1955)

Let $\alpha \in \mathbb{R} \cap \overline{\mathbb{Q}}$ and $\kappa > 2$. Then there exists a constant $c(\alpha, \kappa) > 0$, such that

$$|\xi - \alpha| \geq c(\alpha, \kappa) H(\xi)^{-\kappa}$$

for all $\xi \in \mathbb{Q}$, $\xi \neq \alpha$.

Remark: a) For all $\alpha \in \mathbb{R} \cap \overline{\mathbb{Q}}$ with $\alpha \notin \mathbb{Q}$, we deduce that the approximation exponent $\tau_\alpha = 2$.

b) The constant $c(\alpha, \kappa)$ in Roth's theorem is ineffective, i.e. with the methods of the proof one cannot compute $c(\alpha, \kappa)$ explicitly.

We will prove a weaker result:

Theorem 5.10 (Liouville, 1844)

Let $\alpha \in \mathbb{R} \cap \overline{\mathbb{Q}}$ of degree $d \geq 1$. Then there exists an effectively computable constant $c(\alpha) > 0$, such that

$$|\xi - \alpha| \geq c(\alpha) H(\xi)^{-d}$$

for all $\xi \in \mathbb{Q}$, $\xi \neq \alpha$.

Remark: If $f(X) = a_0 X^d + \cdots + a_d \in \mathbb{Z}[X]$ is a primitive polynomial of α , then

we define the *Mahler measure* of α as

$$M(\alpha) = |a_0| \prod_{j=1}^d \max(1, |a^{(j)}|)$$

and Liouville's theorem can be formulated as

$$|\alpha - \xi| \geq 2^{1-d} M(\alpha) H(\xi)^{-d}$$

for all $\xi \in \mathbb{Q}$, $\xi \neq \alpha$.

Conjecture (Lehmer, 1930s)

There is a constant $c > 0$, such that for all $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$, which is not a root of unity, one has

$$M(\alpha) \geq 1 + c.$$

Roth's theorem and Thue equations

Definition (Square-free binary form)

We call a binary form $F(X, Y) \in \mathbb{Z}[X, Y]$ *binary-free* if it is not divisible by a square $(\alpha X + \beta Y)^2 \in \mathbb{C}[X, Y]$, $\alpha, \beta \in \mathbb{C}$ not both zero.

Theorem 5.11

Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a square-free form of degree $d \geq 3$ and $\kappa > 2$. Then there exists a constant $c(F, \kappa) > 0$, such that for every $(p, q) \in \mathbb{Z}^2$ with $F(p, q) \neq 0$, we have

$$|F(p, q)| \geq c(F, \kappa) \max(|p|, |q|)^{d-\kappa}.$$

Corollary 5.12

Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a binary form, such that $F(X, 1)$ has at least three distinct roots, and $m \in \mathbb{Z} \setminus \{0\}$. Then the equation $F(X, Y) = m$ has at most finitely many integer solutions.

5.3 Schmidt's subspace theorem

Theorem 5.13 (Schmidt, 1972)

Let $L_i(X_1, \dots, X_n) = a_{i1}X_1 + \dots + a_{in}X_n$, $1 \leq i \leq n$, be n linearly independent integer forms with $a_{ij} \in \overline{\mathbb{Q}}$, $1 \leq i, j \leq n$, $n \geq 2$. Let $\delta, C > 0$. Then all integer

solutions $x \in \mathbb{Z}^n$ to the inequality

$$|L_1(x) \cdots L_n(x)| \leq C \|x\|_\infty^{-\delta}$$

are contained in a finite union of proper linear subspaces of \mathbb{Q}^n .

Question: Under which conditions does the inequality

$$|L_1(x) \cdots L_n(x)| \leq C \|x\|_\infty^{-\delta} \quad (5.3)$$

in Schmidt's subspace theorem have only *finitely many* solutions?

Example: Let $L_1(X) = a_{11}X_1 + \dots + a_{1n}X_n$ with $a_{11}, \dots, a_{1n} \in \mathbb{Z}$ and $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in \mathbb{Z}^n \setminus \{0\}$ a solution to $L_1(x^{(0)}) = 0$. Then (5.3) has infinitely many solutions given for example by $\lambda x^{(0)}$, $\lambda \in \mathbb{Z}$.

Question: How about restricting to the inequality

$$0 < |L_1(x) \cdots L_n(x)| \leq C \|x\|_\infty^{-\delta}?$$

Example: Let $n = 3$ and consider $L_1(X) = X_1 + \sqrt{2}X_2 + \sqrt{3}X_3$, $L_2 = X_1 - \sqrt{2}X_2 + \sqrt{3}X_3$, $L_3(X) = X_1 - \sqrt{2}X_2 - \sqrt{3}X_3$. By Dirichlet's theorem (4.3), there are in fact infinitely many solutions $(p, q) \in \mathbb{Z}^2$, $q \neq 0$, to

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

Note that for each solution $\max(|p|, |q|) \leq (1 + \sqrt{2})q$ and

$$\left| L_1(p, q, 0) L_2(p, q, 0) L_3(p, q, 0) \right| \leq (1 + \sqrt{2})^2 |q| |q|^{-2} \leq (1 + \sqrt{2})^3 \max(|p|, |q|)^{-1}.$$

I.e. for $C = (1 + \sqrt{2})^3$, $\delta = 1$, we obtain infinitely many solutions in the subspace $x_3 = 0$.

Lemma 5.14

Let $L_i(X) = a_{i1}X_1 + a_{i2}X_2$, $i = 1, 2$, be two linearly independent forms with $a_{ij} \in \overline{\mathbb{Q}}$, $1 \leq i, j \leq 2$, and $C, \delta > 0$. Then

$$0 < |L_1(x) L_2(x)| \leq C \|x\|_\infty^{-\delta}$$

has only finitely many integer solutions.

Theorem 5.15

Let $L_i(X) = a_{i1}X_1 + \cdots + a_{in}X_n$, $1 \leq i \leq r$, be linear forms with coefficients in $\overline{\mathbb{Q}}$ and $r \geq n \geq 2$. Assume that each n -tuple among L_1, \dots, L_r are linearly independent and let $C, \delta > 0$. Then all integer solutions to the inequality

$$|L_1(x) \cdots L_r(x)| \leq C \|x\|_\infty^{r-n-\delta}$$

are contained in a finite union of proper linear subspaces of \mathbb{Q}^n .

Applications to Diophantine approximations

Lemma 5.16

Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ be linearly independent over \mathbb{Q} . Then there exists a constant $C > 0$, such that the following inequality

$$|\alpha_1 x_1 + \cdots + \alpha_n x_n| \leq C \|x\|_\infty^{1-n}$$

has infinitely many solutions $x \in \mathbb{Z}^n$.

Theorem 5.17

Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ and $C, \delta > 0$. Then there are at most finitely many integer solutions to the inequality

$$0 < |\alpha_1 x_1 + \cdots + \alpha_n x_n| \leq C \|x\|_\infty^{1-n-\delta}$$

Remark: For $n \leq 2$ we get Roth's theorem.

Question: How well can we approximate $\alpha \in \overline{\mathbb{Q}}$ by algebraic numbers ξ of bounded degree?

Theorem 5.18

Let $\alpha \in \overline{\mathbb{Q}}$, $d \in \mathbb{N}$ and $C, \delta > 0$. Then the inequality

$$|\alpha - \xi| \leq CH(\xi)^{-d-1-\delta}$$

has at most finitely many solutions $\xi \in \overline{\mathbb{Q}}$ with $\deg \xi \leq d$.

Conjecture (Wirsing⁴, 1960)

Let $d \in \mathbb{N}$ and $\alpha \in \mathbb{R}$, such that α is not an algebraic number of degree $\leq d$. Then there is a constant $C(\alpha) > 0$ and there are infinitely many integer numbers ξ of degree at most d with

$$|\alpha - \xi| \leq C(\alpha)H(\xi)^{-d-1}.$$

Conjecture (Wirsing, 1961)

The same statement is true for the inequality

$$|\alpha - \xi| \leq C(\alpha)H(\xi)^{-\frac{d+\xi}{2}+\varepsilon}.$$

Applications to norm form equations

Let $\alpha \in \overline{\mathbb{Q}}$ of degree d and with monic polynomial $f(X) \in \mathbb{Q}[X]$. If we write $\alpha^{(1)}, \dots, \alpha^{(d)}$ for the conjugates of α , then $f(X) = \prod_{j=1}^d (X - \alpha^{(j)})$ and the homogenisation

$$F(X, Y) = Y^d f\left(\frac{X}{Y}\right) = \prod_{j=1}^d (X - \alpha^{(j)}Y)$$

is an irreducible binary form in $\mathbb{Q}[X, Y]$.

Definition (Norm form in two variables)

Let $K = \mathbb{Q}(\alpha)$ and extend the definition of the norm map $N_{K/\mathbb{Q}}$ to $K[X, Y]$. Then we can write

$$F(X, Y) = \prod_{\sigma_j: K \hookrightarrow \mathbb{C}} (X - \sigma_j(\alpha)Y) = N_{K/\mathbb{Q}}(X - \alpha Y).$$

We call $F(X, Y)$ a *norm form in two variables*.

Example: • if $[K : \mathbb{Q}] \geq 3$, $m \in \mathbb{Z}_{\neq 0}$ and α and algebraic integer, then the norm form equation $F(X, Y) = m$ is a Thue equation.

• if $d \in \mathbb{N}$, d not a square and $\alpha = \sqrt{d}$, $K = \mathbb{Q}(\sqrt{d})$, then the norm form equation

$$N_{K/\mathbb{Q}}(x - \sqrt{d}y) = x^2 - dy^2 = 1$$

is a Pell equation.

⁴Eduard Wirsing (1931-2022), a German mathematician

Norm form equations in more variables

Definition (Norm form)

Let K be a number field of degree $d \geq 2$, $n \in \mathbb{N}$ with $2 \leq n \leq d$, and $\alpha_1, \dots, \alpha_n \in K$. Let $\sigma_1, \dots, \sigma_d : K \hookrightarrow \mathbb{C}$ be the d distinct embeddings of K . We call the polynomial

$$F(X_1, \dots, X_n) = N_{K/\mathbb{Q}}(\alpha_1 X_1 + \dots + \alpha_n X_n) = \prod_{j=1}^d (\sigma_j(\alpha_1) X_1 + \dots + \sigma_j(\alpha_n) X_n)$$

a (partial) norm form.

Claim: $F(X_1, \dots, X_n) \in \mathbb{Q}[X_1, \dots, X_n]$. Let L be a normal closure of K and $G = \text{Gal}(L/\mathbb{Q})$. Then every $\tau \in G$ permutes the linear factors $\sigma_j(\alpha_1) X_1 + \dots + \sigma_j(\alpha_n) X_n$ and hence

$$\tau(F(X_1, \dots, X_n)) = F(X_1, \dots, X_n) \quad \forall \tau \in G.$$

Definition (Norm form equation)

If $m \in \mathbb{Q} \setminus \{0\}$, then we call the equation

$$N_{K/\mathbb{Q}}(\alpha_1 X_1 + \dots + \alpha_n X_n) = m, \tag{5.4}$$

$X_1, \dots, X_n \in \mathbb{Z}$, a norm form equation.

Theorem 5.19 (Schmidt, 1970s)

Let K be a number field of degree d , L a normal closure of K and $m \in \mathbb{Q} \setminus \{0\}$. Assume that $2 \leq n \leq d$, let $\alpha_1, \dots, \alpha_n \in K$ be \mathbb{Q} -linearly independent and assume that $\text{Gal}(L/\mathbb{Q}) \cong S_d$ ⁵. Then the norm form equation

$$N_{K/\mathbb{Q}}(\alpha_1 X_1 + \dots + \alpha_n X_n) = m$$

has at most finitely many integer solutions.

Lecture 26,
09.02.2024

Recall: Let H be a group, K a field and $\chi_1, \dots, \chi_M : H \rightarrow K^*$ be distinct characters. Then χ_1, \dots, χ_m are linearly independent in the K -vector space of all maps $H \rightarrow K$.

Corollary

Let K be a number field of degree d and let $\alpha_1, \dots, \alpha_d \in K$ be \mathbb{Q} -linearly independent. Let $\sigma_1, \dots, \sigma_d : K \hookrightarrow \mathbb{C}$ be the d distinct embeddings. Then the matrix $(\sigma_i(\alpha_j))_{1 \leq i, j \leq d}$

⁵Here S_d denotes the symmetric group in d elements.

has rank d .

Lemma 5.20

Let the notation be as above and assumptions as in Theorem 5.19. Then the linear forms

$$L_j(X) = \sigma_j(\alpha_1)X_1 + \cdots + \sigma_j(\alpha_n)X_n, \quad 1 \leq j \leq d$$

are in general position, i.e. any subset of n forms is linearly independent.

Example: If K has r real and $2s$ complex embeddings with $r + s - 1 > 0$ and $\alpha_1, \dots, \alpha_d$ is a basis for \mathcal{O}_K , then $N_{K/\mathbb{Q}}(\alpha_1 x_1 + \cdots + \alpha_n x_n) = 1$ has infinitely many integer solutions. (For $u \in \mathcal{O}_K^*$ we have $N_{K/\mathbb{Q}}(u^2) = 1$.) More generally, for $\alpha_1, \dots, \alpha_n \in K$, linearly independent over \mathbb{Q} , write

$$M = \left\{ \sum_{i=1}^n \alpha_i y_i \mid y_i \in \mathbb{Z} \ \forall i \right\}.$$

If there is a subfield $\mathbb{Q} \subseteq L \subseteq K$, which is neither \mathbb{Q} nor an imaginary quadratic field and such that $\mu \mathcal{O}_L \subseteq M$ for some $\mu \in \mathbb{K}^*$, then the norm form equation with $m = N_{K/\mathbb{Q}}(\mu)$ has infinitely many solutions. (as $N_{K/\mathbb{Q}}(u^2 \mu) = m \ \forall u \in \mathcal{O}_L^*$)

Theorem (Schmidt, 1972)

Let K be a number field, $\alpha_1, \dots, \alpha_n \in K$ \mathbb{Q} -linearly independent, and write $M = \{\sum_{i=1}^n \alpha_i y_i \mid y_i \in \mathbb{Z} \ \forall i\}$. Then the following are equivalent:

- (i) there does not exist $\mu \in K^*$ and a subfield $\mathbb{Q} \subset L \subset K$ with $L \neq \mathbb{Q}$ and not imaginary quadratic, such that

$$\mu \mathcal{O}_L \subseteq M$$

- (ii) for every $m \in \mathbb{Q}^*$, the equation

$$N_{K/\mathbb{Q}}(\alpha_1 x_1 + \cdots + \alpha_n x_n) = m$$

has at most finitely many solutions.

Example: Let $K = \mathbb{Q}(\sqrt[6]{2})$, $L = \mathbb{Q}[3]2$, $\alpha_1 = \sqrt[6]{2}$, $\alpha_2 = \sqrt{2}$, $\alpha_3 = (\sqrt[6]{2})^5$. One can show that $1, \sqrt[3]{2}, \sqrt[3]{4}$ is an integral basis for \mathcal{O}_L and $\mathcal{O}_L^* = \{\pm(1 - \sqrt[3]{2})^n \mid n \in \mathbb{Z}\}$. Then $N_{K/\mathbb{Q}}(x_1 \alpha_1 + x_2 \alpha_2 + x_3 \alpha_3) = N_{K/\mathbb{Q}}(\sqrt[6]{2}) = 2$ has infinitely many integer solutions.

More examples of "norm form equations"

Let K be a number field of degree d with a basis $\alpha_1, \dots, \alpha_d$ for the ring of integers \mathcal{O}_K . Let $c_j \in \mathbb{Z} \setminus \{0\}$, $1 \leq j \leq 3$, and consider the equation

$$c_1 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) + c_2 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) = c_3$$

with $x, y \in \mathbb{Q}^d$.

Theorem (Birch⁶, Davenport⁷, Lewis⁸, 1962)

If the homogenised equation

$$c_1 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) + c_2 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) = c_3 z^d \quad (5.5)$$

has non-trivial solutions over \mathbb{R} and modulo all prime powers p^l , then

$$c_1 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) + c_2 N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_d x_d) = c_3$$

has a solution in $x, y \in \mathbb{Q}^d$. Moreover, in that case one can show that (5.5) has infinitely many integer solutions $x, y \in \mathbb{Z}^d$, $z \in \mathbb{Z}$, and count them asymptotically in the box

$$N(B) = \# \left\{ x, y \in \mathbb{Z}^d \mid z \in \mathbb{Z}, (5.5) \text{ holds, } \|x\|_\infty, \|y\|_\infty, |z| \leq B \right\}.$$

The Hardy-Littlewood circle method works similarly asymptotic for $N(B)$ as $B \rightarrow \infty$ of the shape $N(B) \sim cB^{d+1}$.

Linear equations in norm forms

Notation as above. For $x \in \mathbb{Z}^d$ or $x \in \mathbb{Q}^d$, write $N(x) = N_{K/\mathbb{Q}}(\alpha_1 x_1 + \dots + \alpha_n x_n)$. Imagine we want to solve a system of linear equations

$$a_{i1} N(x_1) + \dots + a_{in} N(x_n) = 0, \quad 1 \leq i \leq m,$$

in $x_1, \dots, x_n \in \mathbb{Z}^d$ or \mathbb{Q}^d .

⁶Bryan John Birch (*1931), a British mathematician

⁷Harold Davenport (1907-1969), an English mathematician

⁸Donald John Lewis (1926-2015), an American mathematician

Definitions

Algebraic integer, [2](#)

Approximation exponent, [60](#)

Dedekind domain, [12](#)

Discriminant, [7](#)

House, [53](#)

Ideal

 Ideal class group, [13](#)

 Ideal classes, [13](#)

 Inertia degree, [16](#)

 Norm, [17](#)

 Prime, [15](#)

 Ramification index, [16](#)

 Ramified prime, [19](#)

Lattice, [24](#)

 Determinant, [24](#)

Index, [25](#)

Successive minima, [29](#)

Norm form, [66](#)

 in two variables, [65](#)

Norm form equation, [66](#)

Number field, [1](#)

 Discriminant, [9](#)

 Norm, [4](#)

 Trace, [4](#)

Ring

 integrally closed, [11](#)

 noetherian, [11](#)

 of algebraic integers, [3](#)

Transcendence, [42](#)

 Transcendence degree, [42](#)

Important theorems

Baker's theorem, [48](#)

Dirichlet's unit theorem, [37](#)

Gelfond-Schneider theorem, [48](#)

 weaker version, [54](#)

Jordan's theorem, [31](#)

Lindemann-Weierstraß theorem, [45](#)

 weaker version, [46](#)

Littlewood conjecture, [40](#)

Minkowski's convex body theorems

 first, [27](#)

 second, [29](#)

Roth's theorem, [61](#)

Schanuel's conjecture, [49](#)

Siegel's lemma, [51](#)

Subspace theorem, [41](#), [62](#)