Diophantine geometry

Prof. Gisbert Wüstholz
ETH Zurich

Spring Semester, 2013

Contents

About these notes 2
Introduction 2

1 Algebraic Number Fields **6**

2 Rings in Arithmetic **10**
2.1 Integrality 11
2.2 Valuation rings 14
2.3 Discrete valuation rings 16

3 Absolute Values and Completions **18**
3.1 Basic facts about absolute values 18
3.2 Absolute Values on Number Fields 21
3.3 Lemma of Nakayama 24
3.4 Ostrowski's Theorem 26
3.5 Product Formula for $\mathbb{Q}$ 29
3.6 Product formula for number fields 30

4 Heights and Siegel's Lemma **34**
4.1 Absolute values and heights 34
4.2 Heights in affine and projective spaces 35
4.3 Height of matrices 37
4.4 Absolute height 38
4.5 Liouville estimates 38
4.6 Siegel's Lemma 39

5 Logarithmic Forms **41**
5.1 Historical remarks 41
5.2 The Theorem 43
5.3 Extrapolations 48
5.4 Multiplicity estimates 53

6 The Unit Equation **54**
6.1 Units and S-units 54
6.2 Unit and regulator 56

7 Integral points in $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ **57**

8 Appendix: Lattice Theory **59**
8.1 Lattices 59
8.2 Geometry of numbers 60
8.3 Norm and Trace 62
8.4 Ramification and discriminants 64

About these notes

These are notes from the course on Diophantine Geometry of Prof. Gisbert Wüstholz in the Spring of 2013.

Please attribute all errors first to the scribe[1].

Introduction

The primary goal will be to consider the unit equation and especially its effective solution via linear forms in logarithms.

Lecture 1 introduces the primary object of study in algebraic number theory: an algebraic number field, and Lecture 2 will move on to arithmetic rings, including local rings, valuation rings and Dedekind rings. Dirichlet proved a basic result:

Theorem 0.1 (Dirichlet)**.** *The group of units in a number field is finitely generated.*

Lecture 3 treats the general theory of absolute values.

Both heights and Siegel's Lemma provide the subject matter of Lecture 4. We will begin with the notion of an absolute value for a general algebraic number field and study how they split upon extension of the number field. This information determines the height, a key tool for modern number theory (see, e.g., the six hundred page tome, *Heights*, by Bombieri and Gubler). Siegel's lemma will be treated only in a very elementary form. We will examine the problem of finding integers solutions of equations of the form,

$$x_1 a_1 + \cdots + x_n a_n = 0, \qquad a_1, \ldots, a_n \in \mathbb{Z} \text{ not all zero}$$

[1] Jonathan Skowera, jskowera@gmail.com

The discussion proceeds in Lecture 5 to linear forms in logarithms. Baker received the Fields medal for a version of the following theorem: (the best version, presented here, is due to

Theorem 0.2. *Let $a_1, \ldots, a_n$ be n integers different from 0 and 1, and let $b_1, \ldots, b_n$ be any integers. Then,*

$$\begin{aligned} |a_1^{b_1} \cdots a_n^{b_n} - 1| &> f(A, B, n), \quad A = \max(|a_i|), B = \max(|b_i|, e) \\ &> A^{-C(n) \log B} \end{aligned}$$

where $C(n)$ is an effective constant.[2]

The proof requires much machinery, so we will prove a simpler version. A trivial lower bound would be A^{-CB}.

The theorem includes an absolute value, and for a general algebraic number field, the various embeddings of the number field into the complex numbers $\mathbb{C}$ induce various absolute values.

Fix a finitely generated subgroup $\Gamma \subset K^\times$, then you can exactly determine (in principle) the solutions x and y of the unit equation. Hence the whole chain of results can be solved effectively. Solutions of covers of $\mathbb{P}^1$ follow from this and also an effective form of Siegel's theorem.

Roth received the Field's medal for the following theorem. Note that it is not effective!

Theorem 0.3 (Thue-Siegel-Dyson-Schneider-Roth)**.** *Let α be algebraic and irrational, and let $\epsilon > 0$. Then*

$$|x - \alpha y| < \max(|x|, |y|)^{-(1+\epsilon)}$$

has only finitely many solutions.

In Thue's version from the 19th century, $n = \deg \alpha$ and the exponent is $n/2$. Because this is the non-effective side of the story, we will leave it aside for the course and concentrate on the arithmetic approach.

Theorem 0.4 (Fermat's Last Theorem/Wiles' Theorem)**.** *The equation $x^n + y^n = 1$ has only finitely many solutions $x, y \in \mathbb{Q}$.*

Precursors to the theorem are the work of Masser and Wüstholz.

In Lecture 6, we will briefly touch on lattice theory. Let B be a simply connected domain in $\mathbb{R}^3$ containing a single lattice point: the origin. As

[2]We will call an effective constant any constant which one can, at least in principle, compute.

the domain expands, it grows to contain more and more point. If the expanded domain $\lambda B, \lambda \in \mathbb{R}_{>1}$ contains a lattice point on its boundary, then the number λ is said to be a minima of B.

Beginning with a smaller B generally increases the values of the successive minima, which are a function of the domain, $\lambda_1(B) < \lambda_2(B) < \cdots$ The λ_i depend on two parameters: on the volume of the domain $\mathrm{vol}(B)$ and on the determinant $\mathrm{vol}(\mathbb{R}^3/\Lambda)$ of the lattice. We only touch the surface of this very interesting theory whose modern developments include Hermitian lattices and their algebraic geometry, including generalizations of theorems like Riemann-Roch.

In lecture 7 we will discuss unit equations which are basic tools for solving a large class of diophantine equations and diophantine problems.

Definition 0.5 (Unit equation)**.** Let K be an algebraic number field, and let $\Gamma \subset K^\times$ be finitely generated. The unit equation is $x + y = 1$, where $x, y \in \Gamma$.

In lecture 8, we will examine the Thue equation.[3]

Definition 0.6 (Thue equation)**.** The Thue equation of degree d is of the form

$$f(x, y) = 1$$

for a homogeneous, degree d polynomial f with at least 3 distinct zeros and coefficients in $\mathbb{Q}$.

The unit equation has further ramifications for solutions of $y^n = f(x)$. The projective closure C of the solutions $\widetilde{C} = \{(x, y) \mid y^n = f(x)\}$ form a covering of $\mathbb{P}^1$ of degree n (technically, an algebraic curve C with a morphism $C \to \mathbb{P}^1$). Its genus can be calculated by the Hurwitz formula and defined over $\mathbb{Q}$. It has finitely many solutions in the integers x and y. This fact relies on the unit equation.

Conjecture 0.7 (Mordell conjecture)**.** *If C has genus $g > 1$ over the field $\mathbb{Q}$, then it has only finitely many rational points.*

Faltings proved this fact in 1983 in the more general case where the curve C may be over any finite extension of $\mathbb{Q}$.

Conjecture 0.8 (Shafarevich's Conjecture)**.** *Fix a number field K and a finite set of primes S. Then there are only finitely many elliptic curves (up to isomorphism) over K which have good reduction outside of S.*

[3]Thue's 150th birthday incidentally occurs this year.

The general Shafarevich conjecture makes the same claim for Abelian varieties.

Good reduction means the following: let $y^2 = x^3 + ax + b, a, b \in \mathbb{Z}$ be reduced modulo p. Its discriminant has the form $\Delta = 4a^3 + 27b^2$. When $p \mid \Delta$, the discriminant vanishes modulo p, and hence the curve is not elliptic (it is not smooth). In all others cases, the curve is said to have good reduction at p.

The Mordell conjecture is a consequence of the Shafarevich conjecture according to Parshin.

Faltings proved the Shafarevich conjecture, and hence the Mordell conjecture. The proof reduces the problem to a covering of $\mathbb{P}^1$ which hence has finitely many solutions. For a fixed Δ, the covering is $4a^3 = -27b^2 + \Delta$.

Theorem 0.9 (Siegel's theorem)**.** *Let $f(x, y) = 0$ define a curve of genus at least one with coefficients in $\mathbb{Q}$. Then there are only finitely many solutions $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.*

Certain questions can be generalized to number fields of class number one which are those number fields with rings of integers in which ideals factor uniquely into prime ideals.

Conjecture 0.10 (Gauß conjecture on class numbers)**.** *There are only finitely many imaginary quadratic number fields with class number one.*

Baker demonstrated the truth of this conjecture by showing that

$$|\Delta_K| \leq 163$$

for all $K = \mathbb{Q}(\sqrt{n})$, where $n < 0$.

Lecture 1

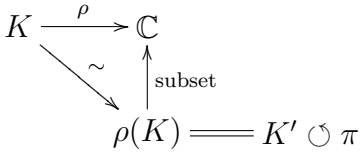
Algebraic Number Fields

Let $K \supseteq \mathbb{Q}$ be a finite extension of fields. Then $K = \sum_{i=1}^{n} \mathbb{Q} \cdot \omega_i$ is in particular a finite dimensional $\mathbb{Q}$-vector space, where $n = [K : \mathbb{Q}]$ is called the degree of the field extension.

The Galois group of the extension $K/\mathbb{Q}$ is $\mathrm{Gal}(K/\mathbb{Q}) = \{\phi : K \xrightarrow{\sim} K \mid \phi \text{ fixes } \mathbb{Q}\}$ (in fact, the condition is vacuous, since homomorphisms send 1 to 1). The extension need not be normal, hence not Galois, but the set,

$$\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C}) = \{\sigma : K \to \mathbb{C} \text{ homomorphism}\}$$

is a measure of the distance of the field from being Galois and gives a critera: the extension $K \supseteq \mathbb{Q}$ is Galois if and only if $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C}) \cong \mathrm{Gal}(K/\mathbb{Q})$. Furthermore, there is an action,



where the rightmost notation means that the Galois group acts on the image ($\pi \in \mathrm{Gal}(K/\mathbb{Q})$). So $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$ is a principal homogeneous space under π.

Theorem 1.1 (Primitive element theorem)**.** *The finite extension $K/\mathbb{Q}$ has a primitive element, i.e., there exists an $\alpha \in K$ such that $K = \mathbb{Q}(\alpha) = \mathbb{Q}[\alpha]$, where the first field is the field of rational functions in α which is equal to the polynomial ring $\mathbb{Q}[\alpha] = \mathbb{Q} \cdot 1 + \mathbb{Q} \cdot \alpha + \cdots + \mathbb{Q} \cdot \alpha^{n-1}$. Hence*

$$K \cong \mathbb{Q}[T]/(f_\alpha(T)) \qquad f_\alpha(T) \text{ irreducible, } f_\alpha(\alpha) = 0.$$

Let $\mathcal{O}_K \subset K$ be the ring of integers of K, i.e.,

$$\mathcal{O}_K = \{\alpha \in K \mid \text{minimal polynomial } f_\alpha = T^n + a_1 T^{n-1} + \cdots + a_{n-1} T + a_n \qquad a_i \in \mathbb{Z}\}$$

It is not direct that this forms a ring, so we must return to this point.

Rings and Ideals

Convention 1.2. *Rings R are always commutative and unitary. Homomorphisms of rings always send 1 to 1.*

$A \subseteq B$ means that A is a subset of B, while $A \subset B$ means that $A \subseteq B$ and $A \neq B$.

Recall that a ring is *entire*[1] if $0 \neq 1$ and the ring is without zero divisors, i.e. $ab = 0$ implies $a = 0$ or $b = 0$.

- A subset $I \subseteq R$ is an *ideal* if an only if I is an R-module.
- An ideal $\mathfrak{p} \subset R$ is *prime* if and only if $rs \in \mathfrak{p}$ implies $r \in \mathfrak{p}$ or $s \in \mathfrak{p}$ for all $r, s \in R$.[2] Equivalently, $\mathfrak{p}$ is prime if and only if $R/\mathfrak{p}$ is entire.
- An ideal $\mathfrak{q} \subset R$ is *primary* if and only if $rs \in \mathfrak{q}$ implies $r \in \mathfrak{q}$ or $s^n \in \mathfrak{q}$ for some n.[3] These are powers of prime ideals in the ring of integers of a number field.
- An ideal $\mathfrak{p} \subset R$ is *maximal* if and only if $R/\mathfrak{p}$ is a field.

Let M be an R-module. Then the annihilator ideal of M in R is

$$\mathrm{Ann}(M) = 0 : M = \{r \in R : rM = 0\} \subseteq R.$$

For example, if $R = \mathbb{Z}$ and $M = \mathbb{Z}/6\mathbb{Z}$, then the annihilator ideal is $\mathrm{Ann}(M) = 6\mathbb{Z}$.

To M are associated prime ideals which occur as annihilators of rank one modules:

$$\mathrm{ass}(M) = \{\mathfrak{p} \subset R \text{ prime ideal} \mid \exists m \in M : \mathfrak{p} = \mathrm{Ann}(Rm)\}.$$

Returning to the viewpoint of K as a $\mathbb{Q}$-vector space, to each $\mathbb{Q}$-basis of K, $\mathcal{B} = \{\omega_1, \ldots, \omega_n\} \subset K$, is associated a discriminant.

$$\Delta(\omega_1, \ldots, \omega_n) = \det(\sigma_i \omega_j)^2,$$

where $1 \leq i, j \leq n$ and $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\}$. Using the primitive element theorem, this can be made more explicit. Let $K = \mathbb{Q}[\alpha] = \mathbb{Q} \cdot 1 + \mathbb{Q} \cdot \alpha + \cdots + \mathbb{Q} \cdot \alpha^{n-1} \subset \mathbb{C}$. Then the primitive element α has a minimal

[1] Sometimes called *integral*

[2] Compare to the case of integers and a prime number p: $mn \in (p) \Rightarrow m \in (p)$ or $n \in (p)$ means $p \mid mn \Rightarrow p \mid m$ or $p \mid n$.

[3] Compare to the case of integers and a prime power p^n: $ab \in (p^n) \Rightarrow a \in (p^n)$ or $b^n \in (p^n)$ means $p^n \mid ab \Rightarrow p^n \mid a$ or $p^n \mid b^n$.

polynomial f_α, the lowest degree polynomial *with integer coefficients* such that $f_\alpha(\alpha) = 0$.

$$\begin{aligned} f_\alpha(T) &= a_0 T^n + a_1 T^{n-1} + \cdots + a_{n-1} T + a_n, \quad a_i \in \mathbb{Z} \\ &= a_0 \prod_{i=1}^{n} (T - \alpha_i) = a_0 \prod_{i=1}^{n} (T - \sigma_i \alpha) \\ \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C}) &= \{\sigma_1, \ldots, \sigma_n\} \qquad \sigma_i(\alpha) = \alpha_i \end{aligned}$$

The discriminant is a polynomial in the coefficients a_i which vanishes exactly when $f_\alpha(T)$ has a multiple root. The exact expression appears as the determinant of a Sylvester matrix, but the situation becomes much simpler if the roots α_i are used instead. In that case, a multiple root means that $\alpha_i = \alpha_j$ for some $i \neq j$, and the Vandermonde determinant calculates the discriminant:

$$\Delta(\mathcal{B}) = \Delta(1, \alpha, \ldots, \alpha^{n-1}) = \begin{vmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ \vdots & & & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{vmatrix}^2 = \prod_{i<j} (\alpha_j - \alpha_i)^2$$

For example, let $K = \mathbb{Q}[i]$. Then $\alpha = i$ and

$$f_\alpha(T) = T^2 + 1 = (T - i)(T + i) = (T - i)(T - \sigma(i))$$

where $\sigma : K \to K$ exchanges i and $-i$. In this case, the discriminant of the associated basis is $\Delta(1, i) = (i - (-i))^2 = (2i)^2 = -4$. Note that this recovers the usual discriminant of a quadratic polynomial from algebra: $b^2 - 4ac$.

Norm and Trace

The norm of an element in a number field is just the product of all the conjugates of it. But the norm and the trace can be defined using linear algebra using the left regular representation of K,

$$\rho : \begin{array}{ccc} K & \hookrightarrow & \mathrm{End}_{\mathbb{Q}}(K) \\ x & \mapsto & (a \mapsto x \cdot a) \end{array}.$$

Via this representation, the *norm* and *trace* on K are defined as the group homomorphisms,

$$N : \begin{array}{ccc} K^\times & \to & \mathbb{Q}^\times \\ x & \mapsto & \det(\rho(x)) \end{array}$$

$$Tr : \begin{array}{ccc} K & \to & \mathbb{Q} \\ x & \mapsto & \mathrm{tr}(\rho(x)) \end{array}$$

Question: Does $N(x) = \mathrm{Norm}_{K/\mathbb{Q}}(x)$? Does $\mathrm{Tr}(x) = \mathrm{Trace}_{K/\mathbb{Q}}(x)$?

We observe that if $B \subset \mathcal{O}_K$ is a basis of the ring of integers, then its discriminant $\Delta(\omega_1, \ldots, \omega_n)$ is an integer.

Theorem 1.3. *Let $\mathcal{B} \subset \mathcal{O}_K$ be a basis such that $|\Delta(\omega_1, \ldots, \omega_n)|$ is minimal among all bases. Then $\mathcal{O}_K = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n$.*

Two such minimal bases differ by an element in $\mathrm{GL}_n(\mathbb{Z})$.

Proof. Let $\alpha \in \mathcal{O}_K$. Then

$$\begin{aligned} \alpha &= a_1\omega_1 + \cdots + a_n\omega_n \qquad a_i \in \mathbb{Q}, 1 \leq i \leq n. \\ &= m_1\omega_1 + \cdots + m_n\omega_n + (r_1\omega_1 + \cdots + r_n\omega_n) \qquad m_i = \lfloor a_i \rfloor, r_i = a_i - m_i. \end{aligned}$$

Assume that not all r_i are zero, and without loss of generality, let us assume that $r_1 \neq 0$. Change basis:

$$\begin{aligned} \omega_1' &:= \alpha - m_1\omega_1 - \cdots - m_n\omega_n \\ \omega_2' &:= \omega_2 \\ &\vdots \\ \omega_n' &:= \omega_n. \end{aligned}$$

So

$$\Delta(\omega_1', \ldots, \omega_n') = \det \begin{pmatrix} r_1 & & & \cdots & \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}^2 \Delta(\omega_1, \ldots, \omega_n) = r_1^2 \Delta(\omega_1, \ldots, \omega_n).$$

Then $\omega_1', \ldots, \omega_n'$ is a basis in $\mathcal{O}_K$ with $|\Delta(\omega_1', \ldots, \omega_n')| < |\Delta(\omega_1, \ldots, \omega_n)|$, a contradiction. $\square$

Lecture 2

Rings in Arithmetic

In this lecture we consider localization, integrality and generalities on valuation rings.

Definition 2.1. Let R be a commutative, unital ring and $S \subset R$ a multiplicative set, i.e., a set with $1 \in S$, such that $x, y \in S$ implies $xy \in S$. Then the *localization* of R with respect to S makes precise the notion of a ring of fractions $\{\frac{r}{s} \mid r \in R, s \in S\}$. In particular, the localization $S^{-1}R$ can be formalized as a ring of (equivalence classes) of pairs (r, s):

$$S^{-1}R := R \times S/\sim \qquad (r,s) \sim (r',s') \Leftrightarrow \exists t \in S : t(rs' - r's) = 0$$

$$\begin{array}{cc} [(r,s)] + [(r',s')] = [(rs' + r's, ss')] & [(r,s)] \cdot [(r',s')] = [(rr', ss')] \\ 0 = [(0,1)] & 1 = [(1,1)] \end{array}$$

Direct calculation shows that addition and multiplication are well-defined on equivalence classes.

The reason for the requirement that S be multiplicative becomes clear; the denominator of the pairs contain terms of the form ss'. These formulas follow from the analogy with fractions, i.e., $(r, s) \leftrightarrow \frac{r}{s}$. For example,

$$\frac{r}{s} = \frac{r'}{s'} \Leftrightarrow rs' = r's \Leftrightarrow rs' - r's = 0$$

The additional t in the definition of $\sim$ deals with the case where R has zero-divisors, i.e., the counter-intuitive case when neither t nor $rs' - r's$ are zero and yet their product is. For entire rings with $0 \notin S$, the t may be left off.

There is a well-defined, canonical morphism,

$$\iota_S : \begin{array}{ccc} R & \to & S^{-1}R \\ r & \mapsto & [(r,1)] \end{array}$$

If R is entire, and $0 \notin S$, then ι_S is injective, and $S^{-1}R$ is entire.

- Let $R = K$ and $S = K^\times$. Then localization has no effect, since a field already contains all fractions, i.e., $S^{-1}R = K$.
- Localization at a multiplicative set containing 0 always produces the zero ring, i.e., if $0 \in S$, then $S^{-1}R = \{0\}$ is the zero ring.
- Let $\mathfrak{p} \subset R$ be a prime ideal. Then $S := R \setminus \mathfrak{p}$ is a multiplicative set by the definition of a prime ideal. The localization at S is denoted by $R_\mathfrak{p} := (R \setminus \mathfrak{p})^{-1}R$. It has a unique maximal ideal $\mathfrak{p}R_\mathfrak{p}$. In fact, there is a bijective correspondence,

$$\begin{array}{rcl} \{I \subset \mathfrak{p} \mid (\forall s \in R \setminus \mathfrak{p}) sa \in I \Rightarrow a \in I\} & \longleftrightarrow & \{I \subset R_\mathfrak{p}\} \\ I & \mapsto & IR_\mathfrak{p} \\ \iota_\mathfrak{p}^{-1}(J) & \leftarrow & J \end{array}$$

  This holds in particular for prime ideals $\mathfrak{q} \subset \mathfrak{p}$.
- Let $R = \mathbb{Z}$ and $S = \mathbb{Z} \setminus \{0\}$. Then $S^{-1}R = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\} = \mathbb{Q}$
- More generally, let R be arbitrary and $S = R \setminus \{0\}$, then $S^{-1}R = \mathrm{Frac}(R)$ is the field of fractions of R.
- Let $R = \mathbb{Z}$ and p be a prime number. Letting $S = \mathbb{Z} \setminus p\mathbb{Z}$ gives the localization $S^{-1}R = \mathbb{Z}_{(p)} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, (p, b) = 1\}$[1] The only ideals are the zero ideal (0) and powers of the maximal ideal (p^n) for $n \in \mathbb{N}$.
- Again let $R = \mathbb{Z}$ and p be a prime. Letting $S = \{p^n \mid n \in \mathbb{N}\}$ and localizing gives the ring of fractions whose denominators are powers of p, i.e., $S^{-1}\mathbb{Z} = \{\frac{a}{p^n} \mid a \in \mathbb{Z}, n \in \mathbb{N}\}$.

Definition 2.2. A ring of the form $R_\mathfrak{p}$ is called a local ring.[2] Equivalently, a local ring is a ring with a unique maximal ideal. They are called *local* in analogy with geometry, where the ring of germs of functions (i.e., locally defined functions) at a fixed point p of a smooth manifold has a unique maximal ideal formed by the functions vanishing at p.

2.1 Integrality

Let R be a commutative, unitary, entire ring. An (associative, commutative, unitary) R-algebra A is an R-module A with a bilinear multiplication $\beta : A \times A \to A$, i.e., $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$. We identify[3] R with its isomorphic image in A under the ring morphism $R \to A$

[1] The notation (p, b) means the greatest common divisor $gcd(p, b)$ of p and b.

[2] That is, R' is local if there exists a ring R and prime ideal $\mathfrak{p} \subset R$ such that $R' \cong R_\mathfrak{p}$.

[3] That is, we abuse notation to write $R \subset A$.

which maps $r \mapsto r \cdot 1_A$.

Definition 2.3. An element $a \in A$ is integral over R if and only if there exists $g(T) \in R[T]$of the form

$$g(T) = T^n + g_1 T^{n-1} + \cdots + g_{n-1}T + g_n$$

such that $g(a) = 0$.

The weakness of this definition is that it does not immediately show why sums and products of integral elements are integral. The below theorem remedies this.

Definition 2.4. An R-algebra A is integral over R if and only if every $a \in A$ is integral over R. For example, R is integral over R, since for any $r \in R$, the polynomial $g(T) = T - r$ has zero $g(r) = 0$.

Theorem 2.5 (Equivalent characterizations of integrality)**.** *There are three equivalent conditions for the integrality of an element of an R-algebra.*

(i) $a \in A$ *is integral over* R.

(ii) $R[a]$ *is a finitely generated* R*-module where* $R[a]$ *is the smallest subalgebra of* A *containing* a*, i.e.,* $R[a] := \cap_{A' \ni a} A'$.

(iii) There exists an $R[a]$*-module* M *which is finitely generated as an* R*-module and satisfies* $\mathrm{Ann}_{R[a]}(M) = 0$.

Proof. $(i) \Rightarrow (ii)$: $M_q := \langle 1, a, a^2, \ldots, a^{n+q} \rangle \subset R[a] = \sum_{n \geq 0} Ra^n$ for $q \geq 0$. We have

$$\begin{aligned} 0 = g(a) &= a^n + g_1 a^{n-1} + \cdots + g_{n-1}a + g_n \\ &\Rightarrow a^n = -g_1 a^{n-1} - \cdots - g_n \\ &\Rightarrow a^{n+q} = -g_1 a^{n+q-1} - \cdots - g_n a^q \in M_{q-1} \\ &\Rightarrow M_0 \subset M_q \subset M_{q-1} \subset \cdots \subset M_0 \\ &\Rightarrow R[a] = M_0 \text{ is finitely generated.} \end{aligned}$$

$(ii) \Rightarrow (iii)$: How can we find an M? We use the only candidate we have: $M = R[a]$. Then $\mathrm{Ann}_{R[a]}(R[a]) = 0$, because the ring $R[a]$ is entire.
$(iii) \Rightarrow (i)$: This follows immediately from a lemma.

Lemma 2.6. *Let* M *be an* $R[a]$*-module which is finitely generated over* R *such that* $\mathrm{Ann}_{R[a]}(M) = 0$ *and let* $I \subset R$ *be an ideal with* $aM \subset IM$*. Then* a *is a zero of a polynomial* $g(T) = T^n + g_1 T^{n-1} + \cdots + g_n$ *with* $g_i \in I$ *for all* i.

Let $M = Rm_1 + \cdots + Rm_n$. Writing multiplication by a in this basis gives $am_i = \mu_{i1}(a)m_1 + \cdots + \mu_{in}(a)m_n$. This defines a matrix $\mu(a) = (\mu_{ij}(a)) \in I \cdot M^{n\times n}(R)$. Let $\Phi := a \cdot I_{n\times n} - \mu(a)$, and then

$$a \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \mu(a) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \quad \Rightarrow \quad (a \cdot I_{n\times n} - \mu(a)) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = 0$$

Multiplying on the left by the adjugate matrix[4] shows that $\det \Phi = 0$:

$$(\det \Phi) \cdot \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = 0 \Rightarrow \det \Phi \cdot m_i = 0 \Rightarrow \det \Phi \cdot M = 0 \Rightarrow \det \Phi \in \mathrm{Ann}(M) = 0 \Rightarrow \det \Phi = 0$$

Hence $g(T) = \det(T \cdot I_{n\times n} - \mu(T))$ is a polynomial of the required form. □

Corollary 2.7. *(i) An element $a \in A$ is integral over R if and only if there exists a finite[5] R-algebra $A' \subset A$ such that $R[a] \subset A'$.*

(ii) The set $\overline{R}$ of elements of A which are integral over R is a subring of A called the integral closure of R in A.

Proof. Part (i). $\Rightarrow$: By (ii) of the above theorem, we may take $A' = R[a]$. Since a is integral over R, we note that this is the same as $R[a]$ being finitely generated. $\Leftarrow$: Let A' be a finite R-algebra and $R[a] \subset A'$. Hence $aA' \subset A'$. So A' is an $R[a]$-module with $\mathrm{Ann}_{R[a]}(A') = 0$, since $1 \in R[a] \subset A'$. Using the $(iii) \Rightarrow (i)$ part of the above theorem shows that a is integral over R.

Part (ii). Let $a, b \in \overline{R}$. Then there exist R-subalgebras $A', A'' \subset A$ such that $R[a] \subset A'$ and $R[b] \subset A''$ and A', A'' are finitely generated R-modules. The product $A'A''$ is also finitely generated.[6] Since $R[a+b], R[ab] \subset R[a,b]$ and $R[a,b] \subset R[a] \cdot R[b] \subset A'A''$, applying Part (i) gives the proof. □

The outlook for all this is the application to the case when $K \supseteq \mathbb{Q} \supset \mathbb{Z}$ for $A = K$ and $R = \mathbb{Z}$, where K is a number field. Then the ring of integers in K is $\mathcal{O}_K := \overline{\mathbb{Z}}$, the integral closure of $\mathbb{Z}$ in K.

Looking forward, we will consider absolute values which correspond roughly to valuations, and then move on to heights.

---

[4]The adjugate matrix $\mathrm{adj}(\Phi)$ always exists and satisfies $\mathrm{adj}(\Phi) \cdot \Phi = \det \Phi$.

[5]We use "finite" to mean "finitely generated as a module".

[6]Proof: $A' = \sum_{i=1}^{n} Ra_i, A'' = \sum_{j=1}^{m} Rb_j \Rightarrow A'A'' = \sum_{i,j} Ra_ib_j$.

2.2 Valuation rings

We again let R be a commutative, unitary ring and let $U(R) = R^\times$ be the units[7] of R.

Theorem 2.8. *Let $I = \cup_{\mathfrak{a} \subset R} \mathfrak{a}$ be the union of all proper ideals[8] of R. Then $I = R \setminus U(R)$, and I is an ideal if and only if R has a unique maximal ideal.*

Proof. The first claim is proved by showing both inclusions.

$I \subseteq R \setminus U(R)$**:** Let $x \in R \setminus U(R)$ and let $\mathfrak{a} = xR$. If $\mathfrak{a}$ were the whole ring, then we would have $1 \in xR$. In other words, there would be an element $y \in \mathfrak{a}$ with $xy = 1$. But then x would be a unit. Contradiction. Hence $x \in \mathfrak{a} \subseteq I$.

$I \supseteq R \setminus U(R)$**:** Let $x \in I$, so that $x \in \mathfrak{a} \subset R$. If x were a unit, then $\mathfrak{a}$ would contain $x^{-1}x = 1$ and hence be the whole ring, a contradiction. So $x \in R \setminus U(R)$.

The second claim is proved by showing both directions of implication.

($\Rightarrow$) Assume I is an ideal. Since I contains all proper ideals, it is by definition maximal. Let $\mathfrak{m}$ be another maximal ideal. Then $\mathfrak{m}$ is a proper ideal, hence I contains $\mathfrak{m}$. Since $\mathfrak{m}$ is maximal and I is a proper ideal, $\mathfrak{m} = I$, i.e., I is unique.

($\Leftarrow$) Assume R has a unique maximal ideal. Let x, y be elements of I. By the first part, they are not units, so $xR, yR \neq R$. Let $\mathfrak{m}$ be the unique maximal ideal which must contain both xR and yR. Thus it contains both x and y. Hence $x - y \in \mathfrak{m}$ and $rx \in \mathfrak{m}$ for all $r \in R$. Since I contains all proper ideas, it also contains $\mathfrak{m}$, so $x - y, rx \in I$, i.e., I is an ideal.

□

Note that the proposition implies that $U(R) = R \setminus I$.

Definition 2.9. Let R be entire with quotient field $K = \operatorname{Frac} R$. The ring R is called a valuation ring if either $r \in R$ or $r^{-1} \in R$ for every $r \in K$.

Definition 2.10. We will denote by $\mathcal{I}(R)$ the set of ideals of R.

[7] An element $r \in R$ is a unit if it is invertible with respect to multiplication, i.e., there exists $r^{-1} \in R$ such that $r^{-1}r = 1_R$.

[8] An ideal $I \subset R$ is a proper ideal if $I \neq R$.

Theorem 2.11. *Let R be a valuation ring. Then the set of ideals, $\mathcal{I}(R)$, is totally ordered by inclusion.*

Proof. Let $I, J \in \mathcal{I}(R)$. If $I = J$, then there is nothing to show, so let $I \neq J$. Without loss of generality, let I contain an x which is not in J. Let $y \in J \setminus \{0\}$. Then

$$\left(\frac{x}{y}\right) \cdot y = x \notin J \text{ and } y \in J \Rightarrow \frac{x}{y} \notin R$$

By hypothesis, R is a valuation ring, hence $y/x = (x/y)^{-1} \in R$ and $y = (y/x) \cdot x \in I$. Since y was arbitrary, $J \subset I$. □

The total ordering on $\mathcal{I}(R)$ immediately implies a couple corollaries.

Corollary 2.12. *Let R be a valuation ring. Then R has a unique maximal ideal.*

Corollary 2.13. *Let R be a valuation ring. Then $U(R) = R \setminus \mathfrak{m}(R)$.*

Corollary 2.14. *Let R be a valuation ring. Then $K \setminus R = \{x \in U(K) \mid x^{-1} \in \mathfrak{m}(R)\}$.*

Proof. **$\subseteq$:** By the definition of a valuation ring, if $x \notin R$, then $x^{-1} \in R$ and also $x^{-1} \notin U(R)$ (otherwise $x = (x^{-1})^{-1} \in U(R) \subset R$.) Hence $x^{-1} \in \mathfrak{m}(R)$.

$\supseteq$: Let $x \in U(K)$, and let $x^{-1} \in \mathfrak{m}(R)$. Since $x^{-1} \in \mathfrak{m}(R)$, $x \notin U(R)$. But if x^{-1} is not in $U(R)$, then neither is x in $U(R)$. Since $x^{-1} \in \mathfrak{m}(R)$, x cannot be in $\mathfrak{m}(R)$, or else $\mathfrak{m}(R)$ would be the whole ring. But if x is neither in $\mathfrak{m}(R)$ nor in $U(R)$, it must not be in R at all, that is, $x \in K \setminus R$.

□

Theorem 2.15. *Let R be a valuation ring. Then R is integrally closed in K.*

Proof. Let $x \in K$ be integral over R. Assume that $x \notin R$. Since R is a valuation ring, this implies $x^{-1} \in R$. By definition of integrality, there is a polynomial relation,

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0.$$

Division by x^n and rearrangement gives,

$$1 = -a_1 x^{-1} - \cdots - a_{n-1} x^{-(n-1)} - a_n x^{-n} = x^{-1}(-a_1 - \cdots - a_{n-1} x^{-(n-2)} - a_n x^{-(n-1)}).$$

So $x^{-1} \in U(R)$. Thus $x^{-1} \in \mathfrak{m}(R)$. In that case, $1 \in \mathfrak{m}(R)$, a contradiction. So the assumption was false and $x \in R$, that is, R is integrally closed. □

2.3 Discrete valuation rings

Let R be an entire ring, and $K = \operatorname{Frac} R$ be its quotient field.

Definition 2.16 (Real valuation and value group)**.** Let K be a field. A real valuation on K is a function $\nu : K \to \mathbb{R} \cup \{\infty\}$ such that

1. $\nu(xy) = \nu(x) + \nu(y)$.
2. $\nu(x + y) \geq \min\{\nu(x), \nu(y)\}$
3. $\nu(x) = \infty \Leftrightarrow x = 0$.

The image $\nu(K \setminus \{0\}) \subseteq \mathbb{R}$ is called the *value group*.

Definition 2.17 (Discrete valuation)**.** Let K be a field. A real valuation on K such that the value group is a subgroup of $\mathbb{Z}$ is called a discrete valuation in which case the value group has the form $e \cdot \mathbb{Z}$ for some $e > 0$.

We introduce the notation,

$$\begin{aligned} R_\nu &:= \{x \in K \mid \nu(x) \geq 0\} \subset K \qquad \text{(subring)} \\ \mathfrak{m}_\nu &:= \{x \in K \mid \nu(x) > 0\} \subset R_\nu \qquad \text{(ideal)} \end{aligned}$$

Then R_ν is a discrete valuation ring. (Let $x \in K \setminus R_\nu$. By the definition of R_ν, this means $\nu(x) < 0$. Hence $\nu(x^{-1}) = -\nu(x) > 0$, from which we conclude that $x^{-1} \in R_\nu$.)

In addition, $\mathfrak{m}_\nu$ is the unique maximal ideal whose existence is guaranteed by Corollary 2.12. (Let $\mathfrak{m}_n := \{x \in R_\nu \mid \nu(x) \geq n\}$ for $n \in \mathbb{Z}_{>0}$. Then

$$\mathfrak{m}_1 = \mathfrak{m}_\nu \supseteq \mathfrak{m}_2 \supseteq \cdots \supseteq \mathfrak{m}_n \supseteq \cdots$$

We take $\pi \in \mathfrak{m}_1$ with $\nu(\pi) = 1$. In fact, π is uniquely determined up to multiplication by a unit in R_ν. Let $\nu(\pi') = 1$. Then

$$\nu\left(\frac{\pi'}{\pi}\right) = 0 \Rightarrow \frac{\pi'}{\pi} \in R_\nu \setminus \mathfrak{m}_\nu = U(R_\nu) \Rightarrow \frac{\pi'}{\pi} = \epsilon \in U(R_\nu) \Rightarrow \pi' = \epsilon\pi.$$

The same argument shows that $\mathfrak{m}_n = \mathfrak{m}^n$ for all n:

$$x \in \mathfrak{m}_n \setminus \mathfrak{m}_{n+1} \Rightarrow \nu(x) = n \Rightarrow \nu\left(\frac{x}{\pi^n}\right) = 0 \Rightarrow x = \epsilon\pi^n \Rightarrow x \in \mathfrak{m}^n \Rightarrow \mathfrak{m}_n = \mathfrak{m}^n.$$

Proposition 2.18. *Let R be a discrete valuation ring. Then R is a principal ideal domain.*

Proof. Let $I \subseteq R$ be an ideal. Then $\nu(I \setminus \{0\}) \subseteq \mathbb{Z}$ is an additive subgroup. Hence $\nu(I) = n\mathbb{Z}$ for some n. Hence $I = \mathfrak{m}^n = (\pi^n)$ is a principal ideal. □

Theorem 2.19. *Let R be a noetherian, entire local ring such that every non-zero prime ideal is maximal. Then the following are equivalent:*

(i) *R is a discrete valuation ring.*

(ii) *R is integrally closed.*

(iii) *$\mathfrak{m}$ is a principal ideal.*

(iv) *$\mathfrak{m}/\mathfrak{m}^2$ is an $R/\mathfrak{m}$-vector space of dimension one.*

(v) *For all ideals $\mathfrak{a} \neq 0$ in R, there is an $n \geq 0$ with $\mathfrak{a} = \mathfrak{m}^n$.*

(vi) *There exists a $\pi \in R$ such that for all ideals $\mathfrak{a} \subseteq R$, $\mathfrak{a} = (\pi)^n$ for some $n \geq 0$.*

Lemma 2.20. *If R is a discrete valuation ring, then the module quotient $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is an $R/\mathfrak{m}$-vector space for all n, and it has dimension one.*

Proof. Since $\mathfrak{m}^{n+1}$ is an R-module, multiplication by r in $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is well-defined:

$$r : x + \mathfrak{m}^{n+1} \mapsto rx + r\mathfrak{m}^{n+1} = rx + \mathfrak{m}^{n+1}.$$

Hence $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is an R-module. Since $\mathrm{Ann}(\mathfrak{m}^n/\mathfrak{m}^{n+1}) = \mathfrak{m}$, it is furthermore an $R/\mathfrak{m}$-module. Finally, the isomorphism

$$\begin{array}{lcl} \mathfrak{m}^n = (\pi^n) : \mathfrak{m}^n/\mathfrak{m}^{n+1} & \xrightarrow{\sim} & R/\mathfrak{m} \\ 0 \neq x \in \mathfrak{m}^n : x = \epsilon\pi^n & \mapsto & \epsilon \end{array}$$

shows that it is a one dimensional vector space over $R/\mathfrak{m}$. □

Let R be a discrete valuation ring. We give it a topology of open sets generated by the basis of open neighborhoods of 0:

$$R \supset \mathfrak{m} \supset \mathfrak{m}^2 \supset \cdots \supset \mathfrak{m}^n \supset \cdots \supset 0$$

The topology is *separated* if

$$\cap_{n \geq 0} \mathfrak{m}^n = (0).$$

If R is noetherian, then by a lemma of Nakayama (to be proved later), $\mathfrak{m}I = I$.

Lecture 3

Absolute Values and Completions

Today we introduce the general theory of absolute values and completions with respect to them. Then we examine the particular case of absolute values on number fields. Finally, we prove Nakayama's Lemma.

3.1 Basic facts about absolute values

Definition 3.1 (Absolute value)**.** Let K be a field. An *absolute value* on K is a function $|\cdot| : K \to \mathbb{R}_{\geq 0}$ such that the following three properties hold.

1. $|x + y| \leq |x| + |y|$.
2. $|xy| = |x| \cdot |y|$.
3. $|x| = 0 \Leftrightarrow x = 0$.

An absolute value is called *non-archimedean* if it additionally satisfies the inequality

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

Lemma 3.2. *Let $(K, |\cdot|)$ be a value field (a field with absolute value). Then $|1| = |-1| = 1$.*

Proof. By multiplicativity, $1 = 1 \cdot 1$ and $1 = (-1)(-1)$ imply that $|1| = |1| \cdot |1|$ and $|-1| = |-1| \cdot |-1|$. Since only 0 has absolute value 0, we may divide by $|1|$ and $|-1|$ respectively, showing that $|1| = |-1| = 1$. □

Proposition 3.3. *The absolute value $|\cdot|$ is non-archimedean if and only if $|\cdot|$ is bounded on $\mathbb{Z} \subset K$.*

Proof. $\Rightarrow$: Let the absolute value be non-archimedean. We proceed by induction on n. In the base cases, $|1| = |1+0| \leq \max(|1|, |0|) \leq 1$. For the general case, let $n \in \mathbb{Z}_{>1}$. Then $|n| = |(n-1)+1| \leq \max(|n-1|, |1|) \leq 1$.

$\Leftarrow$: Now let the absolute value be bounded on $\mathbb{Z}$. Let $x, y \in K$. Consider the binomial formula

$$\begin{aligned} (x+y)^n &= \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k \\ |x+y|^n &\leq \sum_{j=0}^{n} \left| \left(\begin{array}{c} n \\ k \end{array} \right) \right| \cdot M^n, \qquad M = \max(|x|, |y|) \\ &\leq (n+1) c M^n \qquad c = \max_{n \in \mathbb{Z}} |n| < \infty \end{aligned}$$

Hence

$$|x+y| \leq (n+1)^{\frac{1}{n}} c^{\frac{1}{n}} M \qquad \text{for all } n \in \mathbb{N}$$

Taking the limit as n goes to infinity gives $|x+y| \leq M$.

□

An straightforward calculation shows that every absolute value $|\cdot| : K \to \mathbb{R}_{\geq 0}$ induces a metric $d : K \times K \to \mathbb{R}_{\geq 0}$ by the formula

$$d(x, y) = |x - y|.$$

Definition 3.4 (Completion)**.** Let $(K, |\cdot|)$ be a field with absolute value. Let

$$\mathcal{S} = \{(x_n) \in K^{\mathbb{N}} \mid (\forall \epsilon > 0)(\exists N)(\forall m, n > N)(x_m - x_n < \epsilon)\}$$

be the set of Cauchy sequences. The *completion*, $\widehat{K}$, of the field K with respect to $|\cdot|$ is

$$\widehat{K} = \mathcal{S}/\sim \qquad (x_n) \sim (y_n) \Leftrightarrow \lim_{n \to \infty} (x_n - y_n) = 0.$$

with the canonical embedding of K given by the map to constant sequences,

$$\iota : \begin{array}{ccc} K & \to & \widehat{K} \\ x & \mapsto & (x)_{n \in \mathbb{N}} \end{array}$$

The operations are defined element-wise, i.e., $[(x_n)] + [(y_n)] = [(x_n + y_n)]$ and $[(x_n)] \cdot [(y_n)] = [(x_n y_n)]$. These are well-defined and give it the structure of a field.

The following can be easily checked:

- The completion $\widehat{K}$ is complete, i.e., every Cauchy sequence in $\widehat{K}^{\mathbb{N}}$ converges in $\widehat{K}$.
- We extend $|\cdot|$ to $\widehat{|\cdot|}$ on $\widehat{K}$ by taking limits, i.e., $|[(x_n)]| := \lim_{n\to\infty} |x_n|$.

A simple question asks whether an arbitrary field admits an absolute value. In fact, ever field admits at least the trivial absolute value.

Definition 3.5 (Trivial absolute value)**.** Let K be a field. Every field admits at least the trivial absolute value defined by $|x| = 1$ for all $x \neq 0$ in K.

Lemma 3.6. *If $(K, |\cdot|)$ is a finite field with absolute value, then $|x| = 1$ for all $x \in K \setminus \{0\}$.*

Proof. The absolute value must take 0 to the zero element, 0_K, and every non-zero element is a torsion element, so $|x|^q = |x^q| = |1| = 1$. □

Definition 3.7 (Equivalent absolute values)**.** Let K be a field. Two absolute values $|\cdot|_1$ and $|\cdot|_2$ on K are said to be *equivalent* if they define the same topology.

Theorem 3.8. *Let K be a field, and let $|\cdot|_1$ and $|\cdot|_2$ be non-trivial absolute values on K. They are equivalent if and only if $|x|_1 < 1$ implies $|x|_2 < 1$ for all $x \neq 0$. In that case, there exists a $\lambda > 0$ such that*

$$|x|_2 = |x|_1^{\lambda} \qquad \forall x \in K \setminus \{0\}.$$

Proof. Let $x \in K \setminus \{0\}$. Then $|x| < 1$ if and only if $\lim_{n\to\infty} |x^n| = 0$. But equivalent topologies mean that $\lim_{n\to\infty} |x^n|_1 = \lim_{n\to\infty} |x^n|_2$. Hence

$$\begin{aligned} |\cdot|_1 \sim |\cdot|_2 &\Leftrightarrow \left(\lim_{n\to\infty} |x^n|_1 = 0 \Leftrightarrow \lim_{n\to\infty} |x^n|_2 = 0\right) \\ &\Leftrightarrow (|x|_1 < 1 \Leftrightarrow |x|_2 < 1). \end{aligned}$$

This proves the first part.

For the second part, let $x_0 \in K\setminus\{0\}$ such that $|x_0|_1 \neq 1$ (it exists since the absolute values are non-trivial). If $|x_0|_1 < 1$, then replace x_0 by its inverse, so that in either case, $|x_0|_1 > 1$. The goal is to show that $|x_0|_2 = |x_0|_1^{\lambda}$. Solving for λ suggests the definition,

$$\lambda = \frac{\log |x_0|_2}{\log |x_0|_1}.$$

This definition is well-defined, because $|x_0|_1$ is neither 1 nor ∞. It depends *a priori* on x_0. Our task will be to show that it is in fact independent of this choice. Let $x \in K \setminus \{0\}$ be arbitrary, and let $\alpha \in \mathbb{R}$ be such that $|x|_1 = |x_0|_1^\alpha$. Let (s_n) and (r_n) be sequences in $\mathbb{Q}$ whose limits are both α (i.e., $\lim_{n\to\infty} s_n = \lim_{n\to\infty} r_n = \alpha$), but such that $r_n \geq \alpha$ for all n while $s_n \leq \alpha$ for all n. Let $r_n = a_n/b_n$ with $a_n, b_n \in \mathbb{Z}$. Hence

$$|x|_1 = |x_0|_1^\alpha \leq |x_0|_1^{r_n} \Rightarrow \left|\frac{x^{b_n}}{x_0^{a_n}}\right|_1 = \frac{|x|_1^{b_n}}{|x_0|_1^{a_n}} \leq 1 \Leftrightarrow \left|\frac{x^{b_n}}{x_0^{a_n}}\right|_2 \leq 1 \Rightarrow \frac{|x|_2^{b_n}}{|x_0|_2^{a_n}} \leq 1 \Rightarrow |x|_2 \leq |x_0|_2^{r_n}.$$

Taking the limit as n goes to infinity results in the inequality $|x|_2 \leq |x_0|_2^\alpha$. The same argument with (r_n) replaced by (s_n) gives $|x|_2 \geq |x_0|_2^\alpha$. Hence

$$|x|_2 = |x_0|_2^\alpha = |x_0|_1^{\lambda\alpha} = |x|_1^\lambda.$$

for all $x \in K \setminus \{0\}$. □

Definition 3.9 (Discrete absolute value)**.** An absolute value is *discrete* if and only if $|K \setminus \{0\}| \subseteq \mathbb{R}_{\geq 0}$ is a discrete subset.

A (non-trivial) discrete absolute value satisfies $|K \setminus \{0\}| = q^{\mathbb{Z}}$ for some $0 < q < 1$.

Definition 3.10 (Homomorphism of valued fields)**.** A *homomorphism of valued fields*, $\sigma : (K, |\cdot|) \to (K', |\cdot|')$, is a field homomorphism $\sigma : K \to K'$ such that $|\sigma x|' = |x|$. In other words, $\sigma^*(K', |\cdot|') = (K, |\cdot|)$.

Next time we will restrict ourselves to special case of number fields. The set of homomorphisms from the field into the complex numbers which form a valued field. Hence each homomorphism induces an absolute value on the field. By this construction derive all archimedean absolute values on the field. Then we will determine all absolute values on the rationals by the theorem of Ostrowski which gives information about absolute values on number fields more generally. The absolute values on the rationals satisfy the product formula, i.e., that the product of all absolute values of a non-zero rational number gives one. Then we will be prepared to study heights.

3.2 Absolute Values on Number Fields

Let $K \supseteq \mathbb{Q}$ be a number field, and let

$$\Sigma := \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$$

be the $\mathbb{Q}$-vector space of $\mathbb{Q}$-linear field homomorphisms to $\mathbb{C}$. The standard absolute value on $\mathbb{C}$ is $|z| = (z \cdot \overline{z})^{\frac{1}{2}}$ and makes $(\mathbb{C}, |\cdot|)$ into a valued field. Conjugation $F^\infty : z \mapsto \overline{z}$ preserves the standard absolute value and fixes the real numbers. It is known in this context as the *Frobenius morphism at infinity*. The group $\{1, F^\infty\}$ acts on Σ.

Given an embedding $\sigma : K \to \mathbb{C}$, the standard absolute value on $\mathbb{C}$ can be pulled back to give an absolute value on K defined by,

$$|x|_\sigma := |\sigma x|, \qquad x \in K$$

Since F^∞ preserves the absolute value, the pullback $|\cdot|_{F^\infty \sigma}$ gives an equivalent absolute value on K. Such pull-backs give all archimedean absolute values on a number field K.

To illustrate these notions, let us consider a simple example. Let $K = \mathbb{Q}(\sqrt{2})$. Then $\Sigma = \{\sigma, \tau\}$. These are two embeddings in $\mathbb{C}$,

$$\sigma : a + b\sqrt{2} \mapsto a + b\sqrt{2} \qquad \tau : a + b\sqrt{2} \mapsto a - b\sqrt{2}$$

inducing two absolute values,

$$|a + b\sqrt{2}|_\sigma = a + b\sqrt{2} \qquad |a + b\sqrt{2}|_\tau = |a - b\sqrt{2}|$$

On the other hand, if $K = \mathbb{Q}(\sqrt{-2})$, then the embeddings,

$$a + b\sqrt{-2} \mapsto a + b\sqrt{-2} \qquad a + b\sqrt{-2} \mapsto a - b\sqrt{-2},$$

induce the same absolute value ($|a + b\sqrt{-2}| = \sqrt{a^2 + b^2}$), and there is just one (archimedean) absolute value on K.

In general, the set of embeddings can be decomposed into real and complex embeddings as $\Sigma = \Sigma_\mathbb{R} \cup \Sigma_\mathbb{C}$. Let $r = |\Sigma_\mathbb{R}|$ and $2s = |\Sigma_\mathbb{C}|$. Then $r + 2s = [K : \mathbb{Q}]$, and $r + s$ is the total number of archimedean absolute values.

It remains to classify non-archimedean absolute values on K. To that end, let $(K, |\cdot|)$ be a non-archimedean value field. Then

$$\begin{aligned} \mathcal{O} &:= \{x \in K \mid |x| \leq 1\} \supseteq \mathbb{Z} \\ \mathfrak{p} &:= \{x \in K \mid |x| < 1\} \end{aligned}$$

The set $\mathcal{O}$ is a ring. Indeed, if $x, y \in \mathcal{O}$, then both $x - y$ and xy are in $\mathcal{O}$, because $|x - y| < \max(|x|, |-y|) < 1$ and $|xy| = |x||y| < 1$. It is even a valuation ring, because for all $x \in K$, either $|x| \leq 1$ or $|x| > 1$, hence either $x \in \mathcal{O}$ or $x^{-1} \in \mathcal{O}$.

On the other hand, if we begin with $\mathcal{O}_K \subset K$, the integral closure of $\mathbb{Z}$ in K, and $\mathfrak{p} \subset \mathcal{O}_K$ is a non-zero prime ideal, then $\mathfrak{p}$ is maximal. Since $\mathfrak{p} \cap \mathbb{Z} = (p)$ is prime ideal, the number p is prime. Furthermore, the quotient by $\mathfrak{p}$, $\mathcal{O}_K/\mathfrak{p}$, is integral over $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$. Indeed, let $[x] \in \mathcal{O}_K/\mathfrak{p}$ be an equivalence class with $x \in \mathcal{O}_K$ its representative. Since x is integral over $\mathbb{Z}$, it satisfies a monic polynomial relation with integral coefficients which can be reduced modulo $\mathfrak{p}$:

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \Rightarrow [x]^n + [a_1][x]^{n-1} + \cdots + [a_{n-1}][x] + [a_n] = [0]$$

The second relation shows that $[x]$ is integral over $\mathbb{Z}/p\mathbb{Z}$.

Proposition 3.11. *Retaining the above notation, the ring $\mathcal{O}_K/\mathfrak{p}$ is a field. More generally, a ring which is integral over a field is a field.*

Proof. We must show that every element has an inverse. Let $[x] \in \mathcal{O}_K/\mathfrak{p}$. We would like to find a $[y] \in \mathcal{O}_K/\mathfrak{p}$ such that $[x][y] = [1]$. Since [x] is integral over $\mathbb{Z}/\mathfrak{p}\mathbb{Z}$,

$$[1] = -[a_n]^{-1}([x]^{n-1} + [a_1]x^{n-2} + \cdots + [a_{n-1}])[x].$$

Letting $[y] = -[a_n]^{-1}([x]^{n-1} + [a_1]x^{n-2} + \cdots + [a_{n-1}])$ finishes the proof. The same proof holds for the general case. □

Since the quotient by the ideal $\mathfrak{p}$ is a field, $\mathfrak{p}$ must be a maximal ideal.

Proposition 3.12. *The ring of integers $\mathcal{O}_K$ is integrally closed and every nonzero prime ideal of $\mathcal{O}_K$ is maximal.*

Since $\mathcal{O}_K$ is a noetherian domain, the immediate consequence of the proposition is that $\mathcal{O}_K$ is a Dedekind domain, i.e., an integrally closed, noetherian domain of Krull dimension one. Krull dimension one means exactly that every nonzero prime ideal is maximal.

In a Dedekind domain, every nonzero proper ideal factors uniquely into a product of prime ideals. This unique factorization property makes Dedekind domains the most important rings in arithmetic. They also enjoy the property that all primary ideals are powers of prime ideals.

We have begun with a prime ideal $\mathfrak{p}$ in $\mathcal{O}_K$, and we would now like to construct an absolute value from it. To begin, its powers define a filtration,

$$\mathcal{O}_K \supset \mathfrak{p} \supseteq \mathfrak{p}^2 \supseteq \mathfrak{p}^3 \supseteq \cdots \supseteq \mathfrak{p}^n \supseteq \cdots \supseteq (0)$$

We claim the above chain of ideals tend to the zero ideal in the limit, i.e.,

$$I := \cap_{n \geq 0} \mathfrak{p}^n = (0), \tag{3.1}$$

We will show this after proving Nakayama's lemma.

This vanishing implies that the inclusions of the above filtrations are strict and define a basis for a separated topology on $\mathcal{O}_K$. In this case, the filtration determines a valuation by the definition,

$$\begin{array}{rccl} \mathrm{ord}_{\mathfrak{p}} : & \mathcal{O}_K & \to & \mathbb{Z} \cup \{\infty\} \\ & x & \mapsto & \min\{e \in \mathbb{N} \mid x \in \mathfrak{p}^e \setminus \mathfrak{p}^{e+1}\} \end{array}$$

A straightforward calculation shows that this is indeed a valuation. It can be used to define an absolute value on $\mathcal{O}_K$ by choosing $0 < q < 1$ and letting

$$|x|_{\mathfrak{p}} := q^{\mathrm{ord}_{\mathfrak{p}}\, x}$$

This absolute value depends on a choice, namely a choice of real number, q, between 0 and 1. But different choices yield equivalent absolute values.

Definition 3.13 (Place)**.** A *place* of a number field K is an equivalence class of absolute values on K. There are two kinds of absolute values. An *infinite place* is a class containing an archimedean absolute value, while a *finite place* contains non-archimedean absolute values.

Every prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ defines a place by taking the equivalence class of absolute values containing $|\cdot|_{\mathfrak{p}}$. If $p = \mathrm{char}\, \mathcal{O}_K/\mathfrak{p}$, then $\mathrm{ord}_{\mathfrak{p}}(x) \mid p$ for all $x \in K$.

3.3 Lemma of Nakayama

Proposition 3.14. *Let R be a Dedekind domain, and let M bea finitely generated R-module. Let $\mathfrak{a} \subset R$ be an ideal, and let $\phi \in \mathrm{End}_R(M)$[1] be such that $\phi(M) \subseteq \mathfrak{a} \cdot M$[2]. Then ϕ satisfies the equation*

$$\phi^n + a_1\phi^{n-1} + \cdots + a_{n-1}\phi + a_n = 0, \qquad a_i \in \mathfrak{a}$$

Proof. Let $m_1, \ldots, m_k$ be generators for M over R. Then the hypothesis $\phi(M) \subseteq \mathfrak{a} \cdot M$ implies,

$$\phi(m_i) = \sum_{j=1}^{k} a_{ij} m_j, \qquad a_{ij} \in \mathfrak{a}_i.$$

[1]Note that $\mathrm{End}_R(M)$ is an R-module containing R by the canonical morphism $r \mapsto r{\cdot}id$.

[2]The set $\{\phi \in \mathrm{End}_R(M) \mid \phi(M) \subseteq \mathfrak{a} \cdot M\}$ is a right ideal in $\mathrm{End}_R(M)$.

This can be rewritten as,

$$\sum_{j=1}^{k}(\delta_{ij}\phi - a_{ij})m_j = 0, \qquad i = 1, \ldots, k.$$

Letting Ψ be the matrix $(\delta_{ij}\phi - a_{ij})$, the system has the form $\Psi \cdot m = 0$. Multiplying both sides by the adjugate of Ψ gives $dm := \det(\delta_{ij}\phi - a_{ij}) \cdot m = 0$. Hence $dm_i = 0$ for all i. Since the m_i generate M, it must be that $d = 0$. This finishes the proof with the relation $\det(\delta_{ij}\phi - a_{ij}) = 0$ as the desired equation. □

Corollary 3.15. *Let R again be a Dedekind domain with proper ideal $\mathfrak{a} \subset R$, and let M be a finitely generated R-module. If $\mathfrak{a}M = M$ then there exists $x \in R$ such that $x \equiv 1\ (\mathfrak{a})$ and $xM = 0$.*

Proof. Letting $\phi := id$ and applying the previous proposition gives that $id^n + a_1 id^{n-1} + \cdots a_n = 0$ for some $a_i \in \mathfrak{a}$. Let $x = 1 + a_1 + \cdots + a_n$. Then $x \equiv 1\ (\mathfrak{a})$ and $xM = (id^n + \cdots + a + n)M = 0$. □

Lemma 3.16 (Nakayama)**.** *If $\mathfrak{a} \subset R$ is an ideal contained in each maximal ideal, and $\mathfrak{a}M = M$, then $M = 0$.*

Proof. Let x be as in the corollary. Then $x \equiv 1\ (\mathfrak{a})$. Hence $x \equiv 1\ (\mathfrak{m})$ for all maximal ideals $\mathfrak{m}$, i.e., $x - 1 \in \mathfrak{m}$. If x were in $\mathfrak{m}$, then 1 would also be in $\mathfrak{m}$, so x is not in any maximal ideal. Hence x is a unit, so that $x^{-1} \in R$. Thus $xM = 0$ implies that $M = x^{-1}xM = 0$. □

Returning to the condition , the definition of I implies that

$$I = \mathfrak{p} \cdot I \tag{3.2}$$

The inclusion of $\mathfrak{p} \cdot I \subseteq I$ follows straight from the definition. In the other direction, the difficulty is that an element $x \in I = \cap_n \mathfrak{p}^n$ might not factor *uniformly* into a product $x = yz$ for $y \in \mathfrak{p}$ and $z \in \mathfrak{p}^n$ for *all* n. (It does of course factor separately for each n, but this is not sufficient.) Since $\mathcal{O}_K$ is noetherian, this factorization is a consequence of the Artin-Rees Lemma.

Lemma 3.17 (Artin-Rees)**.** *Let $I \subset R$ be an ideal in a Noetherian ring R. Let M be a finitely generated R-module and $N \leq M$, a submodule. There exists an integer $k \geq 1$ such that*

$$I^n M \cap N = I^{n-k}((I^k M) \cap N)$$

for all $n \geq k$.

Applying this to the case where $I := \mathfrak{p}$, $M := \mathcal{O}_K$ and $N := I$ gives that,

$$\mathfrak{p}^n \cap I = \mathfrak{p}^{n-k}(\mathfrak{p}^k \cap I), \qquad n \geq k$$

This can be rewritten as $I = \mathfrak{p}^{n-k}I$. Letting $n = k + 1$ results in the desired equality.

Were the equality 3.2 to hold for a non-zero I, it would seem strange, because the equality could be interpreted as saying that elements of I are infinitely divisible. In other words, one might take $x \in I$ and factor out an element of $\mathfrak{p}$, e.g., $x = p_1 \cdot x_1$. Iteratively applying the equality, an arbitrary number of elements of $\mathfrak{p}$ may be factored out $x = p_1 \cdots p_n x_n$.

More precisely, localizing the exact sequence,

$$0 \longrightarrow \mathfrak{p}I \longrightarrow I \longrightarrow I/\mathfrak{p}I \longrightarrow 0$$

at $\mathfrak{p}$ produces

$$0 \longrightarrow (\mathfrak{p}I)_\mathfrak{p} \xrightarrow{\sim} I_\mathfrak{p} \longrightarrow (I/\mathfrak{p}I)_\mathfrak{p} = 0 \longrightarrow 0.$$

By writing out a generic element, one sees that $(\mathfrak{p}I)_\mathfrak{p} \cong \mathfrak{a}I_\mathfrak{p}$, where $\mathfrak{a} = \mathfrak{p}\mathcal{O}_{K,\mathfrak{p}}$. Localization is exact, so this gives

$$\mathfrak{a}I_\mathfrak{p} \cong I_\mathfrak{p}.$$

Since $\mathfrak{a}$ is the unique maximal ideal in the localization, we may apply Nakayama's lemma to infer that $I_\mathfrak{p} = 0$. Since $\mathcal{O}_K$ has no zero-divisors, this implies that the intersection $I = 0$ which was to be shown.

We have constructed in a very natural way a set of absolute values. It is not yet clear whether these are all. This will be a central point of the next lecture. In the case of $K = \mathbb{Q}$, it will turn out that in principal, we have already found them all. The result must then be extended to arbitrary number fields K. In this direction we will prove the product formula and then proceed to heights.

3.4 Ostrowski's Theorem

To summarize the results so far, if $K \supseteq \mathbb{Q}$ is a number field and $\mathcal{O}_K \supset \mathbb{Z}$ is the integral closure of $\mathbb{Z}$ in K, then the absolute values fall into two categories:

Archimedean absolute values These arise as the pull-backs of $(\mathbb{C}, |\cdot|)$ by the embeddings of K in $\mathbb{C}$, $\mathrm{Hom}_\mathbb{Q}(K, \mathbb{C})$.

Non-archimedean absolute values The ring $\mathcal{O}_K$ is a Dedekind domain. Every prime ideal in $\mathcal{O}_K$ is maximal, and there is a descending chain $\mathcal{O}_K \supset \mathfrak{p} \supset \mathfrak{p}^2 \supset \cdots \supset (0)$ such that $\cap_{n=0}^{\infty} \mathfrak{p}^n = (0)$ (which implies that the inclusions are strict.) The associated absolute value is $|x|_{\mathfrak{p}} := \lambda^{\operatorname{ord}_{\mathfrak{p}} x}$ for $0 < \lambda < 1$.

Each non-archimedean absolute value defines a valuation ring $\mathcal{O}_{\mathfrak{p}} \supset \mathcal{O}_K$ by $\mathcal{O}_{\mathfrak{p}} := \mathcal{O}_{K,\mathfrak{p}}$, the localization at the set $\mathcal{S} = \mathcal{O}_K \setminus \mathfrak{p}$. These local rings intersect to give the original ("global") ring, i.e., $\cap_{\mathfrak{p}} \mathcal{O}_{\mathfrak{p}} = \mathcal{O}_K$.

Remark 3.18. A non-archimedean absolute value can be *normalized*. First a prime number p can be associated to $\mathfrak{p}$ by intersecting $\mathfrak{p} \cap \mathbb{Z} = (p)$ and choosing the positive generator. Let q be chosen so that $q^{-\operatorname{ord}_{\mathfrak{p}} p} = p$, i.e., so that $|p|_{\mathfrak{p}} = |p|$. The correct choice is $q = p^{-1/e_{\mathfrak{p}}}$, where $e_{\mathfrak{p}} = \operatorname{ord}_{\mathfrak{p}} p$. This is the normalization.

Remark 3.19. This works also for $\mathbb{Q}$ replaced by $\mathbb{Q}_p$. There are inclusions $K \supseteq \mathbb{Q}_p \supseteq \mathbb{Z}_p$. The ring of integers, $\mathcal{O}_p \subset K$ is the integral closure of $\mathbb{Z}_p$ in K.

Applying this theory to $\mathbb{Q}$ gives the absolute values

$$\begin{cases} |x|_\infty &= \max(x, -x) & \text{archimedean} \\ |x|_p &= p^{-\operatorname{ord}_p x} & \text{non-archimedean (normalize } \lambda = \frac{1}{p}) \end{cases}$$

These are all absolute values:

Theorem 3.20 (Ostrowski[3])**.** *Up to equivalence these are all the non-trivial absolute values on* $\mathbb{Q}$*.*

Proof. We fix a non-trivial absolute value on $\mathbb{Q}$. Then by Lemma 3.2, $|1| = |-1| = 1$. Let $k \in \mathbb{N}$ be a natural number. Then

$$|k| = |\underbrace{1 + \cdots + 1}_{k}| \leq k|1| = k \quad \Rightarrow \quad |-k| \leq k.$$

For integers $m, n > 1$, define

$$N := \max(1, |n|).$$

We can expand m in terms of powers of n, i.e., $m = m_0 + m_1 n + \cdots + m_d n^d$ where $0 \leq m_i < n$ and $m_d \neq 0$. The necessary power of n may be deteced by

[3]Ostrowski was professor at Basel

the inequality $m \geq n^d$ which rewritten becomes $d \leq \log m / \log n$. Applying absolute values to the expansion of m gives,

$$\begin{aligned} |m| &\leq |m_0| + |m_1||n| + \cdots + |m_d||n|^d \\ &\leq m_0 + m_1|n| + \cdots + m_d|n|^d \\ &< n(1 + |n| + \cdots + |n|^d) \\ &\leq n(1 + N + \cdots + N^d) \leq n(d+1)N^d. \end{aligned}$$

Choose an integer $s > 0$, and replace m by m^s. Then d becomes ds.

$$|m|^s = |m^s| \leq n(ds+1)N^{ds}$$

Taking the s^{th} root gives

$$|m| \leq n^{\frac{1}{s}}(ds+1)^{\frac{1}{s}}N^d$$

The limit as s approaches infinity is $|m| \leq N^d \leq N^{\log m / \log n}$ which rewritten gives

$$|m|^{\frac{1}{\log m}} \leq N^{\frac{1}{\log n}} = \max(1, |n|)^{\frac{1}{\log n}}$$

The maximum divides into two cases according to whether the absolute value is archimedean. In the first case, let $\max(1, |n|) = |n|$ for some integer $n > 1$. Interchanging the role of m and n in the above derivation gives the above inequality with m and n exchanged, i.e., $|n|^{1/\log n} \leq \max(1, |m|)^{1/\log n}$. Concatenating these two inequalities produces an equality,

$$|m|^{\frac{1}{\log m}} = |n|^{\frac{1}{\log n}}.$$

Since the equality holds for all m, it will hold in addition for all n. Independent of n, there is thus a λ such that $|n|^{\frac{1}{\log n}} = e^\lambda$. Hence,

$$|n| = e^{\lambda \log n} = n^\lambda, \qquad n \in \mathbb{Z}_{>1}.$$

In fact, for all rational numbers $r = p/q$, we have

$$|r| = |r|_\infty^\lambda$$

for the standard absolute value $|\cdot|_\infty$. This holds for integers, because $|-n| = |(-1)(n)| = |-1| \cdot |n| = n^\lambda = |n|_\infty^\lambda$, and $|-1| = |1| = 1 = 1^\lambda$ and $|0| = 0 = 0^\lambda$. For a general rational number, $qr = p$ implies $|q||r| = |p|$. Hence

$$|r| = |p|/|q| = |p|_\infty^\lambda / |q|_\infty^\lambda = |p/q|_\infty^\lambda = |r|_\infty^\lambda.$$

By Theorem 3.8, this means that $|\cdot| \sim |\cdot|_\infty$.

On the other hand, let $\max(1, |n|) = 1$ for some n, i.e., $|n| \leq 1$. Then induction on m shows that $|m| \leq 1$ for all integers m. Proposition 3.3 then implies that this is a non-archimedean absolute value bounded by 1. We define,

$$\mathcal{O} := \{x \in \mathbb{Q} \mid |x| \leq 1\} \qquad \mathfrak{p} := \{x \in \mathcal{O} \mid |x| < 1\}.$$

The ring $\mathcal{O}$ is a discrete valuation ring, and the ideal $\mathfrak{p} = (\pi)$ intersects $\mathbb{Z}$ in either $\mathfrak{p} \cap \mathbb{Z} = (p)$ or (0). If $|\cdot|$ is non-trivial, then there exists an $x \neq 0$ with $|x| \neq 1$. Either $|x|$ or $|x^{-1}|$ is less than one, so either $x \in \mathfrak{p}$ or $x^{-1} \in \mathfrak{p}$. So it is impossible that $\mathfrak{p} \cap \mathbb{Z}$ should be (0).

Let therefore p be the generator of $\mathfrak{p} \cap \mathbb{Z}$, and let $|\cdot|_\mathfrak{p}$ be the normalized absolute value corresponding to $\mathfrak{p}$ for which $|p|_\mathfrak{p} = |p|$. Let $n = p^k m$ be an integer factored so that $p \nmid m$, i.e., $m \notin \mathfrak{p} \cap \mathbb{Z}$. Hence $m \in \mathcal{O} \setminus \mathfrak{p}$, so $|m| = 1 = |m|_\mathfrak{p}$. Putting this together gives,

$$|n| = |p^k m| = |p|^k |m| = |p|_\mathfrak{p}^k |m|_\mathfrak{p} = |n|_\mathfrak{p}.$$

The equality may be extended to rational numbers, proving that $|\cdot| \sim |\cdot|_\mathfrak{p}$. □

3.5 Product Formula for $\mathbb{Q}$

Theorem 3.21 (Product Formula)**.** *Let $M_\mathbb{Q}$ be the set of normalized valuations on $\mathbb{Q}$. For $x \in \mathbb{Q}^\times$, we have*

$$\prod_{\nu \in M_\mathbb{Q}} |x|_\nu = 1.$$

Proof. Let $\mathcal{P}(x) := \prod_\nu |x|_\nu$. We show that $\mathcal{P}(x) = 1$. Since $\mathcal{P} : \mathbb{Q}^\times \to \mathbb{R}$ is multiplicative, it suffices to show the formula for prime x, and then apply prime decomposition to an arbitrary integer. Let $x = p$ be prime, and let q be any finite prime. Then

$$\operatorname{ord}_q p = \begin{cases} 1 & q = p \\ 0 & q \neq p \end{cases}$$

Therefore

$$\prod_\nu |p|_\nu = |p|_\infty |p|_p \cdot \prod_{q \neq p} |p|_q = p p^{-1} \cdot \prod_{q \neq p} 1 = 1.$$

□

Remark 3.22. Let $\mathbb{Q}_\nu$ be the completion of $\mathbb{Q}$ at $\nu \in M_\mathbb{Q}$, i.e., $\mathbb{Q}_\nu = \mathbb{Q}_p$ if p is a prime and $\mathbb{Q}_\nu = \mathbb{R}$ if $p = \infty$. Let

$$\mathbb{A}_\mathbb{Q} \subseteq \prod_\nu \mathbb{Q}_\nu$$

be the subset of $(x_\nu)_{\nu\in M_\mathbb{Q}}$ with $|x_\nu|_\nu \leq 1$ for all but finitely many ν, i.e., where x_ν is a p-adic integer for all but finitely many p. Then $\mathbb{A}_\mathbb{Q}$ is a ring called the *ring of Adèles* with an embedding of $\mathbb{Q}$ by

$$\iota : \begin{array}{ccc} \mathbb{Q} & \hookrightarrow & \mathbb{A}_\mathbb{Q} \\ x & \mapsto & (x)_\nu \end{array}$$

By the product formula, the image of $\mathbb{Q}$ lies in the hypersurface cut out by the equation $\prod |x_\nu|_\nu = 1$. This theory was initially developed by A. Weil.

3.6 Product formula for number fields

We now examine the general case of number fields. Let $(K, |\cdot|_\nu)$ be a valued field, and $(K_\nu, |\cdot|_w)$, its completion. Let $L \supseteq K$ be a finite, separable extension. The K-algebra L is a finite, separable field extension. Hence $L = K[x]$ for some $x \in L$, and

$$\pi : \begin{array}{ccc} K[T] & \to & L \\ T & \mapsto & x \end{array}$$

is a ring homomorphism with kernel (f) for an irreducible polynomial f. Its quotient is $L \cong K[T]/(f)$, and changing coefficients to the completion K_ν produces $L \otimes_K K_\nu \cong K_\nu[T]/(f)$. Factoring the polynomial $f = f_1 \cdots f_n$ over the completion K_ν leads to a finite set of pairwise orthogonal idempotents. Namely, by the Chinese Remainder Theorem,

$$K_\nu[T]/(f) = K_\nu[T]/(f_1) \cap \cdots \cap (f_n) \cong \prod_{i=1}^n K_\nu[T]/(f_i) =: \prod_{i=1}^n L_i$$

Hence $E_\nu := L \otimes_K K_\nu \cong \prod_{i=1}^n L_i$ for fields $L_i \supseteq K_\nu$. The projections $\epsilon_i \in \mathrm{End}_{K_\nu}(E_\nu)$ such that $\epsilon_i : E_\nu \to L_i$ are idempotents. They decompose the identity as $1 = 1 \otimes 1 = \sum_i \epsilon_w$ and multiply according to the rule,

$$\epsilon_w \cdot \epsilon_{w'} = \delta_{w,w'}\epsilon_w,$$

for the Kronecker delta, $\delta_{w,w'}$. These two properties can be seen using the diagram,

$$E_\nu \cong \prod_{i=1}^n L_i \underbrace{\xrightarrow{pr_i} L_i \xhookrightarrow{i}}_{e_i} \prod_{j=1}^n L_j \cong E_\nu.$$

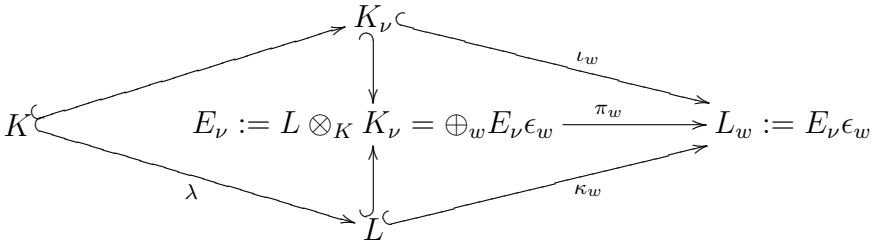
There are only finitely many prime ideals $\mathfrak{p}_w \subset E_\nu$, one for each ϵ_w,

$$\mathfrak{p}_w := \bigoplus_{w'\neq w} E_\nu\epsilon_{w'} = \mathrm{Ann}(E_w\epsilon_w) = \{x = \sum_{w'} x_{w'}\epsilon_{w'} \in E_w \mid x\epsilon_w = 0\}.$$

They are all maximal. Letting $x = \sum_{w'} x_{w'}\epsilon_{w'}$, the condition that $0 = x\epsilon_w$ becomes $0 = \sum_{w'} x_{w'}\epsilon_{w'}\epsilon_w = x_w\epsilon_w$, which happens exactly when $x_w = 0$. Hence the morphism π_w is a quotient by a maximal ideal, and $E_\nu\epsilon_w \cong (E_\nu/\mathfrak{p}_w)\cdot\epsilon_w = L_w$ is a field.

Constructing the absolute values

All absolute values of L will be found as absolute values of L_w in the following manner. The above reasoning may be summarized by the diagram,



By the density of the embedding of K in its completion, the extension of $|\cdot|_\nu$ to K_ν is unique. We have already constructed it above by taking limits. The embedding $L \to L \otimes_K K_\nu$ embeds L as a dense subset, and thus any absolute value on L extends *uniquely* to E_ν. For each norm on L_w which pulls back to $|\cdot|_w$ on K_ν, we can pull it back to L to get a norm extending the given norm $|\cdot|_\nu$ on K.

Each idempotent defines an extension $\|\cdot\|_w$ of the absolute value $\|\cdot\|_\nu$ to E_ν:

$$\|x\|_\nu = \sum_{i=1}^{n} |pr_i x|_\nu = \sum_{i=1}^{n} |x_i|_\nu, \qquad x = (x_1, \ldots, x_n) \in K_\nu^n.$$

This norms makes E_ν into a complete K_ν-vector space. It is an extension: $\iota_w^*\|\cdot\|_w = |\cdot|_\nu$. In turn, it defines an absolute value on L by

$$|x|_w := \|\kappa_w x\|_w.$$

This absolute value extends the original absolute value on K, i.e., $|x|_w = |\lambda x|_\nu$. We have shown,

Lemma 3.23. *Let K be a number field, and let $L \supseteq K$ be a finite, separable extension. In the above notation, every idempotent ϵ_w in E_ν defines an absolute value $|\cdot|_w$ on L which extends $|\cdot|_\nu$.*

Letting $n_{w/v} := [L_w : K_\nu]$ and using the fact that tensoring with a field is flat, we arrive at an important formula,

$$[L : K] = \dim_{K_\nu} E_\nu = \sum_{w|\nu} [L_w : K_\nu] = \sum_{w|\nu} n_{w/v}$$

Lemma 3.24.

$$\|x\|_w^{n_{w/v}} = |N_{L_w/K_\nu}(x)|_\nu$$

Proof. Let $L_w \supseteq K_\nu$. Then every $\sigma \in \mathrm{Gal}(L_w/K_\nu)$ preserves the norm, i.e., $\sigma * \|\cdot\|_w = \|\cdot\|_w$, because the norm on L_w which extends the norm on K_ν is unique. Then

$$\begin{aligned} \|x\|_w^{n_{w/v}} &= \prod_\sigma \|x\|_w = \prod_\sigma \|x^\sigma\|_w = \|\prod_\sigma x^\sigma\|_w \\ &= \|N_{L_w/K_\nu}(x)\|_w \\ &= |N_{L_w/K_\nu}(x)|_\nu. \end{aligned}$$

□

The decomposition of E_ν gives a morphism

$$x \mapsto (T_x : y \mapsto xy) \in \mathrm{End}_{K_\nu} E_\nu \cong \mathrm{End}_{K_\nu}(L_w)$$

The idempotents induce a decomposition, $T_x = \oplus_w T_{x\epsilon_w}$. Each idempotent defines an extension $\|\cdot\|_w$ of the norm of norm $\|\cdot\|_\nu$ to E_ν:

$$\|x\|_\nu = \sum_{i=1}^n |pr_i x|_\nu = \sum_{i=1}^n |x_i|_\nu, \qquad x = (x_1, \ldots, x_n) \in K_\nu^n.$$

This norms makes E_ν into a complete K_ν-vector space. It is an extension: $\iota_w^* \|\cdot\|_w = |\cdot|_\nu$. In turn, it defines an absolute value on L by

$$|x|_w := \|\kappa_w x\|_w.$$

This absolute value extends the original absolute value on K, i.e., $|x|_w = |\lambda x|_\nu$. Then,

$$N_{E_w/K_\nu}(x) = \det T_x = \prod_w \det T_{x\epsilon_w} = \prod_w N_{L_w/K_\nu}(x\epsilon_w).$$

This leads to the theorem,

Theorem 3.25 (Product formula for number fields)**.** *Let $x \in K$ be nonzero. Then*

$$\prod_{w \in M_K} |x|_w^{n_{w/v}} = 1.$$

Proof. By the product formula for $\mathbb{Q}$,

$$\begin{aligned}
\prod_{w \in M_K} |x|_w^{n_{w/v}} &= \prod_{w \in M_K} |N_{L_w/K_\nu}(x)|_\nu \\
&= \prod_{\nu \in M_\mathbb{Q}} \prod_{w/\nu} N_{L_w/K_\nu}(x)|_\nu = \prod_{\nu \in M_\mathbb{Q}} |N_{E_\nu/\mathbb{Q}_\nu}(x)|_\nu \\
&= \prod_{\nu \in M_\mathbb{Q}} |N_{E/\mathbb{Q}}(x)|_\nu = 1.
\end{aligned}$$

□

Lecture 4

Heights and Siegel's Lemma

4.1 Absolute values and heights

The product formula plays an important role in calculating with heights. Let $K \supseteq \mathbb{Q}$ be a number field. The places ν of K have the following form.

- $\nu \mid \infty$: There are two cases. For every, real embedding $\sigma : K \to \mathbb{R}$, there is a real place with absolute value $|x|_\nu := |\sigma x|$. Complex places have the form $|x|_\nu = |\sigma x|$ and correspond to pairs $(\sigma, \overline{\sigma})$ of an embedding $\sigma : K \to \mathbb{C}$ and its conjugate.

- $\nu \mid p$: We classified the finite places in Lemma 3.23. Each has an absolute value of the form $|x|_\nu = p^{-(f_\nu/n_\nu)\,\mathrm{ord}_\nu\, x}$. Here $n_\nu := [K_\nu : \mathbb{Q}]$ is the degree of the completion as an extension of $\mathbb{Q}$. Letting $\mathfrak{p} \subseteq \mathcal{O}_\nu$ be the valuation ideal of ν, we define $f_\nu := \dim_{\mathbb{F}_p} \mathcal{O}_\nu/\mathfrak{p} = [k_\nu : \mathbb{F}_p]$.

What is $|p|_p$? To calculate it, let $e_\nu = \mathrm{ord}_\nu\, p$. Then $n_\nu = e_\nu \cdot f_\nu$, so

$$|p|_p = p^{-\frac{f_\nu}{n_\nu}\,\mathrm{ord}_\nu p} = p^{-\frac{f_\nu e_\nu}{n_\nu}} = p^{-1}.$$

The product formula for K (see below) motivates the definition of the normalization of an absolute value.

Definition 4.1 (Normalization of an Absolute Value)**.** The normalization of an absolute value of a number field $K \supseteq \mathbb{Q}$ is,

$$\|x\|_\nu := |x|_\nu^{n_\nu}$$

where $n_\nu = [K_\nu : \mathbb{Q}_p]$ if ν is a finite place, and

$$n_\nu = \begin{cases} 1 & \sigma \text{ real} \\ 2 & \sigma \text{ complex} \end{cases}$$

when ν is an infinite place.

With this definition in hand, the product formula for K takes the form,

$$\prod_\nu \|x\|_\nu = \prod_\nu |x|_\nu^{n_\nu} = 1, \qquad x \in K, x \neq 0.$$

Definition 4.2. Let $K/\mathbb{Q}$ be a finite extension. The *(logarithmic) height* of an element $x \in K$ is

$$h_K(x) := \sum_\nu \log^+ \|x\|_\nu.$$

where $\log^+ r := \max\{0, \log r\}$.

The height enjoys a number of properties inherited from the norm. It is non-negative, and satisfies a triangle inequality, $h_K(x) \geq 0$ and $h_K(x+y) \leq h_K(x) + h_K(y)$. It also remains invariant under multiplication by $\lambda \in K^\times$.

Example 4.3. Let $K = \mathbb{Q}$, and let $x \in \mathbb{Q}$. We may factor its numerator and denominator:

$$x = \frac{a}{b} = \prod_{p \text{ prime}} p^{\epsilon(p)} = \prod_{\epsilon(p)\geq 0} p^{\epsilon(p)} \prod_{\epsilon(p)<0} p^{\epsilon(p)}$$

The height can be calculated as follows.

$$\begin{aligned} \log^+ |x| &= \begin{cases} \sum_p \epsilon(p) \log p & x \geq 0 \\ 0 & x < 0 \end{cases} \\ \log^+ |x|_p &= \begin{cases} -\epsilon(p) \log p & \epsilon(p) \leq 0 \\ 0 & \epsilon(p) > 0 \end{cases} \end{aligned}$$

If $x = \frac{3}{5}$, then

$$\begin{aligned} \log^+ |x|_\infty &= 0 \\ \log |x|_3 &= 0 \\ \log |x|_5 &= -(-1) \cdot \log 5 = \log 5. \end{aligned}$$

Hence $h_\mathbb{Q}(\frac{3}{5}) = \log 5 = \log(\max(3,5))$.

Exercise 4.4. What is $h(\sqrt{2})$?

4.2 Heights in affine and projective spaces

Let $\mathbb{A}^n(K) = \{(x_1, \ldots, x_n) \mid x_i \in K\}$ be affine n-space.

Definition 4.5. The norm[1] of an point $x \in \mathbb{A}^n(K)$ is defined by

$$\|x\|_\nu := \begin{cases} \max\{\|x_i\|_\nu\} & \nu \text{ non-archimedean} \\ \left(\sum_{i=1}^n \tau(x_i)^2\right)^{1/2} & \nu \text{ real} \\ \sum_{i=1}^n \tau(x_i)\overline{\tau}(x_i) & \nu \text{ complex} \end{cases}$$

Definition 4.6 (Heights in Affine Space)**.** The *(logarithmic) height* of a point $x \in \mathbb{A}^n(K)$ is

$$h_K(x) := \sum_\nu \log \|x\|_\nu.$$

The lack of $\log^+$ means that the height of $x \in K$ might not agree with the height of $x \in \mathbb{A}^1(K)$, i.e., $h_K(x) \neq h_K(x)$.

These heights are compatible. For example, given the embedding

$$i_1 : \begin{array}{ccc} \mathbb{A}^1 & \hookrightarrow & \mathbb{A}^2 \\ x & \mapsto & (1, x) \end{array}$$

the height on $\mathbb{A}^1$ agrees with the restriction of the height on $\mathbb{A}^2$, i.e., $h_K(i_1(x)) = h_K(1, x) = \sum_\nu \log \|x\|_\nu = h_K(x)$.

Recall that projective space is defined as the quotient of affine space by the multiplicative group of the field:

$$\mathbb{P}^n(K) = \mathbb{A}^{n+1}(K)/K^\times$$

We can extend the definition of heights to projective space virtue of a lemma:

Lemma 4.7. *Let h_K be the height on $\mathbb{A}^1(K)$ and $x \in \mathbb{A}^1(K)$. Then for all $\lambda \in K^\times$, $h(\lambda x) = h(x)$.*

Proof. By the product formula for λ,

$$h(\lambda x) = \sum_\nu \log \|\lambda x\|_\nu = \sum_\nu \log \|\lambda\|_\nu \|x\|_\nu = \underbrace{\sum_\nu \log \|\lambda\|_\nu}_{=0} + \sum_\nu \log \|x\|_\nu = h(x).$$

□

Definition 4.8 (Heights in Projective Space)**.** The *(logarithmic) height* of a point $(x_0 : \cdots : x_n) \in \mathbb{P}^n(K)$ can be defined using the definition for affine space, because heights are invariant under the $K^\times$-action:

$$h_K((x_0 : \cdots : x_n)) := h_K(x_0, \ldots, x_n)$$

[1]Note that this does not agree with many standard definitions, where $\max\{\|x_i\|_\nu\}$ is also used for the archimedean case. But our definition agrees with the matrix definition to follow.

4.3 Height of matrices

We will denote by $M_{m,n}(K)$ the K-vector space of $m \times n$-matrices.

There are two ways of assigning height to $X \in M_{m,n}(K)$. The naive definition takes the height of the corresponding point in affine space under the isomorphism $M_{m,n}(K) \cong \mathbb{A}^{m \times n}$ defined by

$$X \mapsto (\ldots, x_{ij}(X), \ldots) \qquad x_{ij}(X) = (i,j)^{th}\text{-entry}$$

The better way is due to Bombieri-Vaaler.

Definition 4.9 (Height of a Matrix)**.** The *height of a matrix* $X \in M_{m,n}(K)$ is defined as follows. Assume $m \leq n$, and let $I \subseteq \{1, \ldots, n\}$ be a choice of a rank m minor (i.e., $|I| = m$). Let X_I denote the minor of coefficients whose columns and rows indices appear in I. Let

$$H_\nu(X) := \begin{cases} \max_{|I|=m} \| \det X_I \|_\nu & \nu \text{ non-archimedean} \\ \| \det X\overline{X}^T \|_\nu^{1/2} & \nu \text{ real} \\ \| \det X\overline{X}^T \|_\nu & \nu \text{ complex} \end{cases}$$

Then the height of X is defined to be

$$h(X) = \sum_\nu \log H_\nu(X).$$

Remark 4.10. A more elegant way to introduce these heights applies the inclusion

$$i: \begin{array}{ccc} M_{m,n}(K) & \hookrightarrow & \mathbb{A}^N \\ X & \mapsto & (\ldots, \det X_I, \ldots) \end{array}$$

for $N = \begin{pmatrix} n \\ m \end{pmatrix}$ and $m \leq n$. Then the case $m = 1$ has $M_{1\times n} = \mathbb{A}^n$, and we recover heights for affine space.

$$h : X \to \begin{cases} \max \| \det X \|_\nu & \nu \nmid \infty \\ \| \det(X\overline{X}^T) \|_\nu^{1/2} & \nu \mid \infty \end{cases}$$

The properties of the heights of matrix include the following.

- $h(X) \geq 0$.
- There are only finitely many $X \in M_{m,n}(X)/K^\times$ such that $h(X) \leq C$ for given C.

- Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ for $X_i \in M_{m_i,n}(K)$, and let $m = m_1 + m_2$. Then

$$h(X) \leq h(X_1) + h(X_2).$$

- Let P be an element of $K[x_1, \ldots, x_n]_d \subseteq K[x_1, \ldots, x_n]$, the polynomials of degree at most d. There is an isomorphism

4.4 Absolute height

Since $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2})$, we now have two definitions of height for an $x \in \mathbb{Q}$: the height on $\mathbb{Q}$ and the height on $\mathbb{Q}(\sqrt{2})$. These do not agree. The height in the $\mathbb{Q}(\sqrt{2})$ will be $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}]$ times the height in $\mathbb{Q}$. This property holds in general and leads to a definition of height preserved by field extension.

Exercise 4.11. Let $p \in \mathbb{Z}$ be prime. Calculate $h_{\mathbb{Q}}(p)$ and $h_{\mathbb{Q}(\sqrt{2})}(p)$.

Definition 4.12. The *absolute height* of a non-zero element $x \in \overline{\mathbb{Q}} \setminus \{0\}$ is given by choosing $K \subset \overline{\mathbb{Q}}$ containing x and defining

$$h(x) := \frac{1}{[K : \mathbb{Q}]} h_K(x).$$

This does not depend on the choice of $K \ni x$.

4.5 Liouville estimates

Let $x = \frac{u}{v} \in \mathbb{Q}$ be non-zero with $u, v \in \mathbb{Z}$. Then

$$|x| \geq \frac{1}{|v|} \geq \frac{1}{\max(|u|, |v|)} = \frac{1}{h(x)}. \tag{4.1}$$

This can be generalized as follows.

Proposition 4.13. *Let $L = \ell_1 T_1 + \cdots \ell_n T_n$ for $\ell_i \in \mathbb{Z}$. Let $K \supseteq \mathbb{Q}$ be a number field, w, an infinite place of K, and $\xi \in K^n$. If $L(\xi) \neq 0$, then*

$$\log \|L(\xi)\|_w \geq -(h(\xi) + h(L)) + \log \|L\|_w + \log \|\xi\|_w$$

where $|L| = (\sum_{i=1}^n \ell_i^2)^{1/2}$.

Proof. By the product formula

$$\begin{aligned} 1 = \prod_{\nu} \|L(\xi)\|_\nu &= \|L(\xi)\|_w \cdot \prod_{\nu \neq w} \|L(\xi)\|_\nu \\ &\leq \|L(\xi)\|_w \prod_{\nu \neq w} \|L\|_\nu \cdot \|\xi\|_\nu \end{aligned}$$

Note that the finite places ν satisfy

$$|L(\xi)|_\nu = |\ell_1\xi_1 + \cdots + \ell_n\xi_n| \leq \max |\ell_i\xi_i|_\nu \leq \max |\ell_i|_\nu \cdot \max |\xi_i|_\nu.$$

so $\|L(\xi)\|_\nu \leq H_\nu(L) \cdot H_\nu(\xi)$. For the rest, one has $\nu \mid \infty$. For the infinite places ν, the Cauchy-Schwarz inequality gives

$$|L(\xi)|_\nu = |\langle \ell, \xi \rangle|_\nu \leq \langle \ell, \ell \rangle^{1/2} \langle \xi, xi \rangle^{1/2}$$

Therefore,

$$\begin{aligned} 1 &\leq \|L(\xi)\|_w \prod_{\nu} \|L\|_\nu \cdot \prod_{\nu} \|\xi\|_\nu \cdot \|L\|_w^{-1} \cdot \|\xi\|_w^{-1} \\ &= \|L(\xi)\|_w \prod_{\nu} h(L)h(\xi) \cdot \|L\|_w^{-1} \cdot \|\xi\|_w^{-1}. \end{aligned}$$

Taking logarithms gives the result. □

This generalizes the equation 4.1 as follows. Let $L = T_2$, and let $\xi = (1, x)$. Then $L(\xi) = x$, $h(L) = 0$ and $h(\xi) = \log \max(|u|, |v|)$, so we conclude by the proposition,

$$\log |x| \geq -\log \max(|u|, |v|) + \log \left(1 + \left|\frac{u}{v}\right|\right) = h(x) + \log \left(1 + \left|\frac{u}{v}\right|\right) \geq -h(x)$$

4.6 Siegel's Lemma

Let $N > M > 0$ be integers. Following Bombieri-Vaaler, we consider a system of M-linear forms in N variables.

$$L_i(T_1, \ldots, T_N) = \sum_{j=1}^{N} \ell_{ij} T_j, \qquad 1 \leq i \leq M$$

We want a small solution of $L_i(T) = 0$ for $1 \leq i \leq M$.

Lemma 4.14 (Siegel's Lemma)**.** *The system has a non-trivial solution* $x = (x_1, \ldots, x_N) \in \mathcal{O}_K^N$ *such that*

$$h(x) \leq \left\{ \left(\frac{2}{\pi} \right)^s \sqrt{D} \right\}^{1/4} \left(\max_i h(L_i) \right)^{\frac{M}{N-M}},$$

where, $d = [K : \mathbb{Q}]$*,* $D = \operatorname{discr}(K)$*, and* s *is the number of real places of* K*.*

Phrasing the system of equations as a matrix equation, previous properties imply a bound on the height of the matrix by the height of the column vectors. Instead of proving this, we prove a light version of Siegel's lemma.

Lemma 4.15 (Siegel's Lemma: Light Version)**.** *Let* $N > M > 0$*, and let*

$$\sum_{j=1}^{N} a_{ij} T_j = 0, \qquad 1 \leq i \leq M, a_{ij} \in \mathbb{Z}.$$

for $V_i = \max_j |a_{ij}|$*. Then the system has a non-tivial solution* $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{Z}^N$ *such that*

$$|\xi_1| < 2 + \left(\prod_j N V_j \right)^{\frac{1}{N-M}}.$$

Proof. We apply Dirichlet's box principle. Let $H > 0$ be an integer, and let C be the cube given by $|x_i| \leq H$ for $1 \leq i \leq N$. It contains $(2H+1)^N$ points. Let $\alpha : \mathbb{R}^N \to \mathbb{R}^M$ be the linear map, $\alpha = (\alpha_1, \ldots, \alpha_M), \alpha_i = \sum_j a_{ij} T_j$. The image $\alpha(C) \subseteq C'$ has $y \in C'$ if and only if $|y_i| \leq N \cdot V_i H$. Then C' contains $(2NV_iH)^M$ points. If H is sufficiently large, then $\alpha | C$ is not injective. Thus there exist $\xi' \neq \xi''$ in C such that $\alpha(\xi') = \alpha(\xi'')$.

Let $\xi := \xi' - \xi'' \neq 0$. Then $|\xi_i| \leq |\xi_i'| + |\xi_i''| \leq 2H$. We choose H such that

$$\left(\prod_j N V_j \right)^{\frac{1}{N-M}} \leq 2H < \left(\prod_j N V_j \right)^{\frac{1}{N-M}} + 2$$

Then

$$(2H+1)^N = (2H+1)^{N-M} \cdot (2H+1)^M > (2H+1)^M \prod_j N V_j \geq \prod (2NV_jH+1) = |C'|$$

In other words, the number of points in C' is greater than the number of points in the image of α. □

Siegel's lemma is essentially the best possible bound.

Lecture 5

Logarithmic Forms

This lecture treats one of the central contributions to number theory of the nineteenth century. It goes back to a problem of Gauß.

Let $\alpha_1, \ldots, \alpha_n \in \overline{Q}^\times$ and $b_1, \ldots, b_n \in \mathbb{Z}$. We define

$$\Gamma := \langle \alpha_1, \ldots, \alpha_n \rangle \subset \overline{\mathbb{Q}}^\times.$$

If $|\alpha - 1| := |\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1| \neq 0$, then we want to know how close this value can approach zero in terms of $B := \max\{|b_1|, \ldots, |b_n|\}$. In other words, we want a lower bound $|\alpha - 1| > C$.

There is a trivial lower bound. Assume that $\alpha_i \in \mathcal{O}_K$. Then $N(\alpha - 1) \geq 1$. Indeed, $N(\alpha - 1) = \prod_{\nu|\infty} |\alpha - 1|_\nu$. Hence

$$|\alpha - 1|_\nu \leq |\alpha|_\nu + 1 \leq (\max H_\nu(\alpha_i))^B + 1 \leq (2H_\nu(\alpha))^B =: C_\nu^B.$$

Hence

$$C^B \geq |\alpha - 1|_\nu \cdot \prod_{\mu \neq \nu} |\alpha - 1|_\mu \geq 1$$

so that $|\alpha - 1| \geq C^{-B}$.

5.1 Historical remarks

We ask whether this can be improved, for example, to a bound of the form

$$|\alpha - 1| > c^{-B^\epsilon}, \qquad \epsilon > 0$$

This would show that diophantine equations of the type,

$$f(x, y) = 1,$$

for a homogeneous, irreducible $f \in \mathbb{Z}[X, Y]$ of degree at least 3, have only finitely many integer solutions. The size can be effectively bounded.

The equation $y^2 = x^3 + ax + b$ for $a, b \in \mathbb{Z}$ has only finitely many solutions $(x, y) \in \mathbb{Z}^2$.

Similarly, the hyperelliptic curve $y^k = F(x)$ is a ramified, cyclic[1] k-fold covering of $\mathbb{P}^1$. The fundamental group $\pi_1(C)$ may regarded as the Galois group of the covering. In general, any rational function $r \in \mathbb{Q}(C)$ defines a ramified cover $C \to \mathbb{P}^1$.

This may be reformulated in terms of logarithms,

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n, \qquad b_i \in \mathbb{Z}$$

Remark 5.1. If the b_i are replaced by algebraic numbers β_i, then $\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ is expected to be transcendental.

Problem 5.2 (Hilbert's 7th Problem, 1900 ICM Paris)**.** Does $\alpha = 0, 1$

Another of Hilbert's problems was the Riemann Hypothesis. Although the Riemann Hypothesis remains unsolved, he considered his seventh problem to be harder.

Theorem 5.3 (Gel'fand, Schneider, 1934)**.** *If $\alpha, \beta \notin \overline{\mathbb{Q}}$, then α^β.*

Fixing $\alpha_1, \ldots, \alpha_n$, we want lower bound for $|\Lambda|$ ($\Lambda \neq 0$) in terms of B. In 1966, Baker showed that

$$\log |\Lambda| > -c(\log B)^k$$

whenever $k > 2n + 1$. He improved the result in 1969 (?) to include all cases where $k > n$. Around 1970, Feldman showed that this holds so long as $k > 1$.

We write $f \gg g$ for functions f and $g$if there exists a constant C such that $f \geq g$. In this case, $|\Lambda$ and B are functions of the b_i. In its final form, the theorem says that

$$\log |\Lambda| \gg -\log B.$$

Remark 5.4. If $\alpha_1, \ldots, \alpha_n$ are rational integers $\alpha_1, \ldots, \alpha_n$ and $\Omega = \log A_1 \cdots \log A_n$. Then $A_n \geq A_i \geq |a_i|$ and $\Omega' = \Omega / \log A_n$. Then Baker showed that $\log |\Lambda| \gg -\log B\Omega \log \Omega'$. Wüstholz improved this to $\log |\Lambda| \gg -\log B\Omega$, which is currently the best form.

Conjecture 5.5.

$$\log |\Lambda| \gg -\log B(\log A_1 + \ldots + \log A_n)$$

(This is closely related to the abc-conjecture, but we omit the relation.)

[1] A cycle covering $C \to \mathbb{P}^1$ is one such that $\pi_1(C)$ is a cyclic group.

5.2 The Theorem

Theorem 5.6. *Let $b_1, \ldots, b_n \in \mathbb{Z}$ and let $B \geq 2$ be a real number bound such that $|b_i| \leq B$ for all i. Let $\alpha_1, \ldots, \alpha_n \in K^\times$ for a number field $K \subseteq \overline{Q}$. If $\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n \neq 0$, then there exists an effectively computable constant $C > 0$ such that*

$$\log |\Lambda| > -C \log B.$$

The constant C depends on $\alpha_1, \ldots, \alpha_n$ and $d = [K : \mathbb{Q}]$.

Remark 5.7. The theorem can also be stated as follows: There exists a $C > 0$ such that if

$$\log |\Lambda| \leq -C \log B$$

then $\Lambda = 0$. In other words, $\alpha_1, \ldots, \alpha_n$ are multiplicatively dependent.

Preparations for the proof

Feldman introduced Δ-functions. For each integer $k \geq 1$, we define a function from $\mathbb{C}$ to $\mathbb{C}$ by,

$$\Delta(z, k) := \frac{(z+1) \cdots (z+k)}{k!}.$$

The Δ functions satisfy a number of properties.

- The Δ functions map integers to integers, i.e., $\Delta(\mathbb{Z}, k) \subseteq \mathbb{Z}$.
- $|\Delta(z, k)| \leq \frac{(|z|+k)^k}{k!} \leq e^{|z|+k}$
- The set $\{\Delta(z, k) \mid k \in \mathbb{N}\}$ is a $\mathbb{Q}$-basis of the polynomial ring $\mathbb{Q}[z]$.
- The $\mathbb{Z}$-submodule $\mathcal{U} = \{f \in \mathbb{Q}[z] \mid f(\mathbb{Z}) \subseteq \mathbb{Z}\} \subseteq \mathbb{Q}[z]$ of functions which preserve integers is generated by the Δ functions.

$$\mathcal{U} = \oplus_{k \geq 0} \mathbb{Z} \Delta(z, k)$$

For all integers $\ell > 0$ and $m \geq 0$, we define

$$\Delta(z, k, \ell, m) = \frac{1}{m!} \left(\frac{d}{dz} \right)^n \Delta(z, k)^\ell = \frac{1}{2\pi i} \int_{|\zeta - z| = 1} \frac{\Delta(\zeta, k)^\ell d\zeta}{(\zeta - z)^{m+1}}$$

Hence we arrive at a bound which does not depend on n.

$$|\Delta(z, k, \ell, m)| \leq e^{(|z|+k)\ell} \tag{5.1}$$

We have sucessfully played the game to find a useful bound. But there are no games without pain, so the question is, where does the pain enter? The answer is that differentiating to define $\Delta(z,k,\ell,m)$ introduces denominators. It no longer maps integers to integers. Let

$$v(k) = \operatorname{lcm}(1,\ldots,k) = \prod_{p \text{ prime}} p^{\log k/\log p}$$

By the Prime Number Theorem, the number of primes less than or equal to k is $k/\log k$. Therefore $v(k) \le 4^k$. By this definition, we have assuaged the pain, since

$$v(k)\Delta(z,k,\ell,m)$$

maps integers to integers.

Proof: The anxiliary function

We introduce the notation $c_1, c_2, c_3, \ldots$ for positive constants which do not depend on B (although perhaps they depend on α_i or K), and which can be effectively computed.

We begin the proof by assuming that $\log|\Lambda| \le -C\log B$. Without loss of generality, we assume that $b_n \ne 0$. Knowing how the proof will end, we define constants

$$L := [k^{n/(n+1)}\log k] \qquad h := [\log kB] \qquad L' := h-1$$

for some large $k \in \mathbb{N}$. We define a function on $\mathbb{C}^n$ by

$$\Phi(z_0,\ldots,z_{n-1}) := \sum_{0\le\lambda_{-1}\le L'} \sum_{0\le\lambda_0,\ldots,\lambda_n\le L} p(\lambda)\Delta(z_0+\lambda_{-1},h,\lambda_0+1,0)\alpha_1^{\gamma_1 z_1}\cdots\alpha_{n-1}^{\gamma_{n-1}z_{n-1}}$$

where the function $p(\lambda)$ is unknown, and

$$\gamma_i = \lambda_i - \frac{b_i}{b_n}\lambda_n \qquad 1 \le i \le n-1.$$

We shall determine $p(\lambda)$ as an element of $K = \mathbb{Q}(\alpha_1,\ldots,\alpha_n)$.

Lemma 5.8. *There exist $p(\lambda) \in K$, not all 0, of height at most c_1^{hk} such that*

$$S := \sum_{\lambda} p(\lambda)\Delta(\ell+\lambda_{-1},h,\lambda_0+1,m_0)\prod_{j=1}^{n}\alpha_j^{\lambda_j\ell}\prod_{j=1}^{n-1}\gamma_j^{m_j} = 0 \tag{5.2}$$

for $0 \le \ell < h$, $m_j \ge 0$ $(0 \le j \le n-1)$ and $m_0+\cdots+m_{n-1} < k$.

Proof. This is a typical application of Siegel's Lemma. Up to a factor of the form $(\log \alpha_1)^{m_1} \cdots (\log \alpha_{n-1})^{m_{n-1}}$, the formula of the lemma can be written as,

$$\frac{1}{m_0!} \left(\frac{d}{dz_0}\right)^{m_0} \left(\frac{d}{dz_1}\right)^{m_1} \cdots \left(\frac{d}{dz_{n-1}}\right)^{m_{n-1}} \Phi(z_0, z_1, \ldots, z_{n-1}).$$

Taking the z_0 derivative of Φ only changes 0 into m_0 in the fourth argument of the Δ function (by the properties of Δ). The rest of the derivatives add a factor of $(\log \alpha_1)^{m_1} \cdots (\log \alpha_{n-1})^{m_{n-1}}$. To make them disappear, we define

$$D_j := (\log \alpha_j)^{-1} \frac{d}{dz_j} \qquad D_0 := \frac{d}{dz_0}$$

The new system is

$$\frac{1}{m_0!} D_0^{m_0} D_1^{m_1} \cdots D_{n-1}^{m_{n-1}} \Phi(z_0, z_1, \ldots, z_{n-1})$$

We reformulate the powers appearing in Φ:

$$\begin{aligned}
\alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}} &= \alpha_1^{\lambda_1 - b_1/b_n \lambda_n z_1} \cdots \alpha_{n-1}^{\lambda_{n-1} - b_{n-1}/b_n \lambda_n z_{n-1}} \\
&= \alpha_1^{\lambda_1 z_1} \cdots \alpha_{n-1}^{\lambda_{n-1} z_{n-1}} \left(\alpha_1^{-b_1/b_n z_1} \cdots \alpha_{n-1}^{-b_{n-1}/b_n z_{n-1}}\right)^{\lambda_n} \\
&= \alpha_1^{\lambda_1 z_1} \cdots \alpha_{n-1}^{\lambda_{n-1} z_{n-1}} \left(\alpha_1^{-b_1/b_n} \cdots \alpha_{n-1}^{-b_{n-1}/b_n}\right)^{\lambda_n z} \\
&= \alpha_1^{\lambda_1 z_1} \cdots \alpha_{n-1}^{\lambda_{n-1} z_{n-1}} (\alpha_n')^{\lambda_n z}
\end{aligned}$$

where we define $\alpha_n' := \alpha_1^{-b_1/b_n} \cdots \alpha_{n-1}^{-b_{n-1}/b_n}$. Recall the Schwarz lemma:

Lemma 5.9 (Schwarz Lemma)**.** *Let $f : D \to D$ be a holomorphic function from the open unit disk, $D = \{z : |z| < 1\} \subset \mathbb{C}$ to itself. If $f(z) = 0$, then $|f(z)| \leq |z|$ for all $z \in D$, and furthermore, $|f'(0)| \leq 1$. If additionally $|f(z)| = |z|$ for some non-zero z or if $|f'(0)| = 1$, then $f(z) = e^{i\theta} z$ for some $\theta \in \mathbb{R}$.*

We estimate the difference $|\alpha_n' - \alpha_n|$. By the power series $e^z = 1 + z + z^2/2! + z^3/3! + \cdots$,

$$|e^z - 1| \leq |z| e^{|z|} \qquad \text{for all } z \in \mathbb{C}$$

Letting $Q = b_n(\log \alpha_n' - \log \alpha_n)$,

$$\begin{aligned} |\alpha_n' - \alpha_n| = |\alpha_n||\alpha_n'\alpha_n^{-1} - 1| &= |\alpha_n| \cdot |e^{\log \alpha_n' - \log \alpha_n} - 1| \\ &= |\alpha_n||e^{Q/b_n} - 1| \\ &\leq |\alpha_n||Q/b_n|e^{|Q/b_n|} \\ &\leq |\alpha_n||Q|e^{|Q|} \\ &\leq |\alpha_n|B^{-C}e^{B^{-C}} \\ &\leq B^{-\frac{3}{4}C} \end{aligned}$$

At the last step, we impose a condition on the constant C. It must satisfy $e^{B^{-C}} \leq 1$. Since $B \geq 2$, this will be true if $e^{2^{-C}} \leq 1$ is satisfied, a condition which does not depend on B.

We now bound the height of 5.2. The proof divides here into two cases. In the first case, we consider archimedean places ν. We use the bound 5.1. Evaluating z at ℓ gives that $|z| \leq h$. Observing that $k = \lambda_0 + 1 \leq L'$ and that $\ell \leq L$ gives that

$$\Delta(\ell, \lambda_{-1}, h, \lambda_0 + 1, m_0) \leq e^{(L'+h)L}$$

Again observing the bounds on the sums, $\prod_{j=1}^n \alpha_j^{\lambda_j \ell} \leq \prod_{i=1}^n H_\nu(\alpha_j)^{Lh}$. Using the definition of γ_i bounds the final product in equation 5.2 to give

$$\begin{aligned} S &\leq c_2^{(L'+h)L} \prod_{i=1}^n H_\nu(\alpha_j)^{Lh} \prod_{j=1}^{n-1} (\max(1, \frac{\|b_j\|_\nu}{\|b_n\|_\nu} 2L)^{m_j} \\ &\leq c_2^{(L'+h)L} \prod_{i=1}^n H_\nu(\alpha_j)^{Lh} (2BL)^{m_1 + \cdots + m_{n-1}} \\ &\leq c_2^{(L'+h)L} \prod_{i=1}^n H_\nu(\alpha_j)^{Lh} (2BL)^k \end{aligned}$$

Second, we consider the case when ν is a finite place. Then

$$S \leq \|\nu(h)\|_\nu^{-m_0} \cdot \prod_{j=1}^n H_\nu(\alpha_j)^{Lh} \|b_n^{-1}\|_\nu^k$$

Altogether this makes

$$H(S) \leq c_3^{hk}$$

For an application of Siegel's Lemma, we need to bound the heights of the coefficients as well as the quantity $\frac{M}{N-M}$, where N is the number of variables, and M is the number of equations in 5.2.

- The coefficients are bounded by c_3^{hk}.
- $N = (L'+1)(L+1)^n = h \cdot (k^n n + 1 \log k)^{n+1} \geq 2hk^n$.
- $M = h \cdot \begin{pmatrix} k+n-1 \\ n \end{pmatrix} \leq hk^n$.

This gives $\frac{M}{N-M} \leq 1$, and now Siegel's Lemma gives the result. □

We look now at the holomorphic functions

$$f_m(z) = \frac{1}{m_0!} D_0^{m_0} D_1^{m_1} \cdots D_{n-1}^{m_{n-1}} \Phi(z_0, \ldots, z_{n-1}) \Big|_{z_0 = \cdots = z_{n-1} = z}.$$

We will sometimes supress the m in the notation and simply write $f(z)$. Similar estimates as in the proof of Lemma 5.8 give

$$|f(z)| \leq c_4^{hk + L|z|} \tag{5.3}$$

for $m_0 + \cdots m_{n-1} < k$.

Exercise 5.10. Prove the estimate of equation 5.3.

Definition 5.11. Let

$$g(\ell) = \sum_{\lambda} p(\lambda) \Delta(\ell + \lambda_{-1}, h, \lambda_0 + 1, 0) \alpha^{\lambda_1 \ell} \cdots \alpha_n^{\lambda_n \ell} \gamma_1^{m_1} \cdots \gamma_n^{m_n}$$

Exercise 5.12. For $0 \leq \ell < hk^{\frac{1}{2(n+1)}} (\log k)^{-1}$,

$$|f(\ell) - g(\ell)| \leq B^{-\frac{1}{2}C}.$$

Lemma 5.13. *For all integers ℓ with $0 \leq \ell < hk^{\frac{1}{2(n+1)}} (\log k)^{-1}$, we either have $g(\ell) = 0$ or*

$$|f(\ell)| > c_6^{-hk - L\ell}.$$

Proof. If $g(\ell) = 0$, then there is nothing to show, so let $g(\ell) \neq 0$. We use the product formula to show that

$$|g(\ell)| > c_7^{-hk}.$$

Then

$$\begin{aligned} |f(\ell)| = |g(\ell) - f(\ell) - g(\ell)| &\geq |g(\ell)| - |f(\ell) - g(\ell)| \\ &\geq |g(\ell)| - B^{-\frac{1}{2}C} \\ &\geq \frac{1}{2}|g(\ell)| \end{aligned}$$

because, as we shall see, we may take C sufficiently large so that

$$B^{-\frac{1}{2}C} < \frac{1}{2}c_7^{-hk}.$$

Since $g(\ell) \neq 0$, by Lemma 5.8, $\ell \geq h$. If $\nu \mid \infty$, then

$$\prod_{j=1}^{n} \alpha_j^{\lambda_j \ell} \leq \max(1, \|\alpha\|_\nu) \qquad \prod_{j=1}^{n} \gamma_j^{m_j} \leq \left(\max\left(1, \frac{\|b_j\|}{\|b_n\|}\right) 2\|L\|_\nu \right)^{m_j}$$

The first part of $g(\ell)$ can also be bounded (we omit this) to give a total bound of

$$\|g(\ell)\|_\nu \leq c_7^{hk} \prod_{j=1}^{n} \max(1, \|\alpha\|_\nu) \left(\max\left(1, \frac{\|b_j\|}{\|b_n\|}\right) 2\|L\|_\nu \right)^{m_j}.$$

If $\nu \nmid \infty$, then $g(\ell)$ has a correspondingly different bound,

$$\|g(\ell)\|_\nu \leq \|\nu(h)\|_\nu^{-m_0} \prod_{j=1}^{n} \max(1, \|\alpha\|_\nu) \left(\max\left(1, \frac{\|b_j\|}{\|b_n\|}\right) 2\|L\|_\nu \right)^{m_j}.$$

Since $L\ell \leq hk$, this implies that

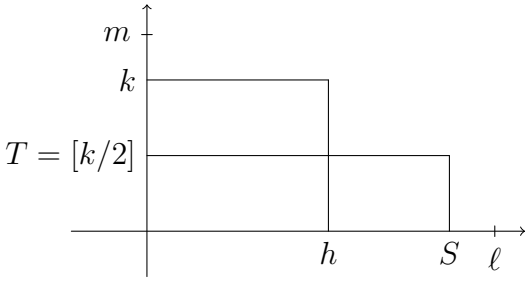
$$1 = \prod_\nu \|g(\ell)\|_\nu \leq c_7^{hk} \|g(\ell)\|$$

Hence $|g(\ell)| > c_7^{-hk}$.

The inequality 5.2 can be acheived for C sufficiently large by a chain of reasoning which produces a list of constants $c_8, \ldots$ imposing a finite number of lower bounds on C. Choosing C greater than all these lower bounds will then satisfy 5.2. □

5.3 Extrapolations

We now turn our attention to the zeros of $f(\ell)$ in the rectangle bounded by h and k:



Recall that $h = \log(kB)$. Our aim is to show that, if $f(\ell)$ has zeros in the $h \times k$ rectangle, then it has zeros in the rectangle $S \times T$, where

$$S := [hk^{\frac{1}{2(n+1)}}(\log k)^{-1}] \qquad T := [k/2].$$

Proposition 5.14. *For $0 \leq \ell < S$ and $m_0, \ldots, m_{n-1} \geq 0$ with $m_0 + \cdots + m_{n-1} < T$, we have $g(\ell) = 0$.*

Proof. We use Cauchy's integral formula. Let

$$F(z) := \prod_{j=0}^{h-1} (z-j)^T$$

and consider the function $\phi(z) := f(z)/F(z)$. We will use the notation

$$|G(z)|_r := \max_{|z|=r} |G(z)|$$

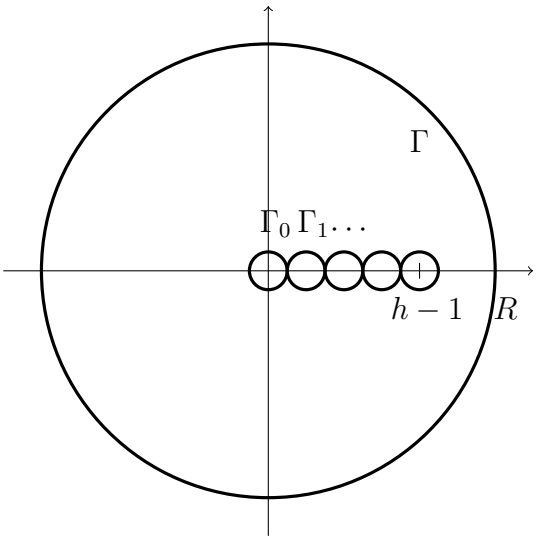
for the maximum of a holomorphic function on the disk of radius r. By the Schwarz lemma 5.9, $|\phi(z)|_S \leq |\phi(z)|_R$ for $R := [k^{\frac{1}{2(n+1)}}S]$. Hence $|f(z)|_S \leq |f(z)|_R \frac{|F(z)|_S}{|F(z)|_R}$ and

$$\frac{S^{hk}}{(R/2)^{hk}} = \left(\frac{z}{R^{\frac{1}{2(n+1)}}}\right)^{hk}$$

Hence $|z-j| \geq |z| - j$ and $R - h \geq R/2$, and

$$|\phi(z)| \leq |\phi(z)|_R \left(\frac{S}{R}\right)^{hk}$$

The function $\phi(z)$ is meromorphic with poles at $0, \ldots, h$. To apply the Cauchy integral formula, we define contours:



Here Γ is the circle of radius R and the Γ_s are circles of radius $1/2$ centered at integral points on the x-axis. Then

$$\frac{1}{2\pi i}\int_\Gamma \frac{\phi(z)}{z-\ell}dz = \phi(\ell) + \frac{1}{2\pi i}\sum_{s=0}^{h-1}\sum_{t=0}^{T-1}\frac{f^{(t)}(s)}{t!}\int_{\Gamma_s}\frac{1}{F(z)}\frac{(z-s)^t}{z-\ell}dz.$$

Multiplying by $F(\ell)$ gives

$$\underbrace{\frac{1}{2\pi i}\int_\Gamma \frac{F(\ell)}{F(z)}\frac{f(z)}{z-\ell}dz}_{I} = f(\ell) + \underbrace{\frac{1}{2\pi i}\sum_{s=0}^{h-1}\sum_{t=0}^{T-1}\frac{f^{(t)}(s)}{t!}\int_{\Gamma_s}\frac{F(\ell)}{F(z)}\frac{(z-s)^t}{z-\ell}dz}_{J}.$$

(Note the definitions of I and J.) We want to exclude the possibility that $|f(\ell)| > c_6^{-hk-L\ell}$. To do this, we show that $I, J < c_8^{-hk}$. We bound $F(\ell)/F(z)$ by first bounding $F(z)$ from below on the disk of radius R, and then bounding $F(\ell)$ from above.

On Γ, the disk of radius R, $|z-j| \geq |z| - j = R - j \geq R/2$. Hence,

$$|F(z)| > 2^{-hT}R^{hT} \qquad |F(\ell)| = \prod_{j=0}^{h-1}|\ell - j|^T \leq S^{hT}$$

Together, this gives,

$$\left|\frac{F(\ell)}{F(z)}\right| \leq 2^{hk}k^{-\frac{hT}{2(n+1)}} \qquad z \in \Gamma. \tag{5.4}$$

We have $L|z| = LR = k^{\frac{n}{h+1}}\log k$ and $k^{\frac{1}{2(n+1)}}S = kh$, so

$$|f(z)| \leq c_9^{hK+L|z|} \leq c_{10}^{2hk} \qquad z \in \Gamma. \tag{5.5}$$

Using equations 5.4 and 5.5, we can now bound the integral I by,

$$I \leq 2\cdot 2^{hk}\left(\frac{S}{R}\right)^{hT} c_{10}^{2hk} = c_{11}^{hk}k^{-\frac{hT}{2(n+1)}} \leq B^{-\frac{1}{4}C}.$$

To bound the double sum of integrals, J, we write the derivatives of f in terms of partial derivatives of Δ,

$$f_m^{(t)}(z) = \left(\frac{\partial}{\partial z_0} + \cdots + \frac{\partial}{\partial z_{n-1}}\right)^t \frac{1}{m_0!}D_0^{m_0}\cdots D_{n-1}^{m_{n-1}}\Phi(z_0,\ldots,z_{n-1})\Big|_{z_0=\cdots=z_{n-1}=z}, \qquad t = m_0+\cdots n$$

Recall that $\frac{\partial}{\partial z_0} = D_0$ and $\frac{\partial}{\partial z_i} = \log \alpha_i D_i$ for $1 \le i \le n-1$. The powers of the derivative are then

$$\left(\frac{d}{dz}\right)^t = (D_0 + \log\alpha_1 D_1 + \cdots + \log\alpha_n D_n)^t = \sum_\tau q(\tau)\frac{1}{\tau_0!}D^\tau$$

where we use the notation $D^\tau = D_0^{\tau_0}\cdots D_{n-1}^{\tau_{n-1}}$. The main point is that these differential operators are not defined over $\overline{\mathbb{Q}}$.

$$\begin{aligned} f_m^{(t)}(z) &= \sum_\tau q(\tau)\frac{1}{m_0!\tau_0!}\, D_0^{m_0+\tau_0}\cdots D_{n-1}^{m_{n-1}+\tau_{n-1}}\Phi(z_0,\ldots,z_{n-1})\big|_{z_0=\cdots=z_{n-1}=z} \\ &= \sum_\tau q^*(\tau)\frac{1}{(m_0+\tau_0)!}\, D_0^{m_0+\tau_0}\cdots D_{n-1}^{m_{n-1}+\tau_{n-1}}\Phi(z_0,\ldots,z_{n-1})\big|_{z_0=\cdots=z_{n-1}=z} \\ &= \sum_{\tau,|m|\le k-1} q^*(\tau) f_{m+\tau}(z) \qquad 0 \le |m| \le T-1 \end{aligned}$$

where

$$q^*(\tau) = \frac{t!}{\tau_1!\cdots\tau_{n-1}!}\begin{pmatrix} \tau_0+m_0 \\ \tau_0 \end{pmatrix}(\log\alpha_1)^{\tau_1}\cdots(\log\alpha_{n-1})^{\tau_{n-1}}.$$

Hence $c^k \cdot \sum \frac{t!}{\tau_1!\cdots\tau_{n-1}!} \le c^k \cdot n^T$. And if $m_0+\cdots+m_{n-1} \le T-1$, then the summation goes over $f_{m+\tau}(z)$ with $m_0+\cdots+m_{n-1} \le k-1$.

$$|f_{m+\tau}(s)| \le \underbrace{|g_{m+\tau}(s)|}_{=0} + \underbrace{|f_{m+\tau}(s)-g_{m+\tau}(s)|}_{<B^{-C/2}}$$

We arrive at an upper bound for the derivatives,

$$|f_m^{(t)}(s)| < c_{12}^{hk}B^{-C/2}, \qquad 0 \le t < T-1, 0 \le s \le h-1 \tag{5.6}$$

Finally, we need a bound on the integrals over Γ_s.

$$|z-j| \ge \begin{cases} 1/2 & s=j \\ (s-j)/2 & 0 \le j < s \\ (j-s)/2 & s \le j \le h-1 \end{cases}$$

Then

$$|F(z)| \ge 2^{-h}s!(h-s)!(h-s)^{-1}.$$

where $s!(h-s)! \le \begin{pmatrix} h \\ s \end{pmatrix} h!$.

Sterling's formula gives that $h! \geq \left(\frac{h}{e}\right)^h \sqrt{2\pi h}$. Finally, $|F(z)| \geq (8e)^{-hT} h^{hT}$. This gives for $z \in \Gamma_s$,

$$\begin{aligned} \left|\frac{F(\ell)}{F(z)}\right| &\leq \frac{S^{hT}}{h^{ST}}(8e)^{hT} \\ &\leq \left(\frac{k^{\frac{1}{2(n+1)}}h}{h}\right)^{hT}(8e)^{hT} \leq (8ek^{\frac{1}{2(n+1)}})^{hT} \end{aligned} \tag{5.7}$$

We are finally ready to bound the double sum of integrals using equation 5.7 and 5.6 to give,

$$\begin{aligned} J &\leq \sum_{s=0}^{h-1}\sum_{t=0}^{T-1}\frac{1}{t!}c^{hk}B^{-\frac{1}{2}C}k^{\frac{hT}{2(n+1)}} \\ &\leq \left(\sum_{s=0}^{h-1}\sum_{t=0}^{T-1}\frac{1}{t!}\right)B^{-\frac{1}{4}C} \leq heB^{-\frac{1}{4}C} \end{aligned}$$

provided that $B^{\frac{1}{4}C} > (ck)^{hk}$.

The bounds on I and J together bound the absolute value of $f(\ell)$:

$$|f(\ell)| = I + J \leq B^{-\frac{1}{4}C}(1+he) \leq B^{-\frac{1}{5}C}$$

Details for abbreviated parts of the proof can be found in Baker and Wüstholz. □

Proposition 5.15. *If* $\log|\Lambda| < -C\log B$ *we have* $g(\ell) = 0$ *for all* $0 \leq \ell < S$ *and all non-negative integers* $m_0, \ldots, m_{n-1}$ *with* $m_0 + \cdots + m_{n-1} \leq T$.

Proof. Let

$$P(X_0, \ldots, X_n) := \sum_{\lambda_{-1}=0}^{L'} \sum_{\lambda_0,\ldots,\lambda_n \leq L} p(\lambda_{-1}, \ldots, \lambda_n)\Delta(X_0+\lambda_{-1}, h, \lambda_0+1, 0)X_1^{\lambda_1}\cdots X_n^{\lambda_n}$$

Then for $m_0, \ldots, m_{n-1}$ with $m_0 + \cdots + m_{n-1} < T$,

$$g(\ell) = \frac{1}{m_0!}D_0^{m_0}\cdots D_{n-1}^{m_{n-1}}P(X_0, \ldots, X_n)\Big|_{(X_0,\ldots,X_n)=(\ell,\alpha_1^\ell,\ldots,\alpha_n^\ell)} = \Phi(\ell, \ldots, \ell)$$

where $D_0 = \frac{\partial}{\partial X_0}$ and $D_i = X_i\frac{\partial}{\partial X_i} - \frac{b_i}{b_n}X_n\frac{\partial}{\partial X_n}$. Hence P satisfies,

$$D_0^{m_0}\cdots D_{n-1}^{m_n}P(\ell, \alpha_1^\ell, \ldots, \alpha_n^\ell) = 0,$$

for all $0 \leq \ell < S$ and all $m_0, \ldots, m_{n-1}$ such that $m_0 + \cdots + m_{n-1} < T$. □

5.4 Multiplicity estimates

The presentation here can be supplemented by the book by Baker and Wüstholz. If

$$S(T/n)^n \geq k^\epsilon L' L^{n+1}$$

for sufficiently large ϵ and k, then $\alpha_1, \ldots, \alpha_n$ are *multiplicatively dependent*, i.e., there exist $X_1, \ldots, X_n \in \mathbb{Z}$ not all zero, such that

$$\alpha_1^{X_1} \cdots \alpha_n^{X_n} = 1. \tag{5.8}$$

This is satisfied for our $P(X_0, \ldots, X_n)$. Siegel's lemma says that $ST^n \leq L' L^n$ and extrapolation says that $ST^n > k^\epsilon L' L^n n^n$.

Theorem 5.16. *If $\alpha_1, \ldots, \alpha_n$ are multiplicatively independent, then*

$$\log |\Lambda| > -C \log B$$

for some effectively computable constant $C > 0$ which depends only on $[K : \mathbb{Q}], h(\alpha_1), \ldots, h(\alpha_n)$ and n.

If the terms are multiplicatively dependent, then taking the logarithm of equation 5.8 gives a linear relation,

$$\log \alpha_n = \frac{1}{X_n}(X_1 \log \alpha_1 \cdots + X_{n-1} \log \alpha_{n-1})$$

The nth case is thereby reduced to the $(n-1)$th case.

Lecture 6

The Unit Equation

6.1 Units and S-units

Let $K \supset \mathbb{Q}$ be a number field, and let $\mathcal{O}_K$ be a ring of integers. Let $\Sigma := \mathrm{Hom}(K, \mathbb{C})$. Then $|\Sigma| = r + 2s$ where r is the number of homomorphisms which factor through $\mathbb{R} \hookrightarrow \mathbb{C}$ and s is the number of pairs of complex conjugate embeddings.

Theorem 6.1 (Dirichlet's Unit Theorem)**.** *With the above notation,*

$$\mathcal{O}_K^\times \cong \mu(K) \times \mathbb{Z}^t, \qquad t = r + s - 1$$

where $\mu(K)$ is the group of roots of unity in K.

Let $\epsilon_1, \ldots, \epsilon_t \in \mathcal{O}_K^\times$ be a basis for $\mathcal{O}_K^\times / \mu(K)$. Then $x \in \mathcal{O}_K^\times \Leftrightarrow x = \zeta \cdot \epsilon_1^{x_1} \cdots \epsilon_t^{x_t}$ for some $x_1, \ldots, x_t \in \mathbb{Z}$ and some $\zeta \in \mu(K)$. In this lecture, we solve the unit equation:

Definition 6.2 (Unit equation)**.** The unit equation is

$$x + y = 1, \qquad x, y \in \mathcal{O}_K^\times. \tag{6.1}$$

Definition 6.3 (S-units)**.** If S is a finite set of primes in $\mathcal{O}_K$, then the S-units are the invertible elements of

$$\mathcal{O}_{K,S} = \left\{ \frac{a}{b} \mid a, b \in \mathcal{O}_K, b \notin \mathfrak{p} \text{ for } \mathfrak{p} \notin S \right\} = \left\{ x \in K \mid |x|_\nu \leq 1 \text{ for } \nu \notin S \right\}.$$

Note that $\mathcal{O}_{K,S} \supset \mathcal{O}_K$.

Definition 6.4 (S-unit equation)**.** The S-unit equation is $x + y = 1$ for $x, y \in \mathcal{O}_{K,S}^\times$. A solution of the S-unit equation gives a solution of $ax + by = 1$ with $x, y \in \mathcal{O}_K$.

Let $x, y \in \mathcal{O}_K^\times$ be solutions of 6.1. Put

$$\begin{aligned} x &= \xi\epsilon_1^{x_1}\cdots\epsilon_t^{x_t}, \qquad \xi \in \mu(K), x_i \in \mathbb{Z} \\ y &= \eta\epsilon_1^{y_1}\cdots\epsilon_t^{y_t}, \qquad \eta \in \mu(K), y_i \in \mathbb{Z} \end{aligned}$$

Let $X := \max(|x_i|)$ and $Y := \max(|y_i|)$. We shall bound $\max(X, Y)$. Without loss of generality, we assume that $X \leq Y$. We assume that there exists $\nu \mid \infty$ such that

$$-\log \|y\|_\nu \gg Y.$$

We will show that this assumption is actually no restriction using the product formula. Equation 6.1 implies that

$$\left\|\frac{x}{y} + 1\right\| = \left\|\frac{1}{y}\right\| = \frac{1}{\|y\|} \leq e^{-\epsilon Y}$$

where the norm is the ν-norm. Now

$$\frac{x}{y} = \frac{\xi}{\eta}\epsilon_1^{x_1-y_1}\cdots\epsilon_t^{x_t-y_t} = \zeta\epsilon_1^{x_1-y_1}\cdots\epsilon_t^{x_t-y_t}$$

so

$$\frac{x}{y} + 1 = \zeta'\epsilon_1^{x_1-y_1}\cdots\epsilon_t^{x_t-y_t} + 1 = e^z + 1$$

where $z = \sum_{i=1}^t (x_i - y_i)\log\epsilon_i + iq\pi$ for some $q \in \mathbb{Q}$ depending on the root of unity ζ'.

Lemma 6.5. *Let $z \in \mathbb{C}$ satisfy*

$$|e^z + 1| \leq \frac{1}{4}$$

Then there exists an integer k such that

$$|z - ik\pi| \leq \frac{4}{3}|e^z + 1|.$$

Proof. We take the principle branch of the logarithm which satisfies $\log e^z = z$ for $0 \leq \arg z < 2\pi$.

Let $\zeta = e^z + 1$, and let w lie in the strip such that $z = w$ modulo $2\pi i$. Then $w = \log e^z = \log(\zeta - 1)$. The hypothesis $|\zeta| = |e^z + 1| \leq \frac{1}{4}$ and the power series formula $\log(\zeta - 1) = -\sum_{n=1}^\infty \frac{\zeta^n}{n}$ together imply that

$$|\log(1 - \zeta)| \leq \frac{4}{3}|\zeta| = \frac{4}{3}|e^z + 1|.$$

□

The Lemma implies that there is some $k \in \mathbb{Q}$ such that

$$|z - ik\pi| = |\Lambda| \leq \frac{4}{3}\left|\frac{x}{y} + 1\right|$$

where $\Lambda = \sum_{i=1}^{t}(x_i - y_i)\log\epsilon_i - k\log(-1)$. Thus $\log|\Lambda| \ll -Y = \max(|x_i|, |y_i|)$. We now have a linear combination of logarithms, and we would like to apply Baker's theorem. This requires,

$$|x_i - y_i| \leq 2\max(|x_i|, |y_i|) = 2Y \quad \text{and} \quad |k| \ll Y.$$

These both hold. Thus $-\log Y \ll \log\Lambda$. But this reasoning depends on the hypothesis that at some place w we have $\log|y|_w \gg Y$.

6.2 Unit and regulator

In this section, we denote the units of K by $U_K = \mathcal{O}_K^\times$. Then $\epsilon \in U_K$ implies that $N(\epsilon) = \pm 1$. Let $\epsilon_1, \ldots, \epsilon_t$ be a set of generators of U_K, and let $\Sigma_0 := \sigma_1, \ldots, \sigma_{r+s} \subseteq \Sigma \subseteq \mathrm{Hom}(K, \mathbb{C})$ be a set of embeddings.

Definition 6.6 (Regulator)**.** Let

$$M := (\log|\sigma_j(\epsilon_i)|)_{\substack{i=1,\ldots,t \\ j=1,\ldots,t+1}} = (\log|\epsilon_i|_w) \in M_{t+1,t}(\mathbb{R}).$$

We take any submatrix R in $M_{t,t}(\mathbb{R})$. Then the *regulator* of K is $\det R$ which is independent of the choice of R.

Taking the logarithm of the absolute value of y gives

$$\log|y|_w = y_1\log|\epsilon_1|_w + \cdots + y_t\log|\epsilon_t|_w, \qquad w \in \Sigma_0.$$

As w ranges through the embeddings Σ_0, this produces a system of equations for $y_1, \ldots, y_t$. Hence

$$Y = \max|y_i| \ll \max(\log|y|_w).$$

So there exists a w such that $\log|y|_w \gg Y$ and finitely many solutions to the unit equation.

Lecture 7

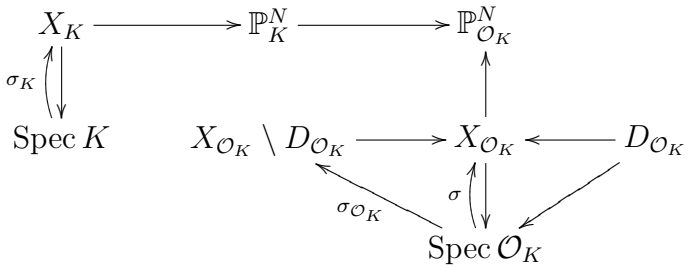
Integral points in $\mathbb{P}^1 \setminus \{0, 1, \infty\}$

Let K be a number field, and let $X = \operatorname{Spec} R$ be an affine variety for $R = K[f_1, \ldots, f_n]$. The variety X embeds in affine space by the morphism $x \mapsto (f_1(x), \ldots, f_n(x)) \in \mathbb{A}^n$.

Definition 7.1 (Integral points)**.** The integral points of X are

$$X(\mathcal{O}_K) = \{P \in X \mid f_1(P), \ldots, f_n(P) \in \mathcal{O}_K\}$$

Let X be projective over K, and let D be a very ample divisor.



The image $\sigma(\operatorname{Spec} \mathcal{O}_K)$ does not intersect $D_{\mathcal{O}_K}$.

Now let $X = \mathbb{P}^1_K \setminus \{0, 1, \infty\}$ be the projective line with three punctures. It can be expressed as the affine variety $X = \mathbb{A}^1 \setminus \{0, 1\}$ whose regular functions include $T, \frac{1}{T}, T-1$ and $\frac{1}{T-1}$. These generate the coordinate ring of $R = K[T, T^{-1}, T-1, (T-1)^{-1}]$.

By definition, a point $P \in X$ is an integral point if and only if the value of these regular functions is integral. Since $1/T(P)$ is defined, $T(P) \in \mathcal{O}_K^\times$. Similarly, $1/(T-1)(P)$ is defined, so $(T-1)(P) \in \mathcal{O}_K^\times$. We conclude that the integral points of $\mathbb{P}^1_K \setminus \{0, 1, \infty\}$ are solutions of the unit equation:

$$T(P) + (T-1)(P) = 1, \qquad T(P), (T-1)(P) \in \mathcal{O}_K^\times$$

Examples

Let E be the elliptic curve in normal form $y^2 = x(x-1)(x-\lambda)$. It forms a ramified cover

$$E \setminus \{E_1, E_2, E_3\} \to \mathbb{P}^1 \setminus \{0, 1, \infty\}$$

Integrality of points in E is intimately related to integrality of points in the image, of which there are finitely many. This allows one to conclude that E also has finitely many integral points.

A super-elliptic curve has normal form

$$y^m = x(x-1)(x-\lambda).$$

On any super-elliptic curve, there are only finitely many integral points, and they can be determined effectively.

This should be compared with Ziegel's theorem that the set of integral points on an algebraic curve of genus greater than one is finite. But Ziegel's theorem is not effective, hence the work we have done.

Lecture 8

Appendix: Lattice Theory

8.1 Lattices

Let $(E, \langle ., . \rangle)$ be a finite dimensional Euclidean $\mathbb{Q}$-vector space, where

$$\langle ., . \rangle : E \times E \to \mathbb{Q}$$

is a positive-definite bilinear form.

A subgroup $\Lambda \subset E$ is called *discrete* if for all bounded $B \subset E$, the intersection $B \cap \Lambda$ is finite.

Theorem A.1. *A subgroup $\Lambda \subset E$ is discrete if and only if*

$$\dim(\mathbb{Q}\Lambda) = \mathrm{rk}(\Lambda)$$

Proof. ($\Leftarrow$) Let $l_1, \ldots, l_r \in \Lambda$ be free generators from $\mathbb{Q}\Lambda (r \leq \dim(E))$. By assumption, these are linear independent over $\mathbb{Q}$ and can therefore be completed to a basis of E. Because $B \subset E$ is bounded, the coordinate functions of $B \cap \Lambda$ are also bounded, and hence $B \cap \Lambda$ is finite.

($\Rightarrow$) We choose a basis $l_1, \ldots, l_r \in \Lambda$ of $\mathbb{Q}\Lambda$ and let

$$F(\Lambda) := \left\{ \sum_{i=1}^{r} x_i l_i \mid x_i \in \mathbb{Q}, -1/2 < x_i \leq 1/2 \right\}.$$

The set $F(\Lambda)$ is bounded and therefore $F(\Lambda) \cap \Lambda$ is finite by assumption. We set $\Lambda' := \sum_i \mathbb{Z} l_i \subset \Lambda$, and then the inclusion $\Lambda/\Lambda' \hookrightarrow \mathbb{Q}\Lambda/\Lambda' \cong F(\Lambda)$ shows that the quotient Λ/Λ' is also finite. Therefore the sublattice $\Lambda' \subseteq \Lambda$ has finite index, so that Λ is finitely generated and torsion free as a subgroup of E. It is thus free and hence

$$\dim(\mathbb{Q}\Lambda) = \mathrm{rk}(\Lambda') = \mathrm{rk}(\Lambda).$$

A discrete subgroup of maximal rank is called a *lattice*. □

We now equip $(E, \langle ., . \rangle)$ with the following measure: given an orthonormal basis $e_1, \ldots, e_n$ of E, we can produce an isomorphism e between E and $\mathbb{Q}^n$. We set then $d\mu = e^* d\mu_L$, where $d\mu_L$ indicates the Lebesgue measure. We obtain in this way a volume,

$$\mathrm{Vol}(F(\Lambda)) = \int_{F(\Lambda)} d\mu.$$

Theorem A.2. *Let $\Lambda \subset E$ be a lattice, and $e_1, \ldots, e_n$, an orthonormal basis of E with $\Lambda = \Phi(\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n)$. Then,*

$$\mathrm{Vol}(F(\Lambda)) = |\det(\Phi)|.$$

Proof. The pullback satisfies $\Phi^* d\mu = \det(\Phi) d\mu$, and therefore,

$$\begin{aligned}
\mathrm{Vol}(F(\Lambda)) &= \int_{F(\Lambda)} d\mu = \int_{F(\Phi(\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n))} d\mu \\
&= \int_{F(\mathbb{Z}e_1 + \cdots \mathbb{Z}e_n)} \Phi^* d\mu \\
&= |\det(\Phi)| \int_{F(\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n)} d\mu \\
&= |\det(\Phi)| \int_{F(\mathbb{Z}^n)} d\mu_L = |\det(\Phi)|.
\end{aligned}$$

□

We note further that $\det(\Phi)^2 = \det(\Phi\Phi^T) = \det(\langle l_i, l_j \rangle)$ for $l_i = \Phi(e_i)$.

8.2 Geometry of numbers

We consider now the lattice $\Lambda \subseteq \mathbb{R}^n$ of the Euclidean space $\mathbb{R}^n$ viewed with a standard scalar product with respect to which the canonical basis is orthonormal. We call a set $\mathcal{S} \subset \mathbb{R}^n$ convex if for every pair of points $x, y \in \mathcal{S}$, the linear combinations $\lambda x + \mu y \in \mathcal{S}$ for all $\lambda + \mu < 1$ with $0 \leq \lambda, \mu \leq 1$. We call the set $\mathcal{S}$ symmetric if $\mathcal{S} = -\mathcal{S}$. For a closed, symmetric and convex set $\mathcal{K} \subset \mathbb{R}^n$, we define the successive minima of the lattice Λ with respect to $\mathcal{K}$ by

$$\lambda_k := \inf\{\lambda \mid \lambda K \text{ contains } k \text{ linear independent lattice points}\}.$$

It clearly holds that

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Writing $d(\Lambda) = \det(\Lambda)$, we formulate the following theorem of Minkowski.

Theorem A.3 (Minkowski)**.** *With the above notation,*

$$\frac{1}{n!} \leq \lambda_1 \cdots \lambda_n \frac{\mathrm{Vol}(\mathcal{K})}{2^n d(\Lambda)} \leq 1.$$

Proof. We refer the reader to J. W. S. Cassels, *An Introduction to the Geometry of Numbers*. □

We now let $\delta(\mathcal{K}, \Lambda) = \frac{\mathrm{Vol}(\mathcal{K})}{2^n d(\Lambda)}$, so that

$$\frac{1}{n!} \leq \lambda_1 \cdots \lambda_n \cdot \delta(\mathcal{K}, \Lambda) \leq 1.$$

In particular,

$$\lambda_1^n \leq \lambda_1 \cdots \lambda_n \leq \delta(\mathcal{K}, \Lambda)^{-1}.$$

Then $\delta(\mathcal{K}, \Lambda) \geq 1$, so that $\lambda_1 \leq 1$ which means that $\mathcal{K}$ contains a nonzero lattice point. This is the content of the following corollary.

Corollary A.4 (Minkowski's Lattice Point Theorem)**.** *If, under the above assumptions, the inequality*

$$\mathrm{Vol}(\mathcal{K}) \geq 2^n d(\Lambda),$$

holds, then $\mathcal{K}$ contains a non-trivial lattice point.

This immediately implies a corollary.

Corollary A.5. *Let $l_1, \ldots, l_n \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$ be linear forms, and let $0 < c_1, \ldots, c_n \in \mathbb{R}$ with $c_1 \cdots c_n \geq \det(l_1, \ldots, l_n) d(\Lambda)$, so there is a $0 \neq u \in \Lambda$ with*

$$|l_j(u)| \leq c_j$$

for $1 \leq j \leq n$.

Proof. We define $\mathcal{K}$ by $|l_j| \leq c_j$ for $1 \leq j \leq n$. Then,

$$\mathrm{Vol}(\mathcal{K}) = 2^n c_1 \cdots c_n / \det(l_1, \ldots, l_n) \geq 2^n d(\Lambda).$$

□

8.3 Norm and Trace

Let L/K be a finite field extension. Then L is in particular a finite dimensional K-vector space. We consider $\xi \in L$ and define $T_\xi \in GL_K(L)$ by $T_\xi(x) = \xi x$ and let

$$\begin{aligned} N_{L/K}(\xi) &:= \det(T_\xi) \\ Tr_{L/K}(\xi) &:= \mathrm{Tr}(T_\xi). \end{aligned}$$

The norm and trace enjoy the following properties:

(i) $N_{L/K}(\xi\eta) = N_{L/K}(\xi)N_{L/K}(\eta)$,

(ii) $N_{L/K}(1) = 1$,

(iii) $\mathrm{Tr}_{L/K}(\xi + \eta) = \mathrm{Tr}_{L/K}(\xi) + \mathrm{Tr}_{L/K}(\eta)$.

Therefore the norm and trace homomorphisms map from L^* to K^* and from L to K respectively.

According to the definition $T_\xi(x) = \xi x$, which means $x \in \ker(\xi id - T_\xi)$ and hence $\det(\xi id - T_\xi) = 0$. Therefore ξ is the root of

$$\begin{aligned} \det(\xi id - T_\xi) &= T^n - \mathrm{tr}_{L/K}(\xi)T^{n-1} + \cdots + (-1)^n N_{L/K}(\xi) \\ &= \prod_{\sigma \in \mathrm{Hom}(L/\overline{K})} (T - \sigma\xi) \end{aligned}$$

Hence a theorem.

Theorem A.6. *For separable extensions L/K we have,*

(i) $\det(\xi id - T_\xi) = \prod_{\sigma \in \mathrm{Hom}(L/\overline{K})}(T - \sigma\xi)$,

(ii) $N_{L/K}(\xi) = \prod \sigma\xi$,

(iii) $\mathrm{Tr}_{L/K}(\xi) = \sum \sigma\xi$.

The following statements are left to the reader as an exercise.

(i) $M \supseteq L \supseteq K \Rightarrow \mathrm{Tr}_{M/K} = \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{M/L}$,

(ii) $N_{M/K} = N_{L/K} \circ N_{M/L}$,

(iii) If L/K is separable, the function

$$\mathrm{Tr}_{L/K} : \begin{array}{ccc} L \times L & \to & K \\ (\xi, \eta) & \mapsto & \mathrm{Tr}(\xi\eta) \end{array}$$

is non-degenerate, symmetric and bilinear.

(iv) If $L \supseteq K_s \supseteq K$, K_s is the separable closure of K and $L \supset K_s$ is purely inseparable, so that

$$\mathrm{Tr}_{L/K}(\xi) = [L : K] \, \mathrm{Tr}_{K_s/K}(\xi)$$

and

$$N_{L/K}(\xi) = N_{K_s/K}(\xi)^{[L:K]}.$$

Now let K be an algebraic number field. Then the norm of an ideal $\mathfrak{a} \in \mathcal{O}_K$ is defined by

$$N_K(\mathfrak{a}) := \mathcal{O}_K : \mathfrak{a}.$$

For $\xi \in \mathcal{O}_K$ we set $N_K(\xi) := N_K(\xi \mathcal{O}_K)$; this definition agrees with the one given earlier.

Since $\mathcal{O}_K$ is a Dedekind domain,

$$\mathfrak{a} \neq \prod \mathfrak{p}^{\nu_\mathfrak{p}(\mathfrak{a})}$$

and by the Chinese Remainder theorem it follows that

$$\mathcal{O}_K/\mathfrak{a} \cong \prod \mathcal{O}_K/\mathfrak{p}^{\nu_\mathfrak{p}(\mathfrak{a})},$$

hence

$$N_K(\mathfrak{a}) = \prod N_K(\mathfrak{p}^{\nu_\mathfrak{p}(\mathfrak{a})})$$

Lemma A.7. *With the above notation,*

$$\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong \mathcal{O}/\mathfrak{p}.$$

Proof. There are isomorphisms,

$$\mathcal{O}/\mathfrak{p} \cong \mathcal{O}_\mathfrak{p}/\mathfrak{p}\mathcal{O}_\mathfrak{p} \cong \mathcal{O}_\mathfrak{p}/\pi\mathcal{O}_\mathfrak{p}$$

and

$$\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong \mathfrak{p}^n\mathcal{O}_\mathfrak{p}/\mathfrak{p}^{n+1}\mathcal{O}_\mathfrak{p} \cong \pi^n\mathcal{O}_\mathfrak{p}/\pi^{n+1}\mathcal{O}_\mathfrak{p}.$$

The final term is an $\mathcal{O}_\mathfrak{p}/\pi\mathcal{O}_\mathfrak{p}$-vector space of dimension one. It follows that

$$\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong \mathcal{O}/\mathfrak{p}.$$

□

The immediate conclusion gives that $N_K(\mathfrak{p}^n) = N_K(\mathfrak{p})^n$, hence

$$N_K(\mathfrak{a}) = \prod N_K(\mathfrak{p})^{\nu_\mathfrak{p}}(\mathfrak{a})$$

8.4 Ramification and discriminants

Let R be a Dedekind domain with fraction field K, and furthermore let $L \supseteq K$ be a finite, separable extension containing R', the integral closure of R in L.

Definition A.8. Let $\mathfrak{P}$ be a nonzero prime ideal in R' and $\mathfrak{P} \cap R = \mathfrak{p}$. We say that $\mathfrak{P}$ is ramified over R if either $e(\mathfrak{P}) > 1$ or $k(\mathfrak{P})$ is non-separable over $k(\mathfrak{p})$. One says the prime ideal $\mathfrak{p}$ is ramified over R' if there is a prime ideal $\mathfrak{P}$ in R' which lies over $\mathfrak{p}$ and is ramified over R.

Let V be a vector space, $B = \{v_1, \ldots, v_n\}$, a basis, and $\beta : V \times V \to K$ a bilinear form. We set

$$d(B) := \det(\beta(v_i, v_j))$$

and call this the *discriminant* of the basis B with respect to β. If L/K is a finite field extension of number fields and $\mathfrak{a} \subset L$ is a finitely generated $\mathcal{O}_K$-module, then we choose a free submodule $a_1\mathcal{O}_K + \cdots + a_n\mathcal{O}_K$ and examine the discriminant

$$d = \det(\mathrm{Tr}_{L/K}(a_i, a_j)).$$

The ideal generated in $\mathcal{O}_K$ by all such d is called the discriminant of $\mathfrak{a}$ and is denoted by $\delta_{L/K}(\mathfrak{a})$. In the particular case when $\mathfrak{a} = \mathcal{O}_L$, one calls $\delta_{L/K} = d(\mathcal{O}_L)$ the discriminant of L/K. It is an ideal in $\mathcal{O}_K$.

Proposition A.9. *Let $\mathcal{S}$ be a multiplicative set in $\mathcal{O}_K$. Then*

$$\delta_{L/K}(\mathfrak{a})_{\mathcal{S}} = \delta_{L/K}(\mathfrak{a}_{\mathcal{S}})$$

Proof. A basis of $\mathfrak{a}$ is a basis of $\mathfrak{a}_{\mathcal{S}}$, and ever basis $B_{\mathcal{S}} \subseteq \mathfrak{a}_{\mathcal{S}}$ defines a basis of $\mathfrak{a}$ by multiplying by an element $s \in \mathcal{S}$. From this follows the conclusion, because

$$d_{sB} = s^{2[L:K]} d_B.$$

□

The relation between ramification and discriminant is given through the following theorem.

Theorem A.10. *The prime ideals of R which are ramified in R' are exactly those ideals which divide the discriminant $\delta_{L/K}$.*

Proof. Let $\mathfrak{p} \in \mathcal{O}_K$ be a prime ideal, and $\mathcal{S} = \mathcal{O}_K \mathfrak{p}$. By Proposition A.9, we can localize at $\mathcal{S}$ and consider the rings $R = \mathcal{S}^{-1}\mathcal{O}_K$ as well as $R' = \mathcal{S}^{-1}\mathcal{O}_L$. Then $\mathfrak{p}$ is ramified in $\mathcal{O}_L$ exactly when $\mathfrak{p}R$ is ramified in R'. The ring R is a discrete valuation ring and therefore a principal ideal ring. For this reason, R' is a free R-module which can be written as $R' = x_1 R \oplus \cdots \oplus x_r R$, etc. (see Gerald J. Janusz, *Algebraic Number Fields*). □

Minkowski Space[1]

Let $F : \mathbb{C} \to \mathbb{C}$ be complex conjugation, i.e., $\mathrm{Gal}(\mathbb{C}, \mathbb{R}) = \{1, F\}$ and $\Sigma = \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$. Then we obtain the embedding

$$j : \begin{array}{ccc} K & \to & \Gamma(\Sigma, \mathbb{C}) \cong K \otimes_{\mathbb{Q}} \mathbb{C} \\ x & \mapsto & j(x) : j(x)(\tau) = \tau(x), \tau \in \Sigma, \end{array}$$

where $\Gamma(\Sigma, \mathbb{C}) = \{z : \Sigma \to \mathbb{C}\}$. The complex conjugation F operates on Σ by $\tau \mapsto F(\tau) = F \circ \tau$ and on $\Gamma(\Sigma, \mathbb{C})$ by $z \mapsto F(z)$ via $F(z)(\tau) = F(z(F(\tau))) = (F \circ z \circ F)(\tau)$. The space $\Gamma(\Sigma, \mathbb{C})$ possesses a non-trivial involution $x \mapsto F(x)$, and we obtain a eigenspace decomposition,

$$\Gamma(\Sigma, \mathbb{C}) = \Gamma(\Sigma, \mathbb{C})^+ \oplus \Gamma(\Sigma, \mathbb{C})^-$$

in the eigenspace with eigenvalues 1 and -1 respectively. For $z \in \Gamma(\Sigma, \mathbb{C})^+$ we have $z(F(\tau)) = F(z(\tau))$, in other words, z and F commute. The embedding j factors through $\Gamma(\Sigma, \mathbb{C})^+$ because

$$\begin{aligned} F(j(x))(\tau) &= (F \circ j(x) \circ F)(\tau) \\ &= F(j(x)(F(\tau))) \\ &= F(F(\tau)(x)) \\ &= (F \circ F \circ \tau)(x) \\ &= \tau(x) \\ &= j(x)(\tau) \end{aligned}$$

for an arbitrary $\tau \in \Sigma$, and therefore $j(x) \in \Gamma(\Sigma, \mathbb{C})^+$. This space together with the Euclidean scalar product that we now introduce is called the Minkowski space.

We define a sesquilinear form on $\Gamma := \Gamma(\Sigma, \mathbb{C})$ by

$$(x, y) \in \Gamma \times \Gamma \mapsto \langle x, y \rangle := \sum_{\tau} x(\tau) F(y(\tau)).$$

It satisfies the property that $F\langle x, y \rangle = \langle y, x \rangle$ and

$$\begin{aligned} \langle Fx, Fy \rangle &= \sum_{\tau} F(x)(\tau) F(F(y)(\tau))) \\ &= \sum_{\tau} (F \circ x \circ F)(\tau)(F(F \circ y \circ F))(\tau) \\ &= F\left(\sum_{\tau} x(\tau) F(y(\tau)) \right) \\ &= F(x, y), \end{aligned}$$

[1] J. Neukirch, *Algebraische Zahlentheorie*, III.4

that means F commutes with Tr, so that,

$$\mathrm{Tr}(x) = \mathrm{Tr}(F(x)) = F(\mathrm{Tr}(x))$$

for $x \in \Gamma(\Sigma, \mathbb{C})^+$, and therefore $j \circ \mathrm{Tr} = \mathrm{Tr}_{K/\mathbb{Q}}$ as claimed.

We now construct a canonical measure on $(\Gamma(\Sigma, \mathbb{C})^+, \langle , \rangle)$. Assume the volume of an orthonormal parallelepiped $\mathcal{P}$ is $\mathrm{Vol}(\mathcal{P}) = 1$. Let $b_1, \ldots, b_n$ be a basis of $\Gamma(\Sigma, \mathbb{C})^+$, and let

$$\mathcal{P}(b_1, \ldots, b_n) = \{x_1 b_1 + \cdots + x_n b_n \mid 0 \leq x_i < 1\},$$

be the parallelepiped defined by the basis. The transformation

$$b_i = \sum_k a_{ik} e_k, 1 \leq i \leq n$$

defines a matrix $A := (a_{ik})$, and we obtain

$$\begin{aligned} \mathrm{Vol}(\mathcal{P}(b_1, \ldots, b_n)) &= \det(\langle b_i, b_j \rangle) \\ &= \det\left(\left\langle \sum_k a_{ik} e_k, \sum_l a_{jl} e_l \right\rangle\right) \\ &= \det\left(\sum_k a_{ik} a_{jk}\right) \\ &= \det(AA^T) \end{aligned}$$

The basis $b_1, \ldots, b_n$ generates the lattice,

$$\Lambda = \mathbb{Z} b_1 + \cdots + \mathbb{Z} b_n \subseteq \Gamma(\Sigma, \mathbb{C})^+$$

and we define the *Lattice volume* to be,

$$d(\Lambda) := \mathrm{Vol}(\mathcal{P}(b_1, \ldots, b_n)).$$

Example A.11. $\mathcal{O}_K$ is a free $\mathbb{Z}$-module of the form

$$\mathcal{O}_K = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n,$$

and therefore

$$j(\mathcal{O}_K) = \mathbb{Z} j(\omega_1) + \cdots \mathbb{Z} j(\omega_n).$$

It follows that

$$d(\mathcal{O}_K)^2 = \det(\langle j\omega_k, j\omega_l \rangle) = d_K = \det(\mathrm{Tr}_{K/\mathbb{Q}}(\omega_k \omega_l)).$$

For $\mathfrak{a} \subset \mathcal{O}_K$ is $\Lambda := j\mathfrak{a} \subseteq \Gamma(\Sigma, \mathbb{C})^+$ a sub-lattice of $j(\mathcal{O}_K$, and one obtains the relation,

$$d(\mathfrak{a}) = N(\mathfrak{a}) d(\mathcal{O}_K) = N(\mathfrak{a}) \sqrt{|d_K|}.$$

The logarithmic Minkowski space[2]

[2]J. Neukirch, *Algebraische Zahlentheorie*, I.7