Abstract

This thesis deals with Diophantine equations, specifically those arising from taking the field norm of certain elements of a number field. For a certain class of these the mathematician, Thoraf Skolem, showed that an infinite number of solutions to an equation belonging to this class gives rise to a certain curve on a local manifold. As a special case of this result he showed that Thue's equation, $f(x, y) = c$, cannot have infinitely many solutions because no such curves can exist in this case. We'll be establishing these two results with a focus on establishing the necessary prerequisites to do that. This includes algebraic structures such as forms, modules and completetions of fields, specifically the $\mathfrak{p}$-adic numbers.

$\mathfrak{p}$-adic Numbers and Skolem's Method

Daniel Grunkin

Contents

1 Introduction

Diophantine equations, named after the greek mathematician Diophantus, are polynomial equations in multiple variables where only integer solutions are of interest, that is equations of the type

$$F(x_1, ..., x_n) = c, \tag{1.1}$$

where F is a multivariate polynomial with integer coefficients and c is a rational number. Some of these equations are known for being notoriously hard to solve and therefore we often ask easier questions like: Does this equation have finitely many solutions or does it have infinitely many? If it has finitely many can we then obtain an upper bound on the number of these? Even these questions become extremely hard to answer so we will restrict our attention to polynomials that occur by applying the field norm to certain elements. Details are given in section 4, but for now lets see an example of one such equation. Consider the real quadratic field $\mathbb{Q}(\sqrt{2})$. This field has basis $1, \sqrt{2}$ over $\mathbb{Q}$ and we have

$$N_{\mathbb{Q}(\sqrt{2})/\mathbb{Q}}(x + y\sqrt{2}) = (x + y\sqrt{2})(x - y\sqrt{2}) = x^2 - 2y^2$$

As such we get, for instance, a Diophantine equation

$$x^2 - 2y^2 = 7$$

and we observe that $(3, 1)$ is a solution. Also, if (x, y) is a solution then so is $(3x+4y, 2x+3y)$ so this equation has infinitely many solutions in integers. The Norwegian mathematician, Axel Thue, showed in 1909 the following theorem, known as Thues theorem.

Theorem 1.1 (Thue)**.** *Suppose $f(x, y)$ is an irreducible homogenous polynomial with integer coefficients of degree $n \geq 3$. Then there are only finitely many integer solutions to the equation $f(x, y) = c$, for any non-zero $c \in \mathbb{Q}$.*

The way Thue went about proving this was by proving that for any algebraic number α of degree ≥ 3 and any $\epsilon > 0$ there are only finitely many coprime integers p, q, where $q > 0$ so that

$$|\alpha - \frac{p}{q}| < q^{-(d+1+\epsilon)/2},$$

see [9]. Thoralf Skolem also proved the theorem but in a different way. He proved a general result that says that an equation like eq. (1.1) which has infinitely many solutions in integers and where F is, among other things, homogenous, then there exists what is called a curve on a certain kind of object called a local manifold. He could then show that the equation $f(x, y) = c$ in theorem 1.1 was a special case of this more general situation if he imposed the extra condition that $f(x, 1)$ has at least one imaginary root. He then proved that no curves could exists in this case, implying only finitely many solutions. This proof strategy is suitably referred to as Skolem's method. The more general result

that Skolem showed takes as its starting point in realizing that the assumptions on F makes it possible to write $F(x_1, ..., x_n) = N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_n\mu_n)$, where K is some number field and $\mu_1, ..., \mu_n \in K$. This is a situation very analogous to the one involving the polynomial $x^2 - 2y^2$ from above. Working in this setting allows us to make use of results from algebraic number theory. For example we shall see that a slightly more general version of Dirichlet's unit theorem stated in [7, Theorem 38] allows for a very concrete description of the elements $x_1\mu_1 + ... + x_n\mu_n$ having norm c. Now the proofs by Thue and Skolem both suffer from not being effective, meaning that they don't provide a procedure for figuring out what the solutions actually are. This was the case for all known proofs until Alan Baker in 1969 overcame this issue by giving a proof based on his theory on linear forms in logarithms, see [1, Theorem 4.1]. Our focus in this thesis however will be in understanding Skolem's method and how to apply it prove Thues theorem, see section 12. All sections preceding that seek to establish the necessary theory for this proof modulo some minor roundabouts here and there.

2 Dual basis

Proposition 2.1. *Suppose that K/k is a field extension of degree n. Given any basis $\mu_1, ..., \mu_n$ for K over k there exists a basis, $\mu_1^*, ..., \mu_n^*$ called the dual basis of $\mu_1, ..., \mu_n$ so that*

$$\mathrm{Tr}_{K/k}(\mu_i^*\mu_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

Proof. For ease of notation we will write Tr instead of $\mathrm{Tr}_{K/k}$. Choose n elements, $c_1, ..., c_n$, in k. We know that the $n \times n$ matrix, $(\mathrm{Tr}(\mu_i\mu_j))$, is non-singular since

$$0 \neq \mathrm{disc}(\mu_1, ..., \mu_n) = |\,\mathrm{Tr}(\mu_i\mu_j)|^2$$

This means that there is a unique solution, $x_1, ..., x_n \in k$, to the n equations

$$\sum_{j=1}^{n} \mathrm{Tr}(\mu_i\mu_j)x_j = \frac{c_i}{n} \quad (i = 1, ..., n)$$

Let $\alpha = \sum_{j=1}^{n} x_j\mu_j$. Using rules of the trace, we get for any i that

$$c_i = \sum_{j=1}^{n} \mathrm{Tr}(x_j\mu_i\mu_j) = \mathrm{Tr}(\sum_{j=1}^{n} x_j\mu_i\mu_j) = \mathrm{Tr}(\alpha\mu_i)$$

Thus, we have demonstrated that for any choice of $c_1, ..., c_n \in k$, there is a unique $\alpha \in K$ such that $\mathrm{Tr}(\alpha\mu_i) = c_i$. For $(i, j) \in \{1, ..., n\}^2$ we now define

$$c_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

For every $i \in \{1, ..., n\}$ we get, by what we have just shown, a unique $\mu_i^* \in K$ such that $\mathrm{Tr}(\mu_i^* \mu_j) = c_{ij}$ for $j \in \{1, ..., n\}$. To show that $\mu_1^*, ..., \mu_n^*$ is a basis, it suffices to show that they are linearly independent over k. So suppose that

$$\sum_{i=1}^{n} x_i \mu_i^* = 0.$$

Multiplying by μ_j and taking the trace, we get

$$0 = \mathrm{Tr}(\sum_{i=1}^{n} x_i \mu_i^* \mu_j) = \sum_{i=1}^{n} x_i \,\mathrm{Tr}(\mu_i^* \mu_j) = x_j,$$

which shows that $x_j = 0$ is zero. Doing this for all $j \in \{1, ..., n\}$ we get $x_1 = ... = x_n = 0$. □

3 Formal power series

Let K be a field. Recall the ring of formal power series, $K[[X]]$. Addition of two elements $f(x) = \sum_{i=0}^{\infty} a_i x^i$ and $g(X) = \sum_{i=0}^{\infty} b_i x^i$ is defined in the obvious way and multiplication is defined as

$$f(X)g(X) = \sum_{i=0}^{\infty} c_i x^i$$

where $c_i = \sum_{j=0}^{i} a_j b_{i-j}$. Thus, multiplication is really just the process of expanding out brackets and collecting like terms. In some cases it is also possible to compose power series. Let $\omega : K[[X]] \to \mathbb{Z} \cup \{\infty\}$ be defined by sending 0 to ∞ and a non-zero power series $\sum_{i=m}^{\infty} a_i x^i$, $a_m \neq 0$, to m. Note that if $f(X) = \sum_{i=1}^{\infty} a_i X^i$ then the sequence $\omega(f^0), \omega(f^1), ...$ is strictly increasing. Hence for any $n \in \mathbb{N}_0$ it is only the first $n+1$ power series in the sequence $f^0, f^1, f^2, ...$ that can contain a term of degree n. Hence we can in this case define, for any $g(X) = \sum_{i=0}^{\infty} b_i X^i$ the composite

$$g(f(X)) = \sum_{i=0}^{\infty} b_i f(X)^i$$

Formal differentiation can be defined on $K[[X]]$, simply by formal differentiation of each term, that is,

$$f'(X) = \sum_{i=1}^{\infty} i a_i X^{i-1}$$

This operation is linear (by definition) and the well known product and chain rules actually hold in $K[[X]]$. Using the product rule and an induction argument

it follows that $(f(X)^n)' = nf(X)^{n-1}f'(X)$. Hence if $f(X) = \sum_{i=1}^{\infty} a_i X^i$ and $g(X) = \sum_{i=0}^{\infty} b_i X^i$ then

$$\begin{aligned}(g(f(X)))' &= \sum_{i=1}^{\infty} ib_i f(X)^{i-1} f'(X) \\ &= (\sum_{i=1}^{\infty} ib_i f(X)^{i-1}) f'(X) \\ &= g'(f(X))f'(X)\end{aligned}$$

We can also consider the field of fractions of $K[[X]]$ and we will denote this field as $K((X))$. This is called the field of Laurant series and each non-zero element in it can be written as $\sum_{i=m}^{\infty} a_i X^i$, for some $m \in \mathbb{Z}$, where $a_m \neq 0$. As such we can extend ω defined above in an obvious way to $K((X))$.

Lemma 3.1. *Suppose K is a field and $f \in K((X))$. If $f \notin K$ then f is transcendental over K. That is, there exists no non-zero polynomial with coefficients in K having f as a root.*

Proof. Suppose that $f \in K((X)) \setminus K$. Then $f \neq 0$ so $\omega(f) \in \mathbb{Z}$. First, let us assume that $\omega(f) > 0$. Then the numbers $\omega(f^0), \omega(f^1), \omega(f^2)...$ constitute a strictly increasing sequence in $\mathbb{Z}$. In particular, these are all different numbers. So if we have a K-linear combination

$$a_0 + a_1 f_1 + ... + a_n f^n = 0,$$

then we must have $0 = a_0 = ... = a_n$. A similar argument can be made when $\omega(f) < 0$ and thus f cannot be algebraic when $\omega(f) \neq 0$. Lastly, suppose that $\omega(f) = 0$ and let $a_0 \in K$ be the constant term of f. If f were algebraic over K then also $f - a_0$ would be algebraic, which is a contradiction since $\omega(f - a_0) > 0$. Hence f is not algebraic. □

4 Forms, modules and orders

4.1 Forms

A form $F(x_1, ..., x_m)$ in m variables of degree k is a homogenous polynomial in the variables $x_1, ..., x_m$ of degree k with coefficients in $\mathbb{Q}$. A form F is called reducible if it can be written as a product of two forms of degree strictly larger than 0. Otherwise it is called irreducible. Two forms of the same degree are called integrally equivalent if one can be obtained from the other and vice versa by a linear change of variables with coefficients in $\mathbb{Z}$. Two forms being integrally equivalent defines an equivalence relation on the set of forms having some fixed degree. If $F(x_1, ..., x_m)$ and $G(y_1, ..., y_l)$ are integrally equivalent forms, then knowing all the solutions to $F(x_1, ..., x_m) = a$ is exactly the same as knowing all the solutions to $G(y_1, ..., y_l) = a$. If a form F splits into linear factors in some extension, K, of $\mathbb{Q}$, then it is called decomposable. The following proposition provides some examples of forms that are always decomposable.

Proposition 4.1. *Let $F(x,y) = \sum_{i=0}^{m} a_i x^{m-i} y^i$ be a form of degree m where $a_0 \neq 0$. Then over the splitting field for $F(x,1)$ we have the factorization*

$$F(x,y) = a_0(x-\alpha_1)...(x-\alpha_m)$$

Proof. Since $a_0 \neq 0$ we must have $y \nmid F(x,y)$. Let $G(t) = F(t,1)$. This is a polynomial in $\mathbb{Q}[t]$ with degree m. We can now write $F(x,y) = y^m F(\frac{x}{y},1) = y^m G(\frac{x}{y})$. As G is just a polynomial we can factor it into linear terms over its splitting field, K.

$$G(t) = a_0(t-\alpha_i)...(t-\alpha_m)$$

This means that

$$F(x,y) = y^m G(\frac{x}{y}) = a_0(x-\alpha_i y)...(x-\alpha_m y)$$

Thus, $F(x,y)$ is decomposable as it splits into linear factors over K. □

Remark 4.1. *Consider the special case where $F(x,y)$ is irreducible and $a_0 = 1$. It then follows that $F(x,1)$ is also irreducible of degree m so adjoining any of its roots α to $\mathbb{Q}$ gives an extension, K, of degree m. The α_i in the factorization of G above are then the conjugates of α so setting $\theta = -\alpha$ we obtain*

$$F(x,y) = N_{K/\mathbb{Q}}(x+y\theta)$$

We shall use this fact once we start looking at Thue's theorem.

4.2 Modules and forms connected to them

Let us assume that K is a number field of degree n over $\mathbb{Q}$ and let $\mu_1, ..., \mu_k$ be elements in K. The set, M, consisting of all $\mathbb{Z}$-linear combinations of these elements is called a module in K and we will write $M = \langle \mu_1, ..., \mu_k \rangle$ for this set. The μ_i's are called generators for the module and we can use them to construct a form, as the following proposition shows.

Proposition 4.2. *For $x_1, ..., x_k \in \mathbb{Z}$ we get a decomposable form $F(x_1, ..., x_k) = N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_n\mu_k)$ of degree n.*

Proof. As K is in particular a finite separable extension of $\mathbb{Q}$, the primitive element theorem states that there is $\theta \in K$ so that $K = \mathbb{Q}(\theta)$. Let now $\sigma_1, ..., \sigma_n$ be the n embeddings from K into $\mathbb{C}$. We obtain a field extension L of K by adjoining all conjugates, $\sigma_1(\theta), ..., \sigma_n(\theta)$ to $\mathbb{Q}$. As the embeddings map into $\mathbb{C}$ we see that L is a subfield of $\mathbb{C}$ and it is a Galois extension of $\mathbb{Q}$ as it is splitting field for the minimal polynomial for θ over $\mathbb{Q}$. Let $G = \mathrm{Gal}(L/\mathbb{Q})$. Let H be the subgroup of G corresponding to K. We have $[G:H] = [L:\mathbb{Q}]$ and we see that we can interpret the embeddings σ_i as a system of coset representatives of

H in G, that is we have a disjoint union, $G = \cup_{i=1}^{n} \sigma_i H$. We observe that

$$\begin{aligned} F(x_1, ..., x_k) &= N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_n\mu_n) \\ &= \prod_{i=1}^{n} \sigma_i(x_1\mu_1 + ... + x_k\mu_k) \\ &= \prod_{i=1}^{n} (x_1\sigma_i(\mu_1) + ... + x_k\sigma_i(\mu_k)) \end{aligned}$$

When expanding out the parentheses we see that F is a homogenous polynomial of degree n in the variables $x_1, ..., x_k$. Let us see that the coefficients belong to $\mathbb{Q}$. Consider the set Γ of all functions from $\{1, ..., n\}$ and consider two functions in Γ to be equivalent if they have the same image. Take any $\tau \in \Gamma$ and let $[\tau]$ be its corresponding equivalence class. We see that the coefficient of $x_{\tau(1)} \dots x_{\tau(n)}$ is

$$\sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma_i(\mu_{\alpha(i)}) \tag{4.1}$$

Let $\sigma \in G$. Then $\sigma\sigma_i \in G$ so there is a unique j such that $\sigma\sigma_i \in \sigma_j H$. Thus, there is a permutation π of $\{1, ..., n\}$ so that $\sigma\sigma_i = \sigma_{\pi(i)} h_i$, for $h_i \in H$. Thus $\sigma\sigma_i(\alpha) = \sigma_{\pi(i)}(\alpha)$ for all $\alpha \in K$ and $i \in \{1, ..., n\}$, since h_i fixes K. Note that $[\tau] = [\tau \circ \pi^{-1}]$ and observe that

$$\begin{aligned} \sigma(\sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma_i(\mu_{\alpha(i)})) &= \sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma\sigma_i(\mu_{\alpha(i)}) \\ &= \sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma_{\pi(i)}(\mu_{\alpha(i)}) \\ &= \sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma_i(\mu_{\alpha \circ \pi^{-1}(i)}) \\ &= \sum_{\alpha \in [\tau \circ \pi^{-1}]} \sum_{i=1}^{n} \sigma_i(\mu_{\alpha(i)}) \\ &= \sum_{\alpha \in [\tau]} \sum_{i=1}^{n} \sigma_i(\mu_{\alpha(i)}) \end{aligned}$$

Thus, every coefficient is fixed under every single $\sigma \in G$ meaning that these all belong to $\mathbb{Q}$ and therefore F is a form of degree n. F is decomposable as it splits into linear factors over L. □

Definition 4.1. *Suppose that M is a module and K and that $M = \langle \mu_1, ..., \mu_m \rangle$. The form $F(x_1, ..., x_m) = N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_m\mu_m)$ is called a norm form and is said to be connected to the module M.*

There are many different forms connected to a single module but it turns out that they are all the same modulo being integrally equivalent.

Proposition 4.3. *All forms connected to a module M are integrally equivalent.*

Proof. Suppose that M is generated by $\mu_1, ..., \mu_m$ and also by $\rho_1, ..., \rho_l$. We get the following two forms which are connected to M;

$$F(x_1, ..., x_m) = N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_m\mu_m)$$
$$G(y_1, ..., y_s) = N_{K/\mathbb{Q}}(y_1\rho_1 + ... + y_l\rho_l)$$

For every $j = 1, ..., l$ we can find integers $c_{j1}, ..., c_{jm}$ so that $\rho_j = \sum_{k=1}^{m} c_{jk}\mu_k$ since $\rho_j \in M$ and the μ_i generate M. We then have

$$\sum_{j=1}^{l} y_j\rho_j = \sum_{j=1}^{l} y_j \sum_{k=1}^{m} c_{jk}\mu_k = \sum_{k=1}^{m}(\sum_{j=1}^{l} c_{jk}y_j)\mu_k = \sum_{k=1}^{m} x_k\mu_k,$$

where we put

$$x_k = \sum_{j=1}^{l} c_{jk}y_j \text{ for } k \in \{1, ..., m\}$$

Taking norms we see that $F(x_1, ..., x_m) = G(y_1, ..., y_l)$ under the linear change of variables right above. A completely symmetrical argument now shows that F and G are integrally equivalent. □

If M is a module and it contains a basis for the vector space $K/\mathbb{Q}$, then it is called a **full module**. Otherwise it is called a **nonfull module**. A form which is connected to a module is correspondingly called full and nonfull. By definition M is a finitely generated abelian group and by the structure theorem, M determines uniquely integers $r, s \geq 0$ and $d_1 \mid ... \mid d_s$, $d_i \geq 2$ such that

$$M \cong \mathbb{Z}^r \oplus \mathbb{Z}/d_1\mathbb{Z} \oplus ... \oplus \mathbb{Z}/d_s\mathbb{Z}$$

But M lives inside a field of characteristic 0 so M must be torsion-free and so $s = 0$. Thus, $M \cong \mathbb{Z}^r$, which means that M is a free abelian group of rank r. The modules in K can therefore be thought of as finitely generated free abelian groups, which means that concepts such as rank and basis now make sense for modules. The full modules inside K are exactly the modules having rank n and the nonfull ones have rank strictly smaller than that.

Proposition 4.4. *If $\mu_1, ..., \mu_m$ is a basis for a module M then these elements are linearly independent over $\mathbb{Q}$.*

Proof. Suppose that we have a $\mathbb{Q}$-linear combination

$$a_1\mu_1 + ... + a_m\mu_m = 0$$

Set c to be the product of all the denominators of the a_i. Then $c \neq 0$ and $ca_i \in \mathbb{Z}$ so we must have $a_i = 0$ as the μ_i are linearly independent over $\mathbb{Z}$. □

4.3 Coefficient rings and orders

In the following we assume that K is an algebraic number field of degree n over $\mathbb{Q}$. An **order** in K is a full module in K which is also a ring with unity. As the ring of integers of K has an integral basis it is in fact a full module, so it is an example of an order. We now show how we can construct other orders. Given a full module M in K, we can consider an element α in K so that $\alpha M \subseteq M$. Such an element is called a **coefficient** of M, and the set of all of these is called the **coefficient ring** of M, which we will denote by $\mathfrak{O}_M$, or simply $\mathfrak{O}$, when it is clear from the context what is meant. It would be strange to call this object a ring, if it wasn't a ring, so let us check that it is. It is clear that $1 \in \mathfrak{O}$ since $1M \subseteq M$ and so $\mathfrak{O}$ is non-empty. Let now α, β in $\mathfrak{O}$ and take any element x in M. We have $\alpha x, \beta x \in M$ so

$$(\alpha - \beta)x = \alpha x - \beta x \in M$$

Thus, $\alpha - \beta \in \mathfrak{O}$. Checking that we have closure under multiplication is similar and so by the subring criterion $\mathfrak{O}$ is a subring of K. In fact, $\mathfrak{O}$ is also a full module. To justify this, we must show that $\mathfrak{O}$ is a module and that it contains a basis for K as a vector space over $\mathbb{Q}$. If γ is any non-zero element of M, then $\gamma\mathfrak{O}$ is a group under addition and we have that $\gamma\mathfrak{O} \subseteq M$. Thus, $\gamma\mathfrak{O}$ is a free abelian group, since it is contained in the free abelian group M which means that it is a module in K. But then also $\mathfrak{O} = \gamma^{-1}\gamma\mathfrak{O}$ is a module. Let us show that $\mathfrak{O}$ is also full. Pick a basis $\mu_1, ..., \mu_n$ for M and let us prove the following lemma.

Lemma 4.1. *For any $\alpha \in K$ we have that $\alpha\mu_i$ is in M for all the μ_i if and only if α is in $\mathfrak{O}$.*

Proof. Take any $x \in M$ and write $x = \sum_{i=1}^{n} a_i\mu_i$ where the a_i are integers. Multiplying by α we get

$$\alpha x = \sum_{i=1}^{n} a_i(\alpha\mu_i)$$

So if the $\alpha\mu_i$ are all in M, this is just a finite sum of elements in M, meaning that the entire sum is in M. Hence, $\alpha M \subseteq M$. The other direction is clear. □

Let us now show that if $\alpha \in K$ then there exists an integer $b \neq 0$ so that $b\alpha \in \mathfrak{O}$. For each μ_i we can find a $\mathbb{Q}$-linear combination

$$\alpha\mu_i = \sum_{j=1}^{n} a_{ij}\mu_j,$$

since the μ_i constitute a basis for K over $\mathbb{Q}$ by proposition 4.4. Choose now an integer, $c \neq 0$ so that ca_{ij} is an integer for all i, j. This implies that $c\alpha\mu_i$ is in M for all the μ_i which by lemma 4.1 means that $c\alpha$ is in $\mathfrak{O}$. From this it is clear that we can pick an integer $b \neq 0$ so that we for all μ_i have that $b\mu_i$ is in $\mathfrak{O}$. The elements $b\mu_1, ..., b\mu_n$ is of course still is a basis for K over $\mathbb{Q}$ so $\mathfrak{O}$ is full so $\mathfrak{O}$ is an order in K. Note that we have in fact also shown the following

Proposition 4.5. *If M is a full module and $\mathfrak{D}$ is its coefficient ring then there is $b \neq 0$ so that $bM \subseteq \mathfrak{D}$.*

In fact $\mathfrak{D}$ is contained in O_K, the ring of integers of K. Because if $\alpha \in \mathfrak{D}$ then we have $\mathbb{Z}[\alpha] \subseteq \mathfrak{D}$ so $\mathbb{Z}[\alpha]$ has rank smaller than or equal to n. Hence $\mathbb{Z}[\alpha]$ is in particular finitely generated and that is the same as saying that $\alpha \in O_K$. Let us summarize what we have found in the following proposition

Proposition 4.6. *The coefficient ring $\mathfrak{D}$ of a full module is an order of K and is a subring of the ring of integers of K.*

4.4 More about full forms

We now present a slight generalization of Dirichlet's unit theorem found in [7, Theorem 38]. Interestingly we can use this theorem to prove this generalized version.

Theorem 4.1 (Dirichlet's unit theorem)**.** *Let $\mathfrak{D}$ be an order in some number field K of degree n and let r and $2s$ be the number of real and complex embeddings, respectively. Then*

$$\mathfrak{D}^* = W \oplus V$$

where W is a finite cyclic group consisting of all roots of unity of $\mathfrak{D}$ and V is a free abelian group of rank $t = r + s - 1$.

Proof. Let O be the ring of algebraic integers in K. For the order O we know that the above theorem holds, so we get

$$O^* = W \oplus V$$

with W being all roots of unity of O and V a free abelian group of rank $t = r + s - 1$. Since $\mathfrak{D}$ is a subring of O we also have $\mathfrak{D}^* \subseteq O^*$. Hence,

$$\mathfrak{D}^* = W' \oplus V'$$

where $W' \trianglelefteq W$ is finite cyclic and $V' \trianglelefteq V$ is free abelian of rank $t' \leq t$. We wish to show two things; That W' does indeed consist of all roots of unity of $\mathfrak{D}$ and that $t' = t$. For the first claim, if we have any root of unity $\xi \in \mathfrak{D}$, then ξ has finite order so it cannot possibly belong to V'. Thus the only possibility is that ξ is in W'. For the second claim, consider the quotient of groups $O/\mathfrak{D}$. Both of these have rank n, so this quotient is finite, and so we know that $f := [O : \mathfrak{D}] < \infty$. Thus, if $x \in O$ then $\overline{fx} = 0$ in $O/\mathfrak{D}$ so $fx \in \mathfrak{D}$, so $fO \subseteq \mathfrak{D}$. Of course fO is also a free abelian group of rank n so $R := O/fO$ is finite. But fO is also an ideal of the ring O, so in fact R is a finite ring. Consider now any unit $\epsilon \in V$. Then ϵ is in O^*, so $\overline{\epsilon} \in R$ is also a unit, since ring maps preserve units. Set now $k = \#R^*$. In R we get the equalities $\overline{\epsilon^k} = \overline{1}$ and $\overline{\epsilon^{-k}} = \overline{(\epsilon^{-1})^k} = \overline{1}$. Together these give us

$$\epsilon^k = 1 + f\alpha$$
$$\epsilon^{-k} = 1 + f\beta$$

where $\alpha, \beta \in O$. But as we argued above, $f\alpha$ and $f\beta$ both belong to $\mathfrak{D}$ and so both ϵ^k and ϵ^{-k} belong to $\mathfrak{D}$ which means that $\epsilon^k \in \mathfrak{D}^*$. Thus, ϵ^k is either in W' or V' and the first option is impossible as that would imply that ϵ^k would be in W. Thus, ϵ^k is in V' so V/V' is finite meaning that $t' = t$. □

Any basis, $\epsilon_1, ..., \epsilon_t$, for V in the above proposition is called a fundamental system of units for $\mathfrak{D}$. We say that two elements, α, β in a full module M are **associated** if there is a unit $\epsilon \in \mathfrak{D}$ so that $\alpha = \epsilon\beta$. Note that when M is equal to its own coefficient ring, this concept is exactly the same as that of being associated in rings. Being associated elements in M defines an equivalence relation on M, and from now on we will denote this relation as $\sim$. Define now for $c \in \mathbb{Z}$ and a module full module M of K the subsets

$$M_c = \{\alpha \in M \mid N_{K/\mathbb{Q}}(\alpha) = c\}$$
$$\overline{M}_c = \{\alpha \in M \mid |N_{K/\mathbb{Q}}(\alpha)| = c\}$$

Note that if $c = 0$ then $M_c = \{0\}$, so let us assume $c \neq 0$.We wish to show that up to associates, the set M_c is finite. This is quite a special feature for the full full modules. Before we can prove this we need the following.

Lemma 4.2. *Consider the field extension K/F and suppose that $\alpha \in K$ is algebraic over F. Write*

$$f(x) = x^n + a_{n-1}x^{n-1} + ... + a + a_0$$

for the minimal polynomial of α over F. Then $N_{F(\alpha)/F}(\alpha) = \pm a_0$. If K/F is finite then $N_{K/F}(\alpha) = \pm a_0^m$ where $m = [K : F(\alpha)]$

Proof. Denote by $\alpha_1, ..., \alpha_n$ the n conjugates of α. Then we can write

$$f(x) = (x - \alpha_1)...(x - \alpha_n),$$

and we see that

$$a_0 = \pm\alpha_1...\alpha_n = N_{F(\alpha)/F}(\alpha)$$

Now if K/F is finite and $m = [K : F(\alpha)]$ then

$$N_{K/F}(\alpha) = N_{F(\alpha)/F}(N_{K/F(\alpha)}(\alpha)) = N_{F/K}(\alpha^m) = N_{F/K}(\alpha)^m = \pm a_0^m$$

□

Let us prove the following theorem.

Theorem 4.2. *Let M be a full module of K. Then the quotient set $\overline{M}_c/\sim$ is finite for any $c \in \mathbb{N}$. In particular $M_c/\sim$ is finite.*

Proof. We first consider the special case where $M = \mathfrak{D}$. As $\mathfrak{D}$ is an order it is a full module so it is a free abelian group of rank n, hence isomorphic to $\mathbb{Z}^n$ so we get the isomorphisms

$$\mathfrak{D}/c\mathfrak{D} \cong \mathbb{Z}^n/c\mathbb{Z}^n \cong (\mathbb{Z}/c\mathbb{Z})^n$$

Now, $\mathbb{Z}/c\mathbb{Z}$ contains c elements, which means that

$$c^n = \#(\mathbb{Z}/c\mathbb{Z})^n = \#\mathfrak{D}/c\mathfrak{D}$$

Denote by $\bar{\alpha}$ as the image of the canonical projection of α in $\mathfrak{D}/c\mathfrak{D}$ and denote by $[\alpha]$ an equivalence class in $\overline{\mathfrak{D}}_c/\sim$, represented by $\alpha \in \overline{\mathfrak{D}}_c$. We show that there is a well-defined surjective function of sets

$$\phi : \overline{\mathfrak{D}}_c/c\mathfrak{D} \twoheadrightarrow \overline{\mathfrak{D}}_c/\sim,$$

given by $\phi(\bar{\alpha}) = [\alpha]$. Suppose $\bar{\alpha}, \bar{\beta}$ are in $\overline{\mathfrak{D}}_c/c\mathfrak{D}$ so that $\bar{\alpha} = \bar{\beta}$, where α, β are in $\overline{\mathfrak{D}}_c$, meaning that $|N_{K/\mathbb{Q}}(\alpha)| = |N_{K/\mathbb{Q}}(\beta)| = c$ and $\alpha, \beta \in \mathfrak{D}$. Let us show that $[\alpha] = [\beta]$. Now, $\overline{\alpha} = \overline{\beta}$ implies that

$$\alpha = \beta + c\gamma = \beta + |N_{K/\mathbb{Q}}(\beta)|\gamma,$$

for some γ in $\mathfrak{D}$. We claim that β divides $N_{K/\mathbb{Q}}(\beta)$ in $\mathfrak{D}$ - the reason for this is as follows. We know, by proposition 4.6, that β is an algebraic integer since it belongs to $\mathfrak{D}$, so its minimal polynomial over $\mathbb{Q}$, $f(x) = x^m + a_{m-1}x^{m-1} + ... + a_0$, has integer coefficients. Now $a_0 = -(\beta^m + a_{m-1}\beta^{m-1} + ... + a_1\beta)$ so $\beta \mid a_0$ in $\mathfrak{D}$ and therefore β also divides a_0^m, where $m = [K : \mathbb{Q}(\alpha)]$. By lemma 4.2 we have $N_{K/\mathbb{Q}}(\beta) = \pm a_0^m$ and the claim follows. We therefore also know that β divides $|N_{K/\mathbb{Q}}(\beta)|$ in $\mathfrak{D}$. Hence, β divides α in $\mathfrak{D}$ and similarly α divides β in $\mathfrak{D}$. Thus, α and β are associates, showing that ϕ is well-defined. It is surjective simply because if $[\alpha] \in \overline{\mathfrak{D}}_c/\sim$, then α is in $\overline{\mathfrak{D}}_c$ so $\phi(\bar{\alpha}) = [\alpha]$. That ϕ is a surjection implies that $\#(\overline{\mathfrak{D}}_c/\sim) \leq \#\overline{\mathfrak{D}}_c/c\mathfrak{D}$, since each element in $\overline{\mathfrak{D}}_c/\sim$ has at least one preimage. Now the inclusion $\overline{\mathfrak{D}}_c/c\mathfrak{D} \subseteq \mathfrak{D}/c\mathfrak{D}$ implies that $\#(\overline{\mathfrak{D}}_c/\sim) \leq \#\overline{\mathfrak{D}}_c/c\mathfrak{D} \leq \#\mathfrak{D}/c\mathfrak{D} = c^n$. We will now prove the general statement. Suppose that M is a full module and that $\mathfrak{D}$ is the coefficient ring of M. Then $\overline{\mathfrak{D}}_c/\sim$ has finitely many elements by what we have just shown. By proposition 4.5 a non-zero integer b so that $bM \subseteq \mathfrak{D}$ so we obtain injective maps

$$M \hookrightarrow bM \hookrightarrow \mathfrak{D},$$

where the first map is multiplication by b and the second one is inclusion. It is clear that if α and β are associated then also $b\alpha$ and $b\beta$ are associated. Hence we get the induced injective maps

$$(\overline{M}_c/\sim) \hookrightarrow (b\overline{M}_c/\sim) \hookrightarrow (\overline{\mathfrak{D}_c}/\sim)$$

which means that

$$\#(\overline{M}_c/\sim) \leq \#(b\overline{M}_c/\sim) \leq \#(\overline{\mathfrak{D}_c}/\sim) \leq c^n$$

The last claim now follows since $M_c \subseteq \overline{M}_c$. □

We now present a result that allows to find all the elements of M_c if we know the elements of $M_c/\sim$ and all the units with norm 1 in $\mathfrak{D}$.

Lemma 4.3. *Assume that the elements of $M_c/\sim$ are $[\gamma_1], ..., [\gamma_k]$ and that $\alpha \in M$. We then have that $\alpha \in M_c$ if and only if there is a uniquely determined i such that $\alpha = \epsilon\gamma_i$ where ϵ is a unit in $\mathfrak{D}$ with norm 1.*

Proof. If $\alpha \in M_c$ then, there is a unique γ_i such that $\alpha \in [\gamma_i]$. This means that $\alpha = \epsilon\gamma_i$ for some unit ϵ in $\mathfrak{D}$. But then

$$c = N(\alpha) = N(\epsilon\gamma_i) = N(\epsilon)N(\gamma_i) = N(\epsilon)c$$

So we must have that $N(\epsilon) = 1$. □

Note that an element $\epsilon \in \mathfrak{D}$ is a unit if and only if $N_{K/\mathbb{Q}}(\epsilon) = \pm 1$ simply because ϵ is an algebraic integer. In light of the above proposition we are now interested in figuring out when ϵ has norm 1.

4.5 Units of norm 1

Lemma 4.4. *Let K be a number field of degree n over $\mathbb{Q}$. Suppose n is odd. Then the only roots of unity in O_K are ± 1 and we have $N(1) = 1$ and $N(-1) = -1$. On the other hand, if n is even, then all the roots of unity in O_K have norm 1.*

Proof. Let ζ be a primitive kth root of unity in O_K. To start with, let us suppose that n is odd. Then

$$\mathbb{Q} \subseteq \mathbb{Q}(\zeta) \subseteq K$$

Denote by φ Euler's totient function. We have $\varphi(k) = [\mathbb{Q}(\zeta) : \mathbb{Q}]$ so $\varphi(k) \mid n$. Thus, $\varphi(k)$ has to be odd. But this happens only when k is 1 or 2 so $\zeta = \pm 1$. We see that $N_{K/\mathbb{Q}}(-1) = (-1)^n = -1$ and the fact that $N_{K/\mathbb{Q}}(1) = 1$ is always true. Next, assume that n is even. It is then clear that both 1 and -1 have norm 1. We have that any embedding $\sigma : K \hookrightarrow \mathbb{C}$ must send ζ to a primitive kth root of unity in $\mathbb{C}$. Suppose that $k \geq 3$. In this case $\sigma(\zeta)$ is an imaginary number and therefore σ must be a complex embedding and so all the embeddings must be complex so $n = 2s$, where s is the number of pairs of complex conjugate embeddings. Thus,

$$N_{K/\mathbb{Q}}(\zeta) = \prod_{i=1}^{s} \sigma_i(\zeta)\overline{\sigma_i}(\zeta) = \prod_{i=1}^{s} |\sigma_i(\zeta)|^2 = 1$$

□

Theorem 4.3. *Let K be a number field of degree $n = 2s + r$ over the rationals and let c be a non-zero integer. Assume further that M is a full module with ring of coefficients $\mathfrak{D}$ and set $t = r + s - 1$. Then there exist independent units $\epsilon_1, ..., \epsilon_t$ in $\mathfrak{D}$, and a finite number of elements $\gamma_1, ..., \gamma_k$ in M_c such that every $\alpha \in M_c$ can be written uniquely as*

$$\alpha = \gamma_i \epsilon_1^{u_1} ... \epsilon_t^{u_t}$$

Proof. By theorem 4.1, we know that we can find a fundamental system of units, $\epsilon_1, ..., \epsilon_t$, where $t = r + s - 1$ such that every element $\epsilon \in \mathfrak{D}^*$ can be written uniquely as

$$\epsilon = \zeta \epsilon_1^{u_1} ... \epsilon_t^{u_t}$$

where the u_i are integers and $\zeta \in \mathfrak{D}$ is a root of unity. Let us find the subgroup of $\mathfrak{D}^*$ containing only elements of norm 1. Suppose first that n is odd. Then we know from lemma 4.4 that there are only two roots of unity, namely 1 and -1 and that $N(-1) = -1$. If need be, let us swap out all ϵ_i that have norm -1 with $-\epsilon_i$. This leaves us with a new system of fundamental units and they generate the subgroup of $\mathfrak{D}$ having norm 1. Suppose now n is even. Then by lemma 4.4 all the roots of unity have norm 1, so if it happens that all the ϵ_i have norm 1, then all units of $\mathfrak{D}$ will have norm 1. So suppose that not all the ϵ_i have norm 1. We can without loss of generality assume that $1 = N(\epsilon_1) = ... = N(\epsilon_q)$ and $-1 = N(\epsilon_{q+1}) = ... = N(\epsilon_t)$, where $q < t$. Define $\mu_i = \epsilon_i$ for $i \in \{1, ..., q\}$ and $\mu_i = \epsilon_i \epsilon_t$ for $i \in \{q+1, ..., t-1\}$. We now have a new fundamental system of units, namely $\mu_1, ..., \mu_{t-1}, \epsilon_t$ and only the last unit, ϵ_t, has norm -1. By setting $\mu_t = \epsilon_t^2$, all units of norm 1 in $\mathfrak{D}$ can now be written as $\zeta \mu_1^{u_1} ... \mu_t^{u_t}$, where ζ is a root of unity in $\mathfrak{D}$. We have now, for any n, found generators for the subgroup of $\mathfrak{D}$ consisting of elements of norm 1. Let now $[\gamma_1'], ..., [\gamma_l']$ be all the elements of $M_c/\sim$ and let $\alpha \in M_c$. By lemma 4.3, there is a unique γ_i' and a unit $\epsilon \in \mathfrak{D}$ of norm 1 so that $\alpha = \gamma_i' \epsilon$. For odd n the statement in the proposition now follows. When n is even, we can consider all products $\zeta \gamma_i'$, where $\zeta \in \mathfrak{D}$ is a root of unity. There are only finitely many of these products since there are only finitely many roots of unity by theorem 4.1. We can denote these products by $\gamma_1, ..., \gamma_k$. All of these belong to M_c since the roots of unity have norm 1 and the statement now follows for when n is even. □

Remark 4.2. *Observe that if $t = 0$, that is if K is $\mathbb{Q}$ or an imaginary quadratic field, then the above theorem says that there can only be finitely many elements in M_c. In this case, M is what is called a degenerate module, see definition 13.1.*

5 Fields with an absolute value

Definition 5.1 (Absolute value)**.** *Let K be a field. A function $|\cdot| : K \to \mathbb{R}$, is called an absolute value if it happens to satisfy the properties*

- $|x| \geq 0$ *for every $x \in K$. (Non-negativity)*
- $|x| = 0$ *if and only if $x = 0$. (Positive definiteness)*
- $|xy| = |x||y|$ *for every $x, y \in K$. (Multiplicativity)*
- $|x + y| \leq |x| + |y|$ *for every $x, y \in K$. (Triangle inequality)*

When the triangle inequality can be upgraded to the stronger condition, referred to as the ultrametric property,

$$|x + y| \leq \max\{|x|, |y|\} \quad \text{for every } x, y \in K,$$

the absolute value is said to be ***non-archimedian****. Otherwise it is called* ***archimedian****. If $|\cdot|$ is indeed an absolute value then the field K is called a valued field. A field with a non-archimedian absolute value is called a non-archimedian valued field and similarly a field with an archimedian absolute value is called an archimedian valued field.*

We could easily have made this definition more general by allowing K to be any ring. In that case, absolute values would only exists on domains. For if we had x, y both non-zero so that $xy = 0$ then $|x|, |y|$ would also be non-zero but $0 = |0| = |xy|$. But $|x||y| \neq 0$, so $|\cdot|$ cannot be an absolute value. The usual absolute value on $\mathbb{R}$, which we will denote by $|\cdot|_\infty$, is an example of an archimedian absolute value, so $(\mathbb{R}, |\cdot|)$ is an example of a valued field. The function sending everything to 1 except for 0 which is sent to 0 is also an example of an absolute value, and it is called the trivial absolute value. Note that for an arbitrary absolute value we have for all $x \in K$ that

- If $x^n = 1$ then $|x| = 1$.
- $|x| = |-x|$

In particular, $1 = |1| = |-1|$ and if K is a finite field then $|\cdot|$ is the trivial absolute value, since every non-zero element in such a field has finite order.

5.1 Topology induced by $|\cdot|$

Let K be a field. Once we have an absolute value on K, we can use it to define a metric, $d : K^2 \to \mathbb{R}$, by setting $d(x, y) = |x - y|$. As such, K can now be thought of as a metric space, and we therefore now have a topology on K, generated by d. We will also refer to this topology as the topology induced by $|\cdot|$. It might very well happen that two absolute values, $|\cdot|_1$ and $|\cdot|_2$, induce the same topology on K. In fact we have the following proposition, see also [8, Proposition 3.3]

Proposition 5.1. *Let $|\cdot|_1$ and $|\cdot|_2$ be two absolute values on K. The following are equivalent*

1. *$|\cdot|_1$ and $|\cdot|_2$ are equivalent*
2. *$|x|_1 = |x|_2^s$ for all $x \in K$ for some $s > 0$, $s \in \mathbb{R}$.*
3. *$|x|_1 < 1 \implies |x|_2 < 1$*

Absolute values being equivalent defines an equivalence relation on the set of absolute values on K, and the equivalence classes are called places of K. The non-archimedian absolute values induce a topology that is quite different compared to what one perhaps might expect. For example, triangles are isosceles - that is to say, any triangle will always have two sides of the same length.

Proposition 5.2. *Suppose $x, y \in K$ and that $|x| \neq |y|$. Then $|x + y| = \max\{|x|, |y|\}$. In particular, all triangles are isosceles.*

Proof. Assume without loss of generality that $|x| > |y|$. Then we have

$$|x + y| \leq \max\{|x|, |y|\} = |x|$$

Also, $|x| = |(x + y) - y| \leq \max\{|x + y|, |y|\} = |x + y|$, since $|x| > |y|$. Thus, $|x| = |x + y|$. For the last claim suppose that also $z \in K$ and consider three sides of a triangle $A = x - y$, $B = y - z$ and $C = x - z$. Suppose that two of these sides, say A and B have different length, that is, $|A| \neq |B|$. We clearly have that $A + B = C$ and so by what we have just shown we have that

$$|C| = |A + B| = \max\{|A|, |B|\}$$

But that means that C has the same length as A or B and we conclude that two sides in the triangle must have the same length. □

We also have the following lemma which does not hold for archimedian absolute values

Lemma 5.1. *Let $|\cdot|$ be a non-achimedian absolute value on a field K. Suppose (x_n) is a sequence in K. Define the sequence $y_n = x_{n+1} - x_n$. The following are equivalent*

1. *x_n is Cauchy.*
2. *$y_n \to 0$ for $n \to \infty$.*

Proof. A sequence being Cauchy clearly implies that y_n converges to 0. On the other hand let $N \in \mathbb{N}$ be so large that $|y_n| < \epsilon$ for all $n \geq N$. Suppose now $n > m > N$. We obtain

$$\begin{aligned} |x_n - x_m| &= |x_n - x_{n-1} + x_{n-1} - ... + x_{m+1} - x_m| \\ &= |y_{n-1} + ... + y_m| \leq \max\{|y_{n-1}|, ..., |y_m|\} < \epsilon \end{aligned}$$

□

This lemma can be really helpful in situations where the sequence (x_n) has the form $x_n = \sum_{i=1}^{n} z_n$, since in this case we can determine if (x_n) is Cauchy simply by checking if z_n converges to 0. This is certainly not something we can do in the archimedian setting. For example if $z_n = \frac{1}{n}$ then x_n is the harmonic series and we know that it diverges with respect to $|\cdot|_\infty$.

Proposition 5.3. *An absolute value $|\cdot|$ on a field K is uniformly continuous.*

Proof. Let $\epsilon > 0$ and pick $\delta = \epsilon$. For any $x, y \in K$ so that $|x - y| < \delta$ we have

$$||x| - |y||_\infty \leq |x - y| < \delta = \epsilon$$

where we have used the reverse triangle inequality. □

Uniformly continuous functions have the nice property that they preserve Cauchy sequences. This implies that if we have some Cauchy sequence, (x_n) in K, then $(|x_n|)$ is a Cauchy sequence in $\mathbb{R}$. We will use this fact later on.

Definition 5.2. *Let K, L be valued fields and suppose that $\phi : K \to L$ is a function. We say that ϕ preserves the absolute value if $|\phi(x)|_L = |x|_K$.*

It is worth pointing out that if ϕ is as above and is also a homomorphism, then ϕ is actually uniformly continuous. This is simply because we have for all $x, y \in K$ that

$$|\phi(x) - \phi(y)|_F = |\phi(x - y)|_F = |x - y|_K$$

5.2 Valuations

We now turn to a notion that is closely related to non-archimedian absolute values - namely valuations.

Definition 5.3 (Valuation)**.** *A valuation on a field K is a function $v : K \to \mathbb{R} \cup \{\infty\}$ with the following properties*

1. *$v(x) = \infty$ if and only if $x = 0$.*
2. *$v(xy) = v(x) + v(y)$ for every $x, y \in K$.*
3. *$v(x + y) \geq \min\{v(x), v(y)\}$ for every $x, y \in K$.*

The symbol ∞ is just an element so that for all $x \in \mathbb{R}$ we have $\infty > x$, $\infty + x = x + \infty = \infty$ and $\infty + \infty = \infty$.

We stipulated above that valuations and non-archimedian absolute values have something to do with each other. Let us see how. Suppose we have access to a valuation, v on a field K. Then for any $q > 1$ we get what we will call a corresponding absolute value on K by setting

$$|x| = q^{-v(x)}$$

This absolute value is in fact non-archimedian because of the third property in definition 5.3 and all absolute values corresponding to v are equivalent. Since we can solve for $v(x)$ in the above equation we see that we can also obtain a valuation if we are given a non-archimedian absolute value and we will similarly call this a corresponding non-archimedian absolute value. We will say that two valuations v_1 and v_2 on K are equivalent if $v_1 = sv_2$ for some $s > 0$. All valuations corresponding to a non-archimedian absolute value will be equivalent in this sense. When working with valuations it is useful to have the following in mind. Assume that K is a valued field with non-archimedian absolute value $|\cdot|$ and v a corresponding valuation. If $a \in K$ and (x_n) is a sequence in K, then $x_n \to a$ with respect to $|\cdot|$ if and only if $v(x_n - a) \to \infty$. We will therefore say that (x_n) converges to a with respect to v if $v(x_n - a) \to \infty$.

Definition 5.4. *Let v be a valuation on a field K. The group $\Gamma_v = v(K^*)$ is called the value group of v and if there is $s > 0$ so that $\Gamma_v = s\mathbb{Z}$ we say that v is discrete. If $s = 1$ then v is called normalized and in this case an element $\pi \in K$ so that $v(\pi) = 1$ is called a prime element.*

We have in fact already seen an example of a field with valuation.

Example 5.1. *Recall from section 3 the field $K((x))$ and the function*

$$\omega : K((x)) \to \mathbb{Z} \cup \{\infty\},$$

that we defined on it. This function is in fact a valuation on $K((x))$ and as $\omega(x) = 1$, it is discrete and normalized. Notice that

$$\{f \in K((x)) \mid \omega(f) \geq 0\},$$

is a local ring since it equals $K[[x]]$ and this ring has unique maximal ideal (x) which can be given as $\{f \in K((x)) \mid \omega(f) > 0\}$

In the example above we saw that we could produce a local ring using the valuation on $K((x))$. This was no coincidence as we will see in the next proposition. Let K be a field with valuation v and let $|\cdot|$ be a corresponding non-archimedian absolute value.

Proposition 5.4. *The ring*

$$O := \{x \in K \mid v(x) \geq 0\} = \{x \in K \mid |x| \leq 1\}$$

is called the valuation ring with respect to v (or we might say with respect to $|\cdot|$), is integrally closed and is a local ring with unique maximal ideal

$$\mathfrak{p} = \{x \in K \mid v(x) > 0\} = \{x \in K \mid |x| < 1\}$$

and we have

$$O^* = \{x \in K \mid v(x) = 0\} = \{x \in K \mid |x| = 1\}$$

The quotient $O/\mathfrak{p}$ is called the residue field of K. If v is discrete and normalized and $\pi \in O$ is a prime element then we have

- $K^* = \{\pi^k \mid k \in \mathbb{Z}\} \times O^*$
- *O is a PID and its non-zero ideals are given by $\mathfrak{p}^n = \{x \in K : v(x) \geq n\}$ for $n \geq 0$, and $\mathfrak{p} = \pi O$.*
- *$O/\mathfrak{p} \cong \mathfrak{p}^n/\mathfrak{p}^{n+1}$ for $n \geq 0$*

Proof. It is clear that O is a ring and that $\mathfrak{p}$ is a proper ideal of O. The elements $x \in O$ for which $v(x) = 0$ are certainly units, which is another way of saying that $O \setminus \mathfrak{p} \subseteq O^*$. This implies that O is a local ring with unique maximal ideal $\mathfrak{p}$ and it follows that $O^* = O \setminus \mathfrak{p}$. O is integrally closed since if $x \in K^*$ is integral over O then we have an equation

$$x^n + a_{n-1}x^{n-1} + ... + a_0 = 0,$$

where the a_i all belong to O and if $x \notin O$ then $v(x) < 0$ so $v(\frac{1}{x}) = -v(x) > 0$ so $x^{-1} \in O$ which means that we can solve for x in the equation above, meaning

that $x \in O$, which is a contradiction. Assume now that v is discrete and normalized and that $x \in K^*$. Let $m = v(x)$. We have $m \in \mathbb{Z}$, since $v(K^*) = \mathbb{Z}$ by assumption, and so $v(x\pi^{-m}) = v(x) - mv(\pi) = 0$ meaning that $x\pi^{-m} \in O^*$. Thus $x\pi^{-m} = u$ for some $u \in O^*$ which means that $K^* = (\pi) \times O^*$. Take now an ideal $I \neq 0$ of O. Since v is discrete we can choose $x \in I$, where $x \neq 0$ so that $v(x) = \min_{z \in I} v(z)$. Let $n := v(x)$. By what we have already shown there is $u \in O^*$ so that $x = u\pi^n$. If now $z \in \pi^n O$ we have that $z = xy$ for some $y \in O$ and since $x \in I$ and I is an ideal also $z \in I$. Hence $\pi^n O \subseteq I$. For the other inclusion, take $z \in I$ and write $z = u\pi^m$ for some $u \in O^*$. We have $v(z) = m \geq n$ so we can write $z = (u\pi^{m-n})\pi^n$ which is an element of $\pi^n O$. Thus, $I = \pi^n O$ and so $\mathfrak{p} = \pi O$. All the non zero ideals of O are therefore powers of $\mathfrak{p}$ and so O is a PID. Consider the homomorphism

$$\begin{aligned} \phi : \mathfrak{p}^n &\to O/\mathfrak{p} \\ a\pi^n &\mapsto \overline{a}, \text{ where } a \in O \end{aligned}$$

This map is well-defined since $a\pi^n = a'\pi^n$ implies $a = a'$ meaning that $\overline{a} = \overline{a}'$. Now $\phi(a\pi^n) = 0$ if and only if $a \in \mathfrak{p}$ if and only if $a\pi^n \in \mathfrak{p}^{n+1}$ so $\ker \phi = \mathfrak{p}^{n+1}$. This map is also clearly surjective and so we get the desired isomorphism. □

It would be nice with some more examples of fields with valuation. The unique factorization of prime ideals in Dedekind domains allows for a way of constructing a valuation on the field of fractions of these rings.

Proposition 5.5. *Let O be a Dedekind domain and K its field of fractions. For any prime $\mathfrak{p}$ of O there is a function, $v_\mathfrak{p} : O \to \mathbb{R} \cup \{\infty\}$ on O so that $v_\mathfrak{p}(0) = \infty$ and that*

$$\forall x \in O \setminus \{0\} : v_\mathfrak{p}(x) = m \iff (x) = \mathfrak{p}^m A, \ A \text{ an ideal of } O \text{ and } \mathfrak{p} \nmid A. \quad (5.1)$$

The function v extends to a discrete valuation on K, which we will also call $v_\mathfrak{p}$, which for $\frac{\alpha}{\beta} \in K$, $\alpha, \beta \in O$ and $\beta \neq 0$ is given by

$$v_\mathfrak{p}(\frac{\alpha}{\beta}) = v_\mathfrak{p}(\alpha) - v_\mathfrak{p}(\beta)$$

Proof. The function $v_\mathfrak{p}$ on O with the mentioned property is well defined since we have unique factorization of prime ideals in Dedekind domains. Let us now see that the function $v_\mathfrak{p}$ we have defined on K is actually well defined. If $\frac{\alpha}{\beta} = \frac{\alpha'}{\beta'}$ then $\alpha\beta' = \alpha'\beta \in O$ and applying $v_\mathfrak{p}$ on both sides and using that it takes multiplication to addition we obtain

$$v_\mathfrak{p}(\alpha) - v_\mathfrak{p}(\beta) = v_\mathfrak{p}(\alpha') - v_\mathfrak{p}(\beta')$$

which shows that it is well defined. If we can show that all the properties in definition 5.3 are satisfied for the restriction of $v_\mathfrak{p}$ to O, we are done. It is immediate that $v_\mathfrak{p}(x) = \infty$ if and only if $x = 0$. Let now $x, y \in O$ and suppose

both x and y are non-zero. We then have $v_{\mathfrak{p}}(x) = m$ and $v_{\mathfrak{p}}(y) = n$ and we can write

$$(x) = \mathfrak{p}^m A, (y) = \mathfrak{p}^n B$$

Thus, $(xy) = (x)(y) = \mathfrak{p}^{m+n}AB$ and $\mathfrak{p} \nmid AB$, so $v_{\mathfrak{p}}(xy) = m + n$. For the last property, let $d = \min\{m, n\}$. Then

$$(x + y) \subseteq (x) + (y) = \mathfrak{p}^d(\mathfrak{p}^{m-d}A + \mathfrak{p}^{n-d}B)$$

From this we see that $\mathfrak{p}^d \mid (x+y)$ and so $d \leq v_{\mathfrak{p}}(x+y)$. Note that if either x or y are 0, then these two properties follow immediately. We conclude that $v_{\mathfrak{p}}$ is a valuation on K. It is discrete and normalized since it maps $O \setminus \{0\}$ into $\mathbb{N}_0$ so it maps K^* into $\mathbb{Z}$. □

Let's see a very concrete and also important instance of the above.

Example 5.2. *Choose a prime $p \in \mathbb{Z}$. By proposition 5.5 we get a valuation $v_{(p)}$ on $\mathbb{Q}$, and we shall denote this valuation simply as v_p. Given $x \in \mathbb{Q}$, $v_p(x)$ tells us "how divisible" x is by p. For example $v_p(p^2) = 2$ and $v_p(\frac{1}{p^3}) = -3$. The valuation ring with respect to v_p is turns out to be the localization of $\mathbb{Q}$ at p. This follows from more general consideration, see proposition 7.4.*

6 Completions of fields

Let $(K, |\cdot|)$ be a valued field. A Cauchy sequence in K is a sequence (x_n) in K so that for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ so that if $m, n \geq N$ then

$$|x_n - x_m| < \epsilon$$

A null sequence is a sequence in K converging to 0.

Definition 6.1. *We say that a metric space X with metric d is complete if every Cauchy sequence in X converges to an element in X with respect to d and a subset $A \subseteq X$ is said to be complete if it is complete as a metric space. A valued field $(K, |\cdot|)$ is called a complete field if it is complete as a metric space with respect to the metric $d(x, y) = |x - y|$.*

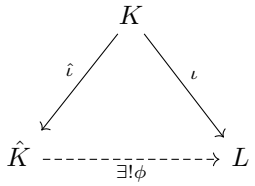
Let K be a field with an absolute value, $|\cdot|$. If this field is not complete could we then find a smallest field extension L of K, extending the absolute value on K, which is complete with resect to this new absolute value? The answer is yes and in this case L is called the completion of K. We'll properly define what the completion is but first, let us recall the following characterization of closed sets in a metric space.

Definition 6.2. *Let X be a metric space and $A \subseteq X$. The set A is closed if it contains all of its limit points. That is, whenever there is a sequence in A converging to some point $a \in X$ with respect to the metric on X then $a \in A$.*

Proposition 6.1. *Suppose X is a complete metric space. Then a subset $A \subseteq X$ is closed if and only if A is complete.*

Proof. Suppose first that A is closed and let (x_n) be any Cauchy sequence in A. Then this sequence converges to some point $a \in X$ since X is complete. But as A is closed in X that means by definition that $a \in A$, meaning that A is complete. On the other hand, assume that A is complete and let (x_n) be a sequence in A converging to some element $a \in X$. Thus (x_n) is in fact a Cauchy sequence in A and as A is complete it follows that $a \in A$. □

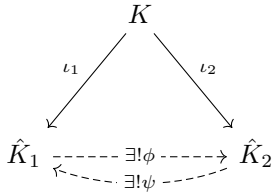
Definition 6.3 (Completion)**.** *Suppose $(K, |\cdot|)$ is a valued field. A completion of K is a complete valued field $\hat{K}$ together with a homomorphism $\hat{\iota} : K \to \hat{K}$ preserving the absolute value, satisfying the following universal property; If L is any other complete valued field and $\iota : K \to L$ is a homomorphism preserving the absolute value, then there is a unique homomorphism $\phi : \hat{K} \to L$, which preserves the absolute value, making the following diagram commute*



Note that the above definition doesn't say that the completion exists. Luckily it does, but it certainly requires some justifying which we will provide soon. Also, the map ϕ in the diagram above is actually uniformly continuous since it preserves absolute values and is a homomorphism. Note also that we talk about 'a' completion, insinuating that there are multiple different ones. This is true, but it turns out that they are all unique up to unique isomorphism. This is a quite strong statement - much stronger than just being unique up to isomorphism.

Proposition 6.2. *Suppose K is a valued field and that $(\hat{K}_1, \iota_1)$ and $(\hat{K}_2, \iota_2)$ are completions of K. Then there is a unique absolute value preserving isomorphism, $\phi : \hat{K}_1 \to \hat{K}_2$ so that $\phi \circ \iota_1 = \iota_2$. In other words, the completion of K is unique up to unique isomorphism.*

Proof. Let us invoke the universal property on both $\hat{K}_1$ and $\hat{K}_2$ to obtain the diagram



By definition both ϕ and ψ preserve the absolute value. Writing out the two commutativity conditions we get

$$\iota_1 = \psi \circ \iota_2$$
$$\iota_2 = \phi \circ \iota_1$$

Combining these we obtain

$$\iota_1 = \psi \circ (\phi \circ \iota_1) = (\psi \circ \phi) \circ \iota_1$$

So $\psi \circ \phi$ has to be the identity on $\hat{K}_1$ and similarly $\phi \circ \psi$ is the identity on $\hat{K}_2$, hence ϕ is an isomorphism and it is the only one satisfying the equation $\iota_2 = \phi \circ \iota_1$. □

Because of this proposition we will more often than not say "the completion" rather than "a completion". Let us now show existence of the completion. This process follows a similar pattern as when showing existence of, for example, the tensor product of two modules - construct a huge object that has some of the properties we want, then sculpt it into the desired object by modding out by a suitable relation. One could show that the completion exists for any absolute value on K, but instead we will show it only for non-archimedian absolute values, as this will suffice for our purposes. At the end of this section we will mention how to deal with the general case. So assume from now on that K is a field with non-archimedian absolute value $|\cdot|$.

Proposition 6.3. *If (x_n) is a Cauchy sequence with respect to $|\cdot|$ and is not a null sequence, then there exists an N so that $|x_n| = |x_m|$ for all $m, n \geq N$. In other words, the sequence must eventually become constant.*

Proof. Since (x_n) does not converge to 0, there exists an $N_1 \in \mathbb{N}$ and $c > 0$ so that $|x_n| \geq c > 0$ for all $n \geq N_1$. As the sequence is Cauchy, we can find $N_2 \in \mathbb{N}$ so that $|x_n - x_m| < c$ for all $n, m \geq N_2$. Set now $N := \max\{N_1, N_2\}$ and suppose $n, m \geq N$. Then

$$|x_n - x_m| < c \leq \min\{|x_n|, |x_m|\}$$

This means that both $|x_n|$ and $|x_m|$ are different from $|x_n - x_m|$ and hence $|x_n| = |x_m|$ since triangles are isosceles by proposition 5.2. □

Theorem 6.1. *Let R be the set of all Cauchy sequences of K wit respect to $|\cdot|$. Then R is a ring and the set $\mathfrak{m}$ of all null sequences of R is a maximal ideal.*

Proof. The addition and multiplication operations on R are defined element wise and it is easily seen that the Cauchy property is preserved under these operations. Rules such as the distributive law are inherited directly from the ring structure on K, and so R is a ring. The set $\mathfrak{m}$ is non-empty, as it most certainly contains the constant sequence $(0, 0, 0, ...)$. It is also contained in R

since every convergent sequence is a Cauchy sequence so in particular the null sequences are Cauchy. Furthermore the difference of two null sequences is again a null sequence and the product of any sequence by a null sequence is also a null sequence. Thus, $\mathfrak{m}$ is an ideal of R. Let us now show that it is a maximal ideal. Take a sequence $(x_n) \in R$ which is not also in $\mathfrak{m}$. Such an element exists since R contains, for example, the constant sequence $1 = (1, 1, 1, ...)$ which is clearly not a null sequence. We wish to show that the ideal $J = (x_n) + \mathfrak{m}$ contains 1, i.e, that this ideal is in fact the entire ring R. By proposition 6.3 the sequence $(|x_n|)$ is eventually constant as it is not a null sequence. Hence there is an N so that if $n \geq N$ then $|x_n| \geq c > 0$. Thus we can define

$$y_n = \begin{cases} \frac{1}{x_n} & \text{if } n \geq N \\ 0 & \text{if } n < N \end{cases}$$

Let now $\epsilon > 0$ be given and choose M so that $|x_n - x_m| < c^2\epsilon$ whenever $n, m \geq M$. For $n, m \geq \max\{N, M\}$ we now have

$$|y_n - y_m| = |\frac{1}{x_n} - \frac{1}{x_m}| = |\frac{x_m - x_n}{x_n x_m}| \leq \frac{1}{c^2}|x_m - x_n| \leq \frac{c^2\epsilon}{c^2} = \epsilon$$

But that means that (y_n) is a Cauchy sequence, meaning that it belongs to R and so $(x_n)(y_n)$ is in J. Denote by 1 the identity in R, i.e, the constant sequence $(1, 1, 1, ...)$. Observe that

$$x_n y_n = \begin{cases} 1 \text{ if } n \geq N \\ 0 \text{ if } n < N \end{cases}$$

Therefore $1 - (x_n)(y_n)$ is a null sequence, meaning that is belongs to J. But then $1 \in J$ and so $\mathfrak{m}$ is a maximal ideal. □

From this it follows that $\hat{K} = R/\mathfrak{m}$ is a field. Define now

$$\hat{\iota}(x) = (x, x, x, ...) + \mathfrak{m} \in \hat{K}$$

for $x \in K$. This map is a homomorphism of fields and it is certainly not the zero map, which implies that it is injective. We can therefore think of K as a subfield of $\hat{K}$. Recall the following result from topology.

Proposition 6.4. *Suppose X and Y are topological spaces and that $f : A \to Y$ is map defined on a dense subset A of X. Then f extends uniquely to a continuous map $\tilde{f} : X \to Y$.*

We need the following result

Lemma 6.1. *Assume (x_n) and (y_n) are sequences in K. If (x_n) is Cauchy and $\lim_{n\to\infty} |x_n - y_n| = 0$ then (y_n) is also Cauchy. If $\lim_{n\to\infty} x_n = a$ then $\lim_{n\to\infty} y_n = a$.*

Proof. Observe that

$$|y_n - y_m| = |x_n - y_n + x_m - y_m + x_n - x_m|$$

A three fold application of the triangle inequality on the right hand side tells us that (y_n) is Cauchy. For the last part, use the triangle inequality on the right hand side of

$$|y_n - a| = |x_n - y_n + x_n - a|$$

and take the limit. □

Theorem 6.2 (Existence of completion)**.** *Suppose $(K, |\cdot|)$ is a valued field and let R the set of all Cauchy sequences of K with respect to $|\cdot|$ and $\mathfrak{m} \subseteq R$ the maximal ideal consitsting of all null sequences. The function $f : R \to \mathbb{R}$, given by*

$$(x_n) \mapsto \lim_{n\to\infty} |x_n|$$

descends to the function, $|\cdot|' : \hat{K} \to \mathbb{R}$ given by

$$|\overline{(x_n)}|' = f((x_n))$$

We have the following properties

- *$(\hat{K}, |\cdot|')$ is a valued field*
- *The homomorphism $\hat{\imath}$ preserves the absolute value.*
- *K is dense in $\hat{K}$*
- *$\hat{K}$ is complete with respect to $|\cdot|'$*

Furthermore, the pair $(\hat{K}, \hat{\imath})$ is the completion of K.

Proof. Let us first show that f is well defined. Assume $(x_n) \in R$. If (x_n) is a null sequence there is nothing to prove, so suppose it is not. Then by proposition 6.3 the sequence $(|x_n|)$ is eventually constant hence converges making f well defined. Assume now x and y are in $\hat{K}$ with $x = y$ and suppose they are represented by the Cauchy sequences (x_n) and (y_n) in K, respectively. By the reverse triangle inequality we have

$$||x_n| - |y_n||_\infty \leq |x_n - y_n|$$

But $(x_n) - (y_n)$ is a null sequence so the right hand side converges to 0. This means that $\lim_{n\to\infty}(|x_n| - |y_n|) = 0$ and therefore

$$\lim_{n\to\infty} |x_n| = \lim_{n\to\infty} |y_n|$$

This shows that the map f descends to the quotient as desired. Let us move on to justifying that $|\cdot|'$ is actually an absolute value on $\hat{K}$. This is essentially inherited from the absolute value on K. So let us just show that $|\cdot|'$ positive definite. Suppose $x = \overline{(x_n)} \in \hat{K}$. Assume first that $|x|' = 0$. We then have

$$0 = |x|' = \lim_{n\to\infty} |x_n|$$

But this just says that (x_n) is a null sequence so $x = 0$. On the other hand, if $x = 0$ then $(x_n) \in \mathfrak{m}$, i.e., (x_n) is a null sequence so $|x|' = \lim_{n\to\infty} |x_n| = 0$. It is also easily seen that $|\cdot|'$ extends $|\cdot|$. For if $x \in K$ then

$$|x| = \lim_{n\to\infty} |x| = f(\hat{\iota}(x)) = |\overline{\hat{\iota}(x)}|'$$

Let us now show that K is dense in $\hat{K}$. What we mean by this is that the image of K under the $\hat{\iota}$ is dense in $\hat{K}$. Let $\epsilon > 0$, and suppose $\lambda \in \hat{K}$ and consider the open ball $B(\lambda, \epsilon)$. Let (x_n) be a Cauchy sequence representing λ and find $N \in \mathbb{N}$ so that $|x_n - x_m| < \epsilon/2$ for all $m, n \geq N$. Define now the constant sequence $y = \hat{\iota}(x_N)$. Using the definition of the absolute value on $\hat{K}$ we get

$$|\lambda - y|' = \lim_{n\to\infty} |x_n - x_N| \leq \epsilon/2 < \epsilon$$

This shows that $y \in B(\lambda, \epsilon)$, meaning that K is dense in $\hat{K}$ as desired. Now we are ready to show that $\hat{K}$ is complete with respect to $|\cdot|'$. To do this, suppose that we have a Cauchy sequence (λ_n) in $\hat{K}$. As we have just seen, K is dense in $\hat{K}$. So for every λ_i, we can find an element $y_i \in K$ so that $|\lambda_i - \hat{\iota}(y_i)|' < \frac{1}{i}$. This way we get a sequence $(\hat{\iota}(y_n))$ in $\hat{K}$ and we observe that

$$\lim_{n\to\infty} |\lambda_n - \hat{\iota}(y_n)|' = 0$$

So by lemma 6.1 $(\hat{\iota}(y_n))$ is a Cauchy sequence in $\hat{K}$. But then also the sequence (y_n) in K is Cauchy since $|\cdot|'$ extends the absolute value on K. But that means that (y_n) represents an element in $\hat{K}$. Call this element λ and let us show that $\lim_{n\to\infty} \lambda_n = \lambda$. To do this it suffices, also by lemma 6.1, to show that $\lim_{n\to\infty} \hat{\iota}(y_n) = \lambda$. As (y_n) is Cauchy, we can find N so that if $n, m \geq N$ then $|y_n - y_m| < \epsilon/2$. For any $n \geq N$ we now have

$$|\hat{\iota}(y_n) - \lambda|' = \lim_{m\to\infty} |y_n - y_m| \leq \epsilon/2 < \epsilon$$

This shows that $\lim_{n\to\infty} \hat{\iota}(y_n) = \lambda$ and as we mentioned this means that $\lim_{n\to\infty} \lambda_n = \lambda$ which shows that $\hat{K}$ is complete with respect to $|\cdot|'$. Moving on to the last statement, let us see why $(\hat{K}, \hat{\iota})$ is the completion of K. So let L be a complete valued field and $\iota : K \to L$ a homomorphism preserving the absolute value. Let us define a function $\phi : \hat{K} \to L$; Pick an element $\lambda \in \hat{K}$. Because $\hat{\iota}(K)$ is dense in $\hat{K}$, there is a Cauchy sequence, (y_n), in $\hat{\iota}(K)$ that converges to λ. Hence there is a sequence (x_n) in K so that $(\hat{\iota}(x_n)) = (y_n)$ and as $\hat{\iota}$ preserves the absolute value the sequence (x_n) is also Cauchy. As ι preserves the absolute value by assumption, we now have a Cauchy sequence $(\iota(x_n))$ in L and the fact that L is complete ensures that it converges to an element, $\lambda' := \lim_{n\to\infty} \iota(x_n)$. By uniqueness of limits we therefore obtain a well defined map $\phi : \hat{K} \to L$ mapping the element λ to λ'. It preserves absolute

values since we have for $x \in \hat{K}$ that

$$\begin{aligned}
|\phi(x)|_L &= |\lim_{n\to\infty} \iota(x_n)|_L \\
&= \lim_{n\to\infty} |\iota(x_n)|_L \\
&= \lim_{n\to\infty} |x_n|_K \\
&= \lim_{n\to\infty} |\hat{\iota}(x_n)|_{\hat{K}} \\
&= |\lim_{n\to\infty} \hat{\iota}(x_n)|_{\hat{K}} = |x|_{\hat{K}}
\end{aligned}$$

We see that ϕ is also a homomorphism because the limit of the sum of two convergent sequences is the sum of the limits of the sequences and similarly for the product. Combining this with the fact that ϕ preserves absolute values, we see that it is uniformly continuous so in particular continuous and it satisfies $\phi \circ \hat{\iota} = \iota$. This means that it is defined on the dense subset $\hat{\iota}(K)$ which makes it uniquely determined by proposition 6.4. Thus the pair $(\hat{K}, \hat{\iota})$ is the completion of K. □

As it was mentioned we have only proven the existence of the completion when $|\cdot|$ is non-archimedian. In the case where $|\cdot|$ is an arbitrary absolute value on K, one can still prove the existence of the completion. One approach to do that is to first show that $\mathbb{R}$ is the completion of $\mathbb{Q}$ with respect to $|\cdot|_\infty$. Once that has been established then for any Cauchy sequence (x_n) in K with respect to $|\cdot|$ we get a Cauchy sequence $(|x_n|_\infty)$ in $\mathbb{R}$ because $|\cdot|_\infty$ is uniformly continuous. Since $\mathbb{R}$ is complete this sequence converges to an element in $\mathbb{R}$. We could then apply this kind of reasoning in theorems 6.1 and 6.2 instead of appealing to proposition 6.3.

7 Complete fields with valuation

Let K be a field with valuation v. We will say that K is complete with respect to v if K is complete with respect to an absolute value corresponding to v. Since valuations and non-archimedian absolute values have so much in common, it would make sense to be able to talk about the completion of K with respect to v. This is something we can definitely do: Let $|\cdot|$ be an absolute value corresponding to v. We can consider the completion, $\hat{K}$, with respect to $|\cdot|$ and as we showed in theorem 6.2, $|\cdot|$ extends to an absolute value on $\hat{K}$, also denoted $|\cdot|$, and is given by sending a representative (x_n) of an element $x \in \hat{K}$ to $\lim_{n\to\infty} |x_n|$. Let us very similarly define the function $\hat{v} : \hat{K} \to \mathbb{R}$ by sending the representative (x_n) to $\lim_{x\to\infty} v(x_n)$. We see that $\hat{v}$ is a valuation, that it extends v and that that $\hat{K}$ corresponds to the extended absolute value on $\hat{K}$ meaning that $\hat{K}$ is complete with respect to $\hat{v}$. Furthermore we see that if $a \in \hat{K}$, then a sequence (x_n) in $\hat{K}$ converges to a with respect to the extended absolute value on $\hat{K}$ if and only if (x_n) converges to a with respect to $\hat{v}$.

Definition 7.1. *The completion of K with respect to v is the field $\hat{K}$ together with the valuation $\hat{v}$.*

We remark the following consequence of lemma 5.1. Assume (x_n) is a sequence in $\hat{K}$. Saying that $\hat{v}(x_{n+1} - x_n) \to \infty$ amount to saying that (x_n) is Cauchy and because $\hat{K}$ is complete every Cauchy sequence converges to an element in $\hat{K}$. Thus a sum $\sum_{n=0}^{\infty} z^n$ converges in $\hat{K}$ if and only if $\hat{v}(z_n) \to \infty$. From the definition of $\hat{v}$ we now see that

Proposition 7.1. *$\Gamma_v = \Gamma_{\hat{v}}$ and $\hat{v}$ is discrete and normalized if v is.*

We also have

Proposition 7.2. *Let O_v and $O_{\hat{v}}$ be the valuation rings with respect to v and $\hat{v}$ respectively and let $\mathfrak{p}_v$ and $\mathfrak{p}_{\hat{v}}$ be their corresponding maximal ideals. If v is discrete then for all $n \geq 1$ we have*

$$O_{\hat{v}}/\mathfrak{p}_{\hat{v}}^n \cong O_v/\mathfrak{p}_v^n$$

Proof. Consider the map

$$O_v \to O_{\hat{v}}/\mathfrak{p}_{\hat{v}}^n,$$

defined by $a \mapsto a \pmod{\mathfrak{p}_{\hat{v}}^n}$. Note that we can write $\mathfrak{p}_v = \pi O_v$ and $\mathfrak{p}_{\hat{v}} = \pi O_{\hat{v}}$ for some $\pi \in O_v$ with $v(\pi) = 1$. This implies that the map above has kernel $\pi^n O_v = \mathfrak{p}_v^n$. Let us show that the map is surjective, assume $x \in O_{\hat{v}}$. As O_v is dense in $O_{\hat{v}}$ we can find $a \in O_v$ so that $\hat{v}(x - a) \geq n$ which means that $x - a \in \mathfrak{p}_{\hat{v}}^n$, i.e. $x \equiv a \pmod{\mathfrak{p}_{\hat{v}}^n}$. Hence the map is surjective and so we get the desired isomorphism. □

7.1 The $\mathfrak{p}$-adic numbers

In example 5.2 we saw how we got a valuation, v_p, on $\mathbb{Q}$ and now that we have the tools we can consider completion with respect to it. This field is known as the so called p-adic numbers. More generally we saw in proposition 5.5 how we can obtain a valuation on the field of fractions of a Dedekind domain. In fact proposition 5.5 can be applied to obtain a valuation on any number field and so we can think about the completion with respect to it as well. First we have the following lemma.

Lemma 7.1. *Let K be a number field. The field of fractions of the ring of integers of K, O_K, is equal to K. In other words, every element of K can be written as $\frac{\alpha}{\beta}$ for some $\alpha, \beta \in O_K$, where $\beta \neq 0$.*

Proof. Let F be the field of fractions of O_K. Then $F \subseteq K$ as F is the smallest field containing O_K. For the other inclusion, suppose for contradiction that $[K : F] > 1$. Take then an element, x, which is in K but not in F and find a non-zero integer d so that $d\alpha \in O_K$. But then $d\alpha \in F$ so $\alpha \in F$, contrary to our initial assumption. Hence $[K : F] = 1$ and so $F = K$. □

Let K be any number field and let O be its ring of integers and take any prime ideal $\mathfrak{p}$ of O. From lemma 7.1 we know that O has field of fractions K and since O is a Dedekind domain we can apply proposition 5.5 to obtain a valuation $v_\mathfrak{p}$ on K. We obtain $K_\mathfrak{p}$, the completion of K with respect to $v_\mathfrak{p}$, and $v_\mathfrak{p}$ extends to a discrete valuation, also denoted $v_\mathfrak{p}$ on $K_\mathfrak{p}$.

Definition 7.2. *The elements of $K_\mathfrak{p}$ are called $\mathfrak{p}$-adic numbers. Its valuation ring, $O_\mathfrak{p}$, is called the ring of $\mathfrak{p}$-adic integers and the valuation $v_\mathfrak{p}$ is called the $\mathfrak{p}$-adic valuation. In the special case where $O = \mathbb{Z}$, $\mathfrak{p} = (p)$ and $K = \mathbb{Q}$ we will define $\mathbb{Q}_p := K_\mathfrak{p}$, $\mathbb{Z}_p := O_\mathfrak{p}$ and $v_p := v_\mathfrak{p}$ and these are called the p-adic numbers, the p-adic integers and the p-adic valuation respectively.*

7.2 Valuation ring, revisited

In the following, let K be a field which is complete with respect to a valuation, v, and let O be the valuation ring with respect to v. Inside O we have the unique maximal ideal, $\mathfrak{p}$. We shall see that we in this case get an alternative description of O. For each n we have the natural projection

$$O \xrightarrow{\pi_n} O/\mathfrak{p}^n.$$

We have the inclusions $\mathfrak{p} \supset \mathfrak{p}^2 \supset ...$, which implies that every element in $\mathfrak{p}^{n+1}$ is sent to zero under π_n, giving us a commutative diagram

$$\begin{array}{ccc} & O & \\ {\scriptstyle \pi_n}\swarrow & & \searrow{\scriptstyle \pi_{n+1}} \\ O/\mathfrak{p}^n & \xleftarrow{\lambda_n} & O/\mathfrak{p}^{n+1} \end{array}$$

Thus, we get a sequence of maps

$$O/\mathfrak{p} \xleftarrow{\lambda_1} O/\mathfrak{p}^2 \xleftarrow{\lambda_2} O/\mathfrak{p}^3 \xleftarrow{\lambda_3} ...$$

We now define

$$\lim_{\leftarrow} O/\mathfrak{p}^n = \{(x_n) \in \prod_{n=1}^{\infty} O/\mathfrak{p}^n : \lambda_n(x_{n+1}) = x_n\}$$

which is seen to be a subring of $\prod_{n=1}^{\infty} O/\mathfrak{p}^n$ and it is called the projective limit of the system of rings, $O/\mathfrak{p}^n$. Because of the commutative diagram above we obtain a ring map

$$O \to \lim_{\leftarrow} O/\mathfrak{p}^n$$

which just sends an element in O to its respective equivalence class. For what follows we refer to [8, p. 128] for details. We equip each of the $O/\mathfrak{p}^n$ with the discrete topology. This way $\lim_{\leftarrow} O/\mathfrak{p}^n$ becomes a closed subset of the product topology on $\prod_{n=1}^{\infty} O/\mathfrak{p}^n$. We have the following theorem.

Theorem 7.1. *The map $O \to \lim_{\leftarrow} O/\mathfrak{p}^n$ is an isomorphism of rings and a homeomorphism.*

7.3 Hensel's Lemma

We keep the same setup as in section 7.2 and let $|\cdot|$ be a non-archimedian absolute value corresponding to v. Denote by $x \mapsto \overline{x}$ the canonical map $O[x] \twoheadrightarrow O/\mathfrak{p}[x]$. We will say that a polynomial $f \in O[x]$ is primitive if $\overline{f} \neq 0$.

Lemma 7.2 (Hensel's Lemma)**.** *Suppose that $f \in O[x]$ is primitive and that there are relatively prime polynomials $g_0, h_0 \in O/\mathfrak{p}[x]$ so that $\overline{f} = g_0 h_0$. Then there are polynomials $g, h \in O[x]$ so that $f = gh$ with $\deg(g) = \deg(g_0)$ and $\overline{g} = g_0$ and $\overline{h} = h_0$*

For a proof, see [8, 4.6]. Let us now consider the following special case of lemma 7.2.

Lemma 7.3. *Suppose $f \in O[x]$ and that $\beta \in O/\mathfrak{p}$ is a root of multiplicity one for $\overline{f}$. Then β lifts uniquely to an element in O which is a root of f.*

Proof. Since β is a root of multiplicity one for $\overline{f}$ we know that f is primitive. We can write

$$\overline{f}(x) = (x - \beta)h_0(x)$$

where $h_0 \in O/\mathfrak{p}$ and where $(x - \beta) \nmid h_0$. Hence $(x - \beta)$ and h_0 are coprime so by lemma 7.2 we can find $g, h \in O$ so that $f = gh$ and where $\overline{g} = (x - \beta)$ and $\overline{h} = h_0$ and $\deg g = \deg(x - \beta) = 1$. Thus we can write $g(x) = x - \zeta$ for some $\zeta \in O$ and we have $\overline{\zeta} = \beta$. If there was another element ζ' in O so that $f(\zeta') = 0$ and $\overline{\zeta'} = \beta$ then $(x - \zeta)(x - \zeta') \mid f(x)$ so $(x - \beta)^2 \mid \overline{f}$ which is impossible since β has multiplicity 1. □

We can now prove the following rather surprising statement.

Proposition 7.3. *Suppose that O has finite residue field and set $q = \#(O/\mathfrak{p})$. Then the ring O contains the $q - 1$'st roots of unity, μ_{q-1}.*

Proof. Let $x \in (O/\mathfrak{p})^*$. Since $(O/\mathfrak{p})^*$ has order $q-1$ we have $x^{q-1} = 1$, showing that x is a $q-1$'st root of unity. Thus $x^{q-1} - 1$ has $q-1$ distinct roots which in $O/\mathfrak{p}$ and so by lemma 7.3 these all lift to distinct roots in O of $x^{q-1} - 1$ meaning that O contains μ_{q-1}. □

From this proposition we get an injective map $(O/\mathfrak{p})^* \to O^*$ defined by sending an element x in $O/\mathfrak{p}$ to the $q - 1$'st root of unity in O that reduces to x. This map is called the Teichmüller lift. Using lemma 7.2 we can derive the following lemma.

Lemma 7.4. *Suppose $f(x) = a_n x^n + ... + a_0 \in K[x]$ is irreducible and has degree n. Then*

$$\max\{|a_n|, |a_{n-1}|, ..., |a_0|\} = \max\{|a_n|, |a_0|\},$$

and if f is monic and $a_0 \in O$ then $f \in O[x]$.

Proof. Among the coefficients of f, one of them must have absolute value larger than or equal to all the other ones. Let a_{k_0} be one such coefficient but with k_0 as small as possible. As $a_n \neq 0$ we must have $|a_{k_0}| \geq |a_n| > 0$ so $a_{k_0} \neq 0$. Let $g(x) = \frac{1}{a_k} f(x)$ and denote the coefficients of g by $b_i = \frac{a_i}{a_{k_0}}$. We have $\max\{|b_1|, ..., |b_n|\} = 1$ meaning that $g \in O[x]$ and we also have $|b_{k_0}| = 1$. As such g is primitive, i.e, $\overline{g} \neq 0$. By the choice of k_0 we have that $b_i \in \mathfrak{p}$ for all $i \in \{1, ..., k_0 - 1\}$ and therefore

$$g(x) \equiv x^{k_0}(b_{k_0} + b_{k_0+1}x + ... + b_n x^{n-k_0}) \pmod{\mathfrak{p}}$$

If $\max\{|b_0|, |b_n|\} < 1$ then $0 < k_0 < n$, which would mean that the above would be a factorization of coprime polynomials in $O/\mathfrak{p}$. But the lemma 7.2 would imply that g is not irreducible. Hence we must have $\max\{|b_0|, |b_n|\} = 1$. We conclude that

$$\max\{|a_n|, |a_{n-1}|, ..., |a_0|\} = \max\{|a_n|, |a_0|\},$$

as desired. If now $a_n = 1$ and $a_0 \in O$ then the above implies that $|a_i| \leq \max\{|a_0|, |a_n|\} = 1$ for all $i \in \{1, ..., n\}$.

□

7.4 Dedekind domains and valuations

Recall that for a commutative ring R with unity we can localize at any given prime ideal $\mathfrak{p} \subseteq R$ and obtain the local ring $R_\mathfrak{p} = \{\frac{\alpha}{\beta} \mid \alpha \in R, \beta \in R \setminus \mathfrak{p}\}$ with two elements $\frac{\alpha}{\beta}$ and $\frac{\alpha'}{\beta'}$ in $R_\mathfrak{p}$ being equal if an only if there exists an element $s \in R \setminus \mathfrak{p}$ so that $s\alpha\beta' = s\alpha'\beta$. The ring $R_\mathfrak{p}$ has unique maximal ideal $\mathfrak{M} = \{\frac{\alpha}{\beta} \mid \alpha \in \mathfrak{p}, \beta \in R \setminus \mathfrak{p}\}$, the localization map is the homomorphism sending an element $r \in R$ to $\frac{r}{1} \in R_\mathfrak{p}$ and the extension of $\mathfrak{p}$ along this map is $\mathfrak{M}$.

Lemma 7.5. $\mathrm{Frac}(R/\mathfrak{p})$ *is canonically isomorphic to* $R_\mathfrak{p}/\mathfrak{M}$. *In particular, if* $\mathfrak{p}$ *is a maximal ideal then* $R/\mathfrak{p} \cong R_\mathfrak{p}/\mathfrak{M}$.

Proof. Firstly, $R/\mathfrak{p}$ is an integral domain as $\mathfrak{p}$ is a prime, so it makes sense to take the field of fractions of this ring. Consider the composition of maps $R \hookrightarrow R_\mathfrak{p} \twoheadrightarrow R_\mathfrak{p}/\mathfrak{M}$, with the first one being the localization map and the second one the natural projection. Then $r \in R$ is sent to zero if and only if $\frac{r}{1} \in \mathfrak{M}$ which is equivalent to $r \in \mathfrak{p}$. Hence we get an induced injective homomorphism, ϕ

$$R/\mathfrak{p} \hookrightarrow R_\mathfrak{p}/\mathfrak{M}$$

which is given by $\phi(r + \mathfrak{p}) = \frac{r}{1} + \mathfrak{M}$. We know that $R_\mathfrak{p}/\mathfrak{M}$ is a field since $\mathfrak{M}$ is a maximal ideal, so it follows from the universal property for the field of fractions that there is a unique injective homomorphism $\sigma : \mathrm{Frac}(R/\mathfrak{p}) \to R_\mathfrak{p}/\mathfrak{M}$, extending ϕ, which is given by $\sigma(\frac{\alpha}{\beta}) = \phi(\alpha)\phi(\beta)^{-1}$. Let us show that this map is surjective. Suppose that $\frac{a}{b} + \mathfrak{M} \in R_\mathfrak{p}/\mathfrak{M}$ with $a \in R$ and $b \notin \mathfrak{p}$. Thus, $b + \mathfrak{p}$ is non-zero in $R/\mathfrak{p}$. We can now evaluate

$$\sigma(\frac{a+\mathfrak{p}}{b+\mathfrak{p}}) = \phi(a+\mathfrak{p})\phi(b+\mathfrak{p})^{-1} = (\frac{a}{1} + \mathfrak{M})(\frac{b}{1} + \mathfrak{M})^{-1} = \frac{a}{b} + \mathfrak{M}$$

For the last statement, suppose that $\mathfrak{p}$ is maximal. Then $R/\mathfrak{p}$ is a field, so it is isomorphic to its field of fractions. □

Let O be a Dedekind domain and K its field of fractions. For a prime ideal $\mathfrak{p}$ of O we get by proposition 5.5 a discrete an normalized valuation $v_\mathfrak{p}$ on K and it extends to the completion $K_\mathfrak{p}$ of K with respect to $v_\mathfrak{p}$, is also denoted by $v_\mathfrak{p}$ and is a discrete valuation by proposition 7.1. Let O_{v_p} be the valuation ring of K and let $\mathfrak{P}_{v_\mathfrak{p}}$ be its unique maximal ideal. Similarly, let $O_\mathfrak{p}$ be the valuation ring of $K_\mathfrak{p}$ and $\hat{\mathfrak{p}}$ its unique maximal ideal.

Proposition 7.4. *We have the following*

1. *The localization of O at $\mathfrak{p}$, that is, the ring*

$$O_{(\mathfrak{p})} = \{\frac{\alpha}{\beta} \in K \mid \alpha \in O, \beta \notin \mathfrak{p}\} = \{\frac{\alpha}{\beta} \in K \mid \alpha \in O, v_\mathfrak{p}(\beta) = 0\}$$

   is the same ring as $O_{v_\mathfrak{p}}$.

2. *We have equalities $\mathfrak{P}_{v_\mathfrak{p}} = \mathfrak{p}O_{(\mathfrak{p})}$ and $\hat{\mathfrak{p}} = \mathfrak{p}O_\mathfrak{p}$ and isomorphisms*

$$O/\mathfrak{p} \cong O_{(\mathfrak{p})}/\mathfrak{P}_{v_\mathfrak{p}} \cong O_\mathfrak{p}/\hat{\mathfrak{p}}$$

Proof. For (1.) it is clear that we have the inclusion $O_{(\mathfrak{p})} \subseteq O_{v_\mathfrak{p}}$. So let $\frac{x}{y} \in O_{v_\mathfrak{p}}$ with $x, y \in O$ both non-zero and let $u = v_\mathfrak{p}(x)$, $v = v_\mathfrak{p}(y)$. Since $v_\mathfrak{p}(\frac{x}{y}) \geq 0$ we have $u \geq v$. We can now write

$$xO = \mathfrak{p}^u A, \text{ and } yO = \mathfrak{p}^v B$$

where A, B are two ideals of O so that $\mathfrak{p} \nmid A$ and $\mathfrak{p} \nmid B$. This means that $\mathfrak{p} \not\supseteq B$, which implies that we can find $c \in B$ such that $c \notin \mathfrak{p}$. But then $v_\mathfrak{p}(c) = 0$ and hence we are done if we can show that there is $b \in O$ so that $xc = yb$ as that would imply that $\frac{x}{y} = \frac{b}{c}$. Now, $xc \in xO$ and also $xc \in B$, meaning that $xc \in xO \cdot B$. But,

$$xO \cdot B = \mathfrak{p}^u AB = \mathfrak{p}^{u-v} A \mathfrak{p}^v B = \mathfrak{p}^{u-v} A \cdot yO$$

So $xc \in \mathfrak{p}^{u-v} A \cdot yO$ which means that we can find $b \in \mathfrak{p}^{u-v} A$ such that $xc = by$. We conclude that $O_{v_\mathfrak{p}} = O_{(\mathfrak{p})}$. As for (2.) it is a general fact that the extension of $\mathfrak{p}$ along the localization map, $O \hookrightarrow O_{(\mathfrak{p})}$ gives the ideal $\mathfrak{P}_{v_\mathfrak{p}}$, that is, $\mathfrak{P}_{v_\mathfrak{p}} = \mathfrak{p}O_{(\mathfrak{p})}$. Since O is a Dedekind domain, $\mathfrak{p}$ is in fact a maximal ideal of O so by lemma 7.5 it follows that $O/\mathfrak{p} \cong O_{(\mathfrak{p})}/\mathfrak{P}_{v_\mathfrak{p}}$. Let us show that $\hat{\mathfrak{p}} = \mathfrak{p}O_\mathfrak{p}$. Let $\alpha \in \hat{\mathfrak{p}}$, that is, $v_\mathfrak{p}(\alpha) > 0$ and suppose that $\alpha \neq 0$. Since $\mathfrak{p}^2 \subsetneq \mathfrak{p}$ we know that there is $\pi \in \mathfrak{p}$ so that $v_\mathfrak{p}(\pi) = 1$. By proposition 7.1 we know that $v_\mathfrak{p}(K_\mathfrak{p}^*) = v_\mathfrak{p}(K^*) = \mathbb{Z}$ so $n := v_\mathfrak{p}(\alpha) \in \mathbb{N}$. Now $v_\mathfrak{p}(\alpha\pi^{-n}) = 0$ so $\alpha\pi^{-n} \in O_\mathfrak{p}$. We now have $\pi\pi^{n-1} \cdot (\alpha\pi^{-n}) \in \mathfrak{p}O_\mathfrak{p}$ and so $\hat{\mathfrak{p}} \subseteq \mathfrak{p}O_\mathfrak{p}$. The other inclusion is clear. The isomorphism $O_{(\mathfrak{p})}/\mathfrak{P}_{v_\mathfrak{p}} \cong O_\mathfrak{p}/\hat{\mathfrak{p}}$ is just proposition 7.2. □

Here is an interesting consequence for the $\mathfrak{p}$-adic numbers.

Proposition 7.5. *Let K be a number field, O its ring of integers and take a prime $\mathfrak{p}$ of O and consider the $\mathfrak{p}$-adic numbers, $K_{\mathfrak{p}}$. We know from algebraic number theory that $O/\mathfrak{p}$ is finite so $K_{\mathfrak{p}}$ has finite residue field by proposition 7.4*

This actually shows that $K_{\mathfrak{p}}$ is what is called a local field, see definition 8.1. We obtain the following special case of proposition 7.2 for the p-adic numbers

Proposition 7.6. *For any $n \geq 1$ we have*

$$\mathbb{Z}_p/p^n\mathbb{Z}_p \cong \mathbb{Z}/p^n\mathbb{Z}$$

7.5 Finite extensions of non-archimedian complete fields

It turns out that the absolute value on a complete non-archimedian field can be extended to any finite field extension in a unique and very concrete way.

Theorem 7.2. *Suppose $(K, |\cdot|)$ is a complete valued non-archimedian field and that L/K is a finite field extension of degree n. Then $|\cdot|$ can be extended in a unique way to a non-archimedian absolute value on L given explicitly as*

$$|\alpha| = \sqrt[n]{|N_{L/K}(\alpha)|}, \text{ for all } \alpha \in L \tag{$\star$}$$

Furthermore $(L, |\cdot|)$ is complete.

Proof. Let us first make sure that the proclaimed function defined in eq. ($\star$) is actually well defined. For all $\alpha \in L$ we know that $N_{L/K}(\alpha) \in K$. Hence it makes sense to use the absolute value on K on it. After doing so we get a real number greater than or equal to 0, and so it makes sense to take the nth root of this number. Let us start with the existence part. Let o be the valuation ring of K and O the integral closure of o in L. To help us establish that eq. ($\star$) is an absolute value on L we will first show that

$$O = \{\alpha \in L \mid N_{L/K}(\alpha) \in o\}. \tag{$\triangle$}$$

First, suppose that $\alpha \in O$. Then α is integral over o and hence so are all the conjugates of α. But then the product of all of the conjugates of α is also integral over o and this product is equal to $N_{L/K}(\alpha)$, which we know is an element of K. Since o is integrally closed in K it now follows that $N_{L/K}(\alpha)$ is in o. Let us now show the other inclusion. So assume that $\alpha \in L^*$ and that $N_{L/K}(\alpha) \in o$. Since L/K is a finite extension, L is algebraic over K and therefore α is in particular algebraic over K. Hence we can consider its minimal polynomial, f, over K. Then by lemma 4.2 we have that

$$N_{L/K}(\alpha) = \pm a_0^m$$

where $m = [L : F]$ and a_0 is the constant term of f. But that means that $a_0^m \in o$ so $|a_0|^m = |a_0^m| \leq 1$, so $|a_0| \leq 1$ so $a_0 \in o$. As f is monic it now follows from lemma 7.4 that $f \in o[x]$ and as $f(\alpha) = 0$ it follows that α is integral over

o meaning that $\alpha \in O$. We have now established eq. ($\triangle$). Suppose now that $\alpha \in L$. We have that $|\alpha| \leq 1$ is the same as saying that $|N_{L/K}(\alpha)| \leq 1$ which is the same as saying that $N_{L/K}(\alpha)$ belongs to the valuation ring o which, by what we have just shown is the same as $\alpha \in O$. This means that

$$O = \{\alpha \in L \mid |\alpha| \leq 1\} \tag{7.1}$$

Note that this means that once we have shown that eq. ($\star$) is a non-archimedian absolute value on L, we can conclude that its valuation ring is O. So let us show this now. First we show that eq. ($\star$) satisfies the ultrametric property. Since O is a ring we know that if $\alpha \in O$ then also $\alpha + 1 \in O$. From eq. (7.1) this immediately implies that if $|\alpha| \leq 1$ then $|\alpha + 1| \leq 1$. Assume that $\alpha, \beta \in L$ and that $|\alpha| \leq |\beta|$. We wish to show that $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$. Note that if $\beta = 0$ there is nothing to show so suppose that $\beta \neq 0$. Then $|\beta| \neq 0$ and we have $\frac{\alpha}{\beta} \in L$ and $|\frac{\alpha}{\beta}| \leq 1$. Thus

$$|\frac{\alpha}{\beta} + 1| \leq 1$$

and after multiplication by $|\beta|$ we are left with the inequality

$$|\alpha + \beta| \leq |\beta|$$

Because we assumed that $|\alpha| \leq |\beta|$ we have $|\beta| = \max\{|\alpha|, |\beta|\}$ and we are done. We still have some things to check before we can conclude that $|\cdot|$ is an absolute value on L, but these are easier to do. Firstly if $\alpha \in L$ then we have $|N_{L/K}(\alpha)| = 0$ if and only if $N_{L/K}(\alpha) = 0$ if and only if $\alpha = 0$. The fact that $|\alpha\beta| = |\alpha||\beta|$ for all $\alpha, \beta \in L$ follows from the fact that all the functions involved in the right hand side of eq. ($\star$) are multiplicative. We conclude that eq. ($\star$) is an absolute value on L. It also extends the absolute value on K since if $\alpha \in K$ then $N_{L/K}(\alpha) = \alpha^n$.

Let us now move on to proving the uniqueness part of the statement. So let $|\cdot|'$ be another non-archimedian absolute value on L which also extends the absolute value on K. Denote by O' the valuation ring with respect to $|\cdot|'$ and let $\mathfrak{p}$ and $\mathfrak{p}'$ be the maximal ideals of O and O' respectively. Let us show that $O \subseteq O'$. Assume for contradiction that $\alpha \in O$ but $\alpha \notin O'$. Then $\alpha \neq 0$ and we have that $|\alpha|' > 1$ and thus $\alpha^{-1} \in O'$. Since $\alpha \notin O'$ then α^{-1} cannot be a unit of O' and hence α^{-1} belongs to $\mathfrak{p}'$. As $\alpha \in O$ and O is the integral closure of o in L, we can find an equation

$$\alpha^d + a_{d-1}\alpha^{d-1} + ... + a_0 = 0$$

where all the a_i belong to o. But o is contained in O' since if $x \in o$ then $x \in K$ so $|x|' = |x| \leq 1$. Hence all the a_i belong to O'. Multiplying the above equation by α^{-d} and moving things around we get

$$1 = -(a_{d-1}\alpha^{-1} + ... + a_0\alpha^{-d})$$

But each term $a_{d-i}\alpha^{-i}$ is in $\mathfrak{p}'$ since $a_{d-i} \in O'$ and $\alpha^{-i} \in \mathfrak{p}'$ combined with the fact that $\mathfrak{p}'$ is an ideal. But that means that $1 \in \mathfrak{p}'$ which is impossible. Hence $O \subseteq O'$. What this means is that if $|\alpha| \leq 1$ then $|\alpha|' \leq 1$ and this implies that $|\cdot|$ and $|\cdot|'$ are equivalent. If they weren't then by proposition 5.1 we could find $\alpha \in K$ so that $|\alpha|' < 1$ and $|\alpha| \geq 1$. The fact that $|\alpha| \geq 1$ implies that $\alpha \neq 0$. Now $|\frac{1}{\alpha}| \leq 1$ so by what we just showed $|\frac{1}{\alpha}|' \leq 1$, i.e, $|\alpha|' \geq 1$ - that's impossible so $|\cdot|$ and $|\cdot|'$ are equivalent. But then again by proposition 5.1 there exists some fixed $s \geq 1$ so that $|x|' = |x|^s$ for all $x \in L$ and as $|\cdot|$ and $|\cdot|'$ are equal on K we must have $s = 1$. Hence $|x|' = |x|$, proving uniqueness. We now only need to show that L is complete with respect to $|\cdot|$. Let $1, \beta_2, ..., \beta_n \in L$ be a basis for L over K. Then for any $z \in L$ there exists unique $x_1, ..., x_n \in K$ so that $z = x_1 + x_2\beta_2 + ... + x_n\beta_n$, and so we get a map, $\phi : L \to \mathbb{R}$ by sending z to $\max\{|x_1|, ..., |x_n|\}$. This defines a non-archimedian absolute value on L and L is complete with respect to it. Observe that if $z \in K$ then we have $0 = x_2 = ... = x_n$, showing that ϕ extends the absolute value on K. Hence by the uniqueness property we have just shown we now have that $\phi = |\cdot|$ and therefore L is complete with respect to $|\cdot|$.

□

In section 6 we tried to avoid making use of the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$ with respect to $|\cdot|_\infty$. Considering the implication (3.) $\implies$ (2.) in proposition 5.1 and inspecting its proof in [8, Proposition 3.3], we see that this fact is actually used here.

8 Local fields

Definition 8.1. *A field K is called a **local field** if it is complete with regards to a discrete valuation and has finite residue field.*

We saw in proposition 7.5 that the $\mathfrak{p}$-adic number fields are examples of local fields. In particular, $\mathbb{Q}_p$ is a local field. Here is another example.

Example 8.1 (Continuation of example 5.1)**.** *It can be shown that $K((x))$ is the completion of $K(x)$ with respect to the discrete valuation ω. Hence $K((x))$ is complete and we have $K((x))/(x) \cong K$ and so $K((x))$ is a local field if and only if K is a finite field.*

The following proposition charaterizes the local fields of characteristic 0 as finite extensions of $\mathbb{Q}_p$.

Proposition 8.1. *A local field K of characteristic 0 is the same thing as a finite extension of $\mathbb{Q}_p$.*

Proof. Suppose first that $K/\mathbb{Q}_p$ is a finite extension of degree n. Then by theorem 7.2 the absolute value on $\mathbb{Q}_p$ extends uniquely to the non-archimedian absolute value $|\alpha| = \sqrt[n]{|N_{K/\mathbb{Q}_p}(\alpha)|}$ on K and K is complete with respect to it. The absolute value on $\mathbb{Q}_p$ is discrete, so this must also be the case for the

absolute value on K. Let κ be the residue field of K. From proposition 7.6 we know that $\mathbb{Q}_p$ has residue field $\mathbb{F}_p$. Let us show that the field extension $\kappa/\mathbb{F}_p$ has finite degree. Suppose that $\overline{x_1}, ..., \overline{x_m} \in \kappa$ are linearly independent over $\mathbb{F}_p$ with $x_1, ..., x_m$ being corresponding preimages in K. These elements must also be linearly indenpendent over $\mathbb{Q}_p$. Assume for contradiction that this is not the case. We would then have an equation

$$\lambda_1 x_1 + ... + \lambda_m x_m = 0$$

with at least one of the $\lambda_i \in \mathbb{Q}$ being non-zero, meaning it has non-zero absolute value. Choose among such non-zero elements the λ_i that has the largest absolute value and divide the above equation by it. We achieve a linear combination where all coefficients have absolute value smaller than or equal to 1 and the ith coefficient equals 1. Reducing to κ we get a non-trivial linear combination of $\overline{x_1}, ..., \overline{x_m}$. We conclude that $x_1, ..., x_m$ must be linearly independent over $\mathbb{Q}$. Thus m can be at most n and so the degree $[\kappa : \mathbb{F}_p]$ can also be at most n and this implies that κ is a finite field, since finite extensions of finite fields are finite. For the remaining implication, see [8, Proposition 5.2].

□

Inspired by the fact that the finite extensions of $\mathbb{Q}$ are called number fields we now define the following.

Definition 8.2. *A local field of characteristic 0 is called a $\mathfrak{p}$-adic number field.*

In the following let K be a local field with valuation v, O the valuation ring corresponding to v and $\mathfrak{p}$ the unique maximal ideal of O. Let us also set $q = \#(O/\mathfrak{p})$ and $U^{(1)} = 1 + \mathfrak{p}$.

Proposition 8.2. *K^* contains μ_{q-1}, the group of $(q-1)$st roots of unity and we have*

$$K^* = \{\pi^k \mid k \in \mathbb{Z}\} \times \mu_{q-1} \times U^{(1)}$$

Proof. Since K is a local field it follows from proposition 7.3 that O^* contains μ_{q-1}. Because of proposition 5.4 we really only need to show that $O^* = \mu_{q-1} \times U^{(1)}$. To do that $U^{(1)}\mu_{q-1} = O^*$ and $U^{(1)} \cap \mu_{q-1} = \{1\}$. Consider the canonical map $\varphi : O^* \to (O/\mathfrak{p})^*$. Saying that $x \in \ker\varphi$ is the same as saying that $1 \equiv x \pmod{\mathfrak{p}}$ which is equivalent to $x \in 1 + \mathfrak{p} = U^{(1)}$, so $\ker\varphi = U^{(1)}$. Thus, if $x \in U^{(1)} \cap \mu_{q-1}$ then $\varphi(x) = 1$ as $x \in U^{(1)}$ and as $x \in \mu_{q-1}$ we must have $x = 1$ since φ maps μ_{q-1} bijectively onto $(O/\mathfrak{p})^*$. So $U^{(1)} \cap \mu_{q-1} = \{1\}$. Suppose that $u \in O^*$. Then $\varphi(u)^{-1} \in (O/\mathfrak{p})^*$ so the Teichmüller lift from underneath the proof of lemma 7.5 gives us $\zeta^{-1} \in \mu_{q-1}$ so that $\varphi(\zeta^{-1}) = \varphi(u)^{-1}$. Now we have $1 = \varphi(u)\varphi(\zeta^{-1})$. But then $u\zeta^{-1} \in \ker\varphi = U^{(1)}$, so we can write $u\zeta^{-1} = v$ for some $v \in U^{(1)}$. Hence $u = v\zeta$ showing that $O^* = U^{(1)}\mu_{q-1}$ and we are done. □

Lemma 8.1. *For any $n \in \mathbb{N}$ we have $\#(O/\mathfrak{p}^n) = q^n$.*

Proof. We prove the statement using induction on n. The base case $n = 1$ is simply by definition of q. So suppose that for given $n \in \mathbb{N}$ we have $\#(O/\mathfrak{p}^n) = q^n$. By the third isomorphism theorem we get

$$(O/\mathfrak{p}^{n+1})/(\mathfrak{p}^n/\mathfrak{p}^{n+1}) \cong O/\mathfrak{p}^n$$

But then by proposition 5.4 we have $\#(\mathfrak{p}^n/\mathfrak{p}^{n+1}) = \#(O/\mathfrak{p}) = q$ and $\#(O/\mathfrak{p}^n) = q^n$ so

$$\#(O/\mathfrak{p}^{n+1}) = \#(O/\mathfrak{p}^n) \cdot \#(\mathfrak{p}^n/\mathfrak{p}^{n+1}) = q^{n+1}$$

□

Theorem 8.1. *The valuation ring of a local field K is compact.*

Proof. The valuation ring can be thought of as the closed unit ball around 0 with respect to the absolute value on K. Hence it is closed and is homeomorphic to $\varprojlim O/\mathfrak{p}^n$ by theorem 7.1 which is then of course also closed. Now $\varprojlim O/\mathfrak{p}^n$ is contained in $\prod_{n=1}^{\infty} O/\mathfrak{p}^n$, which is compact by Tychonoff's theorem since all the $O/\mathfrak{p}^n$ are finite by lemma 8.1 hence compact. It follows that O is compact. □

Remark 8.1. *Since the topology on K is induced by a metric, it means that compactness is equivalent to sequential compactness. Thus every sequence in O has a convergent subsequence.*

9 Formal series viewed as functions

Let K be a complete valued field with valuation v and $|\cdot|$ a corresponding non-archimedian absolute value. Given a formal power series, $f(X) \in K[X]$, we can consider the set C of all points $x \in K$ for which $f(x)$ converges with resepect to $|\cdot|$. This set is called the region of convergence for f, and we get an induced function from C into K, which we also denote f. Let us write $f(X) = \sum_{n=0}^{\infty} a_n X^n$. Suppose that $x \in C$ and consider $f(x) - f(0) = \sum_{n=1}^{\infty} a_n x^n$. By the ultrametric property $|f(x) - f(0)| \leq \max_{n\in\mathbb{N}}\{|a_n x^n|\}$, so we can make $|f(x) - f(0)|$ as small as we want by choosing x to be close enough to 0 and so f is continuous at 0. If we pick any $c \in C$ we see that

$$f(Y + c) = \sum_{n=1}^{\infty} a_n (Y + c)^n = \sum_{n=1}^{\infty} a_n \sum_{k=0}^{n} \binom{n}{k} c^k Y^{n-k},$$

which is just another power series, $f_c(Y)$ and by construction 0 is in the region of convergence for f_c so f_c is continous in this point. Thus, $f_c(y) \to f_c(0) = f(c)$ as $y \to 0$ so $f(x) \to f(c)$ as $x \to c$, meaning that f is also continous in the point c and so f is continous on all of C. We can also consider the ring of formal power series in several variables, $K[[X_1, ..., X_n]]$ and the elements of this ring take the form

$$f(X_1, ..., X_n) = \sum_{a_1 \geq 0, ..., a_n \geq 0} \alpha_{a_1, ... a_n} X_1^{a_i} \dots X_n^{a_n}$$

A series like this also has a region of convergence, that is the set, C, of tuples, $(x_1, ..., x_n)$, so that $f(x_1, ..., x_n)$ converges with respect to $|\cdot|$. We equip K^n with the max norm, $\|(x_1, ..., x_n)\| = \max\{|x_1|, ..., |x_n|\}$ and one can show that f is continuous on C with respect to this absolute value by following a similar approach as in the case for series in one variable.

Sometimes a function can be given locally as a convergent series, $f \in K[X_1, ..., X_n]$. Such a function is called an analytic function. Here is a more precise definition.

Definition 9.1. *Consider an open neighborhood $A \subseteq K^n$ of a point $x_0 \in K^n$ and suppose we have a function $f : A \to K$. If there is a series $g \in K[[X_1, ..., X_n]]$ that converges on A and $g(x_1, ..., x_n) = f(x_1, ..., x_n)$ for every $(x_1, ..., x_n) \in A$ we say that f is given by the series g and in this case f is said to be an analytic function. If A' is another neighborhood of x_0 and $f' : A' \to K$ is also an analytic function, then we consider f and f' equal if they are given by the same series and $A' \subseteq A$ or $A \subseteq A'$.*

By the discussion above, analytic functions are continuous.

10 Logarithms and Exponentials

In this section, let K be a $\mathfrak{p}$-adic field. As such it is by proposition 8.1 a finite extension of $\mathbb{Q}_p$ and by theorem 7.2 the valuation v_p on $\mathbb{Q}_p$ therefore extends uniquely to K. This valuation is discrete so there is $\pi \in K$ so that $v_p(\pi) = \frac{1}{e} > 0$ is the smallest attainable strictly positive value for this function. Hence defining $v_{\mathfrak{p}} = ev_p$ we get a normalized valuation on K. Let us also denote by O the valuation ring with respect to $v_{\mathfrak{p}}$ and $\mathfrak{p}$ its unique maximal ideal and define $U^{(k)} = 1 + \mathfrak{p}^k$ for $k \in \mathbb{N}$. Let us define the following formal power series.

$$\log(1 + X) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{X^k}{k} \tag{10.1}$$

$$\exp(X) = \sum_{k=1}^{\infty} \frac{X^k}{k!} \tag{10.2}$$

We would like to establish the following result.

Proposition 10.1. *The power series defined in eqs. (10.1) and (10.2) induce, for $m > \frac{e}{p-1}$ continuous isomorphisms*

$$U^{(m)} \underset{\exp}{\overset{\log}{\rightleftarrows}} \mathfrak{p}^m$$

and these are inverses of one another.

We'll give the proof of this proposition towards the end of this section. To do this we will need the following equalities in formal series.

$$\exp(\log(1+X)) = 1+X \tag{10.3}$$
$$\log(\exp(X)) = X \tag{10.4}$$

Proofs for these identities can be found here: [6, 4]. We also have the identites,

$$\log(XY) = \log(X) + \log(Y), \quad \exp(X+Y) = \exp(X)\exp(Y)$$

These equalities might lead one to think that the "isomorphism" part of proposition 10.1 follows immediately after establishing the existence of the functions displayed in the diagram in this proposition. However, this kind of reasoning is not valid in general as the following example illustrates.

Example 10.1. *Let $K = \mathbb{Q}_2$ and let $g(X) = 2X^2 - 2X$ and $f(X) = \exp(X)$. Let $h(X) = f(g(X))$ and write $h(X) = \sum_{n=0} a_n X^n$. It is clear that $g(1) = 0$ is in the region of convergence of $f(X)$. It is claimed in [5, Problem 162] that it is possible to show that $v_2(a_n) \geq 1 + \frac{n}{4}$ for $n \geq 2$. Combining this with the fact that $v_2(1) \geq 1$, since $1 \in \mathbb{Z}_2$, we see that $h(1)$ converges. We claim that $h(1) \neq f(g(1))$. Observe first that $f(g(1)) = 1$. We see that $v_2(a_n) \geq 2$ for all $n \geq 4$ which means that $4 \mid a_n$ for all $n \geq 4$. Hence, modulo 4 all the terms beyond and including the fourth term of $h(X)$ vanish. Let us now compute the first 3 terms of $h(X)$*

$$h(X) = 1 - 2X + 4X^2 + \dots$$

So $h(X) \equiv 1 - 2X \pmod 4$ so $h(1) \equiv -1 \pmod 4$ which means that $h(1) \neq f(g(1))$.

In light of this example we would like a result which gives sufficient conditions for when an equality in formal series like $h(X) = f(g(X))$ leads to an equality of induced functions. This is exactly we will build up to now.

Proposition 10.2. *Let $x \in K$ and suppose $f(X)$ and $g(X)$ are formal power series. If both $f(x)$ and $g(x)$ converge then the formal series $(f+g)(x)$ and $(fg)(x)$ converge to $f(x) + g(x)$ and $f(x)g(x)$ respectively.*

We will say that a sequence $b_{ij} \in K$ converges to 0 uniformly in j if for every j and $\epsilon > 0$ there exists an $N \in \mathbb{N}$, which does not depend on j, so that if $i \geq N$ then $|b_{ij}| < \epsilon$. In notation, we will write $\lim_{j\to\infty} b_{ij} = 0$.

Proposition 10.3. *Assume that $b_{ij} \in K$ and assume that*

1. $\lim_{j\to\infty} b_{ij} = 0$
2. $\lim_{i\to\infty} b_{ij} = 0$ *uniformly in j*

Then the following two sums converge and are equal

$$\sum_{i=0}^{\infty}(\sum_{j=0}^{\infty} b_{ij}), \quad \sum_{j=0}^{\infty}(\sum_{i=0}^{\infty} b_{ij}) \tag{10.5}$$

For a proof, see [5, Proposition 5.1.4]. The idea is to first show that the inner sums in eq. (10.5) converge and then use that to show that the double sums converge. To show that they converge to the same thing, look at the absolute value of their difference.

Theorem 10.1. *Suppose we have the power series*

$$f(X) = \sum_{n=0}^{\infty} a_n X^n, \quad g(X) = \sum_{n=1}^{\infty} b_n X^n$$

Let $h(X) = f(g(X))$ and let $x \in K$. Assume

1. *$g(x)$ converges*
2. *$f(g(x))$ converges, that is, $g(x)$ is in the region of convergence of $f(X)$*
3. *$|b_n x^n| \leq |g(x)|$ for all $n \in \mathbb{N}$*

Then $h(x)$ converges and we have $h(x) = f(g(x))$

Proof. We can write $g(X)^m = \sum_{n=m} d_{m,n} X^n$ with

$$d_{m,n} = \begin{cases} \sum_{i_1+\ldots+i_m=n} b_{i_1} \ldots b_{i_m} & \text{if } n \geq m \\ 0 \text{ otherwise} \end{cases}$$

Thus we get

$$h(X) = a_0 + \sum_{n=1}^{\infty} \Big(\sum_{m=1}^{n} a_m d_{m,n} \Big) X^n$$

By assumption we know that $g(x)$ converges so by proposition 10.2 we have that

$$g(x)^m = \sum_{n=m}^{\infty} d_{m,n} x^n$$

Suppose that $n \geq m$. Using the ultrametric property on $|\cdot|$ and the assumption that $|b_n x^n| \leq |g(x)|$ for all $n \in \mathbb{N}$ we obtain

$$\begin{aligned} |d_{m,n} x^n| &= \Big| \sum_{i_1+\ldots+i_m=n} b_{i_1} x^{i_1} \ldots b_{i_m} x^{i_m} \Big| \\ &\leq \max_{i_1+\ldots+i_m=n} \{ |b_{i_1} x^{i_1}| \ldots |b_{i_m} x^{i_m}| \} \\ &\leq |g(x)|^m \end{aligned} \tag{10.6}$$

If instead $n < m$ we have $d_{m,n} = 0$ so we conclude that $|d_{m,n} x^n| \leq |g(x)|^m$ for

all $m, n \in \mathbb{N}$. This means that we now have

$$\begin{aligned} f(g(x)) &= a_0 + \sum_{m=1}^{\infty} a_m g(x)^m \\ &= a_0 + \sum_{m=1}^{\infty} a_m \Big(\sum_{n=m}^{\infty} d_{m,n} x^n \Big) \\ &= a_0 + \sum_{m=1}^{\infty} \sum_{n=m}^{\infty} a_m d_{m,n} x^n \end{aligned}$$

Let n and $\epsilon > 0$ be given. Since $f(g(x))$ converges by assumption we know that $a_m g(x)^m$ converges to 0. Hence we can pick N so that if $m \geq N$ then $|a_m g(x)^m| < \epsilon$. Using eq. (10.6) we now have $|a_m d_{m,n} x^n| \leq |a_m g(x)^m| < \epsilon$. This shows that $\lim_{m\to\infty} a_m d_{m,n} x^n = 0$ uniformly in n. Let now m be given. Since $g(x)^m$ converges we know that $d_{m,n} x^n$ converges to 0 as $n \to \infty$. Thus $a_m d_{m,n} x^n$ converges to 0 as $n \to \infty$. It now follows from proposition 10.3 that the sum

$$h(x) = a_0 + \sum_{n=1}^{\infty} \Big(\sum_{m=1}^{n} a_m d_{m,n} \Big) x^n$$

converges and that it equals $f(g(x))$. □

We will use the above theorem soon but first let us show that log converges on $U^{(1)}$.

Proposition 10.4. *The power series,* $\log(1+z)$*, defined in eq.* (10.1) *converges for all* $1 + z \in U^{(1)}$*.*

Proof. Suppose $(1 + z) \in U^{(1)}$. Then $z \in \mathfrak{p}$ and so $v_p(z) > 0$, which means that $c = p^{v_p(z)} > 0$. Thus we can apply the usual logarithm and get $v_p(z) = \frac{\ln c}{\ln p}$. If k is any natural number, then we always have

$$p^{v_p(k)} \leq k,$$

since $p^{v_p(k)}$ divides k. Applying ln to both sides of this inequality is valid, as both sides are strictly positive and from doing so we get

$$v_p(k) \ln p \leq \ln k$$

and so,

$$v_p(k) \leq \frac{\ln k}{\ln p}$$

Now for any $k \in \mathbb{N}$ we get

$$\begin{aligned} v_p(\frac{z^k}{k}) &= v_p(z^k) - v_p(k) \\ &= kv_p(z) - v_p(k) \\ &\geq k\frac{\ln c}{\ln p} - \frac{\ln k}{\ln p} \\ &= \frac{\ln(c^k/k)}{\ln p} \end{aligned}$$

Clearly, $\ln(c^k/k) \to \infty$ as $k \to \infty$. Hence, $v_p(\frac{z^k}{k}) \to \infty$ as $k \to \infty$ and by lemma 5.1 combined with the fact that K is complete we are now done. □

Lemma 10.1. *(Legendre's formula) Suppose we have $k \in \mathbb{N}$. Then*

$$v_p(k!) = \sum_{i=1}^{\infty} \left\lfloor \frac{k}{p^i} \right\rfloor$$

Proof. First of all, there are only finitely many terms in the sum since $\left\lfloor \frac{k}{p^i} \right\rfloor$ is eventually zero when i is large enough so it converges. For natural numbers q and n we define the function

$$f_q(n) = \begin{cases} 1 & \text{if } q \mid n \\ 0 & \text{otherwise} \end{cases}$$

We then have for any $m \in \mathbb{N}$ that

$$v_p(m) = \sum_{i=1}^{\infty} f_{p^i}(m)$$

Thus,

$$\begin{aligned} v_p(k!) &= \sum_{j=1}^{k} v_p(j) \\ &= \sum_{j=1}^{k} \sum_{i=1}^{\infty} f_{p^i}(j) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{k} f_{p^i}(j) \end{aligned}$$

The swaps in summation are allowed here since all sums involved are finite. But clearly, $\sum_{j=1}^{k} f_{p^i}(j) = \left\lfloor \frac{k}{p^i} \right\rfloor$, so we get the result. □

Proof of proposition 10.1. We have already established that the power series for the logarithm converges on $U^{(1)}$, hence it also converges on $U^{(m)}$ for all $m \in \mathbb{N}$. Our first goal is to show that log maps $U^{(m)}$ into $\mathfrak{p}^m$ for all $m > \frac{e}{p-1}$. To do this we need a little help from the following inequality. Let $k \in \mathbb{N}$ where $k > 1$.

$$\frac{v_p(k)}{k-1} \leq \frac{1}{p-1},$$

Let us show that this actually holds. Write $k = p^a k_0$, where $p \nmid k_0$. Then $v_p(k) = a$ and so the inequality is satisfied when $a = 0$. So suppose $a > 0$. Then we can write

$$p^a - 1 = (p-1)(p^{a-1} + ... + p + 1).$$

We have $a \leq p^{a-1} + ... + p + 1$, since the right side of the inequality consists of a total of a terms all of which are larger than 1. Now

$$\frac{v_p(k)}{k-1} = \frac{a}{p^a k_0 - 1} \leq \frac{a}{p^a - 1} = \frac{1}{p-1} \cdot \frac{a}{p^{a-1} + ... + p + 1} \leq \frac{1}{p-1}$$

Suppose now that $1 + z \in U^{(m)}$ with $m > \frac{1}{p-1}$. Then $z \in \mathfrak{p}^m$ which is the same as saying that $v_\mathfrak{p}(z) \geq m$, since $v_\mathfrak{p}$ is normalized and so $v_\mathfrak{p}(z) > \frac{e}{p-1}$, which means that $v_p(z) > \frac{1}{p-1}$. Let us show that $\log(1+z) \in \mathfrak{p}^m$ by showing that $v_\mathfrak{p}(z) = v_\mathfrak{p}(\log(1+z))$. By what we showed above we have

$$v_p(\frac{z^k}{k}) - v_p(z) = (k-1)v_p(z) - v_p(k) > (k-1)(\frac{1}{p-1} - \frac{v_p(k)}{k-1}) \geq 0$$

Hence $v_\mathfrak{p}(\frac{z^k}{k}) > v_\mathfrak{p}(z)$ for all $k \in \mathbb{N}$ where $k > 1$. Thus, for all $s \in \mathbb{N}$, where $s > 1$, we have

$$v_\mathfrak{p}(A_s) \geq \min\{v_\mathfrak{p}(\frac{z^2}{2}), ..., v_\mathfrak{p}(\frac{z^s}{s})\}$$

with $A_s = \sum_{i=2}^{s} \frac{z^i}{i}(-1)^{(i-1)}$ meaning that $v_\mathfrak{p}(A_s) > v_\mathfrak{p}(z)$. It now follows from proposition 5.2 that

$$v_\mathfrak{p}(z + A_s) = \min\{v_\mathfrak{p}(z), v_\mathfrak{p}(A_s)\} = v_\mathfrak{p}(z)$$

From this it follows that $v_\mathfrak{p}(z + A_s) \to v_\mathfrak{p}(z)$. We also know that $z + A_s \to \log(1+z)$ as $s \to \infty$ and $v_\mathfrak{p}$ is continuous so $v_\mathfrak{p}(z + A_s) \to v_\mathfrak{p}(\log(1+z))$. By uniqueness of limits we now have $v_\mathfrak{p}(z) = v_\mathfrak{p}(\log(1+z))$.

Let us now show that the power series for the exponential function converges on $\mathfrak{p}^m$ for $m > \frac{e}{p-1}$. Using Legendre's formula we get the following inequality

$$v_\mathfrak{p}(k!) = e\sum_{i=1}^{\infty} \lfloor \frac{k}{p^i} \rfloor < ek \sum_{i=1}^{\infty} \frac{1}{p^i} = \frac{ek}{p-1}$$

Let us again assume that $z \in \mathfrak{p}^m$. We now have for any $k \in \mathbb{N}$

$$v_\mathfrak{p}(\frac{z^k}{k!}) = kv_\mathfrak{p}(z) - v_\mathfrak{p}(k!) > k(v_\mathfrak{p}(z) - \frac{e}{p-1})$$

But $v_{\mathfrak{p}}(z) - \frac{e}{p-1} \geq 0$, so $v_{\mathfrak{p}}(\frac{z^k}{k!}) \to \infty$ for $k \to \infty$. Let us now move on to showing that exp maps $\mathfrak{p}^m$ to $U^{(m)}$. Maintaining the assumption that $z \in \mathfrak{p}^m$ we will do this by showing that $v_{\mathfrak{p}}(\exp(z) - 1) = v_{\mathfrak{p}}(z)$ since this implies that $\exp(z) \in U^{(m)}$. Let us show that $v_{\mathfrak{p}}(\frac{z^k}{k!}) > v_{\mathfrak{p}}(z)$ for all $k > 1$. We have

$$v_{\mathfrak{p}}(\frac{z^k}{k!}) - v_{\mathfrak{p}}(z) = (k-1)v_{\mathfrak{p}}(z) - v_{\mathfrak{p}}(k!) > (k-1)v_{\mathfrak{p}}(z) - \frac{ek}{p-1} \geq 0$$

Hence $v_{\mathfrak{p}}(\frac{z^k}{k!}) > v_{\mathfrak{p}}(z)$ for all $k > 1$ and it now follows by an argument similar to one we gave a moment ago that $v_{\mathfrak{p}}(\exp(z) - 1) = v_{\mathfrak{p}}(z)$, which is what we wanted to show. To show that exp and log are inverses of one another observe that we now have the following

1. $\exp(z)$ converges,
2. $\log(\exp(z))$ converges since $\exp(z) \in U^{(m)}$,
3. $v_{\mathfrak{p}}(\frac{z^k}{k!}) \geq v_{\mathfrak{p}}(z) = v_{\mathfrak{p}}(\exp(z) - 1) \geq \min\{v_{\mathfrak{p}}(\exp(z)), v_{\mathfrak{p}}(1)\} = v_{\mathfrak{p}}(\exp(z))$ for all $k \geq 1$.

Combining this with eq. (10.4) and theorem 10.1 we see that we indeed have $\log \exp z = z$. A similar argument shows that we also have $\exp \log(1+z) = 1+z$. Let us now show that exp is indeed a homomorphism. Let $x, y \in \mathfrak{p}^m$. We have

$$\begin{aligned}
\exp(x+y) &= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{n!} \frac{n!}{(n-k)!k!} x^{n-k} y^k \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^{n-k}}{(n-k)!} \frac{y^k}{k!} \\
&= (\sum_{m=0}^{\infty} \frac{x^m}{m!})(\sum_{k=0}^{\infty} \frac{y^k}{k!})
\end{aligned}$$

Now we can show that log is also a homomorphism on $U^{(m)}$. Let $a, b \in U^{(m)}$. We obtain

$$\exp(\log(ab)) = ab = \exp(\log a) \exp(\log b) = \exp(\log a + \log b)$$

As exp is in particular injective on $\mathfrak{p}^m$ we now have $\log(ab) = \log a + \log b$ as desired. The fact that exp and log are continous follows from the fact that these functions are given as power series. □

At this point we are done with what we really wanted to do in this section. Let's keep exploring a little bit, just for curiosity's sake. Note that we have shown that log converges on $U^{(1)}$ but we have only shown that it a homomorphism on $U^{(m)}$, where $m > \frac{e}{p-1}$. In fact we have the following result.

Proposition 10.5. *There is a uniquely determined group homomorphism taking multiplication to addition,*

$$\log : K^* \to K$$

so that $\log p = 0$ *and for* $(1+z) \in U^{(1)}$ *this function is given by the series in eq.* (10.1).

The main obstacle in proving this is showing that log is actually a homomorphism on $U^{(1)}$. One could try to utilize theorem 10.1 but here condition (3.) becomes problematic to verify. Another approach is given in [5, Proposition 5.7.3]. This relies on differentiatiation of power series and the fact that if two series differentiate to the same thing then the original series differ by a constant. We now give the proof of proposition 10.5 given that we already know that log is a homomorphism on $U^{(1)}$.

Proof of proposition 10.5. We only need to show existence and uniqueness. Let $q = \#(O/\mathfrak{p})$. Suppose that $\alpha \in K^*$ and that $\pi \in K$ is an element so that $v_\mathfrak{p}(\pi) = 1$. By proposition 8.2 there exist unique elements $k_\alpha \in \mathbb{Z}$, $\zeta_\alpha \in \mu_{q-1}$ and $u_\alpha \in U^{(1)}$ so that $\alpha = \pi^{k_\alpha}\zeta_\alpha u_\alpha$, and we have $v_\mathfrak{p}(\alpha) = k_\alpha$. $v_\mathfrak{p} = ev_p$. Note that $v_\mathfrak{p}(p) = e$. So as $p \in K^*$ we can write $p = \pi^e \zeta_p u_p$. Let us now define $\log \pi = -\frac{1}{e}\log u_p$. We an extension of log to all of K^* by setting

$$\log \alpha = v_\mathfrak{p}(\alpha)\log(\pi) + \log(u_\alpha)$$

This function is continuous because addition of continuous functions are continuous and it is also a homomorphism by inheritance. Furthermore we have that

$$\log p = e\log(\pi) + \log(u_p) = \frac{-e\log u_p}{e} + \log(u_p) = 0$$

Suppose that we have another extension, $\lambda : K^* \to K$, of log so that $\lambda(p) = 0$ and which is also a homomorphism. Note first that for any $\alpha \in K^*$ we have

$$\lambda(\alpha) = \lambda(\pi^{v_\mathfrak{p}(\alpha)}\zeta_\alpha u_\alpha) = v_\mathfrak{p}(\alpha)\lambda(\pi) + \lambda(\zeta_\alpha) + \lambda(u_\alpha)$$

By assumption we know that $\log(u_\alpha) = \lambda(u_\alpha)$ since $u_\alpha \in U^{(1)}$. Hence

$$\lambda(\alpha) - \log(\alpha) = v_\mathfrak{p}(\alpha)(\lambda(\pi) - \log(\pi)) + \lambda(\zeta_\alpha) \tag{10.7}$$

So to show that $\lambda = \log$ it suffices to show that λ vanishes on μ_{q-1} and that $\lambda(\pi) = \log(\pi)$. Take any $\zeta \in \mu_{q-1}$. Since $\zeta^{q-1} = 1$ and $\lambda(1) = 0$ we see that

$$\lambda(\zeta) = \frac{1}{q-1}\lambda(\zeta^{q-1}) = 0$$

Now setting $\alpha = p$ in eq. (10.7) we get on the left hand side $\lambda(p) - \log(p) = 0$. On the right hand side we see that $\lambda(\zeta_p) = 0$ and $v_\mathfrak{p}(p) = e \neq 0$ so we $\lambda(\pi) - \log(\pi) = 0$ for the equality to hold. □

11 Local manifolds

In this section we define the so called local manifolds which are very similar to varieties known from algebraic geometry. Let K be a non-archimedian field with absolute value $|\cdot|$ and denote by $\boldsymbol{K}$ the algebraic closure of K. Let's refer to the elements of $\boldsymbol{K}^n$ as points and denote the set of points having maximum norm strictly smaller than some $\epsilon \in \mathbb{R}$ with $\epsilon > 0$ as an ϵ-neighborhood of the origin. Let D be the set containing all $f \in K[[X_1, ..., X_n]]$ that converge in some ϵ-neighborhood of the origin. This set is in fact a subring of $K[[X_1, ..., X_n]]$.

Definition 11.1 (Local manifold)**.** *Assume $f_1, ..., f_m \in D$ all of which have zero constant term. The set V of points $(\alpha_1, ..., \alpha_n) \in \boldsymbol{K}^n$ belonging to some ϵ-neighborhood of the origin so that*

$$f_1(\alpha_1, ...\alpha_n) = ... = f_m(\alpha_1, ..., \alpha_n) = 0$$

is called a local manifold. We say that the power series $f_1, ..., f_m$ determine V.

If V and W are two local manifolds we will by $V \subseteq W$ mean that there is an ϵ-neighborhood around the origin, U, so that $U \cap V \subseteq W$. We write $V = W$ if and only if both $V \subseteq W$ and $W \subseteq V$ in the sense described above. So if V consists of a finite number of points then $V = \{(0, ..., 0)\}$. Let $O_{\boldsymbol{K}}$ be the ring of integers of $\boldsymbol{K}$ and let V be a local manifold which is determined by $f_1, ..., f_m \in D$.

Definition 11.2 (Curve)**.** *A curve in $\boldsymbol{K}^n$ is a collection of n power series, $\omega_1(X), ..., \omega_n(X) \in O_{\boldsymbol{K}}[[X]]$ not all identically zero, and with constant term zero. We say the curve lies on V if there is an equality in power series,*

$$f_i(\omega_1(X), ..., \omega_n(X)) = 0 \text{ for all } i \in \{1, ..., m\}$$

For easy reference later on we will here quote two theorems, both of which are from [2]. The first one is Theorem 2, p. 305 and the second one is Theorem 3, p. 307.

Theorem 11.1. *A local manifold V either contains a curve or $V = \{0\}$.*

Theorem 11.2. *Suppose V and V' are local manifolds in $\overline{K}^n$ such that $V \nsubseteq V'$. Then there is a curve on V that does not lie on V'.*

Remark 11.1. *Note that by contraposition, this is the same as saying: If every curve on V is also on V' then $V \subseteq V'$.*

12 Thues Theorem and Skolem's Method

In this section we will finally get to prove theorem 1.1. As we talked about in the introduction we will do this by use of Skolem's method.

Lemma 12.1. *Let K be a field of characteristic 0 and let $n, n_1, n_2 \in \mathbb{N}$ so that $n_1 = n - 2$ and $n_2 \geq 2$ and suppose we have formal power series, $P_1, ..., P_n \in K[[X]]$ with zero constant term so that*

$$\sum_{j=1}^{n} a_{ij} \exp P_j = 0, \text{ for all } i \in \{1, ..., n_1\}$$

$$\sum_{j=1}^{n} b_{ij} P_j = 0, \text{ for all } i \in \{1, ..., n_2\},$$

with the a_{ij} and b_{ij} in K and where both matrices (a_{ij}) and (b_{ij}) have K-linearly independent rows. Then there are two indices $k \neq l$ so that $P_k = P_l$.

Proof. Let $V = \text{span}\{\exp P_1, ..., \exp P_n\}$ and consider the linear map

$$T : K^n \to V$$

given by

$$T(a_1, ..., a_n) = \sum_{i=1}^{n} a_i \exp P_i$$

By assumption we have $n_1 = n - 2$ linearly independent elements in $\ker T$. Hence nullity$(T) \geq n - 2$ so by rank-nullity we get rank$(T) \leq 2$. Thus, it has to be the case that two elements among $\{\exp P_1, ..., \exp P_n\}$ span all of V. After potentially renumbering these elements we can safely say that these elements are $\exp P_{n-1}$ and $\exp P_n$. Thus we can, for any $i \in \{1, ..., n\}$, find elements $a_i, b_i \in K$ so that

$$\exp P_i = a_i \exp P_{n-1} + b_i \exp P_n \tag{12.1}$$

It is easy to see what a_i and b_i should be when $i \in \{n-1, n\}$. So let us restrict ourselves to only look at $i \in \{1, ..., n-2\}$. In this case, let us suppose that $a_i = 0$. Then $\exp P_i = b_i \exp P_n$. But the formal power series, $\exp X$, has constant term 1, so for any power series $F \in K[[X]]$, without constant term, also the composition $\exp F$ has constant term 1. Thus both $\exp P_i$ and $\exp P_n$ have constant term 1 so $b_i = 1$ and so $P_i = P_n$, in which case we are done. We can therefore from now on assume that $a_i \neq 0$ for all $i \in \{1, ..., n_1\}$. Let us now define, for $i = \{1, ...n-1\}$,

$$Q_i = P_i - P_n, \tag{12.2}$$

If we had $Q_i = 0$ we would again be done, so let us assume this never happens. Combining eqs. (12.1) and (12.2) we have for $i \in \{1, ..., n-1\}$ that

$$\begin{aligned} \exp Q_i &= \exp P_i \exp(-P_n) \\ &= (a_i \exp P_{n-1} + b_i \exp P_n) \exp(-P_n) \\ &= a_i \exp(P_{n-1} - P_n) + b_i \\ &= a_i \exp Q_{n-1} + b_i, \end{aligned}$$

where we have used that $\exp(F+G) = \exp(F)\exp(G)$, where $F, G \in K[[X]]$ having non-zero constant term. This identity was essentially shown in the proof of proposition 10.1. Differentiating both sides of this equation and using the chain rule for formal power series we obtain

$$Q_i' \exp Q_i = a_i Q_{n-1}' \exp Q_{n-1}$$

As $\exp Q_i$ has non-zero constant term it is a unit in $K[[X]]$, meaning that we can divide by it. Solving for Q_i' we obtain

$$\begin{aligned} Q_i' &= \frac{(a_i Q_{n-1}' \exp Q_{n-1})}{\exp Q_i} \\ &= \frac{a_i Q_{n-1}' \exp Q_{n-1}}{a_i \exp Q_{n-1} + b_i} \\ &= \frac{Q_{n-1}' \exp Q_{n-1}}{\exp Q_{n-1} + c_i} \end{aligned} \tag{12.3}$$

where $c_i = a_i^{-1} b_i$. In case where $i = n-1$ we choose $a_i = 1$ so that $b_i = 0$. That way we have $c_i = 0$. Let us now deal with the next group of equations. Let $U = \text{span}\{P_1, ..., P_n\}$ and $U' = \text{span}\{P_1 - P_n, ..., P_{n-1} - P_n\}$ and define the linear maps $T : K^n \to U$ and $T' : K^{n-1} \to U'$ as

$$\begin{aligned} T(a_1, ..., a_n) &= \sum_{i=1}^{n} a_i P_i \\ T'(a_1, ..., a_{n-1}) &= \sum_{i=1}^{n-1} a_i Q_i \end{aligned}$$

By assumption, we know that $\text{nullity}(T) \geq 2$. Using rank-nullity on T we have

$$n = \text{rank}(T) + \text{nullity}(T) \geq \text{rank}(T) + 2.$$

So $\text{rank}(T) \leq n-2$. Since $U' \subseteq U$ and both T and T' are surjective, we must have $\text{rank}(T') \leq \text{rank}(T) \leq n-2$ and by rank-nullity used now on T' we get

$$n - 1 = \text{rank}(T') + \text{nullity}(T') \leq n - 2 + \text{nullity}(T')$$

Hence $\text{nullity}(T') \geq 1$, meaning that we can find a non-trivial linear combination,

$$\sum_{i=1}^{n-1} d_i Q_i = 0 \tag{12.4}$$

where the d_i are in K. Differentiating this sum and using eq. (12.3) we obtain

$$\sum_{i=1}^{n-1} d_i Q_i' = Q_{n-1}' \exp Q_{n-1} \sum_{i=1}^{n-1} \frac{d_i}{\exp Q_{n-1} + c_i} = 0$$

It must be the case that $Q'_{n-1} \neq 0$. Otherwise, we would have $Q_{n-1} \in K$, and since the P_i have zero constant term, we would then have $Q_{n-1} = 0$, which is a contradiction. Also, $\exp Q_{n-1} \neq 0$ so we must have

$$\sum_{i=1}^{n-1} \frac{d_i}{c_i \exp Q_{n-1}} = 0 \tag{12.5}$$

Next let us show that this implies that the rational function

$$f(z) = \sum_{i=1}^{n-1} \frac{d_i}{c_i + z}$$

is the zero function on K. As this is a rational function we can find polynomials $\phi, \psi \in K[z]$ so that $f(z) = \frac{\phi(z)}{\psi(z)}$, where $\psi(z)$ is not the zero polynomial. Let us for contradiction suppose that f is not identically zero. Then ϕ cannot be identically 0. But then because of eq. (12.5) we now have

$$0 = f(\exp Q_{n-1}) = \frac{\phi(\exp Q_{n-1})}{\psi(\exp Q_{n-1})},$$

meaning that $\phi(\exp Q_{n-1}) = 0$. But because of lemma 3.1 this is only possible if ϕ is the zero polynomial which is a contradiction. Hence f is identically 0. Now let us define

$$g(z) := f(z) \prod_{j=1}^{n-1} (c_j + z) = \sum_{i=1}^{n-1} d_i \prod_{j \neq i} (c_j + z)$$

Since eq. (12.4) is a non-trival linear combination we can find a coefficient, d_k that is non-zero. The polynomial g is also identically 0 so we get

$$0 = g(-c_k) = d_k \prod_{j \neq k}^{n-1} (c_j - c_k)$$

Since $d_k \neq 0$ we must have $c_j = c_k$ for $j \neq k$, meaning that $a_j b_j^{-1} = a_k b_k^{-1}$, so that $b_k = \frac{b_j a_k}{a_j}$. From eq. (12.1) we get

$$\begin{aligned} \frac{a_k}{a_j} \exp P_j &= \frac{a_k}{a_j}(a_j \exp P_{n-1} + b_j \exp P_n) \\ &= a_k \exp P_{n-1} + \frac{b_j a_k}{a_j} \exp P_n \\ &= \exp P_k \end{aligned}$$

But as we previously mentioned, both $\exp P_k$ and $\exp P_j$ have constant term 1 so we must have $\frac{a_k}{a_j} = 1$ and so $\exp P_j = \exp P_k$. Now, by eq. (10.4) we get

$$P_j = \log(\exp P_j) = \log(\exp P_k) = P_k$$

□

Suppose that F is an irreducible and nonfull form of degree m. This means that there is a number field, K, of degree $n > m$ and a module M of rank m inside K so that F is connected to M. Let $\mu_1, ..., \mu_m \in K$ be a basis for M. In the following it suffices to look at F up to integral equivalence so because of proposition 4.3 we can assume without loss of generality that

$$F(x_1, ..., x_m) = N_{K/\mathbb{Q}}(x_1\mu_1 + ... + x_m\mu_m)$$

Fix now some $c \in \mathbb{Q}$. We see that finding the solutions in integers to the equation

$$F(x_1, ..., x_m) = c \tag{12.6}$$

is exactly the same as finding the elements $\alpha \in M$ so that

$$N_{K/\mathbb{Q}}(\alpha) = c \tag{12.7}$$

Note that if $c = 0$ then $\alpha = 0$, so the above equation is not really interesting in this case. Thus we will from now on assume that $c \neq 0$. By proposition 4.4, the elements $\mu_1, ..., \mu_m$ are linearly independent over $\mathbb{Q}$, so we can extend them to a basis $\mu_1, ..., \mu_m, \mu_{m+1}, ...\mu_n$ for K over $\mathbb{Q}$. We then get a module $M' = \{\mu_1, ..., \mu_n\}$ and we clearly have $M \subseteq M'$. Similar to what we have done previously, we define

$$M'_c = \{\alpha \in M' \mid N_{K/\mathbb{Q}}(\alpha) = c\}$$
$$M_c = \{\alpha \in M \mid N_{K/\mathbb{Q}}(\alpha) = c\}$$

Let now $\sigma_1, ..., \sigma_n$ be the n embeddings of K into $\mathbb{C}$. Let us write $n = 2s + r$ where s is the number of pairs of complex conjugate embeddings and r is the number of real embeddings. By the primitive element theorem we know that $K = \mathbb{Q}(\beta)$ for some $\beta \in K$. Letting $N = \mathbb{Q}(\sigma_1(\beta), ..., \sigma_n(\beta))$ we get a subfield of $\mathbb{C}$ which is splitting field for the minimal polynomial for β over $\mathbb{Q}$ meaning that N is Galois over $\mathbb{Q}$, and it also has finite degree over $\mathbb{Q}$, so N is a number field. Let $\mathfrak{p}$ be a prime ideal of N. By definition 7.2 we get the local field $N_\mathfrak{p}$ with valuation ring $O_\mathfrak{p}$, which has unique maximal ideal $\hat{\mathfrak{p}}$ and so we define $U^{(k)} = 1 + \hat{\mathfrak{p}}^k$.

Lemma 12.2. *Let $\alpha \in N$ and assume $\hat{\mathfrak{p}} \nmid \alpha$. For any $k \in \mathbb{N}$ we can find $q \in \mathbb{N}$ so that $\alpha^q \in U^{(k)}$.*

Proof. Let $q = \#(O_\mathfrak{p}/\hat{\mathfrak{p}}^k)^*$. This is a natural number by lemma 8.1. The condition $\hat{\mathfrak{p}} \nmid \alpha$ means that $\gcd(\alpha O_N, \hat{\mathfrak{p}}) = O_N$ so also $\gcd(\alpha O_N, \hat{\mathfrak{p}}^k) = O_N$. But that means that $\alpha\beta + l = 1$ for some $l \in \hat{\mathfrak{p}}^k$ and $\beta \in O_N$, and so α is a unit in $O_\mathfrak{p}/\hat{\mathfrak{p}}^k$. Thus $\alpha^q \equiv 1 \pmod{\hat{\mathfrak{p}}^k}$ and so $\alpha^q \in U^{(k)}$. □

Using proposition 10.1 there is a natural number k so that the pair of homomorphisms, $\log : U^{(k)} \to \hat{\mathfrak{p}}^k$ and $\exp : \hat{\mathfrak{p}}^k :\to U^{(k)}$ are mutual inverses. Let $\mathfrak{D}$ be the coefficient ring of M'. Using theorem 4.3 we can pick $\phi_1, ..., \phi_t \in \mathfrak{D}$,

$t = r + s - 1$, to be a fundamental system of units for $\mathfrak{D}$, and a finite subset $\Gamma' \subseteq M_c'$ so that any $\alpha \in M_c$ can be written as

$$\alpha = \gamma' \phi_1^{v_1} ... \phi_t^{v_t},$$

for some $\gamma' \in \Gamma'$ and $v_i \in \mathbb{Z}$. Since $\mathfrak{D} \subseteq O_K$ by proposition 4.6 we have that the ϕ_i are also units in O_K. Thus $\sigma_j(\phi_i)$ is a unit in O_N so $\sigma_j(\phi_i)O_N = O_N$ meaning that $\hat{\mathfrak{p}} \nmid \sigma_j(\phi_i)$. Using lemma 12.2 there is a number q so that $\sigma_j(\phi_i)^q \in U^{(k)}$, so setting $\epsilon_i = \phi_i^q$ we now have $\sigma_j(\epsilon_i) \in U^{(k)}$. Let now $\Delta = \{\prod_{i=1}^t \phi_i^{\rho_i} \mid \rho_i \in \{0, ..., q-1\}\}$ and define $\Gamma = \{\delta\gamma \mid (\delta, \gamma) \in \Delta \times \Gamma'\}$. We see that we can now write the element α from above as

$$\alpha = \gamma \epsilon_1^{u_1} ... \epsilon_t^{u_t} \tag{12.8}$$

For an appropriate choice of u_i and $\gamma \in \Gamma$. Also note that $\#\Gamma < \infty$. Consider now the special case where $\alpha \in M_c \subseteq M_c'$ and suppose that it is written as in eq. (12.8). Since $\alpha \in M'$ we can also write $\alpha = x_1\mu_1 + ... + x_n\mu_n$ and since $\alpha \in M$, we must have

$$x_{m+1} = ... = x_n = 0 \tag{12.9}$$

This restriction can be reformulated to something that is more convenient to work with. Let $\mu_1^*, ..., \mu_n^*$ be the dual basis of $\mu_1, ..., \mu_n$. The following computation shows that we can recover the x_i's.

$$\mathrm{Tr}_{K/\mathbb{Q}}(\mu_i^* \alpha) = \mathrm{Tr}_{K/\mathbb{Q}}(\sum_{j=1}^n \mu_i^* x_j \mu_j) = \sum_{j=1}^n x_j \, \mathrm{Tr}_{K/\mathbb{Q}}(\mu_i^* \mu_j) = x_i$$

Therefore the above restriction is the same as the requirement that

$$\mathrm{Tr}_{K/\mathbb{Q}}(\mu_{m+1}^* \alpha) = ... = \mathrm{Tr}_{K/\mathbb{Q}}(\mu_n^* \alpha) = 0$$

Which, by use of eq. (12.8) and the definition of the trace, amounts to the equations

$$\sum_{i=1}^n \sigma_j(\gamma \mu_i^*) \sigma_j(\epsilon_1)^{u_1} ... \sigma_j(\epsilon_t)^{u_t} = 0 \text{ for all } i \in \{1, ..., n\} \tag{12.10}$$

Define now the function

$$\begin{aligned} \varphi : U^{(k)} \times O_{\mathfrak{p}} &\to U^{(k)} \\ (\alpha, u) &\mapsto \exp(u \log \alpha) \end{aligned}$$

This function is well defined since if $\alpha \in U^{(k)}$ and $u \in O_{\mathfrak{p}}$ then $u \log \alpha \in \hat{\mathfrak{p}}^k$ as $\hat{\mathfrak{p}}^k$ is an ideal. Note that if $u \in \mathbb{Z} \subseteq O_{\mathfrak{p}}$ we get $\varphi(\alpha, u) = \exp(\log \alpha^u) = \alpha^u$. For this reason we will simply write $\alpha^u := \varphi(\alpha, u)$. Assume now that all the

u_i are in $O_\mathfrak{p}$. Then eq. (12.10) still makes sense since we have made sure that $\sigma_j(\epsilon_i) \in U^{(k)}$. We can now define

$$L_j(u_1, ..., u_t) = \sum_{i=1}^{t} u_i \log \sigma_j(\epsilon_i), \text{ for } j \in \{1, ..., n\}$$
$$A_{ij} = \sigma_j(\gamma \mu_i^*), \text{ for } (i, j) \in \{1, ..., n\}^2$$

Note that the L_j can be thought of as power series in $P := N_\mathfrak{p}[[u_1, ..., u_t]]$. These have constant term 0 so it makes sense to consider the formal power series $\exp L_j(u_1, ..., u_t) \in P$. This means that we can define

$$G_i(u_1, ..., u_t) = \sum_{j=1}^{n} A_{ij} \exp L_j(u_1, ..., u_t) \text{ for } i \in \{m+1, ..., n\}$$

Note that the G_i all belong to P and also have constant term 0. The G_i also converge for all $(u_1, ..., u_t) \in O_\mathfrak{p}^t$, so these are analytic functions. Observe that we can now rewrite eq. (12.10) to

$$G_i(u_1, ..., u_t) = 0, \text{ for all } i \in \{m+1, ..., n\} \tag{12.11}$$

Assume that there are infinitely many solutions in integers to eq. (12.6). Thus we have infinitely many solutions of the shape eq. (12.8). But as Γ has only finitely many elements, there has to exists some $\gamma \in \Gamma$ giving rise to an infinite subset of solutions,

$$S_\gamma := \{\gamma \epsilon_1^{u_1} ... \epsilon_t^{u_t} \mid (u_1, ..., u_t) \in \mathbb{Z}^t\} \subseteq M_c$$

As the ϵ_i are independent units, we get an injective homomorphism $\iota : S_\gamma \to O_\mathfrak{p}^t$, given by $\gamma \epsilon_1^{u_1} ... \epsilon_t^{u_t} \mapsto (u_1, ..., u_t)$. That S_γ is infinite really just means that we have an injection $\mathbb{N} \hookrightarrow S_\gamma$, meaning that we have a sequence of distinct elements, (α_s), in S_γ. Let us now define the sequence $U_s = \iota(\alpha_s) \in O_\mathfrak{p}^t$. Since ι is injective, the elements of this sequence are also distinct. As the elements α_s all belong to M_c the tuples U_s satisfy eq. (12.11). In other words we have $G_i(U_s) = 0$ for all the G_i and all $s \in \mathbb{N}$. Since $O_\mathfrak{p}$ is compact by theorem 8.1, it follows from Tychonoff's theorem that $O_\mathfrak{p}^t$, equipped with the product topology, is also compact. But $O_\mathfrak{p}$ is a metric space so $O_\mathfrak{p}^t$ is metrizable and therefore it is sequentially compact. This means that the sequence U_s has a convergent subsequence, U_s^*, converging to a point, call it $u^* = (u_1^*, ..., u_t^*) \in O_\mathfrak{p}^t$. As the G_i are analytic functions they are continuous, so we know that $G_i(U_s^*)$ converges to $G_i(u^*)$ for $s \to \infty$. But also $G_i(U_s^*)$ converges to 0 since $G_i(U_s^*) = 0$ for all $s \in \mathbb{N}$. By uniqueness of limits in metric spaces we therefore have $G_i(u^*) = 0$. Also any neighborhood of u^* contains infinitely many solutions to eq. (12.11); Take any open ϵ-ball, B, with center u^*. Because U_s^* converges to u^*, we know that there is an $N \in \mathbb{N}$ so that $U_s^* \in B$ for all $s \geq N$, all of which are distinct and are solutions to eq. (12.11). Let us now write $u_i = u_i^* + v_i$, where $v_i \in O_\mathfrak{p}$.

We get

$$\begin{aligned} G_i(u_1, ..., u_t) &= \sum_{j=1}^{n} A_{ij} \exp L_j(u_1^* + v_1, ..., u_t^* + v_t) \\ &= \sum_{j=1}^{n} A_{ij} \exp L_j(u_1^*, ..., u_t^*) \exp L_j(v_1, ..., v_t) \end{aligned}$$

Setting $A_{ij}^* = A_{ij} \exp L_j(u_1^*, ..., u_t^*)$ and $H_i(v_1, ..., v_t) = \sum_{j=1}^{n} A_{ij}^* L_j(v_1, ..., v_t)$, we get $G_i(u_1, ..., u_t) = H_i(v_1, ..., v_t)$. We get yet another reformulation of eq. (12.9)

$$H_i(v_1, ..., v_t) = 0 \text{ for all } i \in \{m+1, ..., n\} \tag{12.12}$$

As the G_i are formal power series with constant term 0, the H_i are also formal power series with constant term 0. These converge on $O_{\mathfrak{p}}^t$ hence also on $\hat{\mathfrak{p}}^t$, which is an ϵ-neighborhood of the origin (with $\epsilon = 1$). Hence we get a local manifold, V, determined by the system eq. (12.12) and by what we showed above there are infinitely many solutions to this system in any neighborhood of the origin. Therefore, by theorem 11.1, V contains an analytic curve. We summarize what we have found in the following theorem.

Theorem 12.1. *If eq. (12.6) has infinitely many solutions in integers, then there exists $\gamma \in \Gamma$ which determines a local manifold, V, given by eq. (12.12) and V contains an analytic curve.*

Corollary 12.1. *Let $\boldsymbol{N}$ be the algebraic closure of $N_{\mathfrak{p}}$ and let $\omega_1(X), ..., \omega_t(X) \in O_{\boldsymbol{N}_{\mathfrak{p}}}[[X]]$ be any curve on the manifold V determined by eq. (12.12). Define $P_j(X) = L_j(\omega_1(X), ..., \omega_t(X))$ for $j \in \{1, ..., n\}$. There exists a matrix B_{ij} with coefficients in $\boldsymbol{N}$ so that the P_j satisfy the system of equations*

$$\begin{aligned} \sum_{j=1}^{n} A_{ij}^* \exp P_j(X) = 0, \text{ for all } i \in \{m+1, ..., n\} \\ \sum_{j=1}^{n} B_{ij} P_j(X) = 0, \text{ for all } i \in \{1, ..., n-t\}, \end{aligned} \tag{12.13}$$

The matrix (A_{ij}^*) has $\boldsymbol{N}$-linearly independent rows and (B_{ij}) can be chosen so that it also has $\boldsymbol{N}$-linearly independent rows.

Proof. Since $\omega_1(X), ..., \omega_t(X) \in O_{\boldsymbol{N}_{\mathfrak{p}}}[[X]]$ we have $P_j(X) \in \boldsymbol{N}[[X]]$. Because $\omega_1(X), ..., \omega_t(X)$ is a curve on V we have

$$H_i(\omega_1(X), ..., \omega_t(X)) = 0 \text{ for all } i \in \{m+1, ..., n\} \tag{12.14}$$

Considering how the H_i are defined we see that the above is just a different way of spelling out the first equation in eq. (12.13). The matrix (A_{ij}) is invertible because the $\gamma\mu_i^*$ constitute a basis for K over $\mathbb{Q}$ so the discriminant of this basis is non-zero and it equals $|\sigma_j(\gamma\mu_i^*)|^2 = |A_{ij}|^2$ so $|A_{ij}| \neq 0$. Since exp

sends elements to $U^{(k)}$ and $0 \notin U^{(k)}$ we have $\exp L_j(v_1, ..., v_t) \neq 0$ for all $(v_1, ..., v_t) \in O_{\mathfrak{p}}^t$. Since $A_{ij}^* = A_{ij} \exp L_j(u_1^*, ..., u_t^*) \in N_{\mathfrak{p}} \subseteq \boldsymbol{N}$ we now have

$$|A_{ij}^*| = |A_{ij}| \prod_{j=1}^{n} \exp L_j(u_1^*, ..., u_t^*) \neq 0$$

Thus the rows of (A_{ij}^*) are linearly independent over $\boldsymbol{N}$.

Let now $U = \text{span}_{\boldsymbol{N}}\{\omega_1(X), ..., \omega_t(X)\}$ and notice that each P_j is in U. This means we can define the following linear map

$$T : \boldsymbol{N}^n \to U$$
$$T(a_1, ..., a_n) = \sum_{i=1}^{n} a_i P_i$$

The dimension of U is at most t, so the rank of T is also at most t as T is surjective. By rank nullity, the nullity of T is therefore larger than or equal to $n - t$. In other words, there are at least $n - t$ linearly independent elements in the kernel of T and we can therefore construct an $n - t \times n$ matrix, (B_{ij}) with $\boldsymbol{N}$-linearly independent rows. □

We have yet another corollary

Corollary 12.2. *Keep the setup from above theorem 12.1 and consider the special case where $m = 2$, $n \geq 3$ and where at least one of the σ_i is a complex embedding. Then for any curve $\omega_1(X), ..., \omega_t(X)$ on V there exists indices k and l with $k \neq l$ so that $L_k(\omega_1(X), ..., \omega_t(X)) = L_l(\omega_1(X), ..., \omega_t(X))$.*

Proof. The existence of a complex embedding is the same as saying that $2s \geq 1$ which is the same as saying that $n - t \geq 2$. Setting $n_1 = n - m$ and $n_2 = n - t$ the result follows from combining lemma 12.1 and corollary 12.1. □

12.1 Proof of Thue's Theorem

Let us now apply the above to get a proof of Thue's theorem using Skolem's method. We first present the slightly weaker version of theorem 1.1 also mentioned in the introduction.

Theorem 12.2 (Thue)**.** *Suppose $f(x, y)$ is an irreducible form of degree $n \geq 3$ and that $f(x, 1)$ has an imaginary root. Then there are only finitely many integer solutions to the equation $f(x, y) = c$, for any non-zero $c \in \mathbb{Q}$.*

Proof. To prove this, we want to use the results obtained in this section and to do that it suffices to show that f is a nonfull form. Since f is a form of degree n we know that we can write

$$f(x, y) = \sum_{i=0}^{n} a_i x^{n-i} y^i$$

where the a_i are in $\mathbb{Z}$. We must have that $a_0 \neq 0$, since otherwise y would divide $f(x,y)$ which is impossible since f is irreducible. Let $h(x) = f(x,1)$. We have that h is a polynomial of degree n with leading coefficient a_0. The leading term of $a_0^{n-1}h(x)$ is $a_0^n x^n = (a_0 x)^n$, so h can be regarded as a monic polynomial in the variable t using the substitution $a_0 x \mapsto t$. So assume without loss of generality that h is monic, i.e. that $a_0 = 1$. It now follows from remark 4.1 that

$$f(x,y) = N_{K/\mathbb{Q}}(x + \theta y) \tag{12.15}$$

where $K = \mathbb{Q}(\theta)$ is a number field of degree n with embeddings $\sigma_1, ..., \sigma_n$ from K into $\mathbb{C}$. Consider the module M generated by $\mu_1 = 1$ and $\mu_2 = \theta$. These are linearly independent over $\mathbb{Q}$ so the rank, m, of this module is 2 and since $n \geq 3$ we have $m < n$. This shows that f is a nonfull form. Assume now that $f(x,y) = c$ has infinitely many solutions in integers. It now follows from the proof of theorem 12.1 that, using the notation from this proof, that we obtain $\gamma \in \Gamma$, with $\#\Gamma < \infty$ and independent units $\epsilon_1, ..., \epsilon_t \in \mathfrak{D}$ giving rise to a local manifold V determined by eq. (12.14). We get an injective map $\iota : S_\gamma \to O_\mathfrak{p}^t$ and a sequence (α_s) of distinct elements in S_γ so that $U_s^* = \iota(\alpha_s)$ and $U_s^* \to (u_1^*, ..., u_t^*) \in O_\mathfrak{p}^t$. Write $U_s^* = (u_{1s}, ..., u_{ts})$ and $u_{is} = u_i^* + v_{is}$ and define the sequence $V_s = (v_{1s}, ..., v_{ts})$. We have $V_s \in V$ and it converges to the origin. Let W be the local manifold given by

$$\prod_{i \leq k < l \leq n} (L_k(v_1, ...v_t) - L_l(v_1, ..., v_t)) = 0$$

Since $f(x,1)$ has at least one imaginary root, one of the embeddings, σ_i, has to be complex. So if we have any curve, $\tau_1(X), ..., \tau_t(X)$ on V then we can invoke corollary 12.2 to conclude that there exists distinct k and l so that $L_k(\tau_1(X), ..., \tau_t(X)) = L_l(\tau_1(X), ..., \tau_t(X))$ and so $\tau_1(X), ..., \tau_t(X)$ is also a curve on W. Using remark 11.1 we get that $V \subseteq W$, that is there is an ϵ-neighborhood around the origin, U, so that $U \cap V \subseteq W$. Since (V_s) converges to the origin there is an $N \in \mathbb{N}$ so that for all $s \geq N$ we have $V_s \in W$. Thus we can find distinct k and l so that $L_k(V_s) = L_l(V_s)$ for infinitely many $s \geq N$. Therefore there are infinitely many $s_1, s_2 \in \mathbb{N}$, $s_1 \neq s_2$ so that $L_k(V_{s_1}) = L_l(V_{s_2})$. Choose any such s_1 and s_2 and set $\alpha = \alpha_{s_1}$ and $\alpha' = \alpha_{s_2}$. Let us write $\alpha = x + y\theta$ and $\alpha' = x' + y'\theta$. Furthermore, since $\alpha \in S_\gamma$, we can write $\alpha = \gamma \epsilon_1^{u_1} ... \epsilon_t^{u_t}$. Set now

$$\begin{aligned} u_i &= u_i^* + v_i \\ c_j &= \sigma_j(\gamma)\sigma_j(\epsilon_1)^{u_1^*} ... \sigma_j(\epsilon_t)^{u_t^*} \end{aligned}$$

We obtain

$$\sigma_j(\alpha) = c_j \exp L_j(v_1, ..., v_t)$$

And in a similar fashion we get

$$\sigma_k(\alpha) = c_k \exp L_k(v_1, ..., v_t)$$

Since $L_j(v_1, ..., v_t) = L_k(v_1, ..., v_t)$ we have $\exp L_j(v_1, ..., v_t) = \exp L_k(v_1, ..., v_t)$. Also both c_j and c_k are different from 0 since otherwise we would have $\gamma = 0$ which would imply $c = 0$. Thus we have

$$\frac{\sigma_j(\alpha)}{c_j} = \frac{\sigma_k(\alpha)}{c_k}$$

Similarly we obtain

$$\frac{\sigma_j(\alpha')}{c_j} = \frac{\sigma_k(\alpha')}{c_k}$$

Since α and α' have norm $c \neq 0$ they are both non-zero. Hence we now have

$$\frac{x + y\sigma_j(\theta)}{x' + y'\sigma_j(\theta)} = \frac{x + y\sigma_k(\theta)}{x' + y'\sigma_k(\theta)}$$

So $(xy' - x'y)(\sigma_k(\theta) - \sigma_j(\theta)) = 0$. But σ_k and σ_j are distinct embeddings since $k \neq j$ and they are determined by their action on θ. Hence $\sigma_k(\theta) - \sigma_j(\theta) \neq 0$ so we must have

$$xy' - x'y = 0$$

Thus there is $d \in \mathbb{Q}$ so that $\alpha = d\alpha'$, so

$$N_{K/\mathbb{Q}}(\alpha) = N_{K/\mathbb{Q}}(d)N_{K/\mathbb{Q}}(\alpha')$$

But $N_{K/\mathbb{Q}}(\alpha) = N_{K/\mathbb{Q}}(\alpha')$ so $1 = N_{K/\mathbb{Q}}(d) = d^n$. Hence $d = \pm 1$ so $\alpha = \pm\alpha'$. But then it is impossible for all the elements of (α_s) to be distinct contrary to our assumption so S_γ cannot be infinite, meaning that there has to be only finitely many solutions to $f(x, y) = c$. □

13 Discussion of Thue's theorem

A natural question to ask when looking at any theorem is to ask which assumptions can be dropped or altered without changing the conclusion of the theorem too much. This is what we will explore in the case of theorem 12.2.

13.1 The number of variables

The following example shows that the theorem no longer holds if we swap out the form $f(x, y)$ with a form with more variables.

Example 13.1. *Consider the field $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. This field has degree 4 over $\mathbb{Q}$ and*

$$1, \sqrt{2}, \sqrt{3}, \sqrt{6}$$

is a basis for it by [3, p. 526]. Note that K is splitting field for the separable polynomial $(x^2 - 2)(x^2 - 3)$ so K is Galois over $\mathbb{Q}$. The automorphisms on K are determined by their action on $\sqrt{2}$ and $\sqrt{3}$ and these must be sent to $\pm\sqrt{2}$ and $\pm\sqrt{3}$ respectively. There are four ways of doing this, and as K is Galois

all of these possibilities yield an automorphism. Using this information we find that $F(x,y,z) := N_{K/\mathbb{Q}}(x+y\sqrt{2}+z\sqrt{3})$ *is equal to*

$$(x+y\sqrt{2}+z\sqrt{3})(x-y\sqrt{2}+z\sqrt{3})(x+y\sqrt{2}-z\sqrt{3})(x-y\sqrt{2}-z\sqrt{3}) \quad (13.1)$$

which after a (cumbersome) computation turns to be equal to

$$x^4+4y^4+9z^4-4x^2y^2-6x^2z^2-12y^2z^2$$

The module $\{1,\sqrt{2},\sqrt{3}\}$ *has rank 3 as the generators are linearly independent over* $\mathbb{Q}$*, so this module is nonfull and hence* F *is a nonfull form. It is also clearly decomposable as* F *splits into linear factors over* K*. Furthermore* F *is irreducible; Clearly, none of the linear factors in eq. (13.1) are forms. Hence if* F *were reducible, then there has to be a pair of these linear factors that multiply to a form. Consider for example*

$$(x+y\sqrt{2}+z\sqrt{3})(x-y\sqrt{2}+z\sqrt{3}) = x^2-2y^2+2\sqrt{3}xz+3z^2$$

This is certainly not a form as the coefficients are not in $\mathbb{Q}$*. As the isomorphisms in* $\mathrm{Gal}(K/\mathbb{Q})$ *permute the linear factors in eq. (13.1) this shows that no pair of linear factors multiply to a form and therefore* F *must be irreducible. Observe that*

$$\begin{aligned} F(x,y,0) &= x^4+4y^4-4x^2y^2 = (x^2-2y^2)^2 \\ &= N_{\mathbb{Q}(\sqrt{2})/\mathbb{Q}}(x+y\sqrt{2})^2 \end{aligned} \quad (13.2)$$

Consider the full module $\{1,\sqrt{2}\}$ *inside* $\mathbb{Q}(\sqrt{2})$*. Let's look at the equation*

$$N_{\mathbb{Q}(\sqrt{2})/\mathbb{Q}}(x+y\sqrt{2}) = 1$$

Applying lemma 4.3 we have $t = r+s-1 = 2+0-1 = 1$ *so there has to be infinitely many solutions in integers to this equation and therefore, because of eq. (13.2) there are infinitely many solutions to the equation* $F(x,y,z) = 1$*.*

In the above example we got an infinite number of solutions because the module $\{1,\sqrt{2},\sqrt{3}\}$ inside $\mathbb{Q}(\sqrt{2},\sqrt{3})$ contains the submodule $\{1,\sqrt{2}\}$, which is a full module when regarded as a module inside $\mathbb{Q}(\sqrt{2})$. The fact that we get infinitely many solutions can, more abstractly, be attributed to the fact that $\{1,\sqrt{2},\sqrt{3}\}$ is what is called a degenerate module. Let's define what we mean by this

Definition 13.1. *Let* K *be a number field and* M *a module with generators* $\mu_1,...,\mu_m$ *and consider the vector space* $L = \mathrm{span}_{\mathbb{Q}}\{\mu_1,...,\mu_m\}$*. If* L *contains a subspace* L' *so that* $\gamma K' = L'$ *for some subfield* K' *of* K *and* $\gamma \in K$ *and* K' *is neither* $\mathbb{Q}$ *or a quadratic imaginary field then we say that* M *is degenerate. Otherwise the module is called non degenerate.*

We see that the module $\{1, \sqrt{2}, \sqrt{3}\}$ inside $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is certainly degenerate since for example $\mathbb{Q}(\sqrt{2}) = \text{span}_{\mathbb{Q}}\{1, \sqrt{2}\}$. Also all full modules in a field different from $\mathbb{Q}$ or an imaginary quadratic field are degenerate - just consider the span over $\mathbb{Q}$ of any basis of such a module. For an example of a non-degenerate module look at the module $M = \langle 1, \beta \rangle$ from the proof of theorem 12.2. This follows from the following fact combined with the statement of theorem 12.2.

Proposition 13.1. *Assume K is a number field and M is a degenerate module. There exists $c \in \mathbb{Z}$ so that there are infinitely many solutions, $\alpha \in M$, to the equation*

$$N_{K/\mathbb{Q}}(\alpha) = c$$

Proof. By assumption we have that $K' = \gamma L'$, where K', L' and $\gamma \in K$ are as in definition 13.1. Set $M' = L' \cap M$. Then $\gamma^{-1}M'$ is a full module inside K' and since K' is neither $\mathbb{Q}$ or a quadratic imaginary field, we know from theorem 4.3 and remark 4.2 that there exists $c' \in \mathbb{Z}$ so that

$$N_{K'/\mathbb{Q}}(\xi) = c'$$

has infinitely many solutions $\xi \in \gamma^{-1}M'$ (take for example $c' = 1$). Take any such solution, ξ, and also set $c_\bullet = N_{K/\mathbb{Q}}(\gamma)c_\bullet^m$, where $m = [K : K']$. We get

$$N_{K/\mathbb{Q}}(\xi\gamma) = N_{K/\mathbb{Q}}(\xi)N_{K/\mathbb{Q}}(\gamma) = (N_{K'/\mathbb{Q}}(\xi))^m N_{K/\mathbb{Q}}(\gamma) = c$$

Since $\xi\gamma \in M' \subseteq M$ the result now follows. □

This raises the question: If K is a number field of degree n, $F(x_1, ..., x_m)$ is an irreducible form and $M = \langle \mu_1, ..., \mu_m \rangle$ is a non-degenerate module so that $F(x_1, ..., x_m) = N_{K/\mathbb{Q}}(x_1\mu_1, ..., x_m\mu_m)$ is it then always the case that there are only finitely many solutions to $F(x_1, ..., x_m) = c$? Note that since M is non-degenerate it must be non-full so we could try to utilize theorem 12.1.

References

[1] Alan Baker. *Transcendental Number Theory*. Cambride University Press, 1975. ISBN 978-1-009-22994-4.

[2] Z. I. Borevich and I. R. Shafarevich. *Number Theory*. Academic Press, New York, 1966.

[3] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Wiley, Hoboken, NJ, third edition, 2004. ISBN 0-471-43334-9.

[4] Mike Earnest. Composition of formal power series of exp and log, 2022. URL https://math.stackexchange.com/questions/4488730/composition-of-formal-power-series-of-exp-and-log. Accessed: 2024-05-21.

[5] Fernando Q. Gouvêa. *p-adic Numbers: An Introduction*. Springer, 3rd edition, 2020. ISBN 978-3-030-47294-8.

[6] Integrand. How to prove that the exponential function is the inverse of the natural logarithm by power series definition alone, 2016. URL https://math.stackexchange.com/questions/1199411/how-to-prove-that-the-exponential-function-is-the-inverse-of-the-natural-logarit. Accessed: 2024-05-21.

[7] Daniel A. Marcus. *Number Fields*. Universitext. Springer, Cham, second edition, 2018. ISBN 978-3-319-90232-6. doi: 10.1007/978-3-319-90233-3.

[8] Jürgen Neukirch. *Algebraic Number Theory*. Grundlehren der mathematischen Wissenschaften. Springer, 1999. ISBN 3-540-65399-6.

[9] Wikipedia. Thue equation, 2024. URL https://en.wikipedia.org/wiki/Thue_equation. Accessed: 2024-05-21.