Chapter 8

The p-adic Subspace Theorem

Literature:
B. Edixhoven, J.-H. Evertse (eds.), Diophantine Approximation and Abelian Varieties, Introductory Lectures, Lecture Notes in Mathematics 1566, Springer 1993, Chap.IV
J. Neukirch, Algebraic Number Theory, Springer 1999, Chaps. II, III

The p-adic Subspace Theorem deals with Diophantine inequalities in which several different absolute values occur (e.g., the ordinary absolute value and extensions to number fields of the p-adic value $|\cdot|_p$ various primes p). Before we are able to state the p-adic Subspace theorem we have to recall some facts about absolute values on number fields. We give only a brief outline. For more details we refer to Chapters II and III of Neukirch's book mentioned above.

8.1 Absolute values on algebraic number fields

8.1.1 Generalities

The standard absolute value $|\cdot|$ on $\mathbb{C}$ has certain important properties, namely, that it is positive for non-zero complex numbers, that the absolute value of the product of two complex numbers is the product of their absolute values, and that it satisfies the triangle inequality. There is a general concept of absolute values, which can be defined on any field.

Let K be a field. An *absolute value* (or valuation) on K is a function $|\cdot|_* : K \to$

$\mathbb{R}_{\geqslant 0}$ with the following properties:

(AV1) $|0|_* = 0$ and $|x|_* > 0$ for $x \in K^*$;
(AV2) $|xy|_* = |x|_* \cdot |y|_*$ for $x, y \in K$;
(AV3) $|x + y|_* \leqslant |x|_* + |y|_*$ for $x, y \in K$ (triangle inequality).

Notice that (AV1)–(AV3) imply that $|1|_* = 1$.

Examples. 1. The trivial absolute value, given by $|0|_* = 0$ and $|x|_* = 1$ for $x \in K^*$.

2. The standard absolute value $|\cdot|$ on any subfield of $\mathbb{C}$.

3. Let $K := k(X)$ be the field of rational functions over a field k. We define the degree of a rational function f/g with $f, g \in k[X]$ by $\deg(f/g) := \deg f - \deg g$. Then $|\cdot|_{\deg}$, given by $|x|_{\deg} := e^{\deg x}$ for $x \in K^*$ and $|0|_{\deg} := 0$ defines an absolute value on K. Notice that it satisfies something stronger than (AV3), i.e., $|x+y|_{\deg} \leqslant \max(|x|_{\deg}, |y|_{\deg})$ for $x, y \in K$ (verify this).

An absolute value $|\cdot|_*$ on a field K is called *non-archimedean* if instead of (AV3) it satisfies the *strong triangle inequality* or *ultrametric inequality*

(AV3') $|x + y|_* \leqslant \max(|x|_*, |y|_*)$ for $x, y \in K$.

If $|\cdot|_*$ does not satisfy (AV3') it is called *archimedean.*

Two absolute values $|\cdot|_*$, $|\cdot|_{**}$ on K are called *equivalent* if there is $c > 0$ such that

$$|x|_{**} = |x|_*^c \text{ for all } x \in K.$$

An absolute value $|\cdot|_*$ on K defines a topology on K as follows: the open sets are those subsets U of K with the property that for every $a \in U$ there is $\delta > 0$ such that $\{x \in K : |x - a|_* < \delta\} \subseteq U$. It can be shown that two absolute values on K define the same topology on K if and only if they are equivalent.

Let K be a field, L an extension of K, and $|\cdot|_*$ an absolute value on K. A *continuation* of K to L is an absolute value $|\cdot|_{**}$ on L whose restriction to K is $|\cdot|_*$, i.e., $|x|_{**} = |x|_*$ for all $x \in K$. We mention here that such a continuation, if it exists, need not be unique.

Below, we first describe the absolute values on $\mathbb{Q}$, and subsequently their continuations to a number field.

8.1.2 Absolute values on $\mathbb{Q}$

Of course, we have on $\mathbb{Q}$ the standard absolute value, given by

$$|x| := \max(x, -x) \ \text{ for } x \in \mathbb{Q}.$$

Further, for every prime number p there is an associated p-adic absolute value, which is defined as follows. First we put $|0|_p := 0$. If x is a non-zero rational number, then we can write $x = p^k a/b$ where k is an integer and a, b are integers coprime with p, and we put

$$|x|_p := p^{-k}.$$

Verify yourself that $|\cdot|_p$ defines a non-archimedean absolute value on $\mathbb{Q}$.

It will be convenient to denote the standard absolute value on $\mathbb{Q}$ by $|\cdot|_\infty$ and write

$$M_{\mathbb{Q}} := \{\infty\} \cup \{\text{prime numbers}\}.$$

Then the absolute values defined above satisfy the *product formula*

$$\prod_{p \in M_{\mathbb{Q}}} |x|_p = 1 \ \text{ for } x \in \mathbb{Q}^*. \tag{8.1}$$

Indeed, every non-zero rational number x has a unique factorization as a product of prime powers

$$x = \pm p_1^{k_1} \cdots p_t^{k_t}$$

where $k_1, \ldots, k_t$ are non-zero integers and $p_1, \ldots, p_t$ distinct prime numbers. We have $|x|_\infty = p_1^{k_1} \cdots p_t^{k_t}$, $|x|_{p_i} = p_i^{-k_i}$ for $i = 1, \ldots, t$, and $|x|_p = 1$ for every prime p outside $\{p_1, \ldots, p_t\}$.

We state without proof the following

Theorem 8.1 (Ostrowski). *The absolute values $|\cdot|_p$ ($p \in M_{\mathbb{Q}}$) are pairwise inequivalent, and every non-trivial absolute value on $\mathbb{Q}$ is equivalent to one of them.*

8.1.3 Absolute values on number fields

Let K be an algebraic number field. Each of the absolute values $|\cdot|_p$ ($p \in M_{\mathbb{Q}}$) defined above has a continuation (and in general more than one) to K. In most applications, we only need to know that such continuations exist and not how they are defined, but for the interested reader we describe these continuations.

We first describe the continuations of the standard absolute value $|\cdot|_\infty = |\cdot|$ to K. Let $d := [K:\mathbb{Q}]$. Recall that K has precisely d embeddings $K \hookrightarrow \mathbb{C}$. We order these embeddings as

$$\sigma_1, \ldots, \sigma_{r_1},\ \sigma_{r_1+1}, \ldots, \sigma_{r_1+r_2},\ \sigma_{r_1+r_2+1} = \overline{\sigma_{r_1+1}}, \ldots, \sigma_{r_1+2r_2} = \overline{\sigma_{r_1+r_2}},$$

where $d = r_1 + 2r_2$, σ_i $(i = 1, \ldots, r_1)$ are the embeddings with image contained in $\mathbb{R}$, σ_i $(i = r_1 + 1, \ldots, d)$ are the embeddings with images not contained in $\mathbb{R}$, and $\overline{\sigma}(x) := \overline{\sigma(x)}$ for $x \in K$. Denote by $|\cdot|$ the standard absolute value on $\mathbb{C}$. Then

$$|\cdot|_{\sigma_i} := |\sigma_i(\cdot)| \quad (i = 1, \ldots, d)$$

define archimedean absolute values on K, which are clearly continuations of $|\cdot|_\infty$ to K. In fact, we only have to concider these absolute values for $i = 1, \ldots, r_1 + r_2$, since

$$|\cdot|_{\sigma_{r_2+j}} = |\sigma_{r_2+j}(\cdot)| = |\,\overline{\sigma_j(\cdot)}\,| = |\sigma_j(\cdot)| = |\cdot|_{\sigma_j} \ \text{ for } j = r_1 + 1, \ldots, r_1 + r_2.$$

Fact. *The absolute values $|\cdot|_{\sigma_i}$ $(i = 1, \ldots, r_1 + r_2)$ are pairwise inequivalent, and they are precisely the continuations of $|\cdot|_\infty$ to K.*

In order to define the continuations of $|\cdot|_p$ (p prime number) to K, we have to recall some facts about prime ideal decompositions of fractional ideals of K.

Denote by O_K the ring of integers of K. A *fractional ideal* of K is a set $\mathfrak{a} \neq 0$ with the following properties:

(i) if $x, y \in \mathfrak{a}$, then $x - y \in \mathfrak{a}$;
(ii) if $x \in \mathfrak{a}$, $\alpha \in O_K$, then $\alpha x \in \mathfrak{a}$;
(iii) there is $\alpha \in K^*$ such that $\alpha\mathfrak{a} := \{\alpha x : x \in \mathfrak{a}\} \subseteq O_K$.

In particular O_K itself and the non-zero ideals of O_K are fractional ideals of K, and in fact, any fractional ideal of K contained in O_K is a non-zero ideal of O_K. Further, for $\alpha \in K^*$, $\alpha O_K := \{\alpha x : x \in O_K\}$ is a fractional ideal of K.

We define the product of two fractional ideals $\mathfrak{a}$, $\mathfrak{b}$ of K by

$$\mathfrak{a} \cdot \mathfrak{b} := \Big\{ \sum_{\text{finite}} x_i y_i : x_i \in \mathfrak{a},\ y_i \in \mathfrak{b} \Big\}.$$

Further, we define the inverse of a fractional ideal $\mathfrak{a}$ of K by

$$\mathfrak{a}^{-1} := \{x \in K : x\mathfrak{a} \subseteq O_K\}.$$

The following are standard facts from algebraic number theory, which we recall without proof.

Theorem 8.2. *The product of two fractional ideals of K and the inverse of a fractional ideal of K are again fractional ideals of K. With this product and inverse, the fractional ideals of K form an abelian group, with unit element O_K.*

Corollary 8.3. *Let $\mathfrak{a}$, $\mathfrak{b}$ be fractional ideals of K. Then $\mathfrak{a} \subseteq \mathfrak{b}$ if and only if there is a non-zero ideal $\mathfrak{c}$ of O_K such that $\mathfrak{a} = \mathfrak{b} \cdot \mathfrak{c}$.*

Proof. We have the chain of equivalences

$$\begin{aligned}\mathfrak{a} \subseteq \mathfrak{b} \iff & \mathfrak{a} \cdot \mathfrak{b}^{-1} \subseteq \mathfrak{b} \cdot \mathfrak{b}^{-1} = O_K \iff \mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c} \text{ for some non-zero ideal } \mathfrak{c} \text{ of } O_K \\ \iff & \mathfrak{a} = \mathfrak{b} \cdot \mathfrak{c} \text{ for some non-zero ideal } \mathfrak{c} \text{ of } O_K.\end{aligned}$$

□

A prime ideal of O_K is an ideal $\mathfrak{p}$ with the property that whenever a product $\alpha\beta$ with $\alpha, \beta \in O_K$ belongs to $\mathfrak{p}$, then at least one of α, β belongs to $\mathfrak{p}$. Clearly $\{0\}$ is a prime ideal of O_K. It is known that every non-zero prime ideal of O_K is a maximal ideal of O_K.

Theorem 8.4. *(i) Every non-zero ideal of O_K can be expressed uniquely as a product $\mathfrak{p}_1^{k_1} \cdots \mathfrak{p}_t^{k_t}$, where $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$ are distinct non-zero prime ideals of O_K and $k_1, \ldots, k_t$ positive integers.*

(ii) Every fractional ideal of K can be expressed uniquely as a product $\mathfrak{p}_1^{k_1} \cdots \mathfrak{p}_t^{k_t}$, where $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$ are distinct non-zero prime ideals of O_K and $k_1, \ldots, k_t$ non-zero integers.

Now let p be a prime number. Then

$$pO_K = \mathfrak{p}_1^{e(\mathfrak{p}_1)} \cdots \mathfrak{p}_g^{e(\mathfrak{p}_g)},$$

where $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$ are distinct prime ideals of O_K, called the prime ideals dividing p, and $e(\mathfrak{p}_1), \ldots, e(\mathfrak{p}_g)$ are positive integers, called the *ramification indices* of $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$.

We notice that a prime ideal $\mathfrak{p}$ of O_K cannot divide two different primes p, q. For otherwise, $p, q \in \mathfrak{p}$ which would imply $1 = \gcd(p, q) \in \mathfrak{p}$ which is impossible.

Every $x \in K^*$ gives rise to a unique prime ideal decomposition

$$xO_K = \mathfrak{p}_1^{k_1} \cdots \mathfrak{p}_g^{k_g} \cdot \mathfrak{a} \tag{8.2}$$

where $k_1, \ldots, k_g$ are integers and $\mathfrak{a}$ is a fractional ideal of K composed of prime ideals other than $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$, and for $i = 1, \ldots, g$ we set

$$|x|_{\mathfrak{p}_i} := p^{-k_i/e(\mathfrak{p}_i)}.$$

Further, we define $|0|_{\mathfrak{p}_i} := 0$ for $i = 1, \ldots, g$.

Lemma 8.5. *Each $|\cdot|_{\mathfrak{p}_i}$ $(i = 1, \ldots, g)$ defines a non-archimedean absolute value on K, and is a continuation of $|\cdot|_p$ to K.*

Proof. We first show that $|\cdot|_{\mathfrak{p}_i}$ $(i = 1, \ldots, g)$ are continuations of $|\cdot|_p$ to K. Let $x \in \mathbb{Q}^*$; then $x = \pm p^k q_1^{l_1} \cdots q_s^{l_s}$ with $k, l_1, \ldots, l_s$ integers and $q_1, \ldots, q_s$ primes different from p. Thus,

$$xO_K = (\mathfrak{p}_1^{e(\mathfrak{p}_1)} \cdots \mathfrak{p}_g^{e(\mathfrak{p}_g)})^k \mathfrak{a}$$

where $\mathfrak{a}$ is composed of prime ideals of O_K dividing $q_1, \ldots, q_s$, and so of prime ideals other than $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$. Hence $|x|_{\mathfrak{p}_i} = p^{-ke(\mathfrak{p}_i)/e(\mathfrak{p}_i)} = p^{-k} = |x|_p$ for $i = 1, \ldots, g$,

We now show that $|\cdot|_{\mathfrak{p}_i}$ $(i = 1, \ldots, g)$ define non-archimedean absolute values on K. Let $\mathfrak{p} \in \{\mathfrak{p}_1, \ldots, \mathfrak{p}_g\}$. We have to prove that $|xy|_{\mathfrak{p}} = |x|_{\mathfrak{p}} \cdot |y|_{\mathfrak{p}}$, $|x + y|_{\mathfrak{p}} \leqslant \max(|x|_{\mathfrak{p}}, |y|_{\mathfrak{p}})$ for $x, y \in K$. This is clear if one of $x, y, x+y$ is 0. Assume that these numbers are all non-zero. Then $|x|_{\mathfrak{p}} = p^{-k/e(\mathfrak{p})}$, $|y|_{\mathfrak{p}} = p^{-l/e(\mathfrak{p})}$ for certain integers k, l, which means that

$$xO_K = \mathfrak{p}^k \mathfrak{a}, \quad yO_K = \mathfrak{p}^l \mathfrak{b} \tag{8.3}$$

for certain fractional ideals $\mathfrak{a}$, $\mathfrak{b}$ composed of prime ideals other than $\mathfrak{p}$. Clearly, $xyO_K = \mathfrak{p}^{k+l}\mathfrak{a}\mathfrak{b}$ and $\mathfrak{a}\mathfrak{b}$ is also composed only of prime ideals different from $\mathfrak{p}$. So $|xy|_{\mathfrak{p}} = p^{-(k+l)/e(\mathfrak{p})} = |x|_{\mathfrak{p}} \cdot |y|_{\mathfrak{p}}$.

We now prove the inequality for $x+y$. For the moment we assume that $x, y \in O_K$. Then (8.3) holds with $k, l \geqslant 0$ and $\mathfrak{a}$, $\mathfrak{b}$ ideals of O_K. We assume without loss of generality that $|y|_{\mathfrak{p}} \leqslant |x|_{\mathfrak{p}}$, or equivalently, $k \leqslant l$. By Corollary 8.3 we have

$$x \in \mathfrak{p}^k \mathfrak{a} \subseteq \mathfrak{p}^k, \quad y \in \mathfrak{p}^l \mathfrak{b} \subseteq \mathfrak{p}^k,$$

hence $x+y \in \mathfrak{p}^k$. Again from Corollary 8.3 we deduce that $(x+y)O_K = \mathfrak{p}^k\mathfrak{c}$ for some non-zero ideal $\mathfrak{c}$ of O_K. We can write $\mathfrak{c} = \mathfrak{p}^m\mathfrak{d}$ where m is a non-negative integer and $\mathfrak{d}$ is composed of prime ideals other than $\mathfrak{p}$, so $(x+y)O_K = \mathfrak{p}^{k+m}\mathfrak{d}$. This leads to

$$|x+y|_{\mathfrak{p}} = p^{-(k+m)/e(\mathfrak{p})} \leqslant p^{-k/e(\mathfrak{p})} = |x|_{\mathfrak{p}} = \max(|x|_{\mathfrak{p}}, |y|_{\mathfrak{p}}).$$

In case that x, y are not both in O_K choose a positive integer α such that $\alpha x, \alpha y \in O_K$. Then $|\alpha x+\alpha y|_{\mathfrak{p}} \leqslant \max(|\alpha x|_{\mathfrak{p}}, |\alpha y|_{\mathfrak{p}})$, and by dividing by $|\alpha|_{\mathfrak{p}}$ we obtain $|x+y|_{\mathfrak{p}} \leqslant \max(|x|_{\mathfrak{p}}, |y|_{\mathfrak{p}})$. □

Fact. *The absolute values $|\cdot|_{\mathfrak{p}_i}$ $(i=1,\dots,g)$ are pairwise inequivalent, and they are precisely the continuations of $|\cdot|_p$ to K.*

8.2 The p-adic Subspace Theorem and some applications

Recall that $|\cdot|_\infty$ denotes the standard absolute value on $\mathbb{Q}$, and for each prime number p, $|\cdot|_p$ denotes the p-adic absolute value as defined above. In this section, the following notation is used:

$$(8.4) \qquad \left\{ \begin{array}{l} K \text{ is a number field, } p_1,\dots,p_s \text{ are distinct prime numbers,} \\ \text{for each } p \in \{\infty, p_1,\dots,p_s\} \text{ we choose a continuation of } |\cdot|_p \text{ to } K \\ \text{which we denote also by } |\cdot|_p. \end{array} \right\}$$

We start with a generalization of Roth's Theorem. Recall that if $\xi = x/y$ with x, y coprime integers, then $H(\xi) = \max(|x|, |y|)$.

Theorem 8.6. (p-adic Roth's Theorem). *Let $\kappa > 2$, $C > 0$. Further, for each $p \in \{\infty, p_1,\dots,p_s\}$ let $\alpha_p \in K$. Then the inequality*

$$(8.5) \qquad |\alpha_\infty - \xi|_\infty \cdot |\alpha_{p_1} - \xi|_{p_1} \cdots |\alpha_{p_s} - \xi|_{p_s} \leqslant C \cdot H(\xi)^{-\kappa} \text{ in } \xi \in \mathbb{Q}$$

has only finitely many solutions.

Example. Let α be a real, irrational algebraic number, and p a prime number. We show that for every $\varepsilon > 0$, $C > 0$ there are only finitely many pairs of integers (u, y)

such that $u > 0$, y is coprime with p and

$$\left|\alpha - \frac{p^u}{y}\right| \leqslant C \cdot (\max(p^u, |y|))^{-1-\varepsilon} \tag{8.6}$$

or equivalently

$$|\alpha - \xi| \leqslant C \cdot H(\xi)^{-1-\varepsilon},$$

where $\xi = p^u/y$. Notice that Roth's Theorem from Theorem 6 gives only the finiteness for the number of solutions $\xi = p^u/y$ if we replace the exponent on $H(\xi)$ by something smaller than -2; so if we restrict ourselves to solutions $\xi = p^u/y$ we get a significant improvement.

We first observe that if (u, y) is a solution with $|p^u/y| < |\frac{1}{2}\alpha|$, then $0 < |\frac{1}{2}\alpha| < C\cdot(\max(p^u, |y|))^{-1-\varepsilon}$. So there are only finitely many such pairs (u, y). Now consider the solutions (u, y) with $|p^u/y| \geqslant |\frac{1}{2}\alpha|$, i.e., $|y| \leqslant |\dfrac{2}{\alpha}| \cdot p^u$. Then we get

$$\begin{aligned} |\alpha - \xi| \cdot |\xi|_p &= |\alpha - \xi| \cdot p^{-u} \leqslant Cp^{-u} \max(p^u, |y|)^{-1-\varepsilon} \leqslant C' \max(p^u, |y|)^{-2-\varepsilon} \\ &\leqslant C' H(\xi)^{-2-\epsilon} \end{aligned}$$

for some constant $C' > 0$. We apply Theorem 8.6 with $K = \mathbb{Q}(\alpha) \subset \mathbb{R}$, with the single prime p, with $|\cdot|_\infty$ the restriction to K of the absolute value $|\cdot|$ on $\mathbb{R}$, with any continuation of $|\cdot|_p$ to K (which one doesn't matter), with $\alpha_\infty = \alpha$ and with $\alpha_p = 0$. It follows that the latter inequality, and hence (8.6), has only finitely many solutions.

We now formulate the p-adic Subspace Theorem. We keep notation (8.4). The p-adic Subspace Theorem involves for each $|\cdot|_p$ with $p \in \{\infty, p_1, \ldots, p_s\}$ a system of n linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in K.

Theorem 8.7. (p-adic Subspace Theorem, Schlickewei, 1976). *Let $n \geqslant 2$, $\varepsilon > 0$, $C > 0$, and for each $p \in \{\infty, p_1, \ldots, p_s\}$ let $L_{1,p}, \ldots, L_{n,p}$ be linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in K. Consider the inequality*

$$|L_{1,\infty}(\mathbf{x}) \cdots L_{n,\infty}(\mathbf{x})|_\infty \cdot \prod_{j=1}^{s} |L_{1,p_j}(\mathbf{x}) \cdots L_{n,p_j}(\mathbf{x})|_{p_j} \leqslant C \cdot \|\mathbf{x}\|^{-\varepsilon} \quad \textit{in } \mathbf{x} \in \mathbb{Z}^n\,. \tag{8.7}$$

There are a finite number of proper linear subspaces $T_1, \ldots, T_t$ of $\mathbb{Q}^n$ such that all solutions of (8.7) lie in $T_1 \cup \cdots \cup T_t$.

Remark. Similarly as for the Subspace Theorem from Chapter 7, the only available proofs of Theorem 8.7 are ineffective, that is that they do not provide a method to determine the subspaces $T_1, \dots, T_t$.

Theorem 8.7 $\Longrightarrow$ Theorem 8.6. Let ξ be a solution of (8.5). Write $\xi = x/y$ with $x, y \in \mathbb{Z}$, $\gcd(x, y) = 1$. Multiply (8.5) with $A := \big(|y| \cdot |y|_{p_1} \cdots |y|_{p_s}\big)^2$. Notice that $|y|_{p_j} \leqslant 1$ for $j = 1, \dots, s$. Hence $A \leqslant y^2 \leqslant H(\xi)^2$. Let $\varepsilon = \kappa - 2$. Then (8.5) implies

$$|(x - \alpha_\infty y)y|_\infty \cdot \prod_{j=1}^{s} |(x - \alpha_{p_j} y)y|_{p_j} \leqslant C \cdot \max(|x|, |y|)^{-\varepsilon}.$$

The solutions $(x, y) \in \mathbb{Z}^2$ of the latter lie in only finitely many proper one-dimensional linear subspaces of $\mathbb{Q}^2$, and each of these gives rise to a single fraction $\xi = x/y$. So (8.5) has only finitely many solutions. □

Remark. In many applications of Theorem 8.7, we let $K \subset \mathbb{C}$, we choose $|\cdot|_\infty$ to be the restriction to K of the standard absolute value $|\cdot|$ on $\mathbb{C}$, and for $p \in \{p_1, \dots, p_s\}$ we let $L_{i,p}$ be linear forms with coefficients in $\mathbb{Q}$ so that it is irrelevant which continuation of $|\cdot|_p$ to K we choose.

Example. Let $\varepsilon > 0$. We show that the inequality

$$(8.8) \qquad |2^u + 3^v - 5^w| \leqslant \max(|2^u|, |3^v|, |5^w|)^{1-\varepsilon}$$

has only finitely many solutions in non-negative integers u, v, w. We apply the p-adic Subspace theorem with $K = \mathbb{Q}$ and with the primes $2, 3, 5$. Write $x_1 = 2^u$, $x_2 = 3^v$, $x_3 = 5^w$, $\mathbf{x} = (x_1, x_2, x_3)$. We first show that the set of solutions $\mathbf{x}$ lies in the union of finitely many proper linear subspaces of $\mathbb{Q}^3$. Consider for the moment those solutions for which $\|\mathbf{x}\| = |x_3|$. Notice that

$$|x_1x_2x_3|_2 \cdot |x_1x_2x_3|_3 \cdot |x_1x_2x_3|_5 = 2^{-u}3^{-v}5^{-w} = |x_1x_2x_3|^{-1}.$$

In combination with (8.8), this gives

$$|(x_1 + x_2 - x_3)x_1x_2| \cdot |x_1x_2x_3|_2 \cdot |x_1x_2x_3|_3 \cdot |x_1x_2x_3|_5 \leqslant |x_3|^{-1}\|\mathbf{x}\|^{1-\varepsilon} \leqslant \|\mathbf{x}\|^{-\varepsilon}.$$

By Theorem 8.7 with $K = \mathbb{Q}$, the solutions of the latter inequality lie in the union of finitely many proper linear subspaces of $\mathbb{Q}^3$. So the solutions of (8.8) with $\|\mathbf{x}\| = |x_3|$ lie in finitely many proper linear subspaces of $\mathbb{Q}^3$. In a similar way one proves that the solutions with $\|\mathbf{x}\| = |x_1|$ or with $\|\mathbf{x}\| = |x_2|$ lie in finitely many proper linear subspaces of $\mathbb{Q}^3$.

It is left as an exercise to prove that if T is a two-dimensional linear subspace of $\mathbb{Q}^3$ then T contains only finitely many solutions of (8.8). $\square$

Similarly as for the basic Subspace Theorem discussed in Chapter 7, there is a version with linear forms in general position. We keep again notation (8.4).

Theorem 8.8. *Let $\varepsilon > 0$, $C > 0$, and for $p \in \{\infty, p_1, \ldots, p_s\}$ let $L_{1,p}, \ldots, L_{r_p,p}$ ($r_p \geqslant n$) be linear forms in $X_1, \ldots, X_n$ in general position with coefficients in K. Consider the inequality*

$$(8.9)\ |L_{1,\infty}(\mathbf{x}) \cdots L_{r_\infty,\infty}(\mathbf{x})|_\infty \cdot \prod_{j=1}^{s} |L_{1,p_j}(\mathbf{x}) \cdots L_{r_j,p_j}(\mathbf{x})|_{p_j} \leqslant C \cdot \|\mathbf{x}\|^{r_\infty - n - \varepsilon}$$

$$\text{in } \mathbf{x} \in \mathbb{Z}^n \text{ with } gcd(x_1, \ldots, x_n) = 1.$$

Then there are a finite number of proper linear subspaces $T_1, \ldots, T_t$ of $\mathbb{Q}^n$ such that all solutions of (8.9) *lie in $T_1 \cup \cdots \cup T_t$.*

Lemma 8.9. *Let L be a field and $|\cdot|_*$ an absolute value on L. Let $M_1, \ldots, M_n$ be linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in L. Then there is a constant $C' > 0$ such that for $\mathbf{x} = (x_1, \ldots, x_n) \in L^n$ we have*

$$\max_{1 \leqslant k \leqslant n} |x_k|_* \leqslant C' \max_{1 \leqslant j \leqslant n} |M_j(\mathbf{x})|_*.$$

Proof. The same as that of Lemma 7.4 (verify this). $\square$

Proof of Theorem 8.8. For every solution $\mathbf{x} \in \mathbb{Z}^n$ of (8.9) and $p \in \{\infty, p_1, \ldots, p_s\}$, we choose a permutation σ_p of $1, \ldots, r_p$ such that $|L_{\sigma_p(1),p}(\mathbf{x})|_p \leqslant \cdots \leqslant |L_{\sigma_p(r_p),p}(\mathbf{x})|_p$. It clearly suffices to show that the solutions $\mathbf{x}$ corresponding to given permutations σ_p lie in finitely many proper linear subspaces of $\mathbb{Q}^n$.

Consider for instance the solutions $\mathbf{x} \in \mathbb{Z}^n$ of (8.9) such that

$$|L_{1,p}(\mathbf{x})|_p \leqslant \cdots \leqslant |L_{r_p,p}(\mathbf{x})|_p \text{ for } p \in \{\infty, p_1, \ldots, p_s\}.$$

Since $L_{1,p}, \ldots, L_{n-1,p}, L_{i,p}$ are linearly independent, by Lemma 8.9 there is a constant $C_{i,p}$ depending only on the coefficients of the linear forms $L_{j,p}$ ($j = 1, \ldots, n-1, i$) such that

$$\max_k |x_k|_p \leqslant C_{i,p} |L_{i,p}(\mathbf{x})|_p.$$

If $p = \infty$, then $\max_k |x_k|_p = \|\mathbf{x}\|$. If p is one of $p_1, \ldots, p_s$, then since $\gcd(x_1, \ldots, x_n) = 1$ at least one of $x_1, \ldots, x_n$ is not divisible by p, which implies $\max_k |x_k|_p = 1$. So we have

$$\begin{aligned} &\|\mathbf{x}\| \leqslant C_{i,\infty} |L_{i,\infty}(\mathbf{x})|_\infty \ \text{ for } i = n+1, \ldots, r_\infty, \\ &1 \leqslant C_{i,p_j} |L_{i,p_j}(\mathbf{x})|_{p_j} \text{ for } j = 1, \ldots, s, \ i = n+1, \ldots, r_{p_j}. \end{aligned}$$

By combining these inequalities with (8.9), we obtain

$$|L_{1,\infty}(\mathbf{x}) \cdots L_{n,\infty}(\mathbf{x})|_\infty \cdot \prod_{j=1}^{s} |L_{1,p_j}(\mathbf{x}) \cdots L_{n,p_j}(\mathbf{x})|_{p_j} \leqslant C'' \|\mathbf{x}\|^{-\varepsilon}$$

for some constant $C'' > 0$. Now apply Theorem 8.7 to the latter. □

Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a square-free binary form of degree $n \geqslant 3$ and $p_1, \ldots, p_s$ distinct prime numbers. We consider the so-called *Thue-Mahler equation*

$$(8.10) \qquad |F(x,y)| = p_1^{z_1} \cdots p_s^{z_s} \text{ in } x, y, z_1, \ldots, z_s \in \mathbb{Z} \text{ with } \gcd(x,y) = 1.$$

Notice that if we drop the condition $\gcd(x,y) = 1$ it is possible to construct infinitely many solutions from a given solution. We prove the following.

Theorem 8.10. (Mahler, 1933). *Equation* (8.10) *has only finitely many solutions.*

We use the following important fact.

Lemma 8.11. *Let $u \in \mathbb{Q}$. Then $u = \pm p_1^{w_1} \cdots p_s^{w_s}$ for certain integers $w_1, \ldots, w_s$ if and only if $|u| \cdot |u|_{p_1} \cdots |u|_{p_s} = 1$.*

Proof. Trivial. □

Proof of Theorem 8.10. We can factor F as $a_0(X - \alpha_1 Y) \cdots (X - \alpha_n Y)$ with $\alpha_1, \ldots, \alpha_n$ distinct if $F(1,0) \neq 0$, and as $a_0 Y(X - \alpha_1 Y) \cdots (X - \alpha_{n-1} Y)$ with $\alpha_1, \ldots, \alpha_{n-1}$ distinct if $F(1,0) = 0$. So in both cases we have $F(X,Y) = \prod_{i=1}^{n} (\beta_i X - \gamma_i Y)$ where the linear forms $\beta_i X - \gamma_i Y$ $(i = 1, \ldots, n)$ are in general position. Let $K = \mathbb{Q}(\beta_1, \gamma_1, \ldots, \beta_n, \gamma_n)$ and keep notation (8.4). Take ε with $0 < \varepsilon < n - 2$. Then by Lemma 8.11 we have for any solution $(x, y, z_1, \ldots, z_s)$ of (8.10),

$$|F(x,y)| \cdot \prod_{j=1}^{s} |F(x,y)|_{p_j} = 1 \leqslant \max(|x|, |y|)^{n-2-\varepsilon},$$

hence

$$|(\beta_1 x-\gamma_1 y)\cdots(\beta_n x-\gamma_n y)|_\infty\cdot\prod_{j=1}^{s}|(\beta_1 x-\gamma_1 y)\cdots(\beta_n x-\gamma_n y)|_{p_j}\leqslant \max(|x|,|y|)^{n-2-\varepsilon}.$$

By Theorem 8.8, the set of solutions $(x, y) \in \mathbb{Z}^2$ of this inequality lies in the union of finitely many one-dimensional linear subspaces of $\mathbb{Q}^2$. Each such subspace contains only two solutions with $\gcd(x, y) = 1$. This proves that (8.10) has only finitely many solutions. □

Remark. The above proof of the finiteness of the number of solutions of the Thue-Mahler equation is based on the p-adic Subspace Theorem and is therefore ineffective. There is however an alternative, effective proof of Theorem 8.10. There are effective lower bounds for the p-adic absolute value of linear forms in p-adic logarithms of algebraic numbers, similar to those mentioned in Chapter 5. Then one can prove Theorem 8.10, with an effective upper bound for $\max(|x|, |y|)$, by combining estimates for linear forms in 'ordinary logarithms' with estimates for linear forms in p_j-adic logarithms for $j = 1, \ldots, s$.

Recall that in Chapter 5, we considered the unit equation $ax + by = 1$ where the unknowns x, y are taken from the unit group O_K^* of the ring of integers O_K of an algebraic number field K. It was proved that this equation has only finitely many solutions. By Dirichlet's Unit Theorem, the group O_K^* is finitely generated, and we have

$$O_K^* \cong W \times \mathbb{Z}^r$$

where W is the group of roots of unity in K (which is finite), and where r is the unit rank. Recall that $r = r_1 + r_2 - 1$ where r_1 is the number of embeddings $K \to \mathbb{R}$ and r_2 the number of complex conjugate pairs of embeddings $\sigma, \overline{\sigma} : K \to \mathbb{C}$, where $\overline{\sigma}$ is the composition of σ and complex conjugation.

We consider a much more general situation where x, y are taken from an arbitrary finitely generated multiplicative group in an arbitrary field of characteristic 0. For such a finitely generated group Γ we have $\Gamma \cong \Gamma_{\text{tors}} \times \mathbb{Z}^r$ where Γ_{tors} is the (necessarily finite) torsion subgroup of Γ, consisting of roots of unity. Thus,

$$(8.11) \qquad \Gamma = \{\zeta g_1^{u_1}\cdots g_r^{u_r} : \zeta \in \Gamma_{\text{tors}},\ u_1, \ldots, u_r \in \mathbb{Z}\}$$

for certain generators $g_1, \ldots, g_r$.

Theorem 8.12. (Lang, 1960). *Let K be any field of characteristic 0, let a, b be non-zero elements from K, and let Γ be a finitely generated subgroup of the multiplicative group K^* of K. Then the equation*

$$ax + by = 1 \quad \text{in } x, y \in \Gamma \tag{8.12}$$

has only finitely many solutions.

Lang's proof is ineffective.

From Theorem 5.17, that we proved in Chapter 5, one can derive an effective proof of the above theorem in the special case that Γ is a subgroup of $\mathbb{Q}^*$ and that a, b are non-zero elements of $\mathbb{Q}^*$. We now give another, but ineffective proof of this result. Let $g_1, \ldots, g_r$ be a set of generators of Γ as in (8.11). Let $p_1, \ldots, p_s$ be primes such that the numerators and denominators of $a, b, g_1, \ldots, g_r$ are composed of primes from $p_1, \ldots, p_s$. Write $ax = u/w$, $by = v/w$, where u, v, w are integers, necessarily composed of primes from $p_1, \ldots, p_s$, with $\gcd(u, v, w) = 1$ and $u + v = w$. Now clearly, we have

$$|uv(u+v)| = p_1^{z_1} \cdots p_s^{z_s}, \ \gcd(u, v) = 1$$

for certain non-negative integers $z_1, \ldots, z_s$. This is a Thue-Mahler equation. Therefore there are only finitely many possibilities for the pair (u, v), hence for (u, v, w), hence for (x, y). □

Remark. In case that the group Γ is contained in an algebraic number field K, it is possible to give an effective proof of Theorem 8.12, see Theorem 5.18. If the degree of K and the number of generators of Γ are not too large, there is a practical algorithm to determine all solutions.

Example. Let Γ be the multiplicative group generated by $2, 3, 5, 7, 11, 13$ and consider the equation

$$x + y = 1 \ \text{ in } x, y \in \Gamma \text{ with } x \leqslant y. \tag{8.13}$$

We give some solutions:

$$\Big(\frac{1}{2}, \frac{1}{2}\Big), \Big(\frac{3}{7}, \frac{4}{7}\Big), \Big(\frac{2}{13}, \frac{11}{13}\Big), \Big(\frac{3993}{20800}, \frac{16807}{20800}\Big) = \Big(\frac{3 \cdot 11^3}{2^6 \cdot 5^2 \cdot 13}, \frac{7^5}{2^6 \cdot 5^2 \cdot 13}\Big).$$

In his PhD-thesis of 1988, de Weger determined all 545 solutions of (8.13).

8.3 Linear equations in several unknowns

We consider higher dimensional generalizations of equation (8.12). Let K be a field of characteristic 0 and Γ a finitely generated subgroup of K^*. Further, let $n \geqslant 2$ and $\alpha_1, \dots, \alpha_n \in K^*$. We consider the equation

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 1 \text{ in } x_1, \dots, x_n \in \Gamma. \tag{8.14}$$

If $n \geqslant 3$ this equation may have infinitely many solutions. For instance, consider the equation

$$2^u - 2^v + 3^w = 1 \quad \text{in } u, v, w \in \mathbb{Z}.$$

This may be viewed as a special case of (8.14), with Γ the group generated by 2 and 3. This equation has infinitely many solutions with $u = v$ and $w = 0$. More generally, let $2 \leqslant m < n$ and suppose (8.14) has a solution $(x_1, \dots, x_n)$ with

$$\alpha_1 x_1 + \dots + \alpha_m x_m = 1, \quad \alpha_{m+1} x_{m+1} + \dots + \alpha_n x_n = 0\,.$$

Then for every $u \in \Gamma$, the tuple $(x_1, \dots, x_m, ux_{m+1}, \dots, ux_n)$ is also a solution of (8.14). Assuming the group Γ is infinite, we obtain in this way infinitely many solutions of (8.14). More generally, we can construct infinitely many solutions from a given solution $(x_1, \dots, x_n)$ with a *vanishing subsum* $\sum_{i \in I} \alpha_i x_i = 0$ for some non-empty subset I of $\{1, \dots, n\}$.

To make such easy constructions of infinite sets of solutions impossible, we consider only solutions without vanishing subsums.

Definition. A solution $(x_1, \dots, x_n)$ of (8.14) is called *non-degenerate* if

$$\sum_{i \in I} \alpha_i x_i \neq 0 \text{ for each non-empty subset } I \text{ of } \{1, \dots, n\}.$$

Theorem 8.13. (van der Poorten, Schlickewei, Laurent, E., 1980's) *Equation* (8.14) *has only finitely many non-degenerate solutions.*

Roughly speaking, the proof consists of two steps. In the first step one makes a reduction from the general case that K is an arbitrary field of characteristic 0 to the special case that K is an algebraic number field by using techniques from algebraic geometry. To treat the case that Γ is contained in an algebraic number field one has to apply the 'p-adic Subspace Theorem over number fields,' which is a generalization of the p-adic Subspace Theorem which involves absolute values on an algebraic number field and in which the unknowns are algebraic integers of that number field.

The presently known proofs of Theorem 8.13 are all ineffective. For equations (8.14) in two unknowns there is an effectice proof which provides an algorithm to determine all solutions in principle. For equations (8.14) in more than two unknowns no such effective proof is known.

It should be mentioned that there are quantitative versions of Theorem 8.13, giving an explicit upper bound for the number of non-degenerate solutions. Suppose Γ has rank r, i.e., there are $g_1, \dots, g_r$ such that

$$\Gamma = \{\zeta g_1^{u_1} \cdots g_r^{u_r} : \zeta \in \Gamma_{\text{tors}},\, u_1, \dots, u_r \in \mathbb{Z}\}.$$

In 2002, E., Schlickewei, and Schmidt proved that (8.14) has at most $c(n,r) := \exp\big((6n)^{4n}(r+1)\big)$ non-degenerate solutions. One of the tools in the proof is a much refined, quantitative version of the p-adic Subspace Theorem over number fields, giving an explicit estimate for the number of subspaces containing the solutions. In 2009, Amoroso and Viada improved this bound to $c'(n,r) := \exp\big(5n^5 \log(8n)(r+1)\big)$. The importance of the bounds of ESS and AV is that they depend only on the number of unknowns n and the rank r. So whatever group Γ of rank r and coefficients $\alpha_1, \dots, \alpha_n$ we take, we always get an upper bound $c'(n,r)$ for the number of non-degenerate solutions. But it should be mentioned that this upper bound is probably much too large, and one may hope that with better techniques one can improve it further.

Since in these notes we have only the p-adic Subspace Theorem over $\mathbb{Q}$ at our disposal, we assume henceforth

$$\Gamma \subset \mathbb{Q}^*,\ \alpha_1, \dots, \alpha_n \in \mathbb{Q}^*$$

and prove Theorem 8.13 in this special case. It will be convenient to consider instead of (8.14) the homogeneous equation

$$\alpha_0 x_0 + \cdots + \alpha_n x_n = 0 \ \text{ in } x_0, \dots, x_n \in \Gamma, \tag{8.15}$$

where $\alpha_0, \dots, \alpha_n$ are non-zero rational numbers. Solutions $(x_0, \dots, x_n)$ of (8.15) will be called non-degenerate if $\sum_{i \in I} \alpha_i x_i \neq 0$ for each proper, non-empty subset I of $\{0, \dots, n\}$. We prove the following.

Theorem 8.14. *There is a finite set U such that $x_i/x_j \in U$ for each non-degenerate solution $(x_0,\dots,x_n)$ of* (8.15) *and each pair of indices $i,j \in \{0,\dots,n\}$.*

By taking $\alpha_0 = -1$ and considering solutions of (8.15) with $x_0 = 1$ we obtain Theorem 8.13 in the case $\Gamma \subset \mathbb{Q}^*$.

Let H be the linear subspace of $\mathbb{Q}^{n+1}$ given by $\alpha_0 x_0 + \cdots + \alpha_n x_n = 0$.

Lemma 8.15. *There are finitely many proper linear subspaces $T_1,\dots,T_t$ of H such that the set of solutions $(x_0,\dots,x_n)$ of* (8.14) *(non-degenerate or not) lies in $T_1 \cup \cdots \cup T_t$.*

Proof. We use the 'general position version' of the p-adic Subspace Theorem. We start with some preparations.

There are $g_1,\dots,g_r$ of $\mathbb{Q}^*$ such that every element of Γ can be expressed as $\pm g_1^{u_1}\cdots g_r^{u_r}$ with $u_1,\dots,u_r \in \mathbb{Z}$. Let $p_1,\dots,p_s$ be the prime numbers occurring in the numerators and denominators of $\alpha_1,\dots,\alpha_n,g_1,\dots,g_r$. Let φ be the bijective linear map from H to $\mathbb{Q}^n$ given by $(x_0,\dots,x_n) \mapsto (\alpha_1 x_1,\dots,\alpha_n x_n)$.

Take a solution $\mathbf{x} = (x_0,\dots,x_n)$ of (8.15). Let w be a positive rational number such that

$$y_i := w\alpha_i x_i \in \mathbb{Z} \text{ for } i = 1,\dots,n, \quad \gcd(y_1,\dots,y_n) = 1$$

and put $\mathbf{y} = (y_1,\dots,y_n)$. Thus, $\mathbf{y} = \varphi(w\mathbf{x})$. Further, $y_1 + \cdots + y_n = -w\alpha_0 x_0$. Clearly, $y_1,\dots,y_n$ and $y_1 + \cdots + y_n$ are composed of primes from $p_1,\dots,p_s$. This implies that for any ε with $0 < \varepsilon < 1$,

$$(8.16) \quad |y_1\cdots y_n(y_1+\cdots+y_n)| \cdot \prod_{j=1}^{s} |y_1\cdots y_n(y_1+\cdots+y_n)|_{p_j} = 1 \leqslant \|\mathbf{y}\|^{(n+1)-n-\varepsilon}.$$

The linear forms $y_1,\dots,y_n$, $y_1+\cdots+y_n$ are in general position. So by the 'general position-version' of the p-adic Subspace Theorem, the set of solutions $\mathbf{y} = (y_1,\dots,y_n) \in \mathbb{Z}^n$ of (8.16) with $\gcd(y_1,\dots,y_n) = 1$ lies in a union $S_1 \cup \cdots \cup S_t$ of proper linear subspaces of $\mathbb{Q}^n$. Hence the corresponding solutions $\mathbf{x} = (x_0,\dots,x_n)$ of (8.15) lie in $T_1 \cup \cdots \cup T_t$, where $T_i := \varphi^{-1}(S_i)$ is a proper linear subspace of H, for $i = 1,\dots,t$. This proves the lemma. □

Lemma 8.16. *There is a finite set $U' \subset \mathbb{Q}^*$ such that for every solution $(x_1,\dots,x_n)$ of* (8.15) *(non-degenerate or not) there are distinct $i,j \in \{0,\dots,n\}$ with $x_i/x_j \in U'$.*

Proof. We proceed by induction on n. If $n = 1$ we have an equation $\alpha_0 x_0 + \alpha_1 x_1 = 0$ and the lemma is obvious.

Now let $n \geqslant 2$ and assume that the lemma is true for equations of type (8.15) in fewer than $n+1$ unknowns. By the previous lemma, there are proper linear subspaces $T_1, \ldots, T_t$ of H such that the solutions of (8.15) lie in $T_1 \cup \cdots \cup T_t$. Consider the solutions in $T \in \{T_1, \ldots, T_t\}$. The points $\mathbf{x} = (x_0, \ldots, x_n) \in T$ satisfy, apart from the defining equation $\alpha_0 x_0 + \cdots + \alpha_n x_n = 0$ for H, a second equation that is linearly independent of it, say $\gamma_0 x_0 + \cdots + \gamma_n x_n = 0$. By subtracting γ_n/α_n times the first equation from the second, we get an equation

$$\beta_0 x_0 + \cdots + \beta_{n-1} x_{n-1} = 0 \tag{8.17}$$

valid for all $\mathbf{x} \in T$, where at least one of $\beta_0, \ldots, \beta_{n-1}$ is non-zero.

By the induction hypothesis, applied to (8.17) with the terms with $\beta_i = 0$ removed, there is a finite set U_T such that for every solution $(x_0, \ldots, x_n)$ of (8.15) lying in T there are distinct indices $i, j \in \{0, \ldots, n-1\}$ such that $x_i/x_j \in U_T$.

Now the lemma holds with $U' = U_{T_1} \cup \cdots \cup U_{T_t}$. □

Proof of Theorem 8.14. We proceed again by induction on n. For $n = 1$ Theorem 8.14 is trivial. Let $n \geqslant 2$ and suppose Theorem 8.14 is true for equations in fewer than $n+1$ unknowns.

Suppose the set U' from the previous lemma is $\{\beta_1, \ldots, \beta_m\}$. Then the non-degenerate solutions $(x_1, \ldots, x_n)$ of (8.15) can be divided into finitely many sets S_{pqr} ($p, q = 0, \ldots, n$, $p \neq q$, $r = 1, \ldots, m$), where S_{pqr} is the set of solutions with $x_p/x_q = \beta_r$.

Consider for instance the non-degenerate solutions in $S_{n,n-1,1}$, i.e., with $x_n = \beta_1 x_{n-1}$. These solutions satisfy

$$\alpha_0 x_0 + \cdots + (\alpha_{n-1} + \beta_1 \alpha_n) x_{n-1} = 0.$$

Each non-empty subsum of the left-hand side is non-zero, since $(x_0, \ldots, x_n)$ is non-degenerate. By the induction hypothesis, there is a finite set $U_{n,n-1,1}$ such that $x_i/x_j \in U_{n,n-1,1}$ for all solutions $(x_0, \ldots, x_n)$ of (8.15) in $S_{n,n-1,1}$ and all $i, j \in \{0, \ldots, n-1\}$. Using $x_n/x_{n-1} = \beta_1$ we can enlarge $U_{n,n-1,1}$ such that it contains all quotients x_i/x_j with $i = n$ or $j = n$ as well. We get a similar set U_{pqr} for each other triple of indices p, q, r. Now Theorem 8.14 is satisfied with U equal to the union of the sets U_{pqr} with $p, q = 0, \ldots, n$, $p \neq q$ and $r = 1, \ldots, m$. □

8.4 Linear recurrence sequences

A *linear recurrence sequence* (in $\mathbb{C}$) is a sequence $U = \{u_h\}_{h=0}^{\infty}$ with terms in $\mathbb{C}$ given by a linear recurrence

$$u_h = c_1 u_{h-1} + \cdots + c_k u_{h-k} \text{ for } h \geqslant k, \tag{8.18}$$

where $c_1, \ldots, c_k$ are constants in $\mathbb{C}$ and $c_k \neq 0$, and by initial values $u_0, \ldots, u_{k-1}$.

Given a linear recurrence sequence U, there are various linear recurrences which it may satisfy but there is a unique one with minimal length k (exercise). This k is called the *order* of the linear recurrence sequence U, and the polynomial

$$f_U(X) := X^k - c_1 X^{k-1} - \cdots - c_k$$

the *companion polynomial* of U.

Theorem 8.17. *Let $U = \{u_h\}_{h=0}^{\infty}$ be a linear recurrence sequence in $\mathbb{C}$ with companion polynomial $f_U(X) = X^k - c_1 X^{k-1} - \cdots - c_k$. Write*

$$f_U(X) = (X - \theta_1)^{e_1} \cdots (X - \theta_m)^{e_m},$$

where $\theta_1, \ldots, \theta_m$ are distinct complex numbers and $e_1, \ldots, e_m$ positive integers. Then there are polynomials $g_1, \ldots, g_m \in \mathbb{C}[X]$ of degrees at most $e_1 - 1, \ldots, e_m - 1$, respectively, such that

$$u_h = g_1(h)\theta_1^h + \cdots + g_m(h)\theta_m^h \quad \textit{for } h \geqslant 0. \tag{8.19}$$

Conversely, any sequence satisfying (8.19) *is a linear recurrence sequence.*

Proof. Consider the power series

$$y(z) = \sum_{h=0}^{\infty} \frac{u_h}{h!} z^h.$$

One proves easily by induction on h that there is a constant $C > 0$ such that $|u_h| \leqslant C^h$ for all $h \geqslant 0$. Hence $y(z)$ converges for all $z \in \mathbb{C}$, and thus it defines a function that is everywhere analytic on $\mathbb{C}$. Using that the sequence U satisfies recurrence relation (8.18), it follows easily that y satisfies the linear differential equation

$$y^{(k)} = c_1 y^{(k-1)} + \cdots + c_{k-1} y' + c_k y\,.$$

By the theory of linear differential equations, the set of solutions of the latter equation is a complex vector space with basis $\{z^j e^{\theta_i z} : i = 1, \ldots, m,\, j = 0, \ldots, e_i - 1\}$. Hence there are $c_{ij} \in \mathbb{C}$ such that

$$\begin{aligned} y(z) &= \sum_{i=1}^{m} \sum_{j=0}^{e_i-1} c_{ij} z^j e^{\theta_i z} = \sum_{i=1}^{m} \sum_{j=0}^{e_i-1} c_{ij} \sum_{l=0}^{\infty} \theta_i^l \frac{z^{j+l}}{l!} \\ &= \sum_{h=0}^{\infty} \left(\sum_{i=1}^{m} \left\{ \sum_{j=0}^{e_i-1} c_{ij} h(h-1)\cdots(h-j+1)\theta_i^{-j} \right\} \theta_i^h \right) \frac{z^h}{h!}. \end{aligned}$$

This implies that $\{u_h\}_{h=0}^{\infty}$ satisfies (8.19). Conversely, if $\{u_h\}_{h=0}^{\infty}$ satisfies (8.19) then by reversing the above argument one shows that $y(z) = \sum_{h=0}^{\infty}(u_h/h!)z^h$ satisfies a linear differential equation with constant coefficients, and subsequently that $\{u_h\}_{h=0}^{\infty}$ is a linear recurrence sequence. □

Example. Let $U = \{u_h\}_{h=0}^{\infty}$ be given by

$$u_h = 10u_{h-1} - 31u_{h-2} + 30u_{h-3} \ (h \geqslant 3), \quad u_0 = 1, u_1 = 0, u_2 = -12.$$

The companion polynomial of U is given by

$$f_U(X) = X^3 - 10X^2 + 31X - 30 = (X-2)(X-3)(X-5).$$

By Theorem 8.17 there are constants c_1, c_2, c_3 such that $u_h = c_1 2^h + c_2 3^h + c_3 5^h$. Substituting $h = 0, 1, 2$ one obtains $c_1 = 1, c_2 = 0, c_2 = -12$ and

$$u_h = 2^h + 3^h - 5^h.$$

The *zero set* of a linear recurrence sequence $U = \{u_h\}_{h=0}^{\infty}$ is defined by

$$Z_U := \{h \in \mathbb{Z}_{\geqslant 0} : u_h = 0\}$$

and the *zero multiplicity* of U is $N_U := \#Z_U$. With the notation from Theorem 8.17, the set Z_U is the set of solutions of

$$g_1(h)\theta_1^h + \cdots + g_m(h)\theta_m^h = 0 \text{ in } h \in \mathbb{Z}_{\geqslant 0}. \tag{8.20}$$

This is called an *exponential-polynomial equation.*

A linear recurrence sequence $U = \{u_h\}_{h=0}^{\infty}$ is called *non-degenerate* if the zeros of its companion polynomial $\theta_1, \ldots, \theta_m$ are such that none of the quotients θ_i/θ_j $(1 \leqslant i < j \leqslant m)$ is a root of unity.

Theorem 8.18. (Skolem-Mahler-Lech, 1953) *Let U be a non-degenerate linear recurrence sequence. Then its zero set is finite.*
Stated equivalently, if $\theta_1, \dots, \theta_m$ are non-zero complex numbers such that none of the quotients θ_i/θ_j $(1 \leqslant i, j \leqslant m,\ i \neq j)$ is a root of unity and if $g_1(X), \dots, g_m(X)$ are polynomials in $\mathbb{C}[X]$, not all equal to 0, then Eq. (8.20) *has only finitely many solutions.*

There are two very different proofs.

In the first proof, which was the one given by Skolem, Mahler and Lech, one 'maps' the linear recurrence sequence to a sequence with terms in the field $\mathbb{Q}_p$ of p-adic numbers for some suitable prime p, and then uses techniques from p-adic analysis.

The field $\mathbb{Q}_p$ is some sort of analogue of the field $\mathbb{R}$ of real numbers. Recall that $\mathbb{R}$ is the completion of $\mathbb{Q}$, that is the smallest extension of $\mathbb{Q}$ in which every Cauchy sequence has a limit. The elements of the field $\mathbb{R}$ are obtained by taking the collection of Cauchy sequences in $\mathbb{Q}$ and identifying two such sequences if their difference converges to 0. Given a prime number p, one obtains in a similar manner the field of p-adic numbers $\mathbb{Q}_p$, by mimicking the arguments in the construction of $\mathbb{R}$, but replacing everywhere the standard absolute value by $|\cdot|_p$. Several aspects of real analysis such as continuity, differentiability, convergent series, etc., can be carried over to $\mathbb{Q}_p$ and this leads to p-adic analysis. For more information on this we refer to *N. Koblitz, p-adic Numbers, p-adic Analysis, and Zeta-Functions, Springer Graduate Texts in Mathematics.*

In the second proof, one 'maps' the linear recurrence sequence to a sequence with terms in an algebraic number field, and then applies the p-adic Subspace Theorem over number fields.

Here we prove Theorem 8.18 in the special case that the companion polynomial f_U of $U = \{u_h\}_{h=0}^{\infty}$ does not have multiple zeros, i.e., in Theorem 8.17 we have $e_1 = \cdots = e_m = 1$. Then the polynomials $g_i(h)$ in (8.19) have degree 0, so $u_h = \sum_{i=1}^{m} g_i\theta_i^h$ for $h \geqslant 0$ where the g_i are constants. That is, we have to show that the equation

$$g_1\theta_1^h + \cdots + g_m\theta_m^h = 0$$

has finitely many solutions in $h \in \mathbb{Z}_{\geqslant 0}$.

We proceed by induction on m. For $m = 1$ there are no solutions and we are done. Let $m \geqslant 2$ and suppose the theorem is true if we have fewer than m terms.

Let $a_i := -g_i/g_m$, $\beta_i := \theta_i/\theta_m$. Then the equation reduces to

$$a_1\beta_1^h + \cdots + a_{m-1}\beta_{m-1}^h = 1. \tag{8.21}$$

Further, none of the numbers β_i, nor any of the quotients β_i/β_j $(i \neq j)$ is a root of unity.

We apply Theorem 8.13 with the group Γ generated by $\beta_1, \ldots, \beta_{m-1}$. It follows that there are only finitely many integers h which satisfy (8.21) and for which none of the subsums of the left-hand side of (8.21) vanishes, i.e.,

$$\sum_{i\in I} a_i\beta_i^h \neq 0 \text{ for each non-empty subset } I \text{ of } \{1, \ldots, m\}.$$

But by the induction hypothesis, each equation $\sum_{i\in I} a_i\beta_i^h = 0$ has only finitely many solutions h. So altogether, (8.21) has only finitely many solutions h. □

Remark. Using a quantitative version of the p-adic Subspace over number fields, giving an explicit upper bound for the number of subspaces containg the solutions, Schmidt proved the following:

Theorem 8.19. (Schmidt, 2000). *Let U be a non-degenerate linear recurrence sequence with terms in $\mathbb{C}$ of order k. Then for its zero multipicity we have*

$$N_U \leqslant \exp\exp\exp 20k.$$

This has been improved by Amoroso and Viada (2011) to $N_U \leqslant \exp\exp 70k$.

Bavencoffe and Bézivin (Une Famille Remarquable de Suites Récurrentes Linéaires, Monatshefte für Mathematik 120 (1995), 189–203) found examples of non-degenerate linear recurrence sequences U of arbitrarily large order k, having $N_U \geqslant \frac{1}{2}k^2 - \frac{1}{2}k + 1$; no linear recurrence sequences of order k with larger zero multiplicity are known. In fact, let

$$P_k(X) := \frac{X^{k+1} + (-2)^{k-1}X + (-2)^k}{X+2};$$

verify that $P_k(X) \in \mathbb{Z}[X]$. Let $U = \{u_n\}_{n=0}^\infty$ be the linear recurrence sequence with companion polynomial P_k and initial values $u_0 = \cdots = u_{k-2} = 0$, $u_{k-1} = 1$. Bavencoffe and Bézivin proved that U is non-degenerate, and moreover, that $u_n = 0$ for

$$n = l(k+1) + q \text{ with } l \geqslant 0,\ q \geqslant 0,\ l + q \leqslant k-2,$$
$$n = j(2k+1) \text{ with } 1 \leqslant j \leqslant k-1.$$

8.5 Expansions of algebraic numbers

Literature:

Y. Bugeaud, Distribution modulo one and Diophantine approximation, Cambridge tracts in mathematics 193, Cambridge University Press, 2012.

Let b be an integer $\geqslant 2$. Every real number α with $0 < \alpha < 1$ has a b-ary expansion

$$\alpha = \sum_{k=1}^{\infty} c_k \cdot b^{-k} = 0.c_1 c_2 \cdots \quad \text{with } c_k \in \{0, \ldots, b-1\} \text{ for all } k \geqslant 1. \tag{8.22}$$

Recall that in general this expansion is unique, except for numbers of the shape $0.c_1 \cdots c_m 00 \ldots = 0.c_1 \cdots c_{m-1}(c_m - 1)(b-1)(b-1) \ldots$. We exclude those expansions where from some point onwards all digits are $b-1$. Then every real number in the open interval $(0,1)$ has a unique b-ary expansion. We call $[\alpha]_b := c_1 c_2 \cdots$ the *word* associated with α.

Given α with b-ary expansion (8.22) and a positive integer n, consider all blocks of n consecutive digits in the b-ary expansion of α,

$$c_1 \cdots c_n, \quad c_2 \cdots c_{n+1}, \quad c_3 \cdots c_{n+2} \quad \ldots, \quad c_k \cdots c_{k+n-1}, \quad \ldots \tag{8.23}$$

We call α *normal with respect to base* b if for every $n \geqslant 1$, all blocks of n digits from $\{0, \ldots, b-1\}$ occur with the same frequency among the blocks in (8.23), more precisely, if for every $n \geqslant 1$ and every block $a_1 \cdots a_n$ with $a_1, \ldots, a_n \in \{0, \ldots, b-1\}$,

$$\lim_{N \to \infty} \frac{\#\{k \leqslant N : c_k \cdots c_{k+n-1} = a_1 \cdots a_n\}}{N} = b^{-n}.$$

Example. Let $b = 10$. The *Champernowne number* $0.123456789101112\cdots$ is normal with respect to base 10.

Exercise 8.1. *Verify this.*

The foundations of the theory of normal numbers were laid by Émile Borel, who proved the following theorem. Recall that a property is said to hold for almost every real number, if the set of real numbers for which it does not hold has Lebesgue measure 0.

Theorem 8.20. (É. Borel, 1909). *Almost every real number* α *with* $0 < \alpha < 1$ *is normal with respect to every base* $b \geqslant 2$.

A proof can be found in Chapter 4 of Bugeaud's book mentioned above. We should mention that in spite of Theorem 8.20, of no number 'occurring in nature' (e.g., e, π, algebraic numbers) it is known whether it is normal with respect to any base b. Borel made the following bold conjecture, a proof of which seems to be out of reach.

Conjecture 8.21. (É. Borel, 1909). *Let α be a real, irrational algebraic number with $0 < \alpha < 1$. Then α is normal with respect to every base $b \geqslant 2$.*

Given a real number α with $0 < \alpha < 1$ and with b-ary expansion (8.22), we denote by $p(n, \alpha, b)$ the *block complexity of length n* of α with respect to base b, that is the number of distinct blocks that occur among $c_k c_{k+1} \cdots c_{k+n-1}$ ($k = 1, 2, \ldots$). Notice that if we fix α, b, then $p(n, \alpha, b)$ is non-decreasing in n. An immediate consequence of Conjecture 8.21 is the following:

Conjecture 8.22. *Let α be a real, irrational algebraic number with $0 < \alpha < 1$ and b an integer $\geqslant 2$. Then $p(n, \alpha, b) = b^n$ for all $n \geqslant 1$.*

Conjecture 8.22 asserts that for every n, every block of length n occurs at least once among the blocks $c_k \cdots c_{k+n-1}$ ($k = 1, 2, \ldots$). So it is much weaker than Conjecture 8.21 which asserts that all blocks of length n occur with the same frequency. However, also for Conjecture 8.22 there is no clue how to prove it, and what people are interested in at the moment is to get as good as possible lower bounds for $p(n, \alpha, b)$.

We start with a simple result.

Lemma 8.23. *Let α be a real, irrational number with $0 < \alpha < 1$ and b an integer $\geqslant 2$. Then $p(n, \alpha, b) > n$ for all $n \geqslant 1$.*

Proof. Let the b-ary expansion of α be given by (8.22). Suppose there is an integer n with $p(n, \alpha, b) \leqslant n$, and take such n minimal. If $n = 1$ then clearly, $\alpha = 0.ccc\cdots$ with a single digit c, and thus, α is rational. Suppose that $n \geqslant 2$. Then by the minimality of n we have $p(n-1, \alpha, b) \geqslant n$, and so $p(n, \alpha, b) = p(n-1, \alpha, b) = n$ since $p(n, \alpha, b) \geqslant p(n-1, \alpha, b)$. It follows that any block of n consecutive digits in the b-ary expansion of α is determined by the first $n-1$ digits in this block. That is, if $c_k \cdots c_{k+n-2}$ is any block of length $n-1$ occurring in the b-ary expansion of α, then c_{k+n-1} is uniquely determined by $c_k \cdots c_{k+n-2}$. But likewise, c_{k+n} is uniquely determined by $c_{k+1} \cdots c_{k+n-1}$, etc. So in fact, a block $c_k \cdots c_{k+n-2}$ uniquely determines all the subsequent digits $c_{k+n-1}, c_{k+n}, c_{k+n+1}, \ldots$. Now clearly, there are

k, l with $k < l$ such that $c_k \cdots c_{k+n-1} = c_l \cdots c_{l+n-1}$. So in fact, $c_{m+l-k} = c_m$ for all $m \geqslant k$. It follows that the b-ary expansion of α is ultimately periodic, hence that α is rational. □

Remark. There exist irrational so-called *Sturmian numbers* α whose binary expansion satisfies $p(n, \alpha, 2) = n + 1$ for all $n \geqslant 1$. For more information on this we refer to Bugeaud's book mentioned at the beginning of this section.

Adamczewski and Bugeaud obtained the following remarkable result, implying that for irrational real algebraic numbers, $p(n, \alpha, b)$ grows faster than any linear function in n as n tends to ∞. Their proof uses in an essential way the p-adic Subspace Theorem.

Theorem 8.24. (Adamczewski, Bugeaud, 2005). *Let α be a real, irrational algebraic number with $0 < \alpha < 1$ and b an integer $\geqslant 2$. Then*

$$\lim_{n\to\infty} \frac{p(n, \alpha, b)}{n} = \infty.$$

Theorem 8.24 is equivalent to the assertion that for every positive integer r there are only finitely many integers n with $p(n, \alpha, b) \leqslant rn$. The idea is that from such n we construct a good rational approximant to α of a very special form, and then show, using the p-adic Subspace Theorem, that there are only finitely many such approximants. We cannot apply Roth's Theorem on the approximation of algebraic numbers by rationals, since the approximants we construct are good, but not good enough to apply Roth's Theorem. The fact that these approximants are of this special form will help us, and it was the insight of Adamczewski and Bugeaud that the p-adic Subspace Theorem can be used.

Before entering the proof we start with some initial comments. Fix r, and let n be an integer with $p(n, \alpha, b) \leqslant rn$. Let as before

$$\alpha = \sum_{k=1}^{\infty} c_k \cdot b^{-k} = 0.c_1c_2\cdots \quad \text{with } c_k \in \{0, \ldots, b-1\} \text{ for all } k \geqslant 1. \tag{8.22}$$

We consider the block of the first $(r+1)n$ digits of α,

$$c_1 \cdots c_{(r+1)n}.$$

By the box principle, at least two among the blocks of length n,

$$c_1 \cdots c_n, \quad c_2 \cdots c_{n+1}, \; \ldots, \; c_{rn+1} \cdots c_{(r+1)n}$$

must be equal, say

$$c_k \cdots c_{k+n-1} = c_l \cdots c_{l+n-1} \text{ with } 1 \leqslant k < l \leqslant rn+1.$$

For the proof to proceed, we need that these blocks do not overlap, i.e., $l \geqslant k+n$, but we cannot know a priori if this is true. To handle this, we need a lemma on the combinatorics of words. We first introduce some notation.

Denote by Σ^* the collection of finite words in the alphabet $\Sigma = \{0, \ldots, b-1\}$, i.e., the collection consisting of the empty word Λ and of all finite strings $s_1 \cdots s_q$ with $q \geqslant 1$ and $s_1, \ldots, s_q \in \Sigma$. The length q of a word $A = s_1 \cdots s_q$ is denoted by $|A|$. We denote by AB the concatenation of $A, B \in \Sigma^*$, i.e., if $A = s_1 \cdots s_u$, $B = t_1 \cdots t_v$, then $AB = s_1 \cdots s_u t_1 \cdots t_v$. More generally, $A_1 \cdots A_m$ denotes the concatenation of $A_1, \ldots, A_m \in \Sigma^*$, and we denote by A^m the concatenation $A \cdots A$ repeated m times. Given $A, B \in \Sigma^*$ we call B a *prefix* of A if there is $C \in \Sigma^*$ such that $A = BC$; a *suffix* of A if there is $C \in \Sigma^*$ with $A = CB$; and a *subword* of A if there are $C, D \in \Sigma^*$ such that $A = CBD$.

Lemma 8.25. *Suppose a word $A \in \Sigma^*$ has two equal, possibly overlapping subwords of length $n \geqslant 1$. Then A has two equal, non-overlapping subwords of length at least $n/2$.*

Remark. This result is optimal. For let $B \in \Sigma^*$ have length m, let $n = 2m$ and take $A = BBB$. Both the first and second block B, and the second and third block B provide equal, overlapping subwords BB of A of length n, while any two blocks B provide two non-overlapping equal subwords of A of length $n/2$. You may verify yourself that A does not have two non-overlapping equal subwords of length larger than $n/2$ if the cyclic shifts of B are all distinct, i.e., if $B = s_1 \cdots s_m$ with $s_1, \ldots, s_m \in \Sigma$, then $s_1 \cdots s_m, \; s_2 \cdots s_m s_1, \; \ldots, \; s_m s_1 \cdots s_{m-1}$ are distinct.

Proof of Lemma 8.25. Capital letters always indicate elements of Σ^*. If A has two non-overlapping equal subwords of length n we are already done. Suppose that A has two overlapping equal subwords of length n. That is, $A = CBD = EBF$ where $|B| = n$, $|C| < |E| < |C| + |B|$, and $|F| < |D| < |B| + |F|$. In fact, we may write

$E = CG$, $D = HF$, and obtain $A = CBHF = CGBF$. So $BH = GB$. Since $|E| < |C| + |B|$ we have $|G| < |B|$, hence G is a prefix of B. That is, $B = GI = IH$. Let m be the largest non-negative integer such that G^m is a prefix of I; so $m = 0$ if G is not a prefix of I, which is the case if $|I| < |G|$. Then $I = G^m J$, where G is not a prefix of J. This gives $B = G^{m+1}J = G^m JH$ and so $GJ = JH$. If $|G| \leqslant |J|$ then G would be a prefix of J which is impossible. So $|J| < |G|$ and J is a prefix of G, that is, $G = JK$. This leads to

$$I = (JK)^m J, \quad B = GI = (JK)^{m+1}J, \quad H = KJ, \quad BH = GB = (JK)^{m+2}J.$$

Thus, $(JK)^{m+2}J$ is a subword of A.

We consider two cases. First assume that m is even, and let $L := (JK)^{(m+2)/2}$. Then $(JK)^{m+2}J = LLJ$, hence A has two non-overlapping subwords L, and we have

$$\frac{|L|}{|B|} = \frac{\frac{1}{2}(m+2)|J| + \frac{1}{2}(m+2)|K|}{(m+2)|J| + (m+1)|K|} \geqslant \frac{1}{2}.$$

Next assume that m is odd, and let $L := (JK)^{(m+1)/2}J$. Then $(JK)^{m+2}J = LKL$, hence again A has two non-overlapping subwords L, and

$$\frac{|L|}{|B|} = \frac{\frac{1}{2}(m+3)|J| + \frac{1}{2}(m+1)|K|}{(m+2)|J| + (m+1)|K|} \geqslant \frac{1}{2}.$$

So in both cases, $|L| \geqslant n/2$. This proves our lemma. □

The next lemma produces the good approximants to α mentioned before.

Lemma 8.26. *Let r be a positive integer, and n a positive integer such that $p(n, \alpha, b) \leqslant rn$. Then there are integers x, k, h such that*

$$h \geqslant \tfrac{1}{2}n, \quad k \geqslant 0, \quad 0 \leqslant x < b^{h+k}$$

and

$$\left| \alpha - \frac{x}{b^k(b^h - 1)} \right| \leqslant (b^{h+k})^{-1-1/(2r+1)}. \tag{8.24}$$

Proof. As mentioned before, the block $A = c_1 \cdots c_{(r+1)n}$ has two equal, possibly overlapping subwords of length n. Lemma 8.25 implies that A has two non-overlapping equal subwords of length at least $n/2$, i.e., $A = CBDBE$ with $|B| \geqslant n/2$. So the infinite word $[\alpha]_b = c_1c_2\cdots$ consisting of all digits of α is equal to

$$[\alpha]_b = CBDBF,$$

for some infinite word F. We approximate α by the rational number ξ with corresponding ultimately periodic word

$$[\xi]_b = CBDBDBD\cdots$$

(so we replace F by $DBDBD\cdots$). Suppose $|C| = k$, $|B| = l$, $|D| = m$. Then

$$\begin{aligned} \xi &= \sum_{i=1}^{k} c_i \cdot b^{-i} + \big(\sum_{i=k+1}^{k+l+m} c_i b^{-i}\big) \cdot \sum_{j=0}^{\infty} b^{-(l+m)j} \\ &= \frac{y}{b^k} + \frac{z}{b^{k+l+m}(1-b^{-l-m})} = \frac{x}{b^k(b^{l+m}-1)} \end{aligned} \tag{8.25}$$

for certain non-negative integers y, z, x. Here $0 \leqslant x < b^{k+l+m}$ since $0 \leqslant \xi \leqslant 1$.

Recall that both the expansions of α and ξ start with $CBDB$, that is, their expansions are equal up to the first $k + 2l + m$ digits. Further, the digits of the expansions of α and ξ from the $(k + 2l + m + 1)$-th place onwards differ by at most $b - 1$ in absolute value. Thus, we have

$$|\alpha - \xi| \leqslant (b-1)(b^{-k-2l-m-1} + b^{-k-2l-m-2} + \cdots) \leqslant b^{-k-2l-m}. \tag{8.26}$$

Observe that $k + 2l + m \leqslant (r + 1)m$ since $CBDB$ is a prefix of $A = c_1 \cdots c_{(r+1)n}$, and recall that $l = |B| \geqslant \frac{1}{2}n$. Hence

$$\frac{k+2l+m}{k+l+m} = 1 + \frac{l}{(k+2l+m)-l} \geqslant 1 + \frac{\frac{1}{2}n}{(r+1)n - \frac{1}{2}n} = 1 + \frac{1}{2r+1}.$$

Now writing $h := l + m$, and combining (8.25) and (8.26), we arrive at (8.24). As mentioned before, we have $0 \leqslant x < b^{k+l+m} = b^{k+h}$, while $h \geqslant l \geqslant n/2$. □

The next lemma shows that for every positive integer r, inequality (8.24) has only finitely many solutions (x, k, h). Together with Lemma 8.26, and $h \geqslant n/2$, this implies that for every positive integer r there are only finitely many positive integers n such that $p(n, \alpha, b) \leqslant rn$. This implies Theorem 8.24.

Lemma 8.27. *Let r be any positive integer. Then inequality* (8.24) *has only finitely many solutions in integers x, k, h with $h > 0$, $k \geqslant 0$, and $0 \leqslant x < b^{h+k}$.*

Proof. We apply the p-adic Subspace Theorem with $p_1, \dots, p_s$ the distinct primes dividing b.

Let (x, h, k) be a solution of (8.24) with the specified properties, and put $y = b^{h+k}$, $z = b^k$, $\mathbf{x} = (x, y, z)$. With these choices for y, z we have

$$|yz| \cdot |yz|_{p_1} \cdots |yz|_{p_s} = 1.$$

Now on multiplying (8.24) with the quantity $b^k(b^h - 1)$ and then with $|x|_{p_1} \cdots |x|_{p_s}$ which is $\leqslant 1$ since x is an integer, we obtain the inequality

$$|(\alpha y - \alpha z - x)yz| \cdot |xyz|_{p_1} \cdots |xyz|_{p_s} \leqslant (b^{h+k})^{-1/(2r+1)} = \|\mathbf{x}\|^{-1/(2r+1)}.$$

We apply the p-adic Subspace Theorem with $K = \mathbb{Q}(\alpha)$, taking for $|\cdot|_\infty$ the restriction to K of the standard absolute value on $\mathbb{R}$, and choose any continuations of the $|\cdot|_{p_i}$ to K; which ones we choose are irrelevant. It follows that the set of vectors $\mathbf{x} = (x, y, z) = (x, b^{h+k}, b^k)$ lie in a union $T_1 \cup \dots \cup T_t$ of proper linear subspaces of $\mathbb{Q}^3$. We have to prove that each T_i contains only finitely many triples $\mathbf{x}$. Let $T \in \{T_1, \dots, T_t\}$, and consider the triples $\mathbf{x}$ lying in T. Suppose that $a_1 x + a_2 y + a_3 z = 0$ identically on T, where a_1, a_2, a_3 are rational numbers, not all zero. We distinguish between the cases $a_1 \neq 0$ and $a_1 = 0$. First assume that $a_1 \neq 0$. Then for $\mathbf{x} = (x, b^{h+k}, b^k) \in T$ we have $x = b_1 b^{h+k} + b_2 b^k$ with $b_1 = -a_2/a_1$, $b_2 = -a_3/a_1$. Substituting this into (8.24), and noting that $\frac{x}{b^k(b^h-1)} = b_1 + \frac{b_1+b_2}{b^h-1}$, we get

$$A(h) := \left| \alpha - b_1 - \frac{b_1 + b_2}{b^h - 1} \right| \leqslant (b^{h+k})^{-1-1/(2r+1)}.$$

Recall that α is irrational, so $A(h) > 0$ for all h. There is h_0 such that

$$A(h) \geqslant \tfrac{1}{2}|\alpha - b_1| > 0 \quad \text{for } h > h_0.$$

So

$$A(h) \geqslant \min\left(\tfrac{1}{2}|\alpha - b_1|, A(1), \dots, A(h_0)\right) =: C > 0 \quad \text{for all } h \geqslant 1.$$

Hence $C \leqslant (b^{h+k})^{-1-1/(2r+1)}$ for all $\mathbf{x} \in T$. This allows only finitely many possibilities for h, k, hence for x. So T contains only finitely many triples $\mathbf{x} = (x, b^{h+k}, b^k)$.

Now assume that $a_1 = 0$. Then for $\mathbf{x} \in T$ we have $b^h = -a_3/a_2$, and on substituting this into (8.24), we get

$$\left|\alpha - \frac{x}{c \cdot b^k}\right| \leqslant (b^{h+k})^{-1-1/(2r+1)}, \tag{8.27}$$

where $c = (-a_3/a_2) - 1$. Similarly as (8.24), we can transform (8.27) into an inequality to which the two-dimensional case of the p-adic Subspace Theorem is applicable, with solution vectors (x, b^k). The details are left to the reader. This yields that the pairs (x, b^k) with $(x, b^{h+k}, b^k) \in T$ lie in finitely many one-dimensional linear subspaces of $\mathbb{Q}^2$, and so that there are only finitely many possibilities for the fraction x/b^k. Since α is irrational, this implies that the left-hand side of (8.27) is bounded below by a positive number independent of x and k. But then we infer again that there are only finitely many possibilities for h and k, hence for x.

So in all cases, the subspace T contains only finitely many triples $\mathbf{x} = (x, b^{h+k}, b^k)$. This completes the proof of Lemma 8.27, and thus, the proof of Theorem 8.24. □

8.6 Exercises

Exercise 8.2. *Prove that the Thue-Mahler equation* (8.10) *has only finitely many solutions in the following two cases:*
(i) $F(1,0) \neq 0$ *and* $F(X,1)$ *has at least three distinct zeros in* $\mathbb{C}$*;*
(ii) $F(X,Y) = Y^k G(X,Y)$*, where* k *is a positive integer and* G *is a binary form such that* $G(1,0) \neq 0$ *and* $G(X,1)$ *has at least two distinct zeros in* $\mathbb{C}$*.*

Exercise 8.3. *For a finite set of primes* $\mathcal{S} = \{p_1, \dots, p_s\}$*, denote by* $U_{\mathcal{S}}$ *the set of integers of the shape* $\pm p_1^{u_1} \cdots p_s^{u_s} : u_1, \dots, u_s \in \mathbb{Z}_{\geqslant 0}$*.*
Let $\mathcal{S}_0, \dots, \mathcal{S}_n$ *be pairwise disjoint sets of prime numbers, and* $a_0, \dots, a_n$ *non-zero integers. Prove that the equation*

$$a_0 x_0 + \cdots + a_n x_n = 0 \textit{ in } x_0 \in U_{\mathcal{S}_0}, \dots, x_n \in U_{\mathcal{S}_n}$$

has only finitely many solutions.

Exercise 8.4. *Let* $p_1, \dots, p_s$ *be distinct prime numbers,* $A_1, \dots, A_s$ *non-zero integers, and* $C > 0$*,* $\varepsilon > 0$*. Prove that the inequality*

$$|A_1 p_1^{u_1} + \cdots + A_s p_s^{u_s}| \leqslant C \cdot \big(\max(p_1^{u_1}, , \dots, , p_s^{u_s})\big)^{1-\varepsilon}$$

has only finitely many solutions in non-negative integers $u_1, \dots, u_s$.

Hint. *Proceed by induction on s, starting with $s = 1$. In the induction step, apply the p-adic Subspace Theorem with $\mathbf{x} = (x_1, \dots, x_s) = (p_1^{u_1}, \dots, p_s^{u_s})$.*

Exercise 8.5. *Let $f(X) = a_0X^n + a_1X^{n-1} + \cdots + a_n \in \mathbb{Z}[X]$ be a square-free polynomial, i.e., without multiple zeros, and let $p_1, \dots, p_s$ be distinct prime numbers. We consider the equation*

$$|f(\xi)| = p_1^{z_1} \cdots p_s^{z_s} \text{ in } \xi \in \mathbb{Q},\ z_1, \dots, z_s \in \mathbb{Z}. \tag{8.28}$$

(i) Let $(\xi, z_1, \dots, z_s)$ be a solution of (8.28). Prove that $|\xi|_p \leqslant 1$ for every prime p with $p \notin \{p_1, \dots, p_s\}$, $p \nmid a_0$.

(ii) Let $n \geqslant 2$. Prove that (8.28) has only finitely many solutions. What if $n = 1$?

Hint. *Write $\xi = x/y$ with $x, y \in \mathbb{Z}$, $gcd(x, y) = 1$ and reduce (8.28) to a Thue-Mahler equation.*

Exercise 8.6. *let $\mathcal{S} = \{p_1, \dots, p_s\}$ be a set of prime numbers, α a real, irrational algebraic number and $\varepsilon > 0$. Let $U_{\mathcal{S}}$ denote the set from Exercise 8.3.*

(i) Prove that the inequality

$$\left|\alpha - \frac{x}{y}\right| \leqslant \max(|x|, |y|)^{-1-\varepsilon}$$

has only finitely many solutions in integers x, y with $y \in U_{\mathcal{S}}$.

(ii) Prove that the inequality

$$\left|\alpha - \frac{x}{y}\right| \leqslant \max(|x|, |y|)^{-\varepsilon}$$

has only finitely many solutions in integers x, y with $x, y \in U_{\mathcal{S}}$.

(iii) Let c be a non-zero integer. Prove that the inequality

$$\left|\alpha - \frac{x}{y-c}\right| \leqslant \max(|x|, |y|)^{-1-\varepsilon}$$

has only finitely many solutions in integers x, y with $y \in U_{\mathcal{S}}$, $y \neq c$.

(iv) Let c, d be integers. Prove that the inequality

$$\left|\alpha - \frac{x-d}{y-c}\right| \leqslant \max(|x|, |y|)^{-\varepsilon}$$

has only finitely many solutions in integers x, y with $x, y \in U_{\mathcal{S}}$, $y \neq c$.

Exercise 8.7. *Let $\varepsilon > 0$. Prove that the inequality*

$$\left|\left(\frac{3}{2}\right)^n - u\right| \leqslant e^{-\varepsilon n}$$

has only finitely many solutions in non-negative integers n, u.
Hint. *Let $x = 3^n$, $y = u2^n$ and apply in an appropriate way the p-adic Subspace Theorem.*

Exercise 8.8. *Let $U = \{u_h\}_{h=0}^{\infty}$ be a linear recurrence sequence with terms in $\mathbb{C}$.*

(i) Prove that the following two assertions are equivalent:
(a) $u_h = c_1u_{h-1} + \cdots + c_ku_{h-k}$ for all $h \geqslant k$;
(b) $\sum_{h=0}^{\infty} u_hX^h = g(X)/h(X)$, where $h(X) = 1 - c_1X - \cdots - c_kX^k$ and $g(X)$ is a polynomial of degree at most $k - 1$.

(ii) Let I_U be the set of all polynomials $d_0X^m + \cdots + d_m \in \mathbb{C}[X]$ ($m \geqslant 0$, $d_0, \ldots, d_m \in \mathbb{C}$) such that $d_0u_h + d_1u_{h-1} + \cdots + d_mu_{h-m} = 0$ for all $h \geqslant m$. Prove that I_U is an ideal of the ring $\mathbb{C}[X]$, generated by the companion polynomial of U.

(iii) Give a necessary and sufficient condition, in terms of the companion polynomial of U, such that U is periodic (i.e., there is $m > 0$ such that $u_{h+m} = u_h$ for all $h \geqslant 0$).

(iv) Give an example of a non-periodic linear recurrence sequence $U = \{u_h\}_{h=0}^{\infty}$ such that $Z_U = \{h \in \mathbb{Z}_{\geqslant 0} : u_h = 0\}$ is infinite.

Exercise 8.9. *An arithmetic progression is a sequence $a, a+d, a+2d, \ldots$ where a, d are integers with $d > 0$.*
Let $U = \{u_h\}_{h=0}^{\infty}$ be a linear recurrence sequence with terms in $\mathbb{C}$. We do not assume that U is non-degenerate. Assuming the Skolem-Mahler-Lech Theorem, prove that either Z_U is finite, or Z_U is the union of a finite set and a finite number of arithmetic progressions.
Hint. *Assume that U is degenerate and let $\theta_1, \ldots, \theta_m$ be the roots of the companion polynomial of U. Let N be a positive integer such that all roots of unity among the quotients θ_i/θ_j have order dividing N. Consider the sequences $\{u_{hN+i}\}_{h=0}^{\infty}$ ($i = 0, \ldots, N-1$).*

Exercise 8.10. *A linear recurrence sequence $U = \{u_h\}_{h=0}^{\infty}$ is called strongly non-degenerate if for the zeros $\theta_1, \ldots, \theta_m$ of the companion polynomial of U, neither any of the numbers θ_i ($i = 1, \ldots, m$), nor any of the quotients θ_i/θ_j ($1 \leqslant i, j \leqslant m\, i \neq j$) is a root of unity.*

(i) Let U be a strongly non-degenerate linear recurrence sequence with terms in $\mathbb{C}$. Prove that for every $a \in \mathbb{C}$, the set $Z_U(a) := \{h \in \mathbb{Z}_{\geqslant 0} : u_h = a\}$ is finite.

(ii) Let $U = \{u_h\}_{h=0}^{\infty}$ be a linear recurrence sequence with companion polynomial $f(X) = (X - \theta_1)(X - \theta_2)$ where none of $\theta_1, \theta_2, \theta_1/\theta_2$ is a root of unity. Prove that the set

$$T_U := \{(h, l) \in \mathbb{Z}^2 : u_h = u_l,\, 0 < h < l\}$$

is finite.
Hint. *Use Theorem 8.13.*

Remark. One can show that T_U is finite for every strongly non-degenerate linear recurrence sequence U.