3. Heights of Algebraic Numbers

A nonzero rational integer has absolute value at least 1. A nonzero rational number has absolute value at least the inverse of any denominator. Liouville's inequality (§ 3.5) is an extension of these estimates and provides a lower bound for the absolute value of any nonzero algebraic number. More specifically, if we are given finitely many (fixed) algebraic numbers $\gamma_1, \ldots, \gamma_t$, and a polynomial $P \in \mathbb{Z}[X_1, \ldots, X_t]$ which does not vanish at the point $(\gamma_1, \ldots, \gamma_t)$ then we can estimate from below $|P(\gamma_1, \ldots, \gamma_t)|$. The lower bound will depend upon the degrees of P with respect to each of the X_i's, the absolute values of its coefficients as well as some measure of the γ_i's.

In order to obtain such lower bounds, we introduce a notion of *height* for an algebraic number (§ 3.2). There are several such heights (§ 3.4) and they all satisfy the fundamental property that for each fixed d and H, the set of algebraic numbers of degree at most d and height at most H is finite. It follows that there exists a function depending on d and H which bounds from below the absolute value of a nonzero algebraic number of degree at most d and height at most H. Now the problem is to compute explicitly such a function, and also to give an upper bound for the height of $P(\gamma_1, \ldots, \gamma_t)$ in terms of $P \in \mathbb{Z}[X_1, \ldots, X_t]$ and the heights of the γ_j's. From this point of view the so-called *absolute logarithmic height* is more convenient than the others, because it has several equivalent definitions:

- The first one is the integral, on the unit circle, of the logarithm of the modulus of the minimal polynomial of the given algebraic number (§ 3.3),
- The second one involves the absolute values (see § 3.1) of the conjugates and the leading coefficient of the minimal polynomial of the algebraic number,
- The third one is phrased in terms of the absolute values — Archimedean and ultrametric — of the algebraic number.

We study this height with somewhat more details than are strictly necessary, because it is an important tool in many situations. We conclude this chapter with Lehmer's problem and related questions (§ 3.6).

3.1 Absolute Values on a Number Field

We need a little bit of algebraic number theory. There are plenty of references on this subject (see, for example, [Ar 1967]; [Bou 1985] Chap. 6; [FrTa 1991] Chap. 1,2,3; [L 1970]; [L 1978]Chap. 4 § 1, pp. 77–84 and Chap. 7 § 1, pp. 159–162; [L 1983] Chap. 3 § 1, pp. 50–54; [L 1993], Chap. 12; [Neu 1999], Chap. 2; [Sc 1999]; [Ser 1989], Chap. 2 § 1–3, pp. 7–16; [Sil 1986], Chap. VIII § 5).

We explain briefly the basic facts we shall need, detailed proofs can be found in [L 1993], (especially Chap. 12) which we take as basic reference.

3.1.1 p-adic Valuation and p-adic Absolute Values over $\mathbb{Q}$

For $x \in \mathbb{Q}$, $x \neq 0$, we write the decomposition of x into a product of prime factors as follows

$$x = \pm \prod_p p^{v_p(x)}.$$

This defines, for each prime number p, a map v_p from $\mathbb{Q}^\times$ to $\mathbb{Z}$, which we extend by $v_p(0) = \infty$. The map $v_p\colon \mathbb{Q} \longrightarrow \mathbb{Z} \cup \{\infty\}$ thus obtained is the *p-adic valuation* over $\mathbb{Q}$. One can easily prove that it satisfies the following properties:

(i) for $x \in \mathbb{Q}$, $v_p(x) = \infty$ is equivalent to $x = 0$
(ii) for $(x, y) \in \mathbb{Q}^2$, $v_p(xy) = v_p(x) + v_p(y)$
(iii) for $(x, y) \in \mathbb{Q}^2$, $v_p(x + y) \geq \min\{v_p(x), v_p(y)\}$.

To v_p is associated an absolute value $|\cdot|_p$, which is the map from $\mathbb{Q}$ to $\mathbb{Q}$ defined by

$$|x|_p = p^{-v_p(x)}$$

for $x \neq 0$ and $|0|_p = 0$.

The *p-adic absolute value* satisfies the following properties:

(i) for $x \in \mathbb{Q}$, $|x|_p = 0$ is equivalent to $x = 0$
(ii) for $(x, y) \in \mathbb{Q}^2$, $|xy|_p = |x|_p |y|_p$
(iii) for $(x, y) \in \mathbb{Q}^2$, $|x + y|_p \leq \max\{|x|_p, |y|_p\}$.

Such an absolute value is called an *ultrametric absolute value*. It is a *nonarchimedean* absolute value: the set $\{|n|_p\,;\, n \in \mathbb{Z}\}$ is bounded. An important property of an ultrametric absolute value $|\cdot|$ is the fact that $|x| < |y|$ implies $|x + y| = \max\{|x|, |y|\} = |y|$. This property will be used several times.

This p-adic absolute value defines a distance on $\mathbb{Q}$, hence a topology. The ball of radius p^{-r} (with $r \in \mathbb{Z}$) with $a \in \mathbb{Q}$ as its center:

$$\mathcal{D}(a, r) = \{x \in \mathbb{Q}\,;\, |x - a|_p \leq p^{-r}\} = \{x \in \mathbb{Q}\,;\, v_p(x - a) \geq r\}$$

is the set of rational numbers x such that the difference $x - a$ is divisible by p^r, i.e. such that $x - a$ is the product of p^r by a rational number with denominator not divisible by p. For $r \geq 1$, this means that the numerator of $x - a$ (written as a quotient of two coprime integers) is congruent to 0 modulo p^r.

The completion of $\mathbb{Q}$ for the p-adic valuation is *the field* $\mathbb{Q}_p$ *of p-adic numbers*. Each element x of $\mathbb{Q}_p$ can be written as

$$x = \frac{a_{-N}}{p^N} + \frac{a_{-N+1}}{p^{N-1}} + \cdots + a_0 + a_1 p + \cdots + a_n p^n + \cdots,$$

with $a_i \in \{0, \ldots, p-1\}$. Such a series is called the *Hensel's expansion* of x. For $x \neq 0$, the least $n \in \mathbb{Z}$ for which $a_n \neq 0$ is nothing else than $v_p(x)$.

Two absolute values on a field are said to be *equivalent* if they define the same topology on that field.

One can show (Ostrowski's Theorem, see for instance [K 1980] or [Neu 1999], Chap. 2 § 4) that any nontrivial absolute value on $\mathbb{Q}$ is equivalent to either a p-adic absolute value or to the usual absolute value on $\mathbb{Q}$.

If one fixes a nonzero rational number x and takes the product of all these absolute values of x, then something quite interesting occurs. A property known as the *product formula* holds:

$$|x| \prod_p |x|_p = 1 \qquad \text{for all } x \in \mathbb{Q}^\times,$$

which can also be written additively:

$$\sum_p v_p(x) \log p = \log |x| \qquad \text{for all } x \in \mathbb{Q}^\times.$$

The fact that this property holds in many common types of fields is of great importance in algebraic number theory as well as in the study of diophantine and transcendence problems. We shall return to the product formula later in this chapter.

3.1.2 Number Fields

Let α be an algebraic number. The image of the homomorphism $\mathbb{Q}[X] \longrightarrow \mathbb{C}$, which maps $f \in \mathbb{Q}[X]$ onto $f(\alpha)$, is the field $\mathbb{Q}(\alpha)$ generated by α over $\mathbb{Q}$. The kernel of the same homomorphism is a prime (hence maximal) ideal of $\mathbb{Q}[X]$, which has a uniquely defined monic generator. This generator f is called the *irreducible polynomial of* α *over* $\mathbb{Q}$. The degree of f is called *the degree* of the algebraic number α. Two algebraic numbers are called *conjugate* if they have the same irreducible polynomial over $\mathbb{Q}$.

Let a_0 be the least positive integer such that $g = a_0 f$ has integer coefficients. The product $g = a_0 f$, say

$$g(X) = a_0 X^d + \cdots + a_d \in \mathbb{Z}[X],$$

is the *minimal polynomial of* α *over* $\mathbb{Z}$. This polynomial g is irreducible in the factorial ring $\mathbb{Z}[X]$ (see [L 1993], Chap. 4 § 2), which means that g is irreducible in $\mathbb{Q}[X]$ and the rational integers $a_0, \ldots, a_d$ are relatively prime. The number α is called an *algebraic integer* if $a_0 = 1$, a *unit* if $a_0 = 1$ and $a_d = \pm 1$. The set of algebraic integers is a ring in the field $\overline{\mathbb{Q}}$, and the units are the invertible elements of this ring.

A *number field* is a subfield k of $\mathbb{C}$ which, considered as a vector space over $\mathbb{Q}$, is of finite dimension. We denote this dimension by $[k : \mathbb{Q}]$ and we call it the *degree* of k (over $\mathbb{Q}$). For instance, when γ is an algebraic number, then $k = \mathbb{Q}(\gamma)$ is a number field of degree $[\mathbb{Q}(\gamma) : \mathbb{Q}]$ equal to the degree of γ and such a γ is called a *generator* of the number field k.

When k is a number field, each $\gamma \in k$ is algebraic over $\mathbb{Q}$. On the other hand, using the fact that $[k_3 : k_2][k_2 : k_1] = [k_3 : k_1]$ when $k_1 \subset k_2 \subset k_3$ are finite extensions (see [L 1993], Chap. 5 § 1), it follows easily that for each number field k there exist $\alpha_1, \dots, \alpha_n$ in k such that $k = \mathbb{Q}(\alpha_1, \dots, \alpha_n)$.

Let k be a number field of degree d. If $k = \mathbb{Q}(\gamma)$ for some $\gamma \in k$, then there are exactly d distinct embeddings of k into $\mathbb{C}$. Indeed, if $\gamma_1, \dots, \gamma_d$ are the roots of f in $\mathbb{C}$ (these are *the conjugates of* γ), then the d embeddings of k into $\mathbb{C}$ are given by

$$\begin{array}{ccc} k & \longrightarrow & \mathbb{C} \\ \gamma & \longmapsto & \gamma_i \end{array}$$

$(1 \leq i \leq d)$. By induction one deduces that any number field k of degree d has exactly d embeddings into $\mathbb{C}$. Moreover, a number $\gamma \in k$ is a generator of k over $\mathbb{Q}$ if and only if the d images of γ under these embeddings are distinct. From this follows the *Theorem of the primitive element* (see Exercise 3.1): *for each number field k there exists an algebraic number $\gamma \in k$ such that $k = \mathbb{Q}(\gamma)$.*

We shall now study the set of absolute values of a number field. To do this, we have to study how an absolute value can be extended.

We can deal with the trivial absolute value (defined by $|0| = 0$ and $|x| = 1$ for $x \neq 0$) as follows: *if K/k is a finite extension, the unique extension to K of the trivial absolute value on k is the trivial absolute value on K.* Indeed, for $\alpha \in K^\times$, there exist a positive integer d and $a_0, \dots, a_d$ in k with $a_0 a_d \neq 0$ such that $a_0\alpha^d + a_1\alpha^{d-1} + \cdots + a_d = 0$. Since v is trivial on k, we have

$$|a_i\alpha^{d-i}| < \begin{cases} |a_0\alpha^d| & \text{for} \quad 1 \leq i \leq d \qquad \text{if } |\alpha| > 1, \\ |a_d| = 1 & \text{for} \quad 0 \leq i \leq d-1 \quad \text{if } |\alpha| < 1. \end{cases}$$

Notice as well that since v is trivial, it is ultrametric and thus if $|x| < |y|$ then $|x + y| = \max\{|x|, |y|\} = |y|$. Therefore, we conclude that $|\alpha| = 1$.

Let $|\cdot|$ be a nontrivial absolute value on a number field k. The restriction of this absolute value to $\mathbb{Q}$ is equivalent either to the usual absolute value on $\mathbb{Q}$ (in this case the absolute value is *Archimedean*), or else to a p-adic absolute value (in this case the absolute value is said to be *ultrametric*).

In each equivalence class v of nontrivial absolute values, we choose the representative $|\cdot|_v$ which is *normalized* by

$$\begin{cases} |x|_v = x & \text{if } x \in \mathbb{Q},\ x > 0, \text{ and } v \text{ is Archimedean}, \\ |p|_v = \dfrac{1}{p} & \text{if } v \text{ extends the } p\text{-adic valuation of } \mathbb{Q}. \end{cases}$$

We write $v \mid \infty$ if v is Archimedean, and $v \mid p$ if v extends the p-adic valuation. We denote by M_k (resp. M_k^∞) the set of normalized absolute values (resp. Archimedean normalized absolute values) of k. For $v \in M_k$, the completion of k at v will be denoted by k_v.

3.1.3 Archimedean Absolute Values over a Number Field

Let k be a number field, γ a generator of k over $\mathbb{Q}$, and f the irreducible polynomial of γ over $\mathbb{Q}$.

To each complex embedding $\sigma\colon k \longrightarrow \mathbb{C}$ we associate a normalized Archimedean absolute value v_σ defined by $|x|_{v_\sigma} = |\sigma(x)|$ for $x \in k$. Conversely, let v be a normalized Archimedean absolute value on k. The completion k_v of k is an extension of the completion $\mathbb{R}$ of $\mathbb{Q}$. We denote by γ_v the image of γ in k_v. Then $\mathbb{R}(\gamma_v)$ is a finite extension of $\mathbb{R}$ (because γ_v is a root of f), hence is either $\mathbb{R}$ or $\mathbb{C}$. We know which one it is by writing the decomposition of $f \in \mathbb{Q}[X]$ into irreducible factors in $\mathbb{R}[X]$: $f = f_1 \cdots f_r$, where $r = r_1 + r_2$, $f_1, \dots, f_{r_1}$ are of degree $d_1 = \ldots = d_{r_1} = 1$, while $f_{r_1+1}, \dots, f_r$ are of degree $d_{r_1+1} = \ldots = d_r = 2$. If γ_v is root of one of the f_i's of degree 1, then $\mathbb{R}(\gamma_v) = \mathbb{R}$, while if γ_v is root of one of the f_i's of degree 2, then $\mathbb{R}(\gamma_v) = \mathbb{C}$. In any case, we have $k_v = \mathbb{R}(\gamma_v)$, since $\mathbb{R}$ and $\mathbb{C}$ are complete, and we get a complex embedding σ_v of k into $\mathbb{C}$ such that $v_{\sigma_v} = v$. Hence the mapping $\sigma \mapsto v_\sigma$ is surjective.

If $\sigma(\gamma) \in \mathbb{R}$, then $\sigma(k) \subset \mathbb{R}$ and $k_v = \mathbb{R}$. The embedding σ and the absolute value v are called *real*. If $\sigma(\gamma) \notin \mathbb{R}$, then $k_v = \mathbb{C}$. Here the embedding σ and the absolute value v are called *complex*. We denote by d_v the degree $[k_v : \mathbb{R}]$:

$$d_v = \begin{cases} 1 & \text{if } v \text{ is real,} \\ 2 & \text{if } v \text{ is complex.} \end{cases}$$

Let σ_1 and σ_2 be two distinct embeddings of k into $\mathbb{C}$ which give rise to the same Archimedean absolute value v. For any $\alpha \in k$ we have

$$|\sigma_1(\alpha)|_v = |\sigma_2(\alpha)|_v \quad \text{and} \quad |1 - \sigma_1(\alpha)|_v = |1 - \sigma_2(\alpha)|_v,$$

which implies that the complex numbers $\sigma_1(\alpha)$ and $\sigma_2(\alpha)$ are conjugate. Therefore, to a real absolute value v corresponds one and only one real embedding of k, while to a complex absolute value v correspond two (complex conjugate) embeddings of k into $\mathbb{C}$. We deduce that the number of elements in M_k^∞ (i.e. the number of nonequivalent Archimedean absolute values of k) is $r = r_1 + r_2$, where (as before) r_1 is the number of real roots of f, while r_2 is the number of pairs of conjugate complex roots of f, with $d = r_1 + 2r_2$. We can index the irreducible factors of f over $\mathbb{R}$ by $v \in M_k^\infty$ (instead of $1 \le i \le r$):

$$f = \prod_{v \in M_k^\infty} f_v \quad \text{and} \quad d = \sum_{v \in M_k^\infty} d_v \quad \text{with} \quad d_v = \deg f_v.$$

The d-tuple $\big(|\gamma_1|, \ldots, |\gamma_d|\big)$ consists of the elements $|\gamma|_v$ $(v \in M_k^\infty)$, where each $|\gamma|_v$ is repeated d_v times. For instance, supposing that the minimal polynomial of γ over $\mathbb{Q}$ is $f(X) = a_0X^d + \cdots + a_d$, we get

$$\prod_{v\in M_k^\infty} |\gamma|_v^{d_v} = \prod_{i=1}^{d} |\gamma_i| = \left|\frac{a_d}{a_0}\right|$$

and

$$\prod_{v\in M_k^\infty} \max\{1, |\gamma|_v\}^{d_v} = \prod_{i=1}^{d} \max\{1, |\gamma_i|\}.$$

3.1.4 Ultrametric Absolute Values over a Number Field

Let p be a prime number. The absolute value $|\cdot|_p$ on $\mathbb{Q}_p$ has a unique extension to any finite extension K of $\mathbb{Q}_p$. This is due to the fact that $\mathbb{Q}_p$ is complete (see [L 1993], Chap. 12, Prop. 2.5 or [Neu 1999], Chap. 2 Th. 4.8). This extension is given as follows. For $\alpha \in K$, let $\mathrm{N}_{K/\mathbb{Q}_p}(\alpha)$ denote the norm of the $\mathbb{Q}_p$-endomorphism of K which maps x onto αx. If n is the degree of K over $\mathbb{Q}_p$, the extension $|\cdot|_p$ of the p-adic absolute value of $\mathbb{Q}_p$ to K is defined by

$$|\alpha|_p = |\mathrm{N}_{K/\mathbb{Q}_p}(\alpha)|_p^{1/n}.$$

Denote by $\overline{\mathbb{Q}}_p$ the algebraic closure of $\mathbb{Q}_p$, equipped with this absolute value. Then $\overline{\mathbb{Q}}_p$ is not complete (which makes a difference with the Archimedean situation). This is not a serious drawback, and we could take $\overline{\mathbb{Q}}_p$ as the analog of the field of complex number. But we shall prefer to denote by $\mathbb{C}_p$ the completion of $\overline{\mathbb{Q}}_p$ for the absolute value $|\cdot|_p$. This is a complete field in which $\overline{\mathbb{Q}}_p$ is dense, and moreover $\mathbb{C}_p$ is algebraically closed (we shall need only that it contains an algebraic closure of $\mathbb{Q}_p$, hence it also contains an algebraic closure of $\mathbb{Q}$).

Again let $k = \mathbb{Q}(\gamma)$ be a number field of degree d and f the irreducible polynomial of γ over $\mathbb{Q}$. Denote by $\gamma_1^{(p)}, \ldots, \gamma_d^{(p)}$ the roots of f in $\mathbb{C}_p$. There are d distinct embeddings of k into $\mathbb{C}_p$ (each embedding maps a root of f onto another root of the same). These embeddings are given by

$$\begin{array}{ccc} k & \longrightarrow & \mathbb{C}_p \\ \gamma & \longmapsto & \gamma_i^{(p)} \end{array}$$

$(1 \le i \le d)$. To each such embedding $\sigma\colon k \longrightarrow \mathbb{C}_p$ we associate an ultrametric absolute value $v_\sigma \mid p$ defined by $|x|_{v_\sigma} = |\sigma(x)|_p$.

Let v be an absolute value on k which extends the p-adic absolute value of $\mathbb{Q}$. We view the completion k_v of k as an extension of $\mathbb{Q}_p$ and denote by γ_v the image of γ into k_v. Then $\mathbb{Q}_p(\gamma_v)$ is a finite extension of $\mathbb{Q}_p$. But we can say more about the degree of this extension. Consider the decomposition of $f \in \mathbb{Q}[X]$ into irreducible factors [6] in $\mathbb{Q}_p[X]$: $f = f_1 \cdots f_r$ (notice that the number r of irreducible factors varies with p). Since γ_v is a root of f into $\mathbb{C}_p$, there is a unique i, $1 \le i \le r$, such that γ_v is a root of f_i. Therefore f_i has a root in the field $\mathbb{Q}_p(\gamma_v)$, which is an extension of $\mathbb{Q}_p$ of degree $d_v = \deg(f_i)$, and $k_v = \mathbb{Q}_p(\gamma_v)$. This number d_v is called *the local degree* at v. From this it follows that k_v is (isomorphic to) a subfield of $\mathbb{C}_p$, and we get an embedding σ_v of k into $\mathbb{C}_p$ such that $v_{\sigma_v} = v$. Hence the mapping $\sigma \mapsto v_\sigma$ is surjective.

Let σ_1 and σ_2 be two distinct embeddings of k into $\mathbb{C}_p$. They give rise to the same ultrametric absolute value v if and only if $\sigma_1(\gamma)$ and $\sigma_2(\gamma)$ are conjugate over $\mathbb{Q}_p$, which means that they are roots of the same irreducible factor f_i (cf. [L 1993], Chap. 12, Prop. 3.2 or [Neu 1999], Chap. 2, Prop. 8.2). Therefore the number of distinct embeddings σ into $\mathbb{C}_p$ associated to a given absolute value $v \mid p$ is the local degree $d_v = [k_v : \mathbb{Q}_p]$ of v, and the number of elements $v \in M_k$ with $v \mid p$ is the number r of irreducible factors of f over $\mathbb{Q}_p$. This enables us to write

$$f = \prod_{v \in M_k, v \mid p} f_v \quad \text{and} \quad d = \sum_{v \in M_k, v \mid p} d_v \quad \text{with} \quad d_v = \deg f_v.$$

The d-tuple $\left(|\gamma_1^{(p)}|_p, \dots, |\gamma_d^{(p)}|_p\right)$ consists of the elements $|\gamma|_v$ $(v \in M_k, v \mid p)$, where each $|\gamma|_v$ is repeated d_v times. For instance

$$\prod_{v \in M_k, v \mid p} |\gamma|_v^{d_v} = \prod_{i=1}^{d} |\gamma_i^{(p)}|_p = \left|\frac{a_d}{a_0}\right|_p$$

and

$$\prod_{v \in M_k, v \mid p} \max\{1, |\gamma|_v\}^{d_v} = \prod_{i=1}^{d} \max\{1, |\gamma_i^{(p)}|_p\}.$$

The next lemma shows that this last number is $1/|a_0|_p$.

Lemma 3.1. *Let p be a prime number. Let*

$$f(X) = a_0 X^d + a_1 X^{d-1} + \cdots + a_d$$

be a polynomial in $\mathbb{Z}[X]$ with degree d and $\gcd(a_0, \dots, a_d) = 1$. *Denote the roots of f in $\mathbb{C}_p$ by $\alpha_1, \dots, \alpha_d$:*

$$f(X) = a_0 \prod_{i=1}^{d} (X - \alpha_i).$$

Then

$$|a_0|_p \prod_{i=1}^{d} \max\{1, |\alpha_i|_p\} = 1.$$

It follows from Lemma 3.1 that for each subset I of $\{1, \dots, d\}$, the number

[6] Since f is irreducible in $\mathbb{Q}[X]$ the polynomials $f_1, \dots, f_r$ in $\mathbb{Q}_p[X]$ are pairwise distinct.

$$a_0 \prod_{i \in I} \alpha_i$$

is an algebraic integer.

Proof. We may assume $|\alpha_1|_p \leq \cdots \leq |\alpha_d|_p$. Since the numbers a_i are relatively prime, $\max\{|a_0|_p, \ldots, |a_d|_p\} = 1$. Let us write a_i/a_0 as a symmetric function of the α_i:

$$\frac{a_i}{a_0} = (-1)^i \sum_{1 \leq s_1 < \cdots < s_i \leq d} \alpha_{s_1} \cdots \alpha_{s_i} \qquad (1 \leq i \leq d).$$

If $|\alpha_i|_p \leq 1$ for all $i = 1, \ldots, d$, then $|a_i|_p \leq |a_0|_p$ and $\max\{|a_0|_p, \ldots, |a_d|_p\} = |a_0|_p = 1$, which gives the desired result. Otherwise let j $(1 \leq j \leq d)$, be such that

$$|\alpha_1|_p \leq \cdots \leq |\alpha_{j-1}|_p \leq 1 < |\alpha_j|_p \leq \cdots \leq |\alpha_d|_p.$$

Then, using the main property of ultrametric absolute values, we obtain

$$\max\left\{\left|\frac{a_i}{a_0}\right|_p ; 1 \leq i \leq d\right\} = \left|\frac{a_{d-j+1}}{a_0}\right|_p = |\alpha_j \cdots \alpha_d|_p = \prod_{i=1}^{d} \max\{1, |\alpha_i|_p\},$$

hence

$$\max\{|a_1|_p, \ldots, |a_d|_p\} = |a_0|_p \prod_{i=1}^{d} \max\{1, |\alpha_i|_p\}.$$

Since this number is at least $|a_0|_p$, we deduce

$$\max\{|a_0|_p, \ldots, |a_d|_p\} = |a_0|_p \prod_{i=1}^{d} \max\{1, |\alpha_i|_p\},$$

hence the result. □

We have already defined an algebraic integer as an algebraic number whose irreducible polynomial over $\mathbb{Q}$ has coefficients in $\mathbb{Z}$ (which means that the minimal polynomial of α over $\mathbb{Z}$ is monic). From Lemma 3.1 we deduce at once:

Corollary 3.2. *Let α be an algebraic number. The following conditions are equivalent:*
(i) *α is an algebraic integer.*
(ii) *There exists a monic polynomial in $\mathbb{Z}[X]$ which vanishes at α.*
(iii) *For each number field k containing α, and for each ultrametric absolute value v of k, we have $|\alpha|_v \leq 1$.*
(iv) *There exists a number field k containing α such that, for each ultrametric absolute value v of k, we have $|\alpha|_v \leq 1$.*

Remark 1. Let α be an algebraic number with conjugates $\alpha_1, \ldots, \alpha_d$ (with $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ and, say, $\alpha_1 = \alpha$). If $D \in \mathbb{Z}$ is such that

$$\left| D \prod_{i \in I} \alpha_i \right|_v \leq 1$$

for all subsets I of $\{1, \ldots, d\}$ and all ultrametric absolute values v, then

$$|D|_p \prod_{i=1}^{d} \max\bigl\{1, |\alpha_i|_p\bigr\} \leq 1$$

for each prime number p and each embedding of $\mathbb{Q}(\alpha_1, \ldots, \alpha_d)$ into $\mathbb{C}_p$. Hence $|D|_p \leq |a_0|_p$ for each p, which means that a_0 divides D. This shows that a_0 is the positive generator of the ideal of $D \in \mathbb{Z}$ for which, for any subset $\{i_1, \ldots, i_t\}$ of $\{1, \ldots, d\}$, the number $D\alpha_{i_1} \cdots \alpha_{i_t}$ is an algebraic integer.

Remark 2. (M. Laurent). Even if we do not need it, the relation between valuations and prime ideals in a number field is worth mentioning. We just quote one result involving ideals. Let α be a nonzero algebraic number. The ring of integers $\mathcal{O}_k$ of the number field $k = \mathbb{Q}(\alpha)$ is a Dedekind domain. The principal fractional ideal $\bigl(\alpha\bigr)$ can be written $\mathcal{B}/\mathcal{C}$, where $\mathcal{B}$ and $\mathcal{C}$ are nonzero relatively prime integral ideals of k. Let us show that

$$\mathcal{C} = \bigl\{\gamma \in \mathcal{O}_k \, ; \, \gamma\alpha \in \mathcal{O}_k\bigr\} \quad \text{and} \quad \mathrm{N}\mathcal{C} = a_0,$$

where $\mathrm{N}\mathcal{C}$ is the absolute norm of the ideal $\mathcal{C}$.

We write

$$\bigl(\alpha\bigr) = \prod_{\mathcal{P}} \mathcal{P}^{m_{\mathcal{P}}(\alpha)},$$

where $\mathcal{P}$ runs over the set of prime ideals of $\mathcal{O}_k$. Hence

$$\mathcal{B} = \prod_{\mathcal{P}} \mathcal{P}^{\max\{0, m_{\mathcal{P}}(\alpha)\}}, \quad \mathcal{C} = \prod_{\mathcal{P}} \mathcal{P}^{\max\{0, -m_{\mathcal{P}}(\alpha)\}}.$$

Recall that the *absolute norm* of $\mathcal{P}$ is defined by $\mathrm{N}\mathcal{P} = \mathrm{Card}(\mathcal{O}_k/\mathcal{P})$. If $v \in M_k$ is the ultrametric absolute value associated to $\mathcal{P}$ and d_v the local degree, then

$$|\alpha|_v^{d_v} = \mathrm{N}\mathcal{P}^{-m_{\mathcal{P}}(\alpha)}$$

(in view of the product formula below, the product of the left hand side for all ultrametric v, as well as the product of the left hand side for all prime ideals $\mathcal{P}$, is $1/|\mathrm{N}(\alpha)|$, where $\mathrm{N}(\alpha)$ is the absolute norm of α). Indeed, for $\gamma \in \mathcal{O}_k$ and $m \geq 1$, we have

$$\gamma \in \mathcal{P}^m \iff |\gamma|_v^{d_v} \leq \mathrm{N}\mathcal{P}^{-m}.$$

Using Corollary 3.2, we conclude

$$\begin{aligned} \mathcal{C} &= \bigl\{\gamma \in \mathcal{O}_k \, ; \, |\gamma|_v \leq |\alpha|_v^{-1} \text{ for all ultrametric } v \in M_k\bigr\} \\ &= \bigl\{\gamma \in \mathcal{O}_k \, ; \, \gamma\alpha \in \mathcal{O}_k\bigr\} \end{aligned}$$

and

$$\mathcal{B} = \{\gamma\alpha\,;\ \gamma \in \mathcal{C}\}.$$

Further, by the multiplicativity property of N, we deduce from Lemma 3.1:

$$\mathrm{N}\mathcal{C} = \prod_{\mathcal{P}} \mathrm{N}\mathcal{P}^{\max\{0,-m_{\mathcal{P}}(\alpha)\}} = \prod_{v \text{ ultrametric}} \max\{1, |\alpha|_v^{d_v}\} = a_0.$$

3.1.5 The Product Formula

Again let k be a number field of degree d. Let $\alpha \in k$ have minimal polynomial $a_0X^d + \cdots + a_d$ over $\mathbb{Z}$. If v is an ultrametric absolute value of k, say $v \mid p$, with $|\alpha|_v > 1$, then, by Lemma 3.1, we deduce that p divides a_0. On the other hand, if $\alpha \neq 0$, then the minimal polynomial of α^{-1} is $a_dX^d + \cdots + a_0$. Hence, if $|\alpha|_v < 1$, then p divides a_d. As a consequence, for each $\alpha \neq 0$ in k, the set of v in M_k for which $|\alpha|_v \neq 1$ is finite.

The *product formula* reads

$$\prod_{v \in M_k} |\alpha|_v^{d_v} = 1 \quad \text{for} \quad \alpha \in k,\ \alpha \neq 0.$$

We already know this formula holds in the rational case $k = \mathbb{Q}$. The general case readily follows by considering a_d/a_0, which is the *absolute norm* of α, namely $N_{k/\mathbb{Q}}(\alpha)$ with $k = \mathbb{Q}(\alpha)$.

We shall need a generalization of the relations $d = \sum_{v \in M_k^\infty} d_v = \sum_{v|p} d_v$ when the basis field $\mathbb{Q}$ is replaced by a finite extension. For this purpose it will be convenient to write $d_v = d_v(k)$. Let K be a finite extension of k. One can define a map from M_K onto M_k by mapping w onto the restriction v of w on k, in which case one writes $w \mid v$. We claim that for each $v \in M_k$,

$$\sum_{w|v} d_w(K) = [K : k]d_v(k)$$

(see [L 1993], Chap. 12 Prop. 3.3 and [Neu 1999], Chap. 2 § 8). Indeed, for $\gamma \in K$ such that $K = \mathbb{Q}(\gamma)$, we also have $K = k(\gamma)$, and the irreducible polynomial g of γ over k (which is of degree $[K : k]$) can be decomposed into irreducible factors in $k_v[X]$, say $g = \prod_{w|v} g_w$, where g_w is of degree $[K_w : k_v]$. Therefore, for each $v \in M_k$,

$$\sum_{w|v} [K_w : k_v] = [K : k].$$

Since $d_w(K) = [K_w : k_v]d_v(k)$, our claim follows.

An alternate proof of this relation (suggested by Dong Ping Ping) is as follows. For $\alpha \in k$ and $\gamma \in K$ such that $k = \mathbb{Q}(\alpha)$ and $K = \mathbb{Q}(\gamma)$, there exists a polynomial $Q \in \mathbb{Q}[X]$ such that $\alpha = Q(\gamma)$. Let f be the minimal polynomial of γ over $\mathbb{Q}$ and denote by $\alpha_1, \ldots, \alpha_{d_v(k)}$ the conjugates of α (in $\mathbb{C}$ if v is Archimedean, in $\mathbb{C}_p$ if v is ultrametric) which induce the absolute value v on k. Among the $[K : \mathbb{Q}]$ roots of f in $\mathbb{C}$ (resp. in $\mathbb{C}_p$), there are exactly $[K : k]d_v(k)$ roots whose images by Q belong to the set $\{\alpha_1, \ldots, \alpha_{d_v(k)}\}$. These roots are all the roots of f in $\mathbb{C}$ (resp. in $\mathbb{C}_p$) which induce the absolute values w in K over v. Therefore there are precisely $\sum_{w|v} d_w(K)$ such roots.

3.2 The Absolute Logarithmic Height (Weil)

Let k be a number field. For $\alpha \in k$ we define

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

This is a finite product (all but finitely many factors in the right hand side are equal to 1). Let K be a finite extension of k. For $\alpha \in k$ we obtain

$$\begin{aligned} H_K(\alpha) &= \prod_{w \in M_K} \max\{1, |\alpha|_w\}^{d_w(K)} \\ &= \prod_{v \in M_k} \max\{1, |\alpha|_v\}^{\sum_{w|v} d_w(K)} \\ &= H_k(\alpha)^{[K:k]}. \end{aligned}$$

This shows that the number $H_k(\alpha)^{1/[k:\mathbb{Q}]}$ does not depend on the number field k containing α. The logarithm of this number will play an important role. When α is an algebraic number and K a number field which contains α, we define

$$\mathrm{h}(\alpha) = \frac{1}{[K:\mathbb{Q}]} \log H_K(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} D_v \log \max\{1, |\alpha|_v\},$$

where D_v denotes the local degree at $v \in M_K$. This is the (Weil) *absolute logarithmic height* of the number α. It does not depend on the choice of the number field K containing α, but only on α.

Example. For two rational integers a, b which are relatively prime,

$$\mathrm{h}\left(\frac{a}{b}\right) = \log \max\{|a|, |b|\}.$$

Property 3.3. *For algebraic numbers α_1 and α_2,*

$$\mathrm{h}(\alpha_1\alpha_2) \le \mathrm{h}(\alpha_1) + \mathrm{h}(\alpha_2) \tag{3.4}$$

and

$$\mathrm{h}(\alpha_1 + \alpha_2) \le \log 2 + \mathrm{h}(\alpha_1) + \mathrm{h}(\alpha_2). \tag{3.5}$$

Moreover, for any algebraic number $\alpha \neq 0$ and for any $n \in \mathbb{Z}$,

$$\mathrm{h}(\alpha^n) = |n|\mathrm{h}(\alpha). \tag{3.6}$$

Proof. The upper bound (3.4) is a consequence of the estimate

$$\max\{1, xy\} \le \max\{1, x\} \max\{1, y\} \quad \text{for all } x \ge 0,\ y \ge 0,$$

while (3.5) follows from the inequality

$$\max\{1, x+y\} \le 2\max\{1,x\}\max\{1,y\} \quad \text{for all } x \ge 0,\ y \ge 0.$$

Since

$$\max\{1, x^n\} = \max\{1,x\}^n \quad \text{for all } x > 0,\ n \in \mathbb{Z},\ n \ge 0,$$

property (3.6) reduces to $\mathrm{h}(\alpha) = \mathrm{h}(1/\alpha)$ for $\alpha \neq 0$, which follows from the product formula, since $\max\{1,x\} = x\max\{1,1/x\}$ for $x > 0$. □

Remark. The term log 2 in the right hand side of the estimate (3.5) cannot be replaced by a smaller absolute constant, as shown by the following example: $\alpha_1 = q/(q-1)$, $\alpha_2 = q/(q+1)$ with q an even integer. Another example is $\alpha_1 = \alpha_2 = 1$.

The next lemma provides an upper bound for the absolute logarithmic height of an algebraic number which is given as the value of a polynomial in algebraic numbers $\gamma_1, \dots, \gamma_t$.

When $f \in \mathbb{C}[\mathrm{X}_1, \dots, \mathrm{X}_t]$ is a polynomial in t variables, with complex coefficients, we denote by $\mathrm{L}(f)$ its *length*, which is the sum of the modulus of its complex coefficients. The length is very convenient because it satisfies the inequalities

$$\mathrm{L}(f+g) \le \mathrm{L}(f) + \mathrm{L}(g) \quad \text{and} \quad \mathrm{L}(fg) \le \mathrm{L}(f)\mathrm{L}(g)$$

which will be used repeatedly in the transcendence proofs. Indeed, if we write

$$f = \sum_{\underline{\lambda}} p_{\underline{\lambda}} \underline{\mathrm{X}}^{\underline{\lambda}} \quad \text{and} \quad g = \sum_{\underline{\mu}} q_{\underline{\mu}} \underline{\mathrm{X}}^{\underline{\mu}},$$

where $p_{\underline{\lambda}}$ and $q_{\underline{\mu}}$ are complex numbers, $\underline{\lambda} = (\lambda_i)$ and $\underline{\mu} = (\mu_i)$ run over finite subsets of $\mathbb{N}^t$, while $\underline{\mathrm{X}}^{\underline{\lambda}}$ stands for $\prod_{i=1}^t \mathrm{X}_i^{\lambda_i}$, then the length of

$$fg = \sum_{\underline{\nu}} \sum_{\underline{\lambda}+\underline{\mu}=\underline{\nu}} p_{\underline{\lambda}} q_{\underline{\mu}} \underline{\mathrm{X}}^{\underline{\nu}}$$

satisfies

$$\mathrm{L}(fg) = \sum_{\underline{\nu}} \Bigl| \sum_{\underline{\lambda}+\underline{\mu}=\underline{\nu}} p_{\underline{\lambda}} q_{\underline{\mu}} \Bigr| \le \sum_{\underline{\nu}} \sum_{\underline{\lambda}+\underline{\mu}=\underline{\nu}} \bigl| p_{\underline{\lambda}} q_{\underline{\mu}} \bigr| = \mathrm{L}(f)\mathrm{L}(g).$$

We shall prove (as a consequence of Lemma 3.8 below) the following estimate:

Lemma 3.7. *Let $f \in \mathbb{Z}[\mathrm{X}_1, \dots, \mathrm{X}_t]$ be a nonzero polynomial in t variables with rational integer coefficients. Let $\gamma_1, \dots, \gamma_t$ be algebraic numbers. Then*

$$\mathrm{h}\bigl(f(\gamma_1, \dots, \gamma_t)\bigr) \le \log \mathrm{L}(f) + \sum_{i=1}^t \bigl(\deg_{\mathrm{X}_i} f\bigr)\mathrm{h}(\gamma_i).$$

Applying Lemma 3.7 with $f(\mathrm{X}_1, \dots, \mathrm{X}_n) = \mathrm{X}_1 + \cdots + \mathrm{X}_n$, one can deduce a generalization of (3.5):

$$\mathrm{h}(\alpha_1 + \cdots + \alpha_n) \le \log n + \mathrm{h}(\alpha_1) + \cdots + \mathrm{h}(\alpha_n).$$

When p_1/q_1 and p_2/q_2 are two rational numbers with $(p_1, q_1) = (p_2, q_2) = 1$ and $q_i > 0$, then (3.5) (as well as Lemma 3.7) yields

$$\mathrm{h}\left(\frac{p_1}{q_1} + \frac{p_2}{q_2}\right) \le \log 2 + \log\max\{|p_1|, q_1\} + \log\max\{|p_2|, q_2\}.$$

However, it is sometimes more efficient to write $p_1/q_1 = a/c$ and $p_2/q_2 = b/c$ with $\gcd(a, b, c) = 1$ and $c > 0$:

$$\begin{aligned}\mathrm{h}\left(\frac{a}{c} + \frac{b}{c}\right) &\le \log\max\{|a+b|, c\}\\ &\le \log 2 + \log\max\{|a|, |b|, c\}.\end{aligned}$$

This example suggests a refinement of Lemma 3.7, using a notion of simultaneous height for several numbers. Let K be a number field of degree D. Let $\gamma_0, \ldots, \gamma_\nu$ and λ be elements of K with $(\gamma_0, \ldots, \gamma_\nu) \neq (0, \ldots, 0)$ and $\lambda \neq 0$. From the product formula, it follows that the number

$$\frac{1}{D}\sum_{v\in M_K} D_v \log\max\{|\gamma_0|_v, \ldots, |\gamma_\nu|_v\},$$

which is attached to the $(\nu+1)$-tuple $(\gamma_0, \ldots, \gamma_\nu) \in K^{\nu+1}$, is the same as the number

$$\frac{1}{D}\sum_{v\in M_K} D_v \log\max\{|\lambda\gamma_0|_v, \ldots, |\lambda\gamma_\nu|_v\},$$

which is attached to the $(\nu+1)$-tuple $(\lambda\gamma_0, \ldots, \lambda\gamma_\nu) \in K^{\nu+1}$. Therefore this number, which we will denote by $\mathrm{h}(\gamma_0 : \cdots : \gamma_\nu)$, depends only on the class $(\gamma_0 : \cdots : \gamma_\nu)$ of $(\gamma_0, \ldots, \gamma_\nu)$ in the projective space $\mathbb{P}_\nu(K)$. For instance $\mathrm{h}(\alpha) = \mathrm{h}(1 : \alpha)$.

Lemma 3.8. *Let K be a number field and $\nu_1, \ldots, \nu_\ell$ be positive integers. For $1 \le i \le \ell$, let $\gamma_{i1}, \ldots, \gamma_{i\nu_i}$ be elements of K and denote by $\underline{\gamma}$ the point $\big(\gamma_{ij}\big)_{1\le j\le\nu_i, 1\le i\le\ell}$ in $K^{\nu_1+\cdots+\nu_\ell}$. Further, let f be a nonzero polynomial in $\nu_1 + \cdots + \nu_\ell$ variables, with coefficients in $\mathbb{Z}$, of total degree at most N_i with respect to the ν_i variables corresponding to $\gamma_{i1}, \ldots, \gamma_{i\nu_i}$. Then*

$$\mathrm{h}\big(f(\underline{\gamma})\big) \le \log \mathrm{L}(f) + \sum_{i=1}^{\ell} N_i \mathrm{h}(1 : \gamma_{i1} : \cdots : \gamma_{i\nu_i}).$$

Recall that $\mathrm{L}(f)$ denotes the length of f (sum of the absolute values of the coefficients). We deduce Lemma 3.7 from Lemma 3.8 by taking $\nu_i = 1$ for $1 \le i \le \ell$.

Proof. Write

$$f=\sum_{\underline{\lambda}} p_{\underline{\lambda}} \prod_{i=1}^{\ell}\prod_{j=1}^{\nu_i} \mathrm{X}_{ij}^{\lambda_{ij}},$$

where $p_{\underline{\lambda}}$ are rational integers and $\underline{\lambda}=(\lambda_{ij})$ runs over a finite subset of $\mathbb{N}^{\nu_1+\cdots+\nu_\ell}$. Let v be an absolute value of k. If v is ultrametric, then

$$\begin{aligned}\log\max\{1,|f(\underline{\gamma})|_v\} &\le \log\max\left\{1, \max_{\underline{\lambda}} \prod_{i=1}^{\ell}\prod_{j=1}^{\nu_i} |\gamma_{ij}|_v^{\lambda_{ij}}\right\}\\ &\le \sum_{i=1}^{\ell} N_i \log\max\{1,|\gamma_{i1}|_v,\ldots,|\gamma_{i\nu_i}|_v\}.\end{aligned}$$

If v is Archimedean, then

$$\begin{aligned}\log\max\{1,|f(\underline{\gamma})|_v\} &\le \log \mathrm{L}(f)+\log\max\left\{1, \max_{\underline{\lambda}} \prod_{i=1}^{\ell}\prod_{j=1}^{\nu_i} |\gamma_{ij}|_v^{\lambda_{ij}}\right\}\\ &\le \log \mathrm{L}(f)+\sum_{i=1}^{\ell} N_i \log\max\{1,|\gamma_{i1}|_v,\ldots,|\gamma_{i\nu_i}|_v\}.\end{aligned}$$

Using the relation $\sum_{v\in M_k^\infty} D_v = D$, we easily deduce the conclusion. □

3.3 Mahler's Measure

Lemma 3.9. *Let $f\in\mathbb{C}[X]$ be a nonzero polynomial of degree d:*

$$f(X)=a_0X^d+a_1X^{d-1}+\cdots+a_{d-1}X+a_d=a_0\prod_{i=1}^{d}(X-\alpha_i).$$

Then

$$|a_0|\prod_{i=1}^{d}\max\{1,|\alpha_i|\}=\exp\left(\int_0^1 \log|f(e^{2i\pi t})|dt\right).$$

Proof. This is a special case of Jensen's formula for analytic functions. Since both sides of the conclusion of Lemma 3.9 are multiplicative functions of f, it is sufficient to consider the case where f is either a_0 or else $X-\alpha$. In the first case the left hand side is $|a_0|$ and the desired equality plainly holds. In the latter case, the left hand side is $\max\{1,|\alpha|\}$. Therefore Lemma 3.9 is equivalent to the fact that, for any complex number α,

$$\int_0^1 \log|e^{2i\pi t}-\alpha|dt=\log\max\{1,|\alpha|\}.$$

(See for instance [M 1976], pp. 5–6). □

Under the notation of Lemma 3.9, we define *Mahler's measure* of f by

$$\mathrm{M}(f) = |a_0| \prod_{i=1}^{d} \max\{1, |\alpha_i|\}.$$

This is a multiplicative function:

$$\mathrm{M}(f_1 f_2) = \mathrm{M}(f_1)\mathrm{M}(f_2)$$

for f_1 and f_2 in $\mathbb{C}[X]$, a fact which follows immediately from the definition of M.

When α is an algebraic number with minimal polynomial $f \in \mathbb{Z}[X]$ over $\mathbb{Z}$, we define its *Mahler's measure* by $\mathrm{M}(\alpha) = \mathrm{M}(f)$.

Lemma 3.10. *Let α be an algebraic complex number of degree d. Then*

$$\mathrm{h}(\alpha) = \frac{1}{d} \log \mathrm{M}(\alpha).$$

Proof. Denote, as before, by $a_0 > 0$ the leading coefficient of the minimal polynomial of α, by k the number field $\mathbb{Q}(\alpha)$, and, for $v \in M_k$, by d_v the local degree at v. From the definition of $\mathrm{M}(\alpha)$ follows

$$\mathrm{M}(\alpha) = a_0 \prod_{v \in M_k^\infty} \max\{1, |\alpha|_v\}^{d_v}.$$

In Lemma 3.1 we have proved

$$|a_0|_p^{-1} = \prod_{v|p} \max\{1, |\alpha|_v\}^{d_v}.$$

Therefore the product formula

$$a_0 = \prod_p |a_0|_p^{-1}$$

implies

$$a_0 = \prod_{v \notin M_k^\infty} \max\{1, |\alpha|_v\}^{d_v},$$

which provides the desired conclusion. □

3.4 Usual Height and Size

There are several other notions of heights or size (in French: *taille*) for algebraic numbers. We shall give a few examples (see also the appendix to this Chap. 3). One main property of a height is that the set of algebraic numbers of bounded height and degree should be finite. For instance, for any $\nu \geq 1$, $D \geq 1$ and $h \geq 1$, the set of projective points $\underline{\gamma} \in \mathbb{P}_\nu(\overline{\mathbb{Q}})$ with $\mathrm{h}(\underline{\gamma}) \leq h$, and for which there exists a system of projective coordinates $\underline{\gamma} = (\gamma_0 : \cdots : \gamma_\nu)$ satisfying

$$\bigl[\mathbb{Q}(\gamma_0, \ldots, \gamma_\nu) : \mathbb{Q}\bigr] \leq D,$$

is a finite subset of $\mathbb{P}_\nu(\overline{\mathbb{Q}})$. This is *a completely elementary result due to Northcott* ([L 1991], Chap. II, Th. 2.2; see also [Sc 1991], Lemma 7C). To give estimates for the number of elements of such sets is also an interesting question (see the reference to Schanuel's work in [L 1983], [L 1991] and [Sc 1999], Th. 3B).

The *usual height* $\mathrm{H}(f)$ of a polynomial $f(X) = a_0X^d + a_1X^{d-1} + \cdots + a_d \in \mathbb{Z}[X]$ is the maximum of the complex modulus of its coefficients:

$$\mathrm{H}(f) = \max\{|a_0|, \ldots, |a_d|\}.$$

The *usual height* $\mathrm{H}(\alpha)$ of an algebraic number α is the usual height of its minimal polynomial over $\mathbb{Z}$.

The *house* of an algebraic number is the maximum of the modulus of its conjugates in $\mathbb{C}$:

$$\lceil\overline{\alpha}\rceil = \max\{|\alpha_1|, \ldots, |\alpha_d|\}$$

when the minimal polynomial of α is written in $\mathbb{C}[X]$ as

$$f(X) = a_0X^d + \cdots + a_d = a_0 \prod_{i=1}^{d} (X - \alpha_i).$$

The *denominator* $\mathrm{den}(\alpha)$ of α is the positive generator of the ideal of $D \in \mathbb{Z}$ for which $D\alpha$ is an algebraic integer. It is a divisor of a_0.

Among several notions of *size*, one of the most frequently used is

$$s(\alpha) = \log \max\{\mathrm{den}(\alpha)\,;\, \lceil\overline{\alpha}\rceil\}.$$

Lemma 3.11. *For $\alpha \in \overline{\mathbb{Q}}$ of degree d, we have*

$$\frac{1}{d}\log \mathrm{H}(\alpha) - \log 2 \leq \mathrm{h}(\alpha) \leq \frac{1}{d}\log \mathrm{H}(\alpha) + \frac{1}{2d}\log(d+1)$$

and

$$\frac{1}{d}s(\alpha) \leq \mathrm{h}(\alpha) \leq \log \mathrm{den}(\alpha) + \log \max\bigl\{1, \lceil\overline{\alpha}\rceil\bigr\} \leq 2s(\alpha).$$

Proof. The first part of the conclusion can be written

$$2^{-d}\mathrm{H}(\alpha) \le \mathrm{M}(\alpha) \le \mathrm{H}(\alpha)\sqrt{d+1}.$$

The left inequality follows from the identity which relates the coefficients of a polynomial with the roots of this polynomial:

$$a_j = (-1)^j a_0 \sum_{1\le s_1<\cdots<s_j\le d} \alpha_{s_1}\cdots\alpha_{s_j}, \qquad (1 \le j \le d).$$

The number of terms in the sum is $\binom{d}{j} \le 2^d$, and each of these terms is bounded from above by $\mathrm{M}(\alpha)/a_0$.

The right inequality follows from the arithmetico-geometric inequality:

$$\exp\left(\int_0^1 \log\left|f\left(e^{2i\pi t}\right)\right|dt\right) \le \int_0^1 \left|f\left(e^{2i\pi t}\right)\right|dt.$$

Using this bound for f^p, with p positive real, we deduce

$$\mathrm{M}(f) \le \left(\int_0^1 \left|f\left(e^{2i\pi t}\right)\right|^p dt\right)^{1/p}.$$

For $p = 2$ we obtain the desired estimate.

The proof of the second series of inequalities does not involve any difficulty and is left as an exercise. □

Remark 1. Some authors (for instance W. M. Schmidt in [Sc 1991], Ch. I, § 7) prefer another normalization of the absolute height, using the Euclidean norm at the Archimedean places; so the modified logarithmic height of a rational number a/b (with a, b relatively prime) is then $\log\sqrt{a^2+b^2}$ (see Exercise 3.2.b).

Remark 2. The fact that M is a multiplicative function on the ring $\mathbb{C}[X]$:

$$\mathrm{M}(f_1 f_2) = \mathrm{M}(f_1)\mathrm{M}(f_2),$$

combined with the estimates

$$2^{-d}\mathrm{H}(f) \le \mathrm{M}(f) \le \sqrt{d+1}\,\mathrm{H}(f) \tag{3.12}$$

for $d = \deg f$, yields

$$\mathrm{H}(f_1)\mathrm{H}(f_2) \le 2^d\sqrt{d+1}\,\mathrm{H}(f_1 f_2)$$

where $d = \deg(f_1 f_2)$.

Such an upper bound for the product $\mathrm{H}(f_1)\mathrm{H}(f_2)$ in terms of $\mathrm{H}(f_1 f_2)$ already appears in the seminal paper [KoPop 1932] of J. F. Koksma and J. Popken, where the authors give a transcendence measure for e^π (see [FNe 1998], Chap. 2, § 4.2, Th. 27 p. 102). In their paper Koksma and Popken introduce some of the main tools which will enable A. O. Gel'fond, at the end of the 40's, to create his method of algebraic independence (see [G 1952]). Gel'fond established sharp estimates concerning the height of polynomials and extended his investigations to polynomials in several variables. Related results also occur in the book *Diophantine Geometry* of S. Lang in 1962 (see also [L 1983]). The above simple proof, which rests on the multiplicativity of the measure M, is due to K. Mahler [M 1962]. Further references on this topic are given in [Ev 1998].

3.5 Liouville's Inequalities

3.5.1 Introduction

One characteristic of transcendence proofs, and more generally of results of diophantine approximation, is that some variant of the following fact is needed: *if a rational integer is nonzero, then its absolute value is at least* 1. One of the variants of this fact asserts that *if a rational number* p/q (where p and q are relatively prime rational integers and $q > 0$) *is nonzero, then* $|p/q| \ge 1/q$. Now the bound depends on the number considered. In terms of the logarithmic height $\mathrm{h}(p/q) = \log\max\{|p|, q\}$ of p/q (with $(p,q) = 1$ and $q > 0$), the previous inequality yields:

$$\log|x| \ge -\mathrm{h}(x) \quad \text{for all } x \in \mathbb{Q}^{\times}.$$

Liouville's inequality is a generic name for similar lower bounds for nonzero algebraic numbers α.

There is a simple lower bound for the modulus of a nonzero complex algebraic number α in terms of the usual height $\mathrm{H}(\alpha)$:

$$|\alpha| \ge \frac{1}{\mathrm{H}(\alpha)+1}\cdot$$

Since $\mathrm{H}(\alpha^{-1}) = \mathrm{H}(\alpha)$, this lower bound is equivalent to an upper bound $|\alpha| \le \mathrm{H}(\alpha)+1$ (which plainly holds also for $\alpha = 0$). More generally, if α is a complex number which is root of a nonzero polynomial $f(X) = a_0X^n + \cdots + a_n \in \mathbb{Z}[X]$ of degree n with $\max_{0\le i\le n} |a_i| \le H$ (f need not be the minimal polynomial of α), then $|\alpha| \le H+1$. Indeed, this estimate holds trivially if $|\alpha| \le 1$, while if $|\alpha| > 1$, then

$$\begin{aligned}|\alpha| \le |a_0\alpha| &= |a_1 + a_2\alpha^{-1} + \cdots + a_n\alpha^{-n+1}| \\ &\le H\bigl(1 + |\alpha|^{-1} + \cdots + |\alpha|^{-n+1}\bigr) < H\bigl(1 - |\alpha|^{-1}\bigr)^{-1}.\end{aligned}$$

One of the most useful inequalities of Liouville's type is

$$\log|\alpha|_v \ge -[\mathbb{Q}(\alpha)\colon \mathbb{Q}]\mathrm{h}(\alpha) \tag{3.13}$$

for all $\alpha \in \overline{\mathbb{Q}}$, $\alpha \ne 0$, and all absolute values v of $\mathbb{Q}(\alpha)$. For the proof, we first remark that for all $\alpha \in \overline{\mathbb{Q}}$ (including $\alpha = 0$), we have

$$\log|\alpha|_v \le [\mathbb{Q}(\alpha)\colon \mathbb{Q}]\mathrm{h}(\alpha).$$

Further, if $\alpha \neq 0$, then $\mathrm{h}(\alpha) = \mathrm{h}(\alpha^{-1})$ (see (3.6)).

From Lemma 3.8 we now deduce the following statement: *under the hypotheses of Lemma* 3.8, *if the number $f(\underline{\gamma})$ is nonzero, then for all absolute values v of the number field k, we have*

$$\log|f(\underline{\gamma})|_v \geq -D\log \mathrm{L}(f) - D\sum_{i=1}^{\ell} N_i \mathrm{h}(1{:}\,\gamma_{i1}{:}\cdots{:}\,\gamma_{i\nu_i}),$$

where $D = [K : \mathbb{Q}]$.

In the next section we give a slight refinement, where $D\log \mathrm{L}(f)$ is replaced by $(D-1)\log \mathrm{L}(f)$ when v is an Archimedean absolute value.

3.5.2 The Main Lower Bound

Proposition 3.14 *(Liouville's inequality). Let K be a number field of degree D, v be an Archimedean absolute value of K and $\nu_1, \dots, \nu_\ell$ be positive integers. For $1 \leq i \leq \ell$, let $\gamma_{i1}, \dots, \gamma_{i\nu_i}$ be elements of K. Further, let f be a polynomial in $\nu_1 + \cdots + \nu_\ell$ variables, with coefficients in $\mathbb{Z}$, which does not vanish at the point $\underline{\gamma} = \left(\gamma_{ij}\right)_{1\leq j\leq \nu_i, 1\leq i\leq \ell}$. Assume f is of total degree at most N_i with respect to the ν_i variables corresponding to $\gamma_{i1}, \dots, \gamma_{i\nu_i}$. Then*

$$\log|f(\underline{\gamma})|_v \geq -(D-1)\log \mathrm{L}(f) - D\sum_{i=1}^{\ell} N_i \mathrm{h}(1{:}\,\gamma_{i1}{:}\cdots{:}\,\gamma_{i\nu_i}).$$

The simplest case $\ell = 1$, $\nu_1 = 1$ can be written as follows: *for a polynomial $f \in \mathbb{Z}[X]$ of degree at most N and an algebraic number $\alpha \in \mathbb{C}$ of degree d which is not a root of f, we have*

$$|f(\alpha)| \geq \mathrm{L}(f)^{1-d} e^{-dN\mathrm{h}(\alpha)}.$$

(We deduce this estimate from Proposition 3.14 by taking for v the Archimedean absolute value associated with the given embedding of $\mathbb{Q}(\alpha)$ in $\mathbb{C}$).

Proof. We write the product formula for $f(\underline{\gamma}) \neq 0$:

$$D_v \log|f(\underline{\gamma})|_v = -\sum_{w\neq v} D_w \log|f(\underline{\gamma})|_w,$$

where w runs over the absolute values of K distinct from v. If w is Archimedean we have

$$\log|f(\underline{\gamma})|_w \leq \sum_{i=1}^{\ell} N_i \log\max\left\{1, |\gamma_{i1}|_w, \dots, |\gamma_{i\nu_i}|_w\right\} + \log \mathrm{L}(f).$$

The sum of D_w for w Archimedean and $w \neq v$ is $D - D_v \leq D - 1$. If w is ultrametric, the same estimate holds without the term $\log \mathrm{L}(f)$. We conclude the proof by using the bound

$$\sum_{w \neq v} D_w \sum_{i=1}^{\ell} N_i \log \max\bigl\{1, |\gamma_{i1}|_w, \ldots, |\gamma_{i\nu_i}|_w\bigr\} \leq D \sum_{i=1}^{\ell} N_i \mathrm{h}(1 : \gamma_{i1} : \cdots : \gamma_{i\nu_i}).$$

□

3.5.3 Further Lower Bounds

Using inequality (3.13) for $\alpha = \beta - (p/q)$ (or, alternatively, if v is Archimedean, using Proposition 3.14 for the polynomial in a single variable $f(\mathrm{X}) = q\mathrm{X} - p$), we deduce that for each algebraic number β, there exists a constant $c(\beta) > 0$ such that for all $p/q \in \mathbb{Q}$ with $q > 0$ and $p/q \neq \beta$, and for any absolute value v of $\mathbb{Q}(\beta)$, we have

$$\left|\beta - \frac{p}{q}\right|_v \geq \frac{c(\beta)}{\max\{|p|\,,\, q\}^d}$$

with $d = [\mathbb{Q}(\beta) : \mathbb{Q}]$ (and $|p|$ is the usual absolute value of p). An admissible value for $c(\beta)$ is $2^{-d} e^{-dh(\beta)}$.

Finally, the *size inequality*

$$\begin{cases} \log|\alpha|_v \geq -(d-1)\log\lceil\overline{\alpha}\rceil - d\log \mathrm{den}\alpha & \text{if } v \text{ is Archimedean} \\[2ex] \log|\alpha|_v \geq -d\log\lceil\overline{\alpha}\rceil - d\log \mathrm{den}\alpha & \text{if } v \text{ is ultrametric} \end{cases}$$

for all $\alpha \in \overline{\mathbb{Q}}$, $\alpha \neq 0$ is proved by writing

- that the norm over $\mathbb{Q}$ of the product $\alpha \cdot \mathrm{den}(\alpha)$ is a nonzero rational integer if v is Archimedean,
- the product formula for $\alpha \cdot \mathrm{den}(\alpha)$ if v is ultrametric.

A *Liouville number* is a real number ϑ such that, for any $\kappa > 0$, there exists $p/q \in \mathbb{Q}$ with $q \geq 2$ and

$$0 < \left|\vartheta - \frac{p}{q}\right| \leq \frac{1}{q^\kappa}\cdot$$

From Liouville's inequality one deduces that a Liouville number is transcendental.

3.5.4 Proof of Lemma 2.1

From Proposition 3.14 one deduces the following result:

Given algebraic numbers $\gamma_1, \ldots, \gamma_m$ and a polynomial $f \in \mathbb{Z}[\mathrm{X}_1, \ldots, \mathrm{X}_m]$ which does not vanish at the point $(\gamma_1, \ldots, \gamma_m)$, we have

$$|f(\gamma_1, \ldots, \gamma_m)| \geq e^{-cT}$$

where $T = \deg f + \log \mathrm{H}(f)$,

$$c = D\bigl(2 + \mathrm{h}(\gamma_1) + \cdots + \mathrm{h}(\gamma_m)\bigr) \quad \textit{and} \quad D = [\mathbb{Q}(\gamma_1, \ldots, \gamma_m) : \mathbb{Q}].$$

Lemma 2.1 easily follows.

3.5.5 Estimates for Determinants

Most often we shall use Proposition 3.14 for a polynomial given by a determinant. We need to produce upper bounds for the degrees and heights of this polynomial; such estimates are given by the following simple lemma:

Lemma 3.15. *Let L be a positive integer and* $p_{\lambda\mu}$ $(1 \le \lambda, \mu \le L)$ *be* L^2 *polynomials in* $\nu_1 + \cdots + \nu_\ell$ *variables* X_{ij} $(1 \le j \le \nu_i, 1 \le i \le \ell)$*, with coefficients in* $\mathbb{Z}$*. Define, for* $1 \le \lambda \le L$,

$$M_\lambda = \max_{1 \le \mu \le L} \mathrm{L}(p_{\lambda\mu})$$

and

$$N_{i\lambda} = \max_{1 \le \mu \le L} \deg_{\underline{\mathrm{X}}_i} p_{\lambda\mu} \quad (1 \le i \le \ell),$$

where $\deg_{\underline{\mathrm{X}}_i}$ *denotes the total degree with respect to the set of variables* $\mathrm{X}_{i1}, \ldots, \mathrm{X}_{i,\nu_i}$*. Then*

$$\Delta = \det\Bigl(p_{\lambda\mu}\Bigr)_{1 \le \lambda, \mu \le L}$$

is a polynomial in $\mathbb{Z}[\underline{\mathrm{X}}_1, \ldots, \underline{\mathrm{X}}_\ell]$ *of length bounded by*

$$\mathrm{L}(\Delta) \le L! \prod_{\lambda=1}^{L} M_\lambda$$

and degrees bounded by

$$\deg_{\underline{\mathrm{X}}_i} \Delta \le \sum_{\lambda=1}^{L} N_{i\lambda} \quad (1 \le i \le \ell).$$

Consequently if γ_{ij} $(1 \le j \le \nu_i, 1 \le i \le \ell)$ are algebraic numbers in a number field of degree $\le D$ such that the polynomial Δ does not vanish at the point

$$\underline{\gamma} = \bigl(\gamma_{ij}\bigr)_{\substack{1 \le j \le \nu_i \\ 1 \le i \le \ell}} \in \mathbb{C}^{\nu_1 + \cdots + \nu_\ell},$$

then

$$\log|\Delta(\underline{\gamma})| \ge$$
$$-(D-1)\left(\log(L!) + \sum_{\lambda=1}^{L} \log M_\lambda\right) - DL \sum_{i=1}^{\ell} \left(\mathrm{h}(1 : \gamma_{i1} : \cdots : \gamma_{i\nu_i}) \sum_{\lambda=1}^{L} N_{i\lambda}\right).$$

Remark. In § 2.2.1 we introduced the ring $\mathbb{Z}[\mathrm{X}_1^{\pm 1}, \ldots, \mathrm{X}_k^{\pm 1}, \mathrm{Y}_1, \ldots, \mathrm{Y}_{\ell-k}]$. Let $\Delta = \det\bigl(p_{\lambda\mu}\bigr)$ be the determinant of a $L \times L$ matrix with coefficients in this ring. Assume

$$\max_{1\le\mu\le L} \deg_{\mathrm{X}_i^{\pm 1}} p_{\lambda\mu} \le N_{i\lambda} \quad (1 \le i \le k),$$

$$\max_{1\le\mu\le L} \deg_{\mathrm{Y}_j} p_{\lambda\mu} \le N'_{j\lambda} \quad (1 \le j \le \ell - k)$$

and

$$\mathrm{L}(p_{\lambda\mu}) \le M_\lambda$$

for $1 \le \lambda \le L$. We can apply Lemma 3.15 with ℓ replaced by $\ell + k$, with $\nu_j = 1$ for all j and with

$$\mathrm{X}_{i1} = \begin{cases} \mathrm{X}_i & \text{for } 1 \le i \le k, \\ \mathrm{X}_{i-k}^{-1} & \text{for } k < i \le 2k, \\ \mathrm{Y}_{i-2k} & \text{for } 2k < i \le \ell + k. \end{cases}$$

We deduce

$$\deg_{\mathrm{X}_i^{\pm 1}} \Delta \le \sum_{\lambda=1}^{L} N_{i\lambda} \quad (1 \le i \le k),$$

$$\deg_{\mathrm{Y}_j} \Delta \le \sum_{\lambda=1}^{L} N'_{j\lambda} \quad (1 \le j \le \ell - k).$$

Further, let $\underline{\gamma} = (\gamma_1, \ldots, \gamma_\ell)$ be a ℓ-tuple of algebraic numbers in a number field of degree D with $\gamma_j \ne 0$ for $1 \le j \le k$. Assume $\Delta(\underline{\gamma}) \ne 0$. We can apply Proposition 3.14, but one must carefully add the contributions of $\deg_{\mathrm{X}_i}$ and $\deg_{\mathrm{X}_i^{-1}}$:

$$\log|\Delta(\underline{\gamma})| \ge$$
$$(D-1)\sum_{\lambda=1}^{L} \log M_\lambda - (D-1)\log(L!) - 2D\sum_{i=1}^{k}\sum_{\lambda=1}^{L} N_{i\lambda} - D\sum_{j=1}^{\ell-k}\sum_{\lambda=1}^{L} N'_{j\lambda}.$$

See for instance Exercise 3.8.

3.6 Lower Bound for the Height

We quoted Northcott's Theorem in § 3.4 as a fundamental property of any height. Another important property of the absolute logarithmic height (which distinguishes this height from most other ones) is that for $\alpha \in \overline{\mathbb{Q}}^\times$, $\mathrm{h}(\alpha) = 0$ if and only if α is a root of unity (i.e. a torsion point in the multiplicative group $\mathbb{G}_{\mathrm{m}}(\overline{\mathbb{Q}}) = \overline{\mathbb{Q}}^\times$). This raises the important problem of estimating $\mathrm{h}(\alpha)$ from below when it does not vanish.

3.6.1 Kronecker's Theorem

By definition the values of Mahler's measure M are ≥ 1.

If a nonzero algebraic number α satisfies $\mathrm{M}(\alpha) < 2$, then α is an algebraic integer, and α^{-1} also, which means that α is a unit. In other terms the (absolute logarithmic) height of an algebraic number which is not a unit is at least $(\log 2)/d$.

Let α be nonzero algebraic integer. Assume $\mathrm{M}(\alpha) = 1$, which means that all conjugates of α have modulus at most 1. *Then α is a root of unity.* Indeed, if α has degree d, then each α^ℓ with $\ell \geq 1$ is a root of a monic polynomial, with rational integer coefficients, of degree d, whose coefficients have usual absolute values at most 2^d. The set of such polynomials is finite, hence so is the set of α^ℓ $(\ell \geq 1)$. The conclusion plainly follows.

Using Corollary 3.2, one deduces the following statement, due to L. Kronecker [Kr 1857]: *if k is a number field and α a nonzero element of k such that $|\alpha|_v \leq 1$ for all $v \in M_k$, then α is a root of unity.*

Therefore the only algebraic numbers α which satisfy $\mathrm{h}(\alpha) = 0$ are 0 and the roots of unity. The other ones satisfy $\mathrm{h}(\alpha) > 0$. To give a sharp lower bound for $\mathrm{h}(\alpha)$, when α is a unit but not a root of unity, in terms of the degree of α is an interesting and difficult problem (see [L 1991], Chap. IX, § 7).

Remark. For an algebraic number α of degree d, since $\mathrm{h}(\alpha) = (1/d)\log \mathrm{M}(\alpha)$, the conditions $\mathrm{h}(\alpha) > 0$ and $\mathrm{M}(\alpha) > 1$ are plainly equivalent.

If $\kappa > 0$ and $\alpha \in \overline{\mathbb{Q}}$ satisfy $\mathrm{h}(\alpha) \geq \kappa$, then from the inequality

$$e^{\kappa d} > 1 + \kappa d$$

with $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ we deduce $\mathrm{M}(\alpha) > 1 + \kappa d$.

Conversely, if $\alpha \in \overline{\mathbb{Q}}$ has degree d and if $\kappa > 0$, $\epsilon > 0$ satisfy

$$\mathrm{M}(\alpha) \geq 1 + \kappa d \quad \text{and} \quad \kappa d < \epsilon,$$

then (using Exercise 1.1.a) we deduce

$$\mathrm{h}(\alpha) > c\kappa \quad \text{with} \quad c(\epsilon) = \frac{1}{\epsilon}\log(1+\epsilon).$$

Notice that $c(\epsilon) \to 1$ as $\epsilon \to 0$. More precisely we have

$$1 < c(\epsilon) < 1 + \epsilon.$$

3.6.2 Lehmer's Problem

In 1933 D. H. Lehmer [Le 1933] asked whether it is true that for every positive ϵ there exists an algebraic integer α for which $1 < \mathrm{M}(\alpha) < 1+\epsilon$. The answer is not yet known but it is easy to see (Exercise 3.9) that for each positive integer d there exists a positive number $c(d)$ such that, for any nonzero algebraic number α which is not a root of unity and is of degree at most d, the inequality $\mathrm{h}(\alpha) \geq c(d)$ is valid. The example $\alpha = 2^{1/d}$ shows that such a function $c(d)$ must satisfy $c(d) \leq (\log 2)/d$. It is widely believed that there exists a positive absolute constant c_0 such that $c(d) \geq c_0/d$. This problem is known as *Lehmer's problem* (see Chap. 7 of [BerDGPS 1992]) and an answer would have various applications. The first one is due to D. H. Lehmer himself in [Le 1933]: he introduced the subject while looking for large prime numbers. Next, following A. Schinzel, C. Pinner and J. Vaaler related the Mahler measure of a polynomial and the number of its irreducible non-cyclotomic factors. Polynomials with small measure also occur in ergodic theory and dynamical systems (works of Ia. Sinaï, W. Lawton, E. Bombieri and J. E. Taylor). We refer to M. J. Mossinghoff's thesis [Mos 1995] for further references (see also [Ev 1998]).

The smallest known value $\mathrm{M}(\alpha) > 1$ for an algebraic number α is the root $1.1762808183\ldots$ of the reciprocal[7] polynomial of degree 10 :

$$X^{10} + X^9 - X^7 - X^6 - X^5 - X^4 - X^3 + X + 1 = X^5 Q\left(X + \frac{1}{X}\right)$$

with

$$Q(Y) = (Y+1)^2(Y-1)(Y+2)(Y-2) - 1$$

This example is due to D. H. Lehmer [Le 1933].

In 1980 and then in 1989, D. Boyd developed an algorithm for searching polynomials with small Mahler's measure. He found all polynomials of degree at most 20 and Mahler's measure at most 1.3. In his thesis M. J. Mossinghoff [Mos 1995] listed 1560 irreducible non-cyclotomic polynomials with Mahler's measure less than 1.3 and degree at most 64. None of these has Mahler's measure less than Lehmer's degree 10 example reported in 1933.

The first result in the direction of Lehmer's problem is due to A. Schinzel and H. Zassenhaus [SZa 1965]: when $\alpha \neq 0$ is an algebraic integer of degree $d \geq 2$ which is not a root of unity, then

$$\lceil\overline{\alpha}\rceil > 1 + 4^{-s-2}$$

where $2s$ is the number of nonreal conjugates of α. Therefore

$$\mathrm{M}(\alpha) > 1 + \frac{c}{2^d}$$

for some absolute constant $c > 0$.

In 1971, by means of an averaging technique in Fourier analysis, P. E. Blanksby and H. L. Montgomery [BlMon 1971] refined this result and proved, for an algebraic integer of degree $d > 1$ which is not a root of unity,

$$\mathrm{M}(\alpha) > 1 + \frac{1}{52d\log(6d)}\cdot$$

A consequence is the estimate

$$\lceil\overline{\alpha}\rceil > 1 + \frac{1}{30d^2\log(6d)}\cdot$$

[7] A polynomial $f \in \mathbb{C}[X]$ of degree d is *reciprocal* if $f(X) = X^d f(1/X)$.

Also in that year Smyth [Sm 1971] used Parseval's formula to prove, under the same assumptions, that if α^{-1} is not a conjugate of α, then $\mathrm{M}(\alpha) \geq 1.3247179572\ldots$, this number being the real root of $X^3 - X - 1$ and the smallest PV-number [8]. An interesting consequence of his result is that it solves Lehmer's problem when d is *odd* (one can use $c_0 = 0.2811\ldots$ in this case). In 1978, C. L. Stewart [Ste 1978] introduced a method from transcendental number theory to prove

$$\mathrm{M}(\alpha) > 1 + \frac{1}{10^4 d \log d}$$

for $d \geq 2$. This is marginally weaker than the previous result of Blanksby-Montgomery, but the interest lies in the method.

3.6.3 Dobrowolski's Theorem

In 1979, E. Dobrowolski [Do 1979] succeeded to extend Stewart's argument and to obtain the following statement: for each $\epsilon > 0$, there exists an integer $d_0(\epsilon)$ such that, for any $d > d_0(\epsilon)$ and any nonzero algebraic number α of degree $\leq d$ which is not a root of unity,

$$\mathrm{h}(\alpha) > \frac{1-\epsilon}{d}\left(\frac{\log\log d}{\log d}\right)^3,$$

which can be written

$$\mathrm{M}(\alpha) > 1 + (1-\epsilon)\left(\frac{\log\log d}{\log d}\right)^3 \cdot$$

In 1981, independently, D. C. Cantor and E. G. Straus [CaStr 1982] and U. Rausch [Ra 1985] simplified Dobrowolski's proof by introducing a determinant and replaced $1-\epsilon$ by $2-\epsilon$. Finally R. Louboutin [Lo 1983] reached $(9/4)-\epsilon$ by a modification of this determinant. The same result with $(9/4)-\epsilon$ has been also obtained by M. Meyer [Me 1988], using a construction of an auxiliary function (like Dobrowolski), but with Thue-Siegel lemma replaced by a refinement due to E. Bombieri and J. Vaaler [BoVa 1983].

Dobrowolski's result is effective: by [Do 1979], for all $d \geq 2$,

$$\mathrm{M}(\alpha) > 1 + \frac{1}{1200}\left(\frac{\log\log d}{\log d}\right)^3 \cdot$$

P. Voutier [Vou 1996] improved this bound: for $d \geq 2$,

$$\mathrm{h}(\alpha) > \frac{1}{4d}\left(\frac{\log\log d}{\log d}\right)^3 \cdot$$

[8] A *Pisot-Vijayaraghavan number*, or *PV-number*, is a real algebraic integer > 1 all of whose other conjugates lie inside the open unit disc. A *Salem number* is a real algebraic integer > 1 all of whose other conjugates lie inside the closed unit disc, with at least one conjugate on the unit circle. See [BerDGPS 1992].

Let α be a nonzero algebraic integer of degree $\leq d$ with $d \geq 2$. Since $\log\lceil\overline{\alpha}\rceil \geq \mathrm{h}(\alpha)$, we deduce, if α is not a root of unity,

$$\log\lceil\overline{\alpha}\rceil > \frac{1}{4d}\left(\frac{\log\log d}{\log d}\right)^3.$$

The estimate (see [Du 1993])

$$\lceil\overline{\alpha}\rceil > 1 + \left(\frac{64}{\pi^2} - \epsilon\right)\frac{1}{d}\left(\frac{\log\log d}{\log d}\right)^3 \quad \text{for} \quad d > d_0(\epsilon)$$

is sharper for large d, while

$$\log\lceil\overline{\alpha}\rceil > \frac{\log\bigl(d + (1/2)\bigr)}{d^2} \quad \text{for} \quad d \geq 1$$

(see [Mat 1991]) is stronger for small d. From the latter one deduces that a nonzero algebraic integer α satisfying $\mathrm{h}(\alpha) < 2/(3d^3)$ is a root of unity (compare with Theorem 3.16). Another uniform estimate is [Vou 1996]:

$$\lceil\overline{\alpha}\rceil > 1 + \frac{1}{2d}\left(\frac{\log\log d}{\log d}\right)^3 \quad \text{for} \quad d \geq 2.$$

Our aim in the rest of this section is twofold. On one hand we wish to establish a lower bound which will be useful later (namely in Chap. 7, proof of Lemma 7.19):

Theorem 3.16. *Let d be a positive integer and α be a nonzero algebraic number of degree $\leq d$ which is not a root of unity. Then*

$$\mathrm{h}(\alpha) > \frac{1}{11d^3}.$$

On the other hand we wish to give a further example of a *transcendence proof* using an interpolation determinant. This will produce a sharpening of Theorem 3.16, but only for sufficiently large d:

Theorem 3.17. *There exists a positive integer d_0 such that, for any integer $d \geq d_0$ and a nonzero algebraic number α of degree $\leq d$ which is not a root of unity,*

$$\mathrm{h}(\alpha) \geq \frac{1}{250d}\left(\frac{\log\log d}{\log d}\right)^3.$$

From the previous discussion it is clear that $d_0 = 2$ is an admissible value for the constant in Theorem 3.17, and in fact this would follow from the argument given below. But assuming that d is sufficiently large will simplify the estimates (we insist that d is only an *upper bound* for the degree of α, and not the actual degree).

It may be useful for the reader if we repeat that there is no loss of generality, in the proofs of Theorems 3.16 and 3.17, to assume that α is an algebraic integer. Indeed, we have seen that the result is obvious unless α is a unit.

3.6.4 Fermat's Little Theorem

One main tool is Fermat's little theorem, which is used as follows:

Lemma 3.18. *Let p be a prime number and $f \in \mathbb{Z}[X_1, \dots, X_k]$ a polynomial in k variables with integer coefficients. Then there exists $g \in \mathbb{Z}[X_1, \dots, X_k]$ such that*

$$f(X_1^p, \dots, X_k^p) - f(X_1, \dots, X_k)^p = pg(X_1, \dots, X_k).$$

Proof. For simplicity write $\underline{X}$ for $(X_1, \dots, X_k)$ and $\underline{X}^p$ for $(X_1^p, \dots, X_k^p)$, so that the conclusion is just

$$f(\underline{X}^p) - f(\underline{X})^p = pg(\underline{X}).$$

The result holds for a monomial $f(\underline{X}) = aX_1^{i_1} \cdots X_k^{i_k}$:

$$f(\underline{X}^p) - f(\underline{X})^p = (a - a^p)X_1^{pi_1} \cdots X_k^{pi_k}$$

and p divides the integer $a - a^p$. If the result holds for f_1 and for f_2, namely if

$$f_1(\underline{X}^p) - f_1(\underline{X})^p = pg_1(\underline{X}) \quad \text{and} \quad f_2(\underline{X}^p) - f_2(\underline{X})^p = pg_2(\underline{X}),$$

then it holds for $f = f_1 + f_2$, because the coefficients of the polynomial

$$(f_1 + f_2)^p - f_1^p - f_2^p = \sum_{h=1}^{p-1} \binom{p}{h} f_1^h f_2^{p-h}$$

are rational integers which are all divisible by p. Therefore one may choose

$$g = g_1 + g_2 - \sum_{h=1}^{p-1} \frac{(p-1)!}{h!(p-h)!} f_1^h f_2^{p-h}.$$

Lemma 3.18 plainly follows. □

Remark. An explicit expression for g can be given. For instance for $k = 1$ if the given polynomial f is $f(X) = a_0X^d + a_1X^{d-1} + \cdots + a_d$, then one can write:

$$g(X) = \sum_{i=0}^{d} \frac{a_{d-i} - a_{d-i}^p}{p} X^{pi} - \sum_{\substack{i_0+\cdots+i_d=p \\ 0 \le i_h < p, (0 \le h \le d)}} \frac{(p-1)!}{i_0! \cdots i_d!} a_d^{i_0} \cdots a_0^{i_d} X^{i_1+2i_2+\cdots+di_d}.$$

We now give a very simple proof of the following estimate, once more due to E. Dobrowolski [Do 1978]: *if a nonzero algebraic integer α of degree $\le d$ is not a root of unity, then*

$$\lceil\alpha\rceil > 1 + \frac{1}{4ed^2}$$

Let α be a nonzero algebraic integer of degree d. Denote by Σ the set of embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$, so that

$$\{\sigma\alpha\,;\,\sigma\in\Sigma\}=\{\alpha_1,\dots,\alpha_d\}$$

is the set of conjugates of α. For any positive integer h, the value of the *Newton sum*

$$S_h=\sum_{\sigma\in\Sigma}\sigma\alpha^h$$

(which is the *trace* of α^h from $\mathbb{Q}(\alpha)$ to $\mathbb{Q}$) is a rational integer. Let p be a prime number. Fermat's little Theorem gives the congruence $S_h\equiv S_h^p \pmod p$. On the other hand, using Lemma 3.18 with $k=d$ for the polynomial $f=X_1^h+\cdots+X_d^h$, we can write $S_{hp}-S_h^p=pg(\alpha_1,\dots,\alpha_d)$, for some $g\in\mathbb{Z}[X_1,\dots,X_d]$. Now $g(\alpha_1,\dots,\alpha_d)$ is an algebraic integer, and since it is a rational number, we get $S_{hp}\equiv S_h^p \pmod p$. This shows that the three numbers S_{hp}, S_h and S_h^p are congruent modulo p. For any $h\ge 1$ we have

$$|S_h|\le d\,\lceil\alpha\rceil^h.$$

We now assume $\lceil\alpha\rceil\le 1+1/(4ed^2)$. By the so-called *Bertrand's Postulate* (which was proved by Chebishev in 1850 – see [HaWr 1938], Chap. 22 and [GLin 1962], Th. 3.5.1), there exists a prime number p in the range $2ed<p<4ed$. For $1\le h\le d$, the estimates

$$|S_h|\le d\left(1+\frac{1}{4ed^2}\right)^d\le de\quad\text{and}\quad |S_{hp}|\le d\left(1+\frac{1}{4ed^2}\right)^{4ed^2}\le de$$

hold. Therefore $|S_h-S_{hp}|\le 2de<p$, which implies $S_h=S_{hp}$ for $1\le h\le d$. This means that α and α^p have the same minimal polynomial, i.e. that they are conjugates. One deduces from the following lemma that α is a root of unity.

Lemma 3.19. *Let α be a nonzero algebraic number. Assume that there exist two distinct positive rational integers h and ℓ such that α^h and α^ℓ are conjugate. Then α is a root of unity.*

Proof. Let K be the splitting field of α over $\mathbb{Q}$: if Σ denotes the set of embeddings of the field $\mathbb{Q}(\alpha)$ in $\mathbb{C}$, then K is the field generated over $\mathbb{Q}$ by $\{\sigma\alpha\,;\,\sigma\in\Sigma\}$. From the assumption that α^h and α^ℓ are conjugate, we deduce that there exists an element φ in the Galois group of K over $\mathbb{Q}$ such that $\varphi(\alpha^h)=\alpha^\ell$. By induction, for any $n\ge 1$, we deduce $\varphi^n(\alpha^{h^n})=\alpha^{\ell^n}$. Let m be the order of φ in the Galois group. Then $\alpha^{h^m}=\alpha^{\ell^m}$. Since $h\neq\ell$, we conclude that α is a root of unity. □

Proof of Theorem 3.16. We first notice that the inequality

$$\left(1+\frac{1}{4ed^2}\right)^{11d^2}>e$$

holds for $d\ge 2$. Hence for $d\ge 2$ we deduce

$$\mathrm{h}(\alpha) \geq \frac{1}{d} \log \lceil\overline{\alpha}\rceil \geq \frac{1}{d} \log\left(1 + \frac{1}{4ed^2}\right) > \frac{1}{11d^3}\cdot$$

□

Remark. Using the same arguments, E. Dobrowolski [Do 1978] also proves that a nonzero algebraic integer α which is a not root of unity satisfies

$$\lceil\overline{\alpha}\rceil \geq 1 + \frac{\log d}{6d^2}\cdot$$

As we have seen, sharper results [Mat 1991], [Du 1993], [Vou 1996] are now available.

In order to prove Theorem 3.17, we need some preparation.

The proof of the next lemma involves the *norm* $\mathrm{N}_{k/\mathbb{Q}}\colon k \to \mathbb{Q}$ of a number field k: for $\alpha \in k$, the norm of α with respect to the extension $k/\mathbb{Q}$ is the determinant of the endomorphism $x \mapsto \alpha x$ of the $\mathbb{Q}$-vector space k. If we denote again by Σ the set of embeddings of k into $\mathbb{C}$, then

$$\mathrm{N}_{k/\mathbb{Q}}(\alpha) = \prod_{\sigma\in\Sigma} \sigma\alpha.$$

The *absolute norm* of an algebraic number α is $\mathrm{N}_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha) \in \mathbb{Q}$. When α is an algebraic integer, we have $\mathrm{N}_{k/\mathbb{Q}}(\alpha) \in \mathbb{Z}$ (the absolute value of this integer is nothing else than the absolute norm of the principal ideal (α) in the ring of integers of k). The following relation holds for any element α in a number field k:

$$\mathrm{N}_{k/\mathbb{Q}}(\alpha) = \bigl(\mathrm{N}_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha)\bigr)^{[k\,:\,\mathbb{Q}(\alpha)]}.$$

Lemma 3.20. *Let p be a prime number and α an algebraic integer of degree d. Denote by Σ the set of embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$. Then the number*

$$\prod_{\sigma\in\Sigma}\prod_{\tau\in\Sigma}(\tau\alpha^p - \sigma\alpha)$$

is a rational integer which is divisible by p^d.

Proof. The minimal polynomial $f \in \mathbb{Z}[X]$ of α over $\mathbb{Z}$ can be written

$$f(X) = \prod_{\sigma\in\Sigma}(X - \sigma\alpha).$$

Hence

$$\prod_{\sigma\in\Sigma}\prod_{\tau\in\Sigma}(\tau\alpha^p - \sigma\alpha) = \prod_{\tau\in\Sigma} f(\tau\alpha^p) = \mathrm{N}_{\mathbb{Q}(\alpha)/\mathbb{Q}}\bigl(f(\alpha^p)\bigr).$$

Since α is an algebraic integer and $f \in \mathbb{Z}[X]$, the number $f(\alpha^p)$ is also an algebraic integer, hence its norm is a rational integer.

Using Lemma 3.18 with $k = 1$, we write $f(X^p) - f(X)^p = pg(X)$ for some $g \in \mathbb{Z}[X]$. Since $f(\tau\alpha) = 0$ for all $\tau \in \Sigma$, we deduce $f(\tau\alpha^p) = pg(\tau\alpha)$ and

$$\prod_{\tau\in\Sigma} f(\tau\alpha^p) = p^d \prod_{\tau\in\Sigma} g(\tau\alpha).$$

Finally, we observe that the number

$$\prod_{\tau\in\Sigma} g(\tau\alpha) = \mathrm{N}_{\mathbb{Q}(\alpha)/\mathbb{Q}}\bigl(g(\alpha)\bigr)$$

is a rational integer. □

Remark. Using the product formula in place of the norm would enable us to deal with algebraic numbers in place of algebraic integers. Compare with Lemma 3.23 below.

Lemma 3.21. *Let α be an algebraic integer of degree d, measure* $\mathrm{M}(\alpha)$ *and length* $\mathrm{L}(\alpha)$. *Let p be a prime number. If α is neither* 0 *nor a root of unity, then*

$$\mathrm{M}(\alpha) \ge \left(\frac{p}{\mathrm{L}(\alpha)}\right)^{1/p}.$$

Proof. From Lemma 3.19 (with $h = 1$ and $\ell = p$) we deduce that α and α^p are not conjugate. If $f(X) = X^d + a_1X^{d-1} + \cdots + a_d$ is the minimal polynomial of α, then the number $f(\alpha^p)$ is a nonzero algebraic integer. Hence its norm $N_{\mathbb{Q}(\alpha)/\mathbb{Q}}\bigl(f(\alpha^p)\bigr)$ over $\mathbb{Q}$ is a nonzero rational integer which, by Lemma 3.20, is divisible by p^d. A trivial upper bound for the absolute value of this number

$$N_{\mathbb{Q}(\alpha)/\mathbb{Q}}\bigl(f(\alpha^p)\bigr) = \prod_{\sigma\in\Sigma}\sum_{j=0}^{d} a_{d-j}\sigma\alpha^{pj}$$

(where $a_0 = 1$) is obtained as follows:

$$\bigl|N_{\mathbb{Q}(\alpha)/\mathbb{Q}}\bigl(f(\alpha^p)\bigr)\bigr| \le \prod_{\sigma\in\Sigma}\sum_{j=0}^{d} |a_{d-j}| \max\{1, |\sigma\alpha|^{pj}\} \le \mathrm{L}(\alpha)^d\mathrm{M}(\alpha)^{pd}.$$

Therefore

$$p^d \le \mathrm{L}(\alpha)^d\mathrm{M}(\alpha)^{pd}.$$

□

Remark. If we use the fact that there is a prime number p in the interval $[2\mathrm{L}(\alpha), 4\mathrm{L}(\alpha)]$, together with the estimate $2\log x \ge x\log 2$ which holds in the range $2 \le x \le 4$, we deduce

$$\mathrm{M}(\alpha) \ge \left(\frac{p}{\mathrm{L}(\alpha)}\right)^{1/p} \ge 2^{1/2\mathrm{L}(\alpha)} \ge 1 + \frac{\log 2}{2\mathrm{L}(\alpha)}\cdot$$

3.6.5 Dobrowolski's Proof of Theorem 3.17

In order to obtain a lower bound for $\mathrm{M}(\alpha)$ which depends only on the degree of α, one main idea of Dobrowolski's is to use the same outline, not for the minimal polynomial f of α itself, but for a suitable multiple of f^T, where T is a large integer.

The argument is as follows (see [Do 1979] and [Sc 1999], § 6). All throughout the proof we assume d is sufficiently large. Select two parameters L and T (depending on d) with $L > dT$. The first step establishes the existence of a nonzero polynomial F in $\mathbb{Z}[X]$, of degree $\leq L$ and length

$$\mathrm{L}(F) \leq L^{2dT^2/L},$$

which satisfies

$$\left(\frac{d}{dX}\right)^t F(\alpha) = 0 \quad \text{for} \quad 0 \leq t < T.$$

The second step is the zero estimate: there exists a prime number p in the range

$$2 \leq p \leq p_0 \quad \text{where} \quad p_0 := \frac{3}{2} \cdot \frac{L}{d} \log \frac{L}{d}$$

such that $F(\alpha^p) \neq 0$.

Assuming for the moment these two preliminary steps, we complete the proof of Theorem 3.17.

We deduce from Lemma 3.20 that the number

$$N = N_{\mathbb{Q}(\alpha)/\mathbb{Q}} F(\alpha^p) = \prod_{\tau \in \Sigma} F(\tau\alpha^p)$$

is a nonzero rational integer which is a multiple of p^{dT}. Hence

$$|N| \geq p^{dT}.$$

On the other hand the estimate

$$\prod_{\tau \in \Sigma} |F(\tau\alpha^p)| \leq \mathrm{L}(F)^d \prod_{\tau \in \Sigma} \max\{1, |\tau\alpha|\}^{pL}$$

gives

$$|N| \leq \mathrm{L}(F)^d \mathrm{M}(\alpha)^{pL}.$$

Therefore

$$\mathrm{M}(\alpha) \geq p^{dT/(pL)} \mathrm{L}(F)^{-d/(pL)} \geq p^{dT/(pL)} L^{-2d^2T^2/(pL^2)}.$$

We wish now to choose the parameters T and L such that the quantity

$$\min_{p \leq p_0} \left\{ \frac{dT}{pL} \log p - \frac{2d^2T^2}{pL^2} \log L \right\}$$

is *large*: its value will provide a lower bound for $\log \mathrm{M}(\alpha)$. Choose for instance the parameters as follows:

$$T = \left[5 \frac{\log d}{\log \log d} \right] \quad \text{and} \quad L = dT^2.$$

Since

$$\frac{dT^2}{L} > \frac{p}{2T^2 \log p} + \frac{2d^2T^3 \log L}{L^2 \log p},$$

one deduces the lower bound

$$\log \mathrm{M}(\alpha) \geq \frac{1}{2T^3},$$

which yields the conclusion of Theorem 3.17.

Let us come back to the first step: the construction of F. Using Dirichlet's box principle (see Exercise 3.12), Dobrowolski ([Do 1979], Lemma 1) shows the existence of $F \in \mathbb{Z}[\mathrm{X}]$ with a zero of multiplicity $\geq T$ at α and with the following upper bound for its height:

$$\mathrm{H}(F) \leq \left(\left(2^{3/2}(L+1)L^{(T-1)/2}\right)^{dT} \mathrm{M}(\alpha)^{TL} \right)^{1/(L-dT)}.$$

In order to deduce the desired upper bound for the length of F, it suffices to check

$$2^{3dT/2}(L+1)^L \mathrm{M}(\alpha)^{TL} \leq L^{(3dT^2/2)-(2d^2T^3/L)+(dT/2)}.$$

Given our choice of parameters (recall that d is sufficiently large), this estimate is satisfied as soon as $\log \mathrm{M}(\alpha) \leq (1/11) \log \log d$, an assumption which of course does not involve any loss of generality for the proof of Theorem 3.17.

In order to complete the proof of Theorem 3.17, we only need to prove the zero estimate of the second step: *one at least of the numbers $F(\alpha^p)$, with p prime in the range $2 \leq p \leq p_0$, is not zero.* The number of primes in this range is $> L/d$. We are going to check that the set

$$\left\{ \sigma\alpha^p \,;\, \sigma \in \Sigma,\ p \leq p_0 \right\}.$$

has more than L elements. It will follow that the nonzero polynomial $F \in \mathbb{Z}[\mathrm{X}]$ cannot vanish at all points in this set, which is what we want.

By Lemma 3.19, for $p_1 \neq p_2$ and for any σ and τ in Σ, we have $\sigma\alpha^{p_1} \neq \tau\alpha^{p_2}$. If there is a prime p for which the elements $\sigma\alpha^p$ $(\sigma \in \Sigma)$, are not pairwise distinct, then α^p is of degree $< d$, and we complete the proof of our claim by means of an inductive argument, thanks to the following lemma (compare with Lemma 3 of [Ra 1985]):

Lemma 3.22. *Let α be a nonzero algebraic integer which is not a root of unity. Define $d = [\mathbb{Q}(\alpha) \colon \mathbb{Q}]$. Assume that there exists a positive integer n such that $[\mathbb{Q}(\alpha^n) \colon \mathbb{Q}] < d$. Then there exists an algebraic integer β such that*

$$[\mathbb{Q}(\beta) \colon \mathbb{Q}] \leq \frac{d}{2} \quad \textit{and} \quad \mathrm{M}(\beta) \leq \mathrm{M}(\alpha).$$

Proof. Define $k = \mathbb{Q}(\alpha^n)$ and notice that α is a root of $X^n - \alpha^n \in k[X]$. Hence the irreducible polynomial g of α over k is a divisor of $X^n - \alpha^n$ in $k[X]$. It follows that the constant term, say $\beta \in k$, of g, can be written $\zeta\alpha^r$, where $r = [\mathbb{Q}(\alpha) \colon k]$ is the degree of g, while ζ is a n-th root of unity. Therefore we have

$$\mathrm{h}(\beta) = r\mathrm{h}(\alpha)$$

and

$$[\mathbb{Q}(\beta) \colon \mathbb{Q}] \le [k \colon \mathbb{Q}] = \frac{d}{r} = \frac{1}{r}[\mathbb{Q}(\alpha) \colon \mathbb{Q}],$$

hence

$$\mathrm{M}(\beta) \le \mathrm{M}(\alpha).$$

□

It is interesting to compare the previous sketch of proof with the usual one in transcendental number theory: a nonzero number is constructed and its absolute value is estimated from above and from below. But here, in place of a sharp analytic upper bound (Schwarz' Lemma) and a weak arithmetic lower bound (Liouville's inequality), we have a sharp arithmetic lower bound (coming from Fermat's little Theorem) and a trivial upper bound. However it is possible to give the proof in a way which is closer to the usual one, involving a sharp (ultrametric) upper bound together with the product formula. Here is the needed p-adic estimate.

Lemma 3.23. *Let α be an algebraic number of degree $d = [\mathbb{Q}(\alpha) \colon \mathbb{Q}]$. Let p be a prime number, $F \in \mathbb{Z}[X]$ a polynomial of degree L which vanishes at α with multiplicity at least T and v a place of $\mathbb{Q}(\alpha)$ which extends the p-adic valuation of $\mathbb{Q}$. Then*

$$|F(\alpha^p)|_v \le p^{-T} \max\{1, |\alpha|_v\}^{pL}.$$

Proof. Consider first the case where F is the minimal polynomial f of α over $\mathbb{Z}$. Using Lemma 3.18 with $k = 1$, we can write $f(X^p) = f(X)^p + pg(X)$ for some $g \in \mathbb{Z}[X]$ of degree $\le pd$. Hence

$$|g(\alpha)|_v \le \max\{1, |\alpha|_v\}^{pd}$$

and

$$|f(\alpha^p)|_v = |pg(\alpha)|_v \le p^{-1} \max\{1, |\alpha|_v\}^{pd},$$

which is what we wanted.

In the general case, F is divisible by f^T: let $G \in \mathbb{Z}[X]$ satisfy $F = f^T G$. Hence G has degree $L - dT$ and

$$|G(\alpha^p)|_v \le \max\{1, |\alpha|_v\}^{p(L-dT)}.$$

Therefore

$$|F(\alpha^p)|_v = |f(\alpha^p)|_v^T |G(\alpha^p)|_v \le p^{-T} \max\{1, |\alpha|_v\}^{pL}.$$

□

Remark 1. From Lemma 3.23 one deduces the following lower bound, which could be used in the proof of Theorem 3.17 in place of the one which we derived from Lemma 3.20 : *Under the assumptions of Lemma 3.23, assume that α is an algebraic integer and $F(\alpha^p) \neq 0$. Let Σ the set of embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$. Then*

$$\prod_{\tau\in\Sigma} |F(\tau\alpha^p)| \geq p^{dT}.$$

This follows from the product formula

$$\left(\prod_{\tau\in\Sigma} |F(\tau\alpha^p)|\right)\left(\prod_{v} |F(\alpha^p)|_v\right) = 1$$

where v runs over the set of ultrametric absolute values of $\mathbb{Q}(\alpha)$, using the upper bound $|F(\alpha^p)|_v \leq 1$ for any ultrametric absolute value v of $\mathbb{Q}(\alpha)$ such that $|p|_v = 1$.

Remark 2. Under the hypotheses of Lemma 3.23, assume further that α is an integer. Then one can derive the conclusion in the general case from the special case $F = f$ by means of an ultrametric Schwarz' lemma as follows.

Let Σ_v be the set of embeddings of $\mathbb{Q}(\alpha)$ into an algebraically closed field $\mathbb{C}_v$ containing the completion of $\mathbb{Q}(\alpha)$ at v. The analytic function $z \mapsto F(z)$ on $\mathbb{C}_v$ vanishes at the points $z = \sigma\alpha$ $(\sigma \in \Sigma_v)$ with multiplicity $\geq T$. Since $|\sigma\alpha|_v \leq 1$ for any $\sigma \in \Sigma_v$, we deduce

$$\left|\frac{F(w)}{\displaystyle\prod_{\sigma\in\Sigma_v}(w-\sigma\alpha)^T}\right|_v \leq R^{-dT} \sup_{|z|_v=R} |F(z)|_v$$

for any $w \in \mathbb{C}_v$ with $|w|_v \leq 1$ and any $R > 1$. Let $R \to 1$: for the same $w \in \mathbb{C}_v$, we obtain

$$|F(w)|_v \leq \prod_{\sigma\in\Sigma_v} |w-\sigma\alpha|_v^T = |f(w)|_v^T.$$

Remark 3. In [AmD 1999], Th. 3.1, F. Amoroso and S. David prove a multidimensional generalization of Lemma 3.23. They first prove the corresponding estimate when α is an integer, and deduce the general case by means of the strong approximation Theorem.

3.6.6 Proof of Theorem 3.17 Following Cantor-Straus and Rausch

We shall now provide the details of the proof of Theorem 3.17 by means of the idea of Cantor, Straus and Rausch which does not use Dirichlet's box principle, but replaces the auxiliary function by a determinant.

We proceed by induction on the degree $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. We may assume that d is sufficiently large, and that the conclusion of Theorem 3.17 holds for any algebraic integer of degree $< d$. Let α be an algebraic integer of degree d. By Lemma 3.22, we may assume that for any prime number p, the number α^p has degree d over $\mathbb{Q}$.

Let $P = \{p_1, \ldots, p_r\}$ be a set of r distinct primes. Define $L = d(T + r)$. As we have seen, Dobrowolski's original proof involved the construction of a nonzero auxiliary polynomial f of degree $< L$ which vanishes at the points $\sigma_1\alpha, \ldots, \sigma_d\alpha$ with multiplicity $\geq T$. The statement that *not all of the numbers* $f(\sigma_i\alpha^{p_j})$ *are zero* will be called *the zero estimate*.

In place of this construction, Cantor, Straus and Rausch consider the system of equations which occurs in the zero estimate, namely

$$\begin{cases} \dfrac{1}{t!}\left(\dfrac{d}{dX}\right)^t f(\sigma_i\alpha) = 0, & (0 \leq t < T, \quad 1 \leq i \leq d) \\ f(\sigma_i\alpha^{p_j}) = 0, & (1 \leq j \leq r, \quad 1 \leq i \leq d), \end{cases}$$

where the unknowns are the coefficients of $f \in \mathbb{Z}[X]$ with $\deg f < L$. The number of unknowns (the coefficients of f) is L, which is also the number of equations. Let Δ be the determinant of this system. We are going to write down Δ explicitly.

We consider the set $\{\zeta_1, \zeta_2, \ldots, \zeta_L\}$ of complex numbers defined by

$$\begin{cases} \zeta_{(i-1)T+j} = \sigma_i\alpha & \text{for } 1 \leq i \leq d \text{ and } 1 \leq j \leq T, \\ \zeta_{dT+d(j-1)+i} = \sigma_i\alpha^{p_j} & \text{for } 1 \leq i \leq d \text{ and } 1 \leq j \leq r. \end{cases}$$

This means that

- each of the d numbers $\sigma_1\alpha, \ldots, \sigma_d\alpha$ is repeated T times,
- each of the dr numbers $\sigma_i\alpha^{p_j}$ $(1 \leq i \leq d, 1 \leq j \leq r)$ occurs just once.

Next define nonnegative integers $\{t_1, \ldots, t_L\}$ by

$$\begin{cases} t_{(i-1)T+j} = j - 1 & \text{for } 1 \leq i \leq d \text{ and } 1 \leq j \leq T, \\ t_{dT+i} = 0 & \text{for } 1 \leq i \leq dr. \end{cases}$$

Therefore, for $1 \leq \lambda \leq L$,

$$t_\lambda = \mathrm{Card}\{\mu\,;\, 1 \leq \mu < \lambda,\ \zeta_\mu = \zeta_\lambda\} < \begin{cases} T & \text{if } \zeta_\lambda \text{ is of the form } \sigma_i(\alpha), \\ 1 & \text{if } \zeta_\lambda \text{ is of the form } \sigma_i\alpha^{p_j} \end{cases}$$

and we have

$$\Delta = \det\left(\binom{\mu-1}{t_\lambda}\zeta_\lambda^{\mu-1-t_\lambda}\right)_{1\leq\lambda,\mu\leq L},$$

where the binomial coefficient $\binom{m}{n}$ is defined as 0 for $n > m$.

We first check (zero estimate) $\Delta \neq 0$. Indeed, otherwise, there would exist a nonzero polynomial of degree $< L$ vanishing at $\sigma_1\alpha, \dots, \sigma_d\alpha$ with multiplicity $\geq T$, and with a root at each point $\sigma_i\alpha^{p_j}$ $(1 \leq i \leq d, 1 \leq j \leq r)$. Since the $d(r+1)$ numbers

$$\sigma_1\alpha, \dots, \sigma_d\alpha \quad \text{and} \quad \sigma_i\alpha^{p_j}, \quad (1 \leq i \leq d,\ 1 \leq j \leq r)$$

are pairwise distinct (recall Lemmas 3.19 and 3.22), and since a nonzero polynomial of degree $< L$ has not more than $L - 1$ roots (counting multiplicities), that is not possible.

We now invoke Fermat's little Theorem (Lemma 3.20) in order to get a lower bound for the absolute value of the interpolation determinant Δ.

Lemma 3.24. *We have*

$$|\Delta| \geq \prod_{j=1}^{r} p_j^{dT}.$$

Proof. For $T_1, \dots, T_m$ positive integers with $T_1 + \cdots + T_m = L$, consider the determinant $D \in \mathbb{Z}[X_1, \dots, X_m]$ of the following $L \times L$ matrix

$$\boldsymbol{M} = (\, \boldsymbol{M}_1 \quad \boldsymbol{M}_2 \quad \cdots \quad \boldsymbol{M}_m \,)$$

where, for $1 \leq j \leq m$, $\boldsymbol{M}_j$ denotes the $L \times T_j$ block

$$\boldsymbol{M}_j = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ X_j & 1 & \cdots & 0 \\ X_j^2 & 2X_j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ X_j^{\mu-1} & (\mu-1)X_j^{\mu-2} & \cdots & \binom{\mu-1}{T_j-1}X_j^{\mu-T_j} \\ \vdots & \vdots & \ddots & \vdots \\ X_j^{L-1} & (L-1)X_j^{L-2} & \cdots & \binom{L-1}{T_j-1}X_j^{L-T_j} \end{pmatrix}.$$

The square of D is a symmetric polynomial in $X_1, \dots, X_m$. Moreover for $1 \leq i < j \leq m$ the polynomial D is divisible, in the ring $\mathbb{Z}[X_1, \dots, X_m]$, by $(X_i - X_j)^{T_iT_j}$.

Choose $m = d(r+1)$, $T_1 = \cdots = T_d = T$, $T_{d+1} = \cdots = T_m = 1$. If we define, for $0 \leq j \leq r$ and $1 \leq i \leq d$,

$$\xi_{jd+i} = \sigma_i\alpha^{p_j}$$

where $p_0 = 1$, then we have $\Delta = \pm D(\xi_1, \dots, \xi_m)$. It follows from Lemma 3.20 that Δ^2 is a rational integer, which is divisible by p_j^{2dT} for $1 \leq j \leq r$. □

Next we produce an upper bound for the absolute value of Δ:

Lemma 3.25. *We have*

$$|\Delta| \le L^{d(T^2+r)/2}\mathrm{M}(\alpha)^{L(T+p_1+\cdots+p_r)}.$$

Proof. We use the so-called *Hadamard's inequality* (see for instance [F 1982], Appendix C):

- *the determinant* Δ *of a* $L \times L$ *matrix* $\left(a_{\lambda\mu}\right)_{1\le\lambda,\mu\le L}$ *satisfies the inequality*

$$|\Delta|^2 \le \prod_{\lambda=1}^{L}\sum_{\mu=1}^{L}|a_{\lambda\mu}|^2.$$

Here we get

$$|\Delta|^2 \le \prod_{\lambda=1}^{L}\sum_{\mu=1}^{L}\binom{\mu-1}{t_\lambda}^2\max\{1,|\zeta_\lambda|\}^{2(\mu-1)}.$$

We split the product on λ in two parts: the first one is

$$\begin{aligned}\prod_{\lambda=1}^{dT}\sum_{\mu=1}^{L}\binom{\mu-1}{t_\lambda}^2\max\{1,|\zeta_\lambda|\}^{2(\mu-1)} &= \prod_{i=1}^{d}\prod_{j=1}^{T}\sum_{\mu=1}^{L}\binom{\mu-1}{j-1}^2\max\{1,|\sigma_i\alpha|\}^{2L}\\ &\le \prod_{i=1}^{d}\prod_{t=0}^{T-1}L^{2t+1}\max\{1,|\sigma_i\alpha|\}^{2L}\\ &\le L^{dT^2}\mathrm{M}(\alpha)^{2LT}\end{aligned}$$

and the second

$$\begin{aligned}\prod_{\lambda=dT+1}^{L}\sum_{\mu=1}^{L}\binom{\mu-1}{t_\lambda}^2\max\{1,|\zeta_\lambda|\}^{2(\mu-1)} &= \prod_{i=1}^{d}\prod_{j=1}^{r}\sum_{\mu=1}^{L}\max\{1,|\sigma_i\alpha^{p_j}|\}^{2L}\\ &\le L^{dr}\mathrm{M}(\alpha)^{2L(p_1+\cdots+p_r)}.\end{aligned}$$

□

We now complete the proof of Theorem 3.17. From Lemmas 3.24 and 3.25 we derive

$$\prod_{j=1}^{r}p_j^{dT} \le L^{d(T^2+r)/2}\mathrm{M}(\alpha)^{L(T+p_1+\cdots+p_r)}.$$

Recall that $L = d(T+r)$. Take for $p_1,\ldots,p_r$ the first r primes with $r = T^2 - T$, and define

$$T = \left[5\frac{\log d}{\log\log d}\right]\cdot$$

From the prime number Theorem ([HaWr 1938], Th. 6 and Chap. 22) one deduces

$$\sum_{j=1}^{r} \log p_j \simeq r \log r \quad \text{and} \quad \sum_{j=1}^{r} p_j \simeq \frac{1}{2} r^2 \log r$$

as $r \to \infty$. Since, as soon as d is sufficiently large, we have

$$\frac{dT^2}{L} > \frac{dT \log L}{2L \log T} + \frac{4}{5},$$

we deduce

$$\log \mathrm{M}(\alpha) > \frac{1}{T^3}.$$

□

3.6.7 Further Related Questions and Results

Several related results are worth mentioning.

A. Schinzel, E. Dobrowolski and W. Lawton gave lower bounds for the height for an algebraic number α (which is neither zero nor a root of unity) in terms of the number of non-vanishing coefficients of a polynomial $f \in \mathbb{Z}[X]$ such that $f(\alpha) = 0$.

In [Mi 1979], M. Mignotte gave a lower bound for $|\alpha - 1|$ which is stronger than Liouville's one in terms of the degree; this is specially interesting when α has a small Mahler's measure. The proof involved an auxiliary polynomial. This estimate was improved in [MiW 1994] by means of an interpolation determinant. A p-adic analogue has been obtained by Y. Bugeaud [Bu 1998a]. Further refinements are due to E. M. Matveev [Mat 1996b], and F. Amoroso [Am 1996] and [Am 1998]. A remarkable connection with Riemann's hypothesis is described in [Am 1996] , while [Am 1998] contains a survey of such results.

Higher dimensional generalizations of Kronecker's Theorem, Lehmer's problem and Dobrowolski's estimate have been considered from different points of view.

In [AmD 1999], F. Amoroso and S. David extend Dobrowolski's result to simultaneous approximation. Lehmer's Problem is related to the multiplicative group $\mathbb{G}_{\mathrm{m}}$. Here is a generalization to $\mathbb{G}_{\mathrm{m}}^n$ suggested in [AmD 1999].

Conjecture 3.26. *For each positive integer $n \geq 1$ there exists a positive number $c_1(n)$ such that, if $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$ is a n-tuple of multiplicatively independent algebraic numbers and if $\omega(\underline{\alpha})$ denotes the minimum degree of a nonzero polynomial in $\mathbb{Q}[X_1, \ldots, X_n]$ which vanishes at $\underline{\alpha}$, then*

$$\mathrm{h}(1{:}\,\alpha_1{:}\cdots{:}\,\alpha_n) \geq \frac{c_1(n)}{\omega(\underline{\alpha})}. \tag{3.27}$$

A partial result is proved in [AmD 1999]:

$$\mathrm{h}(1{:}\,\alpha_1{:}\cdots{:}\,\alpha_n) \geq \frac{c_2(n)}{\omega(\underline{\alpha})\bigl(1 + \log \omega(\underline{\alpha})\bigr)^{\kappa(n)}}$$

for some positive constants $c_2(n)$ and $\kappa(n)$ which depend only on n.

A consequence (see Exercise 3.13) of Conjecture 3.26 is the following open problem:

(?) *For each positive integer $n \geq 1$ there exist a positive number $c_3(n)$ having the following property. Let $\alpha_1, \ldots, \alpha_n$ be multiplicatively independent algebraic numbers. Define $D = [\mathbb{Q}(\alpha_1, \ldots, \alpha_n) : \mathbb{Q}]$. Then*

$$\prod_{i=1}^{n} \mathrm{h}(\alpha_i) \geq \frac{c_3(n)}{D}. \tag{3.28}$$

A weaker estimate of the form

$$\prod_{i=1}^{n} \mathrm{h}(\alpha_i) \geq \frac{c_4(n)}{D(1+\log D)^{n\kappa(n)}} \tag{3.29}$$

is proved in [AmD 1999].

Another kind of higher dimensional *Lehmer type* problem arose from a work of S. Zhang in 1992 on positive line bundles on arithmetic varieties. He showed that *if $\mathcal{V}$ is a curve of a linear torus which is not a translate of a subtorus of positive dimension by a torsion point, then there exists a positive constant c such that $\mathcal{V}$ has only finitely many algebraic points of height $\leq c$.* Hence for sufficiently small c these points have a vanishing height. A more elementary proof of this result for the special case of the curve $x + y = 1$ in the torus $\mathbb{G}_{\mathrm{m}}^2$ was given by D. Zagier, with the best possible value for the constant: *any solution $(x, y) \in \overline{\mathbb{Q}}^2$ of the equation $x + y = 1$ with $x \neq 0$ and $x^6 \neq 1$ satisfies*

$$\mathrm{h}(x) + \mathrm{h}(y) \geq \frac{1}{2} \log \left(\frac{1+\sqrt{5}}{2} \right).$$

F. Beukers and D. Zagier gave sharp explicit results on this question and mentioned a number of applications. W. M. Schmidt, then E. Bombieri and U. Zannier, and then W. M. Schmidt again, extended Zagier's elementary argument to higher dimensional subvarieties of $\mathbb{G}_{\mathrm{m}}^n$. A survey on this topic is given by W. M. Schmidt in [Sc 1999].

Algebraic units u such that $1 - u$ is also a unit are sometimes called *exceptional units*. See [Sil 1996] for a survey of this topic.

One main tool is the notion of height for subvarieties of an affine or projective space. So far we have considered only the height of an algebraic point, which has dimension 0.

Subvarieties of codimension 1 are hypersurfaces . Lemma 3.9 suggested to Mahler a natural extension of his measure to polynomials in several variables (see [S 1999]):

$$\log \mathrm{M}(f) = \int_0^1 \cdots \int_0^1 \log |f(e^{2i\pi t_1}, \ldots e^{2i\pi t_n})| dt_1 \cdots dt_n.$$

The name *generalized cyclotomic polynomial* is sometimes used for a polynomial f (in several variables) which defines a hypersurface $\mathcal{V}$ of a torus containing a translate of a subtorus by a torsion point. By a result of D. Boyd, W. Lawton and C. Smyth, an irreducible polynomial f which is not a generalized cyclotomic polynomial has $\mathrm{M}(f) > 1$.

For a subvariety of any dimension, H. Gillet, C. Soulé, G. Faltings and P. Philippon introduced closely related notions of height[9]. For instance (see [DP 1999]) the canonical height of a hypersurface defined by $F = 0$ (where F is an irreducible polynomials with coefficients in $\mathbb{Z}$) is nothing else than $\log \mathrm{M}(F)$ where M is Mahler's measure (in several variables).

In the case (which we are interested in) of a subvariety of a torus, a *canonical height* can be defined (à la Néron-Tate), which vanishes exactly for the subvarieties containing a translate of a subtorus by a torsion point. For a hypersurface V of $\mathbb{G}_{\mathrm{m}}^n$ defined over $\mathbb{Q}$, say $f = 0$, which is not an algebraic subgroup of $\mathbb{G}_{\mathrm{m}}$, a lower bound for $\mathrm{h}(V)$ (that is for $\log \mathrm{M}(f)$) has been given in [AmD 2000].

An interesting related topic (see for instance [DP 1999]) is then to compare the height of the variety $\mathcal{V}$ with the minimum height of algebraic points on $\mathcal{V}$. The limit distribution of small points on algebraic tori has been studied by Y. Bilu.

These problems are the multiplicative analogues of a conjecture of F. A. Bogomolov about the discreteness of algebraic points on an algebraic curve of genus at least 2 with respect to the distance induced by the Néron-Tate height on the Jacobian. We do not deal here with Abelian varieties, and we shall only refer to work by L. Szpiro, J-F. Burnol, S. Zhang, E. Bombieri and U. Zannier, E. Ullmo, S. David and P. Philippon (extensions to semi-abelian varieties have also been considered by B. Poonen and A. Chambert-Loir).

There are further lower bounds for the height of algebraic numbers. For instance A. Schinzel and E. Dobrowolski got estimates which depend on the number of nonzero coefficients of the minimal polynomial. In [Mat 1996a], E. M. Matveev proves, for some classes of algebraic integers, a sharper estimate than Dobrowolski's one including the discriminant Δ of α: he replaces the degree d of α by $d/\Delta^{1/d}$.

Other estimates are due to M. Langevin; on one hand he proves [La 1986]: *Let $\mathcal{V}$ be a neighborhood of a point of the unit circle. There exists two effectively computable constants $c > 1$ and $D_0 > 0$ such that for any nonzero algebraic number α of degree $D \geq D_0$, all of whose conjugates are outside $\mathcal{V}$, the inequality $\mathrm{M}(\alpha) > c^D$ holds.* On the other hand, after a joint work with E. Reyssat and G. Rhin, he answered two questions of Favard by proving lower bounds for the *diameter* of an algebraic integer α, which is defined as

$$\mathrm{diam}(\alpha) = \max_{1 \leq i \neq j \leq d} |\alpha_i - \alpha_j|,$$

where $\{\alpha_1, \ldots, \alpha_d\}$ is the set of conjugates of α. These lower bounds are

$$\mathrm{diam}(\alpha) \geq \sqrt{3} \quad \text{for } d = [\mathbb{Q}(\alpha) : \mathbb{Q}] \geq 2$$

[9] A special case was already considered by W. M. Schmidt (see [Sc 1980]) who defined the height of a vector subspace by considering the Plücker coordinates of the corresponding point in a Grassmanian.

and, for any $\epsilon > 0$,

$$\operatorname{diam}(\alpha) \geq 2 - \epsilon \qquad \text{for } d \geq d_0(\epsilon).$$

For further results and references on the height of algebraic numbers, see Chap. 7 of [BerDGPS 1992], [S 1999] and [Sc 1999].

Open Problems

1. (Lehmer's problem [Le 1933] — see § 3.6). Does there exist an absolute constant $c_0 > 0$ such that, for any nonzero algebraic number which is not a root of unity, $d\mathrm{h}(\alpha) \geq c_0$?

2. (Conjecture of Schinzel and Zassenhaus [SZa 1965]). Does there exist an absolute constant $c > 0$ such that, for any nonzero algebraic integer of degree d which is not a root of unity, $\lceil\alpha\rceil \geq 1 + (c/d)$?

Since, for any algebraic integer α of degree d, we have $\mathrm{h}(\alpha) \leq \log\lceil\alpha\rceil \leq d\mathrm{h}(\alpha)$, the conjecture of Schinzel and Zassenhaus would follow from a positive answer to Lehmer's problem 1 above.

3. (A. Dubickas) Check that for any nonzero algebraic integer α of degree d which is not a root of unity,

$$\log\max\left\{\lceil\alpha\rceil\,;\,\lceil\alpha^{-1}\rceil\right\} \geq \frac{1}{d}\log 2.$$

4. (D. Boyd [Boy 1980]) The minimal value for $\lceil\alpha\rceil$ when α is a nonzero algebraic integer of degree d which is not a root of unity should be reached for the roots of

$$X^d + X^{2d/3} - 1$$

with d multiple of 3. An example is $X^3 + X^2 - 1$.

5. In the case $f(X) = qX - p$ with p and q rational integers, Liouville's inequality (Theorem 1.1) gives an estimate for the approximation of algebraic numbers by rational numbers. In this special case this lower bound is not the best known (Theorem 1.10 of Thue-Siegel-Roth-Schmidt; see [Sc 1980]). Is it possible to improve the estimate in the general case of Proposition 3.14? Even an ineffective result might be useful.

Exercises

Exercise 3.1. Let $\alpha_1, \ldots, \alpha_s$ be algebraic numbers. Define $k = \mathbb{Q}(\alpha_1, \ldots, \alpha_s)$ and $d = [k : \mathbb{Q}]$. Show that there exist rational integers $a_2, \ldots, a_s$ with $0 \le a_i \le d(d-1)/2$ such that the number $\gamma = \alpha_1 + a_2\alpha_2 + \cdots + a_s\alpha_s$ satisfies $k = \mathbb{Q}(\gamma)$.

Hint. *See* [MiW 1977] *Lemme 3.*

Exercise 3.2.
a) For $f \in \mathbb{C}[\mathrm{X}_1, \ldots, \mathrm{X}_t]$, we denote by $|f|_1$ the upper bound of $|f(\underline{z})|$ on the unit polydisc:

$$|f|_1 = \sup\left\{|f(z_1, \ldots, z_t)|\,;\ \underline{z} \in \mathbb{C}^t,\ |z_i| = 1,\ 1 \le i \le t\right\}.$$

Hence $|f|_1 \le \mathrm{L}(f)$. Show that in Lemmas 3.7, 3.8 and Proposition 3.14, one can replace $\log \mathrm{L}(f)$ by $\log |f|_1$.

Hint. *Start by proving the following statement: if* $a_0, \ldots, a_N$, y *are complex numbers, then*

$$\left|\sum_{i=0}^{N} a_i y^i\right| \le \sup_{|z|=1} \left|\sum_{i=0}^{N} a_i z^i\right| \cdot \max\bigl(1, |y|\bigr)^N.$$

When $|y| \le 1$, *this inequality follows from the maximum modulus principle for* $a_0 + a_1 z + \cdots + a_N z^N$. *When* $|y| > 1$, *perform the change of variables* $z' = 1/z$.

Deduce by induction: for a polynomial $f \in \mathbb{C}[\mathrm{X}_1, \ldots, \mathrm{X}_t]$, *when* $y_1, \ldots, y_t$ *are complex numbers,*

$$|f(y_1, \ldots, y_t)| \le |f|_1 \prod_{i=1}^{t} \max(1, |y_i|)^{\deg_{\mathrm{X}_i} f}.$$

b) For an algebraic number γ of degree d and minimal polynomial

$$a_0 X^d + \cdots + a_d = a_0 \prod_{i=1}^{d} (X - \gamma_i),$$

define a modified Mahler's measure by

$$\widetilde{\mathrm{M}}(\gamma) = a_0 \prod_{i=1}^{d} \sqrt{1 + |\gamma_i|}$$

and a modified absolute logarithmic height by

$$\widetilde{\mathrm{h}}(\gamma) = \frac{1}{d} \log \widetilde{\mathrm{M}}(\gamma).$$

Check, under the assumptions of Lemma 3.7,

$$\widetilde{\mathrm{h}}\bigl(f(\gamma_1, \ldots, \gamma_t)\bigr) \le \log \mathrm{H}(f) + \sum_{i=1}^{t} \bigl(\deg_{\mathrm{X}_i} f\bigr)\widetilde{\mathrm{h}}(\gamma_i).$$

Hint. *Compare with* [Sc 1991], *Chap. I, § 7, Lemma 7D.*

c) Let k be a number field of degree d. For $\underline{\gamma} = (\gamma_0 \colon \cdots \colon \gamma_\nu) \in \mathbb{P}_\nu(k)$, define

$$\widetilde{\mathrm{h}}(\underline{\gamma}) = \frac{1}{d} \sum_{v \in M_k} d_v \log \|\underline{\gamma}\|_v,$$

where

$$\|\underline{\gamma}\|_v = \begin{cases} \max\{|\gamma_0|_v, \ldots, |\gamma_\nu|_v\} & \text{for } v \text{ ultrametric,} \\ \sqrt{|\gamma_0|_v^2 + \cdots + |\gamma_\nu|_v^2} & \text{for } v \text{ Archimedean.} \end{cases}$$

Check that one can replace the height h by this modified height $\widetilde{\mathrm{h}}$ and at the same time the length L by the usual height H in Lemmas 3.7, 3.8 and 3.14.

Exercise 3.3.

a) Let N and M be positive integers and $\vartheta_1, \ldots, \vartheta_N, \theta_1, \ldots, \theta_M$ algebraic numbers. Check that

$$\mathrm{h}(1 \colon \vartheta_1 \colon \cdots \colon \vartheta_N \colon \theta_1 \colon \cdots \colon \theta_M) \le \mathrm{h}(1 \colon \vartheta_1 \colon \cdots \colon \vartheta_N) + \mathrm{h}(1 \colon \theta_1 \colon \cdots \colon \theta_M).$$

Deduce, for algebraic numbers $\vartheta_0, \ldots, \vartheta_s$, not all of which are zero,

$$\mathrm{h}(\vartheta_0 \colon \cdots \colon \vartheta_s) \le \sum_{i=0}^{s} \mathrm{h}(\vartheta_i).$$

b) Let $a_1, \ldots, a_n$ be rational integers, $b_1, \ldots, b_n$ be non-vanishing integers and $\beta_1, \ldots, \beta_n$ algebraic numbers. Define

$$N = \max\{|a_1|, |b_1|, \ldots, |a_n|, |b_n|\}$$

and

$$\gamma = \frac{a_1}{b_1}\beta_1 + \cdots + \frac{a_n}{b_n}\beta_n.$$

Then

$$\mathrm{h}(\gamma) \le n(n+1)\log N + \log n + \sum_{i=1}^{n} \mathrm{h}(\beta_i).$$

Hint. *This is Lemma 2.7 of* [W 1980].

c) Let $L_1, \ldots, L_k, N_1, \ldots, N_k$ and M be positive integers. For $1 \le i \le k$, let $\gamma_{0i}, \ldots, \gamma_{N_i i}$ be algebraic numbers. Assume that for each $i = 1, \ldots, k$, at least one of the numbers $\gamma_{0i}, \ldots, \gamma_{N_i i}$ is nonzero and denote by γ_i the point in $\mathbb{P}_{N_i}(\overline{\mathbb{Q}})$ with projective coordinates $(\gamma_{0i} \colon \cdots \colon \gamma_{N_i i})$. We will also write $\underline{\gamma}$ for the point $(\gamma_{\nu i})_{0 \le \nu \le N_i, 1 \le i \le k}$ in $\overline{\mathbb{Q}}^{N_1 + \cdots + N_k + k}$. Furthermore, let $F_0, \ldots, F_M$ be polynomials in $N_1 + \cdots + N_k + k$ variables, with coefficients in $\mathbb{Z}$, each of which is homogeneous of degree L_i with respect to the $N_i + 1$ variables $\mathrm{X}_{0i}, \ldots, \mathrm{X}_{N_i i}$. Assume that one at least of the $M + 1$ numbers $\theta_\mu = F_\mu(\underline{\gamma})$ $(0 \le \mu \le M)$ is nonzero, and define θ as the point in $\mathbb{P}_M(\overline{\mathbb{Q}})$ with projective coordinates $(\overline{\theta}_0 \colon \cdots \colon \theta_M)$. Then

$$\mathrm{h}(\theta) \le \log \max_{0 \le \mu \le M} \mathrm{L}(F_\mu) + \sum_{i=1}^{k} L_i \mathrm{h}(\gamma_i).$$

d) Let $P \in \overline{\mathbb{Q}}[\mathrm{X}_0, \ldots, \mathrm{X}_n, Y]$ be a homogeneous polynomial in $n + 2$ variables such that $P(0, \ldots, 0, 1) \not= 0$. Let $(\alpha_0 \colon \cdots \colon \alpha_n \colon \beta) \in \mathbb{P}_{n+1}(\overline{\mathbb{Q}})$ satisfy $P(\alpha_0 \colon \cdots \colon \alpha_n \colon \beta) = 0$. Then

$$\mathrm{h}(\alpha_0 \colon \cdots \colon \alpha_n \colon \beta) \le \mathrm{h}(\alpha_0 \colon \cdots \colon \alpha_n) + \mathrm{h}(p) + \log N,$$

where $N+1$ is the number of monomials in P and p is the projective point which is defined by the sequence of coefficients of P.
(Compare with [Ser 1989], § 2.3, N°4, Prop. 14.)
e) For any polynomial $F \in \mathbb{Z}[X,T]$, there exists a constant $c > 0$ such that, if α and β are algebraic numbers with $F(\alpha,\beta)=0$, and if the polynomial $F(\alpha,T) \in \mathbb{Q}(\alpha)[T]$ is not zero, then $\mathrm{h}(\beta) \le c \max\{1, \mathrm{h}(\alpha)\}$.

Exercise 3.4. For $f \in \mathbb{Z}[\mathrm{X}]$ a nonzero polynomial, define $t(f) = \deg f + \log \mathrm{H}(f)$. For an algebraic number α with minimal polynomial $f_\alpha \in \mathbb{Z}[\mathrm{X}]$, define $t(\alpha) = t(f_\alpha)$. Check the following *Liouville's inequality*:

If $f \in \mathbb{Z}[\mathrm{X}]$ and $\alpha \in \overline{\mathbb{Q}}$ satisfy $f(\alpha) \neq 0$, then

$$|f(\alpha)| \ge e^{-t(f)t(\alpha)}.$$

Hint. *One may use Proposition 3.14 together with the estimates*

$$(N+1)^{d-1}(d+1)^{N/2} \le e^{dN} \quad \textit{for integers } d \ge 1 \textit{ and } N \ge 0.$$

Remark. Another way of proving a lower bound for $|f(\alpha)|$ is to use the fact that the resultant of f and f_α is a nonzero rational integer - see [Bor 1899].

Exercise 3.5. Show that in Proposition 3.14, if the Archimedean absolute value v is not real, then the conclusion can be refined as

$$\log|f(\underline{\gamma})|_v \ge -\left(\frac{d}{2}-1\right)\log|f|_1 - \frac{d}{2}\sum_{i=1}^{t} N_i \mathrm{h}(1{:}\,\gamma_{i1}{:}\cdots{:}\,\gamma_{i\nu_i}).$$

Show also that if v is an ultrametric absolute value of k, then

$$\log|f(\underline{\gamma})|_v \ge -\frac{d}{d_v}\left(\log|f|_1 + \sum_{i=1}^{t} N_i \mathrm{h}(1{:}\,\gamma_{i1}{:}\cdots{:}\,\gamma_{i\nu_i})\right).$$

where d_v is, as usual, the local degree at v.

Hint. *Use Exercise 3.2.a.*

Exercise 3.6. Let $f \in \mathbb{Z}[X]$ be a nonzero polynomial of degree d with leading coefficient $a_0 > 0$ and let $\alpha \in \mathbb{C}$ be a zero of f.
a) Let p/q be a rational number with $q > 0$ such that $f(p/q) \neq 0$. Show that

$$\left|\alpha - \frac{p}{q}\right| \ge \frac{\max\{1,|\alpha|\}}{q(|p|+q)^{d-1}\mathrm{M}(f)}.$$

b) Deduce that for an algebraic number α of degree d, if we set

$$c(\alpha) = \begin{cases} \dfrac{1}{2^{d-1}\mathrm{M}(\alpha)} & \text{if } |\alpha| \le 1, \\ \dfrac{|\alpha|}{(2+|\alpha|)^{d-1}\mathrm{M}(\alpha)} & \text{if } |\alpha| > 1, \end{cases}$$

then for all $p/q \in \mathbb{Q}$ with $p/q \neq \alpha$ we have

$$\left|\alpha - \frac{p}{q}\right| \ge \frac{c(\alpha)}{q^d}\cdot$$

c) Show that, for each $\kappa > |f'(\alpha)|$, there are only finitely many $p/q \in \mathbb{Q}$ with

$$\left|\alpha - \frac{p}{q}\right| \le \frac{1}{\kappa q^d}\cdot$$

Example: Let α be a real quadratic number, which is root of a polynomial $aX^2 + bX + c$ of discriminant $\Delta = b^2 - 4ac > 0$. Then for each $\kappa > \sqrt{\Delta}$ there exist $q_0 > 0$ such that, for $p/q \in \mathbb{Q}$ with $q > q_0$,

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{\kappa q^2}\cdot$$

Exercise 3.7.
a) Let β be a nonzero algebraic number and λ a nonzero logarithm of an algebraic number. Define $\alpha = e^\lambda$ and $D = [\mathbb{Q}(\alpha, \beta) : \mathbb{Q}]$. Then

$$|\beta\lambda| > \left(2e^{\mathrm{h}(\alpha)+\mathrm{h}(\beta)}\right)^{-D}.$$

Hint. *Using (3.13), deduce* $|\beta| \ge e^{-D\mathrm{h}(\beta)}$. *Using Proposition 3.14, show that* $|\alpha - 1| \ge 2\left(2e^{\mathrm{h}(\alpha)}\right)^{-D}$ *if* $\alpha \neq 1$. *From Exercise 1.1, deduce* $\min\{|\alpha - 1|, 1\} < 2|\lambda|\}$.

b) Let $\lambda_1, \dots, \lambda_m$ be logarithms of algebraic numbers and $b_1, \dots, b_m$ rational integers. Let D be the degree of a number field containing the m algebraic numbers $\alpha_j = \exp(\lambda_j)$ $(1 \le j \le m)$. If the number

$$\Lambda = b_1\lambda_1 + \cdots + b_m\lambda_m$$

is nonzero, then

$$|\Lambda| \ge 2^{-D} \exp\left\{-D\sum_{j=1}^{m} |b_j|\mathrm{h}(\alpha_j)\right\}.$$

Exercise 3.8. Let $\alpha_1, \dots, \alpha_{n+1}$ be nonzero algebraic numbers and $\beta_1, \dots, \beta_n$ be algebraic numbers. Denote by D the degree of the number field

$$\mathbb{Q}(\alpha_1, \dots, \alpha_{n+1}, \beta_1, \dots, \beta_n).$$

Let T_0, T_1, $S_1, \dots, S_{n+1}$ be positive rational integers. Define $L = \binom{T_0+n}{n}(2T_1 + 1)$ and $S^* = \max\{S_1, \dots, S_{n+1}\}$. Further let $\underline{s}^{(1)}, \dots, \underline{s}^{(L)}$ be any elements in the set $\mathbb{Z}^{n+1}(\underline{S})$ of $\underline{s} = (s_1, \dots, s_{n+1}) \in \mathbb{Z}^{n+1}$ which satisfy $|s_i| \le S_i$ $(1 \le i \le n + 1)$. Let Δ be the determinant of the $L \times L$ matrix

$$\left((s_1^{(\mu)} + s_{n+1}^{(\mu)}\beta_1)^{\tau_1} \cdots (s_n^{(\mu)} + s_{n+1}^{(\mu)}\beta_n)^{\tau_n} \left(\alpha_1^{s_1^{(\mu)}} \cdots \alpha_{n+1}^{s_{n+1}^{(\mu)}}\right)^t \right)_{\substack{(\tau,t)\\ 1\le\mu\le L}},$$

where (τ, t) ranges over the set of elements $(\tau_1, \dots, \tau_n, t) \in \mathbb{N}^n \times \mathbb{Z}$ for which $\tau_1 + \cdots + \tau_n \le T_0$ and $|t| \le T_1$. Assume $\Delta \neq 0$. Prove

$$\frac{1}{L}\log|\Delta| \ge -(D-1)\big(T_0\log(2S^*) + \log L\big) - D(T_1+1)\sum_{i=1}^{n+1} S_i\mathrm{h}(\alpha_i) - DT_0\mathrm{h}(1 : \beta_1 : \cdots : \beta_n).$$

Hint. *Use Lemma 3.15 with* $\ell = 2n + 3$,

$$\nu_1 = \cdots = \nu_{2n+2} = 1, \quad \nu_{2n+3} = n,$$

$$p_{\lambda\mu} = \prod_{j=1}^{n} \bigl(s_j^{(\mu)} + s_{n+1}^{(\mu)} \mathrm{X}_{2n+3,j}\bigr)^{\tau_j} \cdot \prod_{i=1}^{n+1} \bigl(\mathrm{X}_{i1}^{\max\{ts_i^{(\mu)},0\}} \mathrm{X}_{n+1+i,1}^{\max\{-ts_i^{(\mu)},0\}}\bigr)$$

where (τ, t) *corresponds to the index* λ,

$$N_{i\lambda} = N_{n+1+i,\lambda} = t \max |s_i^{(\mu)}| \le tS_i \quad (1 \le i \le n+1)$$

and $N_{2n+3,\lambda} \le T_0$, *so that* $L(p_{\lambda\mu}) \le (2S^*)^{T_0}$,

$$\sum_{\lambda=1}^{L} N_{i\lambda} = \sum_{\lambda=1}^{L} N_{n+1+i,\lambda} \le \frac{1}{2} L(T_1 + 1) S_i \quad (1 \le i \le n+1)$$

and

$$\sum_{\lambda=1}^{L} N_{2n+3,\lambda} \le LT_0$$

Next apply Proposition 3.14 with $\gamma_{i1} = \alpha_i$ *for* $1 \le i \le n+1$, $\gamma_{i1} = \alpha_{i-n-1}^{-1}$ *for* $n+2 \le i \le 2n+2$ *and* $\gamma_{2n+3,j} = \beta_j$ *for* $1 \le j \le n$.

Exercise 3.9.

a) Check that for a nonzero algebraic number α, of degree $d \in \{1, 2, 3, 4, 5\}$, which is not a root of unity, the number $d\mathrm{h}(\alpha) = \log \mathrm{M}(\alpha)$ is bounded from below by the value given in table 3.30 (the last column provides a polynomial which yields the minimum).

Table 3.30

$d =$	$d\mathrm{h}(\alpha) \ge$	minimum for
1	$\log 2 = 0.6931\ldots$	$X - 2$
2	$\log\bigl((1+\sqrt{5})/2\bigr) = 0.4812\ldots$	$X^2 - X - 1$
3	$0.2811\ldots$	$X^3 - X - 1$
4	$0.3223\ldots$	$X^4 - X - 1$
5	$0.2998\ldots$	$X^5 - X^4 + X^3 - X + 1$

b) Show that the proof (see § 3.6) of Kronecker's result is effective: if d is a positive integer, there exists a positive number $c(d)$ such that, for any nonzero algebraic numbers α which is not a root of unity and is of degree at most d, the inequality $\mathrm{h}(\alpha) \geq c(d)$ is valid.

Hint. *Let α be an algebraic number of degree at most d. Assume that there exists a positive integer ℓ such that*

$$\mathrm{M}(\alpha)^{\ell} < 1 + 2^{-d} \quad \textit{and} \quad \ell \geq d(2^{d+1} + 1)^{d+1}.$$

Check $\mathrm{H}(\alpha^{j}) \leq 2^{d}$ *for* $0 \leq j \leq \ell$ *and deduce that the numbers* $1, \alpha, \ldots, \alpha^{\ell}$ *are not pairwise distinct.*

c) Let A and d be two positive integers, H and C two positive real numbers, and α a nonzero algebraic number of degree d. Assume

$$d\mathrm{h}(\alpha) \leq \frac{1}{H}, \quad \frac{1}{C^2} = \left(\frac{\pi}{A}\right)^2 + \left(\frac{2A-1}{H}\right)^2$$

and

$$C > 2^{d} e^{(2A-1)/H}.$$

Show that α is a root of unity of order $< 2A$.

Hint. *Show that there exists an integer r in the range $1 \leq r \leq 2A - 1$ such that $|\log(\alpha^r)| \leq 1/C$. Deduce $|\alpha^r - 1| < 2/C$. Use Liouville's inequality (3.14) for $f(X) = X^r - 1$ and conclude.*

d) Deduce from c) that a suitable value for $c(d)$ in question b) above is 2^{-2d-4} (compare with [SZa 1965]).

Hint. *Choose $A = 2^{d+2}$, $H = A^2$.*

Exercise 3.10. (see [CaStr 1982] and [Ra 1985]). Let a, b, c be positive real numbers and α an algebraic integer of degree $\leq d$ which is not a root of unity. Assume that there is a prime p in the range $ad < p \leq bd$. Assume also

$$\left(1 + \frac{c}{d^2}\right)^{bd^2} \leq \frac{a}{2}.$$

Deduce

$$\overline{|\alpha|} > 1 + \frac{c}{d^2}.$$

See also [Do 1978] for a much stronger estimate.

Exercise 3.11. Show that the polynomial D which occurs in the proof of Lemma 3.24 (namely the so-called *confluent Vandermonde determinant*) is equal to

$$\pm \prod_{1 \leq i < j \leq m} (X_i - X_j)^{T_i T_j}$$

Exercise 3.12.

a) Let ν, μ, ℓ be positive integers, p_{ij} $(1 \leq i \leq \nu, 1 \leq j \leq \mu)$ polynomials in $\mathbb{Z}[\mathrm{X}_1, \ldots, \mathrm{X}_\ell]$

and $\underline{\gamma} = (\gamma_1, \dots, \gamma_\ell)$ a tuple of algebraic numbers in a number field of degree D. Define, for $1 \leq j \leq \mu$ and $1 \leq k \leq \ell$,

$$N_{kj} = \max_{1 \leq i \leq \nu} \deg_{X_k} p_{ij}$$

and

$$V_j = \left(\sum_{i=1}^{\nu} \mathrm{L}(p_{ij}) \right) \prod_{k=1}^{\ell} e^{N_{kj} \mathrm{h}(\gamma_k)}.$$

Assume $\nu > D\mu$. Show that there exist $x_1 \dots, x_\nu$ in $\mathbb{Z}$ satisfying

$$\sum_{i=1}^{\nu} x_i p_{ij}(\underline{\gamma}) = 0 \qquad (1 \leq j \leq \mu)$$

and

$$0 < \max_{1 \leq i \leq \nu} |x_i| \leq 2 + \left(2^\mu \left(V_1 \cdots V_\mu \right)^D \right)^{1/(\nu - \mu D)}.$$

Hint. *Use Dirichlet's box principle as in the proof of Lemmas 4.11 and 4.12. See also Lemma 4 of* [MiW 1978] *and Lemma 1 of* [Do 1979].

b) Deduce the existence of F for the first step of the proof of Dobrowolski's Theorem given in § 3.6.5.

Exercise 3.13. Let $\alpha_1, \dots, \alpha_n$ be nonzero algebraic numbers. Denote by $[\mathbb{Q}(\underline{\alpha}) \colon \mathbb{Q}]$ the degree of the number field they generate.
a) Recall the notation $\omega(\underline{\alpha})$ in Conjecture 3.26. Check

$$\omega(\underline{\alpha}) \leq n [\mathbb{Q}(\underline{\alpha}) \colon \mathbb{Q}]^{1/n}.$$

Hint. *Using linear algebra, for any integer δ satisfying*

$$\binom{\delta + n}{n} > [\mathbb{Q}(\underline{\alpha}) \colon \mathbb{Q}],$$

show that there exists a nonzero polynomial $P \in \mathbb{Q}[\underline{X}]$ of total degree $\leq \delta$ such that $P(\underline{\alpha}) = 0$.

b) Assume $\alpha_1, \dots, \alpha_n$ are multiplicatively independent. Let $\epsilon > 0$. Show that there exist multiplicatively independent algebraic numbers $\gamma_1, \dots, \gamma_n$ such that

$$[\mathbb{Q}(\underline{\alpha}) \colon \mathbb{Q}] \mathrm{h}(\alpha_1) \cdots \mathrm{h}(\alpha_n) \geq (1 - \epsilon) [\mathbb{Q}(\underline{\gamma}) \colon \mathbb{Q}] \mathrm{h}(1 \colon \gamma_1 \colon \cdots \colon \gamma_n)^n.$$

Hint. *Since $\alpha_1, \dots, \alpha_n$ are multiplicatively independent, we have $\mathrm{h}(\alpha_i) > 0$ for $1 \leq i \leq n$. Let N be a sufficiently large integer. Define $A_i = [N \mathrm{h}(\alpha_i)]$ and select $\gamma_i = \alpha_i^{1/A_i}$.*

c) Deduce that Conjecture 3.28 is a consequence of Conjecture 3.26
d) Check also that (3.29) follows from (3.27).

Hint. *Use Theorem 3.16.*

Exercise 3.14. Let $P \in \mathbb{C}[X]$ be a nonzero polynomial of degree $\leq d$ and let p be a prime number. Check

$$\prod_{\zeta} |P(\zeta)| \leq p^d \mathrm{M}(P)^{p-1},$$

where ζ (in the product of the left hand side) ranges over the set with $p - 1$ elements of primitive p-th roots of unity.

Annex to Chapter 3. Inequalities Between Different Heights of a Polynomial - From a Manuscript by Alain Durand

Let $f \in \mathbb{C}[X]$ be a nonzero polynomial with complex coefficients of degree d:

$$f = a_0X^d + a_1X^{d-1} + \cdots + a_d = a_0\prod_{i=1}^{d}(X - \alpha_i).$$

There are several notions of *height* for f. For instance we have Mahler's measure of f (see § 3.3):

$$\mathrm{M}(f) = |a_0|\prod_{i=1}^{d}\max\{1, |\alpha_i|\},$$

the usual height of f (see § 3.4):

$$\mathrm{H}(f) = \max\{|a_0|, |a_1|, \ldots, |a_d|\},$$

the Euclidean norm of f:

$$\mathrm{L}_2(f) = \bigl(|a_0|^2 + |a_1|^2 + \cdots + |a_d|^2\bigr)^{1/2} = \left(\int_0^1 |f(e^{2i\pi t})|^2 dt\right)^{1/2},$$

the sup norm on the unit disc (or on the unit circle, which is the same by the maximum modulus principle):

$$|f|_1 = \sup_{|z|\le 1}|f(z)| = \sup_{|z|=1}|f(z)|,$$

and finally the length of f (see § 3.2):

$$\mathrm{L}(f) = |a_0| + |a_1| + \cdots + |a_d|.$$

Inequalities like

$$\boxed{(d+1)^{-1/2}\mathrm{M}(f) \le \mathrm{H}(f) \le \mathrm{L}_2(f) \le |f|_1 \le \mathrm{L}(f) \le 2^d\mathrm{M}(f)}$$

relate these functions. Table 3.31 below (due to the late Alain Durand) provides an upper bound for the quotient of one of the norms (left column) by another one (first row). In each case but two, below the upper bound is displayed one polynomial for which the estimate is optimal (where f_d denotes the polynomial $1 + X + \cdots + X^d$). There are two exceptions where the optimal result is not known:
(1) By (3.12),

$$\mathrm{M}(f) \le \sqrt{d+1}\,\mathrm{H}(f).$$

There are examples which show that for d sufficiently large, there exist polynomials f of degree d with $\mathrm{H}(f) = 1$ and $\mathrm{M}(f) \ge \sqrt{d+1} - (\log d)/2$ (see [Dur 1990], p.56).
(2) One can also prove that

$$\mathrm{L}(f) \le \sqrt{d}|f|_1$$

Table 3.31

	$\mathrm{M}(f)$	$\mathrm{H}(f)$	$\mathrm{L}_2(f)$	$\|f\|_1$	$\mathrm{L}(f)$
$\mathrm{M}(f) \leq$	1	$\sqrt{d+1}$ (1)	1 X^d	1 X^d	1 X^d
$\mathrm{H}(f) \ \leq$	$\binom{d}{[d/2]}$ $(X+1)^d$	1	1 X^d	1 X^d	1 X^d
$\mathrm{L}_2(f) \leq$	$\binom{2d}{d}^{1/2}$ $(X+1)^d$	$\sqrt{d+1}$ f_d	1	1 X^d	1 X^d
$\|f\|_1 \ \leq$	2^d $(X+1)^d$	$d+1$ f_d	$\sqrt{d+1}$ f_d	1	1 X^d
$\mathrm{L}(f) \ \leq$	2^d $(X+1)^d$	$d+1$ f_d	$\sqrt{d+1}$ f_d	$\sqrt{d}$ (2)	1

and give examples of polynomials f of degree d with $|f|_1 = 1$ and $\mathrm{L}(f) \geq \sqrt{d} - 3d^{1/6}$ for d sufficiently large (again see A. Durand, op. cit., 64–65, or J-P. Kahane, Sur les polynômes à coefficients unimodulaires, Bull. London Math. Soc., **12** (1980), 321–342).