Chapter 1

Heights on Projective and Affine Spaces

1.1 Absolute values

1.1.1 Basic notions

Definition 1.1.1. *An **absolute value** on a field K is a real valued function $|\cdot|$ on K such that*

(a) $|x| \geq 0$ *and* $|x| = 0$ *if and only if* $x = 0$.

(b) $|xy| = |x||y|$.

(c) $|x + y| \leq |x| + |y|$ *(triangle inequality).*

If furthermore $|\cdot|$ satisfies instead of (c) the stronger condition

(c') $|x + y| \leq \max\{|x|, |y|\}$ *(ultrametric triangle inequality),*

*then it is called **non-archimedean**. If (c') fails to hold for some $x, y \in K$, then the absolute value is called **archimedean**.*

Example 1.1.2. *(i) The **trivial absolute value**:* $|0| = 0$ *and* $|x| = 1$ *for all* $x \in K^*$.

(ii) $K = \mathbb{Q}$

- *An archimedean absolute value defined by*

$$|x|_\infty = \max\{x, -x\}.$$

- *A non-archimedean absolute value for each prime number p defined as follows. For any nonzero rational number $x \in \mathbb{Q}$, there exists a unique integer $\mathrm{ord}_p(x)$ such that x can be written in the form*

$$x = p^{\mathrm{ord}_p(x)} \frac{a}{b} \quad \text{with } a, b \in \mathbb{Z} \text{ and } p \nmid ab.$$

*If $x = 0$, then we set $\mathrm{ord}_p(x) = +\infty$. The **$p$-adic absolute value** of $x \in \mathbb{Q}$ is the quantity*

$$|x|_p = p^{-\mathrm{ord}_p(x)}.$$

Intuitively, x is p-adically small if it is divisible by a large power of p.

Each absolute value $|\cdot|$ on K induces a topology via the metric defined by $\mathrm{disc}(x,y) = |x-y|$. If two absolute values define the same topology, they are called *equivalent*. Here is a basic property.

Proposition 1.1.3. *Two absolute values $|\cdot|_1$ and $|\cdot|_2$ are equivalent if and only if there exists a positive real number s such that*

$$|x|_1 = |x|_2^s$$

for each $x \in K$.

In practice, it is more convenient to study equivalence classes of absolute values.

Definition 1.1.4. *A **place** v is an equivalent class of non-trivial absolute values. By $|\cdot|_v$ we denote an absolute value in the equivalence class determined by the place v.*

*We say that a place v is **(non-)archimedean** if $|\cdot|_v$ is.*

As an example, $\mathbb{Q}$ has a unique archimedean place and there is a natural bijection

$$\{\text{non-archimedean places of } \mathbb{Q}\} \leftrightarrow \{\text{all prime numbers}\};$$

see Example 1.1.2 (ii) for $|\cdot|_v$ with each place v of $\mathbb{Q}$.

Consider a field extension K/K_0. For a place v of K, the restriction of $|\cdot|_v$ to K_0 is an absolute value of K_0, and hence is a representative of a place of K_0. We write $v|v_0$ if and only if the restriction of $|\cdot|_v$ to K_0 is a representative of $v_0 \in M_{K_0}$. In this case, we say that v *divides* v_0 or v *lies over* v_0 or v *extends* v_0.

Before moving on, let us look at the example of an arbitrary number field K.

Example 1.1.5. *By definition, $K/\mathbb{Q}$ is a finite field extension, and hence any place v of K lies over some place of $\mathbb{Q}$. There are two possibilities: either $v|p$ for a prime number p, or $v|\infty$ for the unique archimedean place ∞ of $\mathbb{Q}$. It can be then checked that*

$$\{\textit{non-archimedean places of } K\} \leftrightarrow \{\textit{all prime ideals of } \mathcal{O}_K\}$$

and

$$\{\textit{archimedean places of } K\} \leftrightarrow \{\textit{equivalence classes of embeddings } K \hookrightarrow \mathbb{C}\}.$$ [1]

We will come back to this with a more precise description of the bijections in Example 1.1.11.

We close this subsection with the following discussion. Let K be a field with a non-archimedean place v. The *valuation ring* of v is defined to be

$$R_v := \{x \in K : |x|_v \leq 1\}.$$

The definition is clearly independent of the choice of $|\cdot|_v$. It can be checked that R_v is local ring with unique maximal ideal $\mathfrak{m}_v := \{x \in K : |x|_v < 1\}$. The *residue field* $k(v)$ is defined to be $R_v/\mathfrak{m}_v$. The quotient map $R_v \to k(v)$, $x \mapsto \overline{x}$ is called the *reduction*.

For example when $K = \mathbb{Q}$ and v corresponds to the prime number p, we have $R_v = \{x \in \mathbb{Q} : p^{-\mathrm{ord}_p(x)} \leq 1\} = \{x \in \mathbb{Q} : \mathrm{ord}_p(x) \geq 0\} = \{\frac{a}{b} : a \text{ and } b \text{ coprime, } p \nmid b\}$ and $\mathfrak{m}_v = \{x \in \mathbb{Q} : \mathrm{ord}_p(x) > 0\} = \{\frac{a}{b} : a \text{ and } b \text{ coprime, } p|a\} = pR_v$. The residue field is $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

The place v is called *discrete* if the value group $|K^*|_v$ is cyclic. Then $\mathfrak{m}_v$ is a principal ideal and any principal generator is called a *local parameter*. This is the case for the example above [2] and a local parameter is p.

[1] Two embeddings $\sigma_1, \sigma_2 \colon K \hookrightarrow \mathbb{C}$ are equivalent if and only if they are conjugate (*i.e.* $\sigma_2(x) = \overline{\sigma_1(x)}$ for all $x \in K$).

[2] This holds true for any number field and a non-archimedean place.

1.1.2 Normalized absolute values

For each place v of K, we would like assign a well-chosen absolute value. In this subsection we do this.

Definition-Proposition 1.1.6. *For a place v of K, there exists a unique (up to isometric isomorphisms) pair (K_v, w) with K_v/K an extension and w a place of K_v satisfying the following properties:*

(a) $w|v$.

(b) The topology of K_v induced by w is complete.

(c) K is dense in K_v in the above topology.

*This K_v is called the **completion** of K with respect to v. By abuse of notation, we shall denote the unique place w also by v.*

As an example, the field $\mathbb{Q}_p$ of p-adic numbers is the completion of $\mathbb{Q}$ with respect to the place p, and the completion of $\mathbb{Q}$ with respect to the archimedean place is $\mathbb{R}$. In general, we have:

Theorem 1.1.7 (Ostrowski)**.** *The only complete archimedean fields are $\mathbb{R}$ and $\mathbb{C}$.*

An elementary result of the local and global degrees is the following equality. It can be proved using the primitive element theorem.

Lemma 1.1.8. *Let K_0 be a field with a place v_0, and let K/K_0 be a finite separable extension. Then*

$$\sum_{v|v_0}[K_v : K_{0,v_0}] = [K : K_0].$$

With these preparations in hand, we are ready to state the following result about the uniqueness of the extension of absolute values.

Proposition 1.1.9. *Let K_0 be a field which is complete with respect to an absolute value $|\cdot|_{v_0}$ (i.e. $K_0 = K_{0,v_0}$) and let K/K_0 be a finite extension. Then there exists a unique extension of $|\cdot|_{v_0}$ to an absolute value $|\cdot|_v$ of K. Furthermore, for each $x \in K$, we have*

$$|x|_v = |N_{K/K_0}(x)|_{v_0}^{1/[K:K_0]}$$

where N_{K/K_0} is the norm. Moreover, K is complete with respect to $|\cdot|_v$, i.e. $K = K_v$.

Inspired by this proposition, we make the following constructions. Let K_0 be a field with a non-trivial absolute value $|\cdot|_{v_0}$. Let K/K_0 be a finite separate extension with a place v such that $v|v_0$. For any $x \in K$, define

$$|x|_v := |N_{K_v/K_{0,v_0}}(x)|_{v_0}^{1/[K_v:K_{0,v_0}]} \tag{1.1.1}$$

and

$$\|x\|_v := |N_{K_v/K_{0,v_0}}(x)|_{v_0}. \tag{1.1.2}$$

The following statements are easy to verify.

- The $|\cdot|_v$ defined above is an absolute value representing v by Proposition 1.1.9.

- The $\|\cdot\|_v$ defined above is an absolute value representing v *unless* $K_{0,v_0}=\mathbb{R}$ and $K_v=\mathbb{C}$.

In practice, it is however often more practical to use $\|\cdot\|_v$ than $|\cdot|_v$.

Lemma 1.1.10. *Under the assumptions and notation above, we have*

$$\sum_{v|v_0}\log\|x\|_v=[K:K_0]\log|x|_{v_0}\qquad \text{for all } x\in K_0^*,$$
$$\sum_{v|v_0}\log\|y\|_v=\log|N_{K/K_0}(y)|_{v_0}\qquad \text{for all } y\in K^*.$$

Example 1.1.11. *Again, let us look at the case of number fields. When $K=\mathbb{Q}$, set*

$$M_{\mathbb{Q}}:=\{|\cdot|_p : p \text{ prime number or } p=\infty\}$$

normalized as in Example 1.1.2.(ii).

In general for an arbitrary number field K.

- *Each place v of K lying over p corresponds to a unique prime ideal $\mathfrak{p}$ dividing p.*

  For each $x\in K^*$, the fractional ideal $x\mathcal{O}_K$ can be uniquely factorized into a finite product $\prod_{\mathfrak{p}}\mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}}(x)}$ with $\mathfrak{p}$ running over all the prime ideals of $\mathcal{O}_K$. This defines a homomorphism $\mathrm{ord}_{\mathfrak{p}}\colon K^*\to\mathbb{Z}$ for each $\mathfrak{p}$ (and set $\mathrm{ord}_{\mathfrak{p}}(0):=+\infty$).

  Set $|x|_{\mathfrak{p}}:=p^{-[k(\mathfrak{p}):\mathbb{F}_p]\mathrm{ord}_{\mathfrak{p}}(x)/[K_v:\mathbb{Q}_p]}=p^{-\mathrm{ord}_{\mathfrak{p}}(x)/e_{\mathfrak{p}}}$.[3] *Notice that $|p|_{\mathfrak{p}}=p^{-1}$.*

  We show that $|\cdot|_{\mathfrak{p}}$ is precisely the absolute value $|\cdot|_v$ from (1.1.1). Indeed, we have for (1.1.2)

$$\|x\|_v=|N_{K_{\mathfrak{p}}/\mathbb{Q}_p}(x)|_p=|N_{K_{\mathfrak{p}}/\mathbb{Q}_p}\mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}}(x)}|_p=|p^{[k(\mathfrak{p}):\mathbb{F}_p]\mathrm{ord}_{\mathfrak{p}}(x)}|_p=p^{-[k(\mathfrak{p}):\mathbb{F}_p]\mathrm{ord}_{\mathfrak{p}}(x)}.$$

  It is a standard fact from Algebraic Number Theory that $[K_{\mathfrak{p}}:\mathbb{Q}_p]=e_{\mathfrak{p}}[k(\mathfrak{p}):\mathbb{F}_p]$. Thus $|x|_v=\|x\|_v^{1/[K_{\mathfrak{p}}:\mathbb{Q}_p]}$ equals $|x|_{\mathfrak{p}}$ defined above.

  Now we set

$$M_K^0:=\{|\cdot|_{\mathfrak{p}} : \mathfrak{p} \text{ prime ideal of } \mathcal{O}_K\}. \tag{1.1.3}$$

  Then each element M_K^0 is a representative of a non-archimedean place of K, and all non-archimedean places of K arises in this way.

- *For an archimedean place v of K, it lies over the unique archimedean place of $\mathbb{Q}$ which gives rise to the embedding $\mathbb{Q}\hookrightarrow\mathbb{R}$. Consider all the embeddings $\sigma\colon K\hookrightarrow\mathbb{C}$; there are exactly $[K:\mathbb{Q}]$ of them. Each such embedding defines an absolute value on K*

$$|x|_\sigma:=|\sigma(x)|_\infty$$

  where $|z|_\infty$ is the usual absolute value on $\mathbb{R}$ or $\mathbb{C}$. It can be shown that all archimedean places of K arise in this way.

  Among the embeddings $K\hookrightarrow\mathbb{C}$ there are two kinds: r_1 real embeddings with $\sigma(K)\subseteq\mathbb{R}$ (call them $\rho_1,\dots,\rho_{r_1}$) and r_2 complex embeddings with $\sigma(K)\not\subseteq\mathbb{R}$ (call them $\tau_1,\overline{\tau}_1,\dots,\tau_{r_2},\overline{\tau}_{r_2}$). The complex embeddings come in pairs under the complex conjugation. We have $[K:\mathbb{Q}]=r_1+2r_2$. One can show that two embeddings $K\hookrightarrow\mathbb{C}$ give rise to equivalent absolute values if and only if they are conjugate.

  In summary, there are r_1+r_2 archimedean places of K. Set

$$M_K^\infty:=\{|\cdot|_\sigma\}_{\sigma\in\{\rho_1,\dots,\rho_{r_1},\tau_1,\dots,\tau_{r_2}\}}. \tag{1.1.4}$$

[3] Recall the standard definition $e_{\mathfrak{p}}=\mathrm{ord}_{\mathfrak{p}}(p)$ from Algebraic Number Theory.

Now set

$$M_K = M_K^0 \bigcup M_K^\infty. \tag{1.1.5}$$

From now on, for a number field K we will always use M_K to denote the set from (1.1.5). Moreover, the following convention on the notation M_K for a number field K will always be used in this course.

Notation 1.1.12. *It is sometimes more convenient to work with $\|\cdot\|_v$ than $|\cdot|_v$, and so we will also use the following notation. By $v \in M_K$ for a number field K, we always use $|\cdot|_v$ to denote the corresponding absolute value in the set from (1.1.5), and use $\|\cdot\|_v$ to denote $|\cdot|_v^{[K_v:\mathbb{Q}_p]}$ for $v|p$ and $|\cdot|_v^{[K_v:\mathbb{R}]}$ for $v|\infty$. Notice that when $K = \mathbb{Q}$, $\|\cdot\|_v$ and $|\cdot|_v$ coincide.*

We finish this subsection by the following Product Formula.

Theorem 1.1.13 (Product Formula)**.** *Let K be a number field. Then*

$$\sum_{v \in M_K} \log \|x\|_v = 0 \quad \text{for each } x \in K^*.$$

Proof. Let $x \in K^*$. We start with the case $K = \mathbb{Q}$. In this case, $x = \prod_p p^{\mathrm{ord}_p(x)}$ with p running over all prime numbers. Then

$$\prod_{v \in M_\mathbb{Q}} |x|_v = |x|_\infty \prod_p |x|_p = |x|_\infty \prod_p p^{-\mathrm{ord}_p(x)} = 1.$$

So $\sum_{v \in M_\mathbb{Q}} \log |x|_v = 0$.

For arbitrary K, apply Lemma 1.1.10 to $K/\mathbb{Q}$ and $v|v_0$ with v_0 a place of $M_\mathbb{Q}$. Then we obtain $\sum_{v|p} \log \|x\|_v = \frac{1}{[K:\mathbb{Q}]} \log |N_{K/\mathbb{Q}}(x)|_{v_0}$. So

$$\sum_{v \in M_K} \log \|x\|_v = \sum_{v_0 \in M_\mathbb{Q}} \sum_{v|v_0} \log \|x\|_v = \sum_{v_0 \in M_\mathbb{Q}} \log |N_{K/\mathbb{Q}}(x)|_{v_0},$$

which equals 0 from the case $K = \mathbb{Q}$. Hence we are done. □

1.2 Height on projective spaces

In the whole section, we will use K to denote a number field.

1.2.1 Definition and basic properties

Let us start with the simplest case. Let $x \in \mathbb{P}^1(\mathbb{Q})$. There is a unique way to write x as $[a : b]$ with $a, b \in \mathbb{Z}$ such that we are in one of the following two cases:

- $a = 0$, $b = 1$ or $a = 1$, $b = 0$;
- $a > 0$ and $b \neq 0$ are coprime.

Set

$$H(x) := \max\{|a|, |b|\}.$$

Notice that $H(x) \geq 1$ by definition. Also notice that any rational number x can be identified with $[x : 1] \in \mathbb{P}^1(\mathbb{Q})$, so we can set $H(x) := H([x : 1])$.

In Height Theory, it turns out to be more convenient to work with the *logarithmic height*. On $\mathbb{P}^1(\mathbb{Q})$ it is $h(x) := \log H(x) = \log \max\{|a|, |b|\}$. Then we have $h(x) \geq 0$ for each $x \in \mathbb{P}^1(\mathbb{Q})$.

For more general number fields, we will use the absolute values introduced in the previous section (Example 1.1.11) to define the height. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$.

Definition 1.2.1. *Let $\mathbf{x} = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$. The **(absolute logarithmic Weil) height** of x is defined to be*

$$h(\mathbf{x}) := \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} \log \max\{\|x_0\|_v, \ldots, \|x_n\|_v\},$$

where $K \subseteq \overline{\mathbb{Q}}$ is a number field such that $x_j \in K$ for all j.

*We also set $H(\mathbf{x}) := e^{h(\mathbf{x})}$ to be the **multiplicative height**.*

One can check that this definition coincides with the one for $\mathbb{P}^1(\mathbb{Q})$ above. More generally, we have the following lemma.

Lemma 1.2.2. *Let $\mathbf{x} = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\mathbb{Q})$. Suppose the x_j's are all integers and are coprime. Then*

$$h(\mathbf{x}) = \log \max\{|x_0|, \ldots, |x_n|\}$$

with the usual absolute value.

Proof. Exercise class. □

Lemma 1.2.3. *The height function defined above satisfies the following properties.*

(i) It is independent of the choice of K.

(ii) It is independent of the choice of the homogeneous coordinates.

(iii) $h(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{P}^n(\overline{\mathbb{Q}})$.

Proof. Let $\mathbf{x} = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$.

For (i): Assume that each x_j is in K and L for two number fields $K, L \subseteq \overline{\mathbb{Q}}$. We may assume $K \subseteq L$. Then

$$\begin{aligned}
\sum_{w \in M_L} \log \max_j \|x_j\|_w &= \sum_{v \in M_K} \sum_{w|v} \log \max_j \|x_j\|_w \\
&= \sum_{v \in M_K} \sum_{w|v} \log \max_j \|N_{L_w/K_v}(x_j)\|_v \\
&= \sum_{v \in M_K} \sum_{w|v} \log \max_j \|x_j\|_v^{[L_w : K_v]} \\
&= \sum_{v \in M_K} \sum_{w|v} [L_w : K_v] \log \max_j \|x_j\|_v \\
&= \sum_{v \in M_K} [L : K] \log \max_j \|x_j\|_v \quad \text{by Lemma 1.1.8.}
\end{aligned}$$

This establishes (i).

For (ii): Let $[x_0 : \cdots : x_n]$ and $[y_0 : \cdots : y_n]$ be two homogeneous coordinates for a point $\mathbf{x} \in \mathbb{P}^n(\overline{\mathbb{Q}})$. By part (i), we may and do assume that all coordinates are in the same number field K. Then there exists $\lambda \in K^*$ such that $y_j = \lambda x_j$ for each $j \in \{0, \ldots, n\}$. We have then

$$\sum_{v \in M_K} \log \max_j \|y_j\|_v = \sum_{v \in M_K} \log \max_j \|x_j\|_v + \sum_{v \in M_K} \log \|\lambda\|_v = \sum_{v \in M_K} \log \max_j \|x_j\|_v,$$

where the last equality follows from the Product Formula (Theorem 1.1.13). This establishes (ii).

Part (iii) follows from part (ii) because we can always choose homogeneous coordinates for x such that some coordinate is 1. □

Lemma 1.2.4. *The action of the Galois group* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *on* $\mathbb{P}^n(\overline{\mathbb{Q}})$ *leaves the height invariant. More precisely, for any* $\mathbf{x} \in \mathbb{P}^n(\overline{\mathbb{Q}})$ *and any* $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$*, we have* $h(\sigma(\mathbf{x})) = h(\mathbf{x})$.

Proof. Exercise class. □

The following theorem is of fundamental importance for the Height Machine.

Theorem 1.2.5 (Northcott Property)**.** *For each* $B \geq 0$ *and* $D \geq 1$*, the set*

$$\{\mathbf{x} \in \mathbb{P}^n(\overline{\mathbb{Q}}) : h(\mathbf{x}) \leq B,\ [\mathbb{Q}(\mathbf{x}) : \mathbb{Q}] \leq D\}$$

is a finite set.

Proof. We start with the case $D = 1$. Then the set in question becomes

$$\{\mathbf{x} \in \mathbb{P}^n(\mathbb{Q}) : h(\mathbf{x}) \leq B\}.$$

It is not hard to check that this set is finite by Lemma 1.2.2.

For general D. Write $\mathbf{x} = [x_0 : \cdots : x_n] \in \mathbb{P}^n(K)$ such that at least one coordinate equals 1. Then for each $v \in M_K$, we have

$$\max\{\|x_0\|_v, \ldots, \|x_n\|_v\} \geq \max\{\|x_i\|_v, 1\}$$

for each $i \in \{0, \ldots, n\}$. So $B \geq h(\mathbf{x}) \geq h(x_i)$ for each $i \in \{0, \ldots, n\}$. Moreover, $x_i \in K$, and hence $\mathbb{Q}(x_i) \subseteq \mathbb{Q}(x)$ and hence $[\mathbb{Q}(x_i) : \mathbb{Q}] \leq [\mathbb{Q}(x) : \mathbb{Q}] \leq D$.

It suffices to prove that there are finitely many choices for x_i for each $i \in \{0, \ldots, n\}$. Thus it suffices to establish the following simpler finiteness result.

Claim: For each numbers $B \geq 0$ and $d \geq 1$, the set

$$\{x \in \overline{\mathbb{Q}} : h(x) \leq B,\ [\mathbb{Q}(x) : \mathbb{Q}] = d\}$$

is finite.

Let us prove this claim. Write $K = \mathbb{Q}(x)$, and write $x_1 = x, \ldots, x_d$ for the Galois conjugates of x over $\mathbb{Q}$. The minimal polynomial of x over $\mathbb{Q}$ is

$$F(T) = \prod_{j=1}^{d} (T - x_j) = \sum_{r=0}^{d} (-1)^r s_r(x) T^{d-r}$$

with $s_r(x)$ the r-th symmetric polynomial in $x_1, \dots, x_d$. Denote by $s_r = s_r(x)$; it is a number in $\mathbb{Q}$. For each $v \in M_K$ we have

$$\begin{aligned} |s_r|_v &= \left| \sum_{1 \le i_1 < \cdots < i_r \le d} x_{i_1} \cdots x_{i_r} \right|_v \\ &\le \epsilon(v,r,d) \max_{1 \le i_1 < \cdots < i_r \le d} |x_{i_1} \cdots x_{i_r}|_v \quad \text{triangular inequality} \\ &\le \epsilon(v,r,d) \max_{1 \le i \le d} |x_i|_v^r. \end{aligned}$$

Here one can take $\epsilon(v,r,d) = 1$ if v is non-archimedean and $\epsilon(v,r,d) = \binom{d}{r} \le 2^d$ if v is archimedean.

Thus we have $\|s_r\|_v \le \max_{1 \le i \le d} \|x_i\|_v^r$ if v is non-archimedean, and $\|s_r\|_v \le 2^{d[K_v:\mathbb{R}]} \max_{1 \le i \le d} \|x_i\|_v^r$ if v is archimedean.

Consider the point $s := [s_0 : \cdots : s_d : 1] \in \mathbb{P}^{d+1}(\mathbb{Q})$. We have

$$\begin{aligned} [K:\mathbb{Q}]h(s) &= \sum_{v \in M_K} \log \max_{0 \le r \le d} \{\|s_r\|_v, 1\} \\ &= \sum_{v \in M_K} \max_{0 \le r \le d} \{\log \|s_r\|_v, 0\} \\ &\le \sum_{v \in M_K} \max_{0 \le r \le d} \max_{1 \le i \le d} \{r \log \|x_i\|_v, 0\} + d \sum_{v \mid \infty} [K_v : \mathbb{R}] \log 2 \\ &\le \sum_{v \in M_K} d \max_{1 \le i \le d} \{\log \|x_i\|_v, 0\} + d \sum_{v \mid \infty} [K_v : \mathbb{R}] \log 2 \\ &\le d \sum_{1 \le i \le d} \sum_{v \in M_K} \max\{\log \|x_i\|_v, 0\} + d \sum_{v \mid \infty} [K_v : \mathbb{R}] \log 2 \\ &= d \sum_{1 \le i \le d} [K:\mathbb{Q}] h(x_i) + d \sum_{v \mid \infty} [K_v : \mathbb{R}] \log 2 \\ &= d[K:\mathbb{Q}] \cdot dh(x) + d[K:\mathbb{Q}] \log 2 \quad \text{by Lemma 1.2.4.} \end{aligned}$$

So $h(s) \le d^2 h(x) + d \log 2 \le d^2 B + d \log 2$ is bounded. But $s \in \mathbb{P}^{d+1}(\mathbb{Q})$, so by the case $D = 1$ there are only finitely many choices for s. So there are only finitely many choices for $s_0, \dots, s_d$, and therefore only finitely many choices for the minimal polynomial of x over $\mathbb{Q}$. Thus there are only finitely many choices for x, and this is exactly the desired claim. We are done. □

1.2.2 Height on affine spaces

In the proof of the Northcott property, we computed the height of $[s_0 : \cdots : s_d : 1] \in \mathbb{P}^{d+1}(\overline{\mathbb{Q}})$. This point lies in $\mathbb{A}^{d+1}(\overline{\mathbb{Q}})$, viewed as the complement of the hypersurface with last homogeneous coordinate being 0. It is then convenient to introduce the following notions.

Notation 1.2.6. *Set*

$$\log^+(x) := \max\{\log x, 0\} = \log \max\{x, 1\}$$

for each $x > 0$.

Definition 1.2.7. *For each point* $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{A}^n(\overline{\mathbb{Q}})$, *define*

$$h(\mathbf{x}) := h([\mathbf{x} : 1])$$

with $[\mathbf{x}:1] := [x_1 : \cdots : x_n : 1]$ *viewed as a point in* $\mathbb{P}^n(\overline{\mathbb{Q}})$.
We also set $H(\mathbf{x}) := e^{h(\mathbf{x})}$.

We have

$$h(\mathbf{x}) = \max_{1\le j\le n}\{h(x_j)\} = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K}\log^+\max\{\|x_1\|_v,\dots,\|x_n\|_v\}. \tag{1.2.1}$$

The next proposition discusses the height of the sum of algebraic numbers. It will be seen again in the discussion for heights of polynomials.

Proposition 1.2.8. *Let* $P_1,\dots,P_r \in \mathbb{A}^n(\overline{\mathbb{Q}})$. *Then*

$$h(P_1+\cdots+P_r) \le h(P_1)+\cdots+h(P_r)+\log r.$$

In the case $n=1$, the left hand side is the sum of r algebraic numbers.

Proof. Write, for each $k\in\{1,\dots,r\}$, $P_k = (x_1^{(k)},\dots,x_n^{(k)})$. Assume all the P_k's are in a number field K. Then

$$[K:\mathbb{Q}]h(P_1+\cdots+P_r) = \sum_{v\in M_K}\max_{1\le j\le n}\log^+\|x_j^{(1)}+\cdots+x_j^{(r)}\|_v.$$

If v is not archimedean, then $\|\cdot\|_v$ is an absolute value and hence

$$\|x_j^{(1)}+\cdots+x_j^{(r)}\|_v \le \max_{1\le k\le r}\|x_j^{(k)}\|_v.$$

If v is archimedean, then the triangular inequality for the absolute value $|\cdot|_v$ yields $|x_j^{(1)}+\cdots+x_j^{(r)}|_v \le |r|_v\max_{1\le k\le r}|x_j^{(k)}|_v$. Hence raising both sides to the power of $[K_v:\mathbb{R}]$ we get

$$\|x_j^{(1)}+\cdots+x_j^{(r)}\|_v \le \|r\|_v\max_{1\le k\le r}\|x_j^{(k)}\|_v$$

Thus

$$\begin{aligned}
[K:\mathbb{Q}]h(P_1+\cdots+P_r) &\le \sum_{v\in M_K}\max_{j,k}\log^+\|x_j^{(k)}\|_v + \sum_{v|\infty}\log\|r\|_v\\
&\le \sum_{1\le k\le r}\sum_{v\in M_K}\max_{j}\log^+\|x_j^{(k)}\|_v + \sum_{v|\infty}\log\|r\|_v\\
&= \sum_{1\le k\le r}[K:\mathbb{Q}]h(P_k) + [K:\mathbb{Q}]\log r \quad \text{by Lemma 1.1.10.}
\end{aligned}$$

Hence we are done. □

1.2.3 Liouville's inequality

Lemma 1.2.9. $h(1/\alpha) = h(\alpha)$ *for any* $\alpha\in K^*$.

Proof. By definition, $h(1/\alpha) = h([1/\alpha:1])$ with $[1/\alpha:1]\in\mathbb{P}^1(K)$ and similarly $h(\alpha)=h([\alpha:1])$. So

$$h(1/\alpha) = h([1/\alpha:1]) = h([1:\alpha]) = h([\alpha:1]) = h(\alpha),$$

with $h([1:\alpha]) = h([\alpha:1])$ following directly from the definition of height. □

Alternatively, one can check

$$\log|\alpha|_v = \log^+|\alpha|_v - \log^+|1/\alpha|_v \tag{1.2.2}$$

and use the Product Formula to prove this lemma.

Proposition 1.2.10 (Fundamental Inequality)**.** *Let $S \subseteq M_K$ be a finite set. For each $\alpha \in K^*$, we have*

$$h(\alpha) \geq \frac{1}{[K:\mathbb{Q}]} \sum_{v\in S} \log\|\alpha\|_v \tag{1.2.3}$$

and

$$h(\alpha) \geq -\frac{1}{[K:\mathbb{Q}]} \sum_{v\in S} \log\|\alpha\|_v. \tag{1.2.4}$$

Proof. By the definition $h(\alpha) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K}\log^+\|\alpha\|_v$ and noticing that $\log^+$ takes non-negative values, we get the first inequality.

To prove second inequality, we apply the first inequality to $1/\alpha$ and use Lemma 1.2.9. □

Example 1.2.11. *Consider $K = \mathbb{Q}$ and $\alpha = p$ is a prime number. Then $h(p) = \log p$, $|p|_\infty = p$ and $|p|_p = p^{-1}$. Now (1.2.3) attains equality for $S = \{\infty\}$, and (1.2.4) attains equality for $S = \{p\}$.*

Now we are ready to state Liouville's inequality. The classical formulation is in terms of the multiplicative height $H(\cdot) = e^{h(\cdot)}$.

In the statement of Liouville's Inequality, let K_0 be a number field.

Theorem 1.2.12 (Liouville's Inequality)**.** *Fix $\beta \in K_0$. Let K/K_0 be a finite extension and consider a finite set $S \subseteq M_K$. For any $\alpha \in K$ with $\alpha \neq \beta$, we have*

$$\prod_{v\in S} \|\alpha-\beta\|_{v,K_0} \geq (2H(\alpha)H(\beta))^{-[K:K_0]},$$

where $\|\cdot\|_{v,K_0} := \|\cdot\|_v^{1/[K_0:\mathbb{Q}]}$.

Before moving on to its proof, let us look at the following corollary which is closer to the classical statement of this inequality. It has a flavor of approximating algebraic numbers by rational numbers.

Corollary 1.2.13. *Let $\alpha \in \mathbb{R}$ be an algebraic number of degree $r > 1$, i.e. $[\mathbb{Q}(\alpha):\mathbb{Q}] = r$. Then there exists a constant $c(\alpha) > 0$ such that for the usual absolute value $|\cdot|$ on $\mathbb{R}$, we have*

$$|\alpha-\beta| \geq c(\alpha)H(\beta)^{-r} \quad \text{for all } \beta \in \mathbb{Q}.$$

This corollary follows immediately from Theorem 1.2.12 applied to $K_0 = \mathbb{Q}$, $K = \mathbb{Q}(\alpha)$ and S the archimedean place given by the natural inclusion $\mathbb{Q}(\alpha) \subseteq \mathbb{R}$. Notice that if $\alpha \in \mathbb{C}\setminus\mathbb{R}$, then the same conclusion holds true with $|\cdot|$ replaced by $|\cdot|^2$.

Proof of Theorem 1.2.12. Apply Proposition 1.2.8 to $n = 1$, $r = 2$, $P_1 = \alpha$ and $P_2 = -\beta$. Then we get $h(\alpha-\beta) \leq h(\alpha) + h(\beta) + \log 2$. So $H(\alpha-\beta) \leq 2H(\alpha)H(\beta)$.

Apply the Fundamental Inequality (1.2.4) to $\alpha - \beta$. Then we get

$$h(\alpha-\beta) \geq -\frac{1}{[K:\mathbb{Q}]}\sum_{v\in S}\log\|\alpha-\beta\|_v = \frac{1}{[K:\mathbb{Q}]}\log\Big(\prod_{v\in S}\|\alpha-\beta\|_v\Big)^{-1}.$$

From this, we get

$$\prod_{v\in S}\|\alpha-\beta\|_{v,K_0}=(\prod_{v\in S}\|\alpha-\beta\|_v)^{1/[K_0:\mathbb{Q}]}\geq(e^{h(\alpha-\beta)})^{-[K:K_0]}=H(\alpha-\beta)^{-[K:K_0]}.$$

Now we can conclude because we have seen $H(\alpha-\beta)\leq 2H(\alpha)H(\beta)$. □

1.2.4 The change of height under geometric operations

In this section, we go back to the height function on $\mathbb{P}^n(\overline{\mathbb{Q}})$. We will consider several geometric operations concerning projective spaces and see how the heights change.

Consider the *Segre embedding*

$$S_{n,m}\colon \mathbb{P}^n\times\mathbb{P}^m\to\mathbb{P}^{(n+1)(m+1)-1},\quad (\mathbf{x},\mathbf{y})\mapsto\mathbf{x}\otimes\mathbf{y}:=(x_iy_j)_{i,j} \tag{1.2.5}$$

and the *d-uple embedding*

$$\Phi_d\colon \mathbb{P}^n\to\mathbb{P}^N,\quad \mathbf{x}\mapsto[M_0(\mathbf{x}):\cdots:M_N(\mathbf{x})] \tag{1.2.6}$$

with $N=\binom{n+d}{n}-1$ and $\{M_0(\mathbf{x}),\dots,M_N(\mathbf{x})\}$ the complete collection of monomials of degree d in the variables $x_0,\dots,x_n$.

Proposition 1.2.14. *We have*

(i) $h(\mathbf{x}\otimes\mathbf{y})=h(\mathbf{x})+h(\mathbf{y})$ *for all* $\mathbf{x}\in\mathbb{P}^n(\overline{\mathbb{Q}})$ *and* $\mathbf{y}\in\mathbb{P}^m(\overline{\mathbb{Q}})$.

(ii) $h(\Phi_d(\mathbf{x}))=dh(\mathbf{x})$ *for all* $\mathbf{x}\in\mathbb{P}^n(\overline{\mathbb{Q}})$.

Proof. Part (i) in Exercise class, by using $\max_{i,j}|x_iy_j|_v=\max_i|x_i|_v\cdot\max_j|y_j|_v$.

We prove part (ii). Each $M_i(\mathbf{x})$ is a monomial of degree d in the variables $x_0,\dots,x_n$. It is clear that $|M_j(\mathbf{x})|_v\leq\max_i|x_i|_v^d$ for each $0\leq j\leq N$. Moreover since the particular monomials $x_0^d,\dots,x_n^d$ appear in the collection, we have

$$\max_{0\leq j\leq N}|M_j(\mathbf{x})|_v=\max_{0\leq i\leq n}|x_i|_v^d.$$

From this we can conclude. □

We finish this section by a discussion on the change of heights under linear maps.

Theorem 1.2.15. *Let* $\phi\colon\mathbb{P}^n\to\mathbb{P}^m$ *be a linear map defined over* $\overline{\mathbb{Q}}$, *i.e.* $\phi=[L_0(\mathbf{x}):\cdots:L_m(\mathbf{x})]$ *for some linear forms on* $\mathbb{P}^n$. *Let* $Z\subseteq\mathbb{P}^n$ *be the common zero of the* L_i*'s.*

Let $X\subseteq\mathbb{P}^n$ *be a closed subvariety such that* $X\cap Z=\emptyset$. *Then*

$$h(\phi(\mathbf{x}))=h(\mathbf{x})+O(1)\quad \text{for all } \mathbf{x}\in X(\overline{\mathbb{Q}}).$$

More precisely, the conclusion means that there exists a constant $c=c(\phi,X)>0$ depending only on ϕ and X such that

$$|h(\phi(\mathbf{x}))-h(\mathbf{x})|\leq c$$

for all $\mathbf{x}\in X(\overline{\mathbb{Q}})$. We remark that this bound[4] does not hold true on the whole $\mathbb{P}^n\setminus Z$, but on any closed subvariety disjoint from Z.

[4] Or more precisely, "half" of the bound does not hold true on the whole $\mathbb{P}^n\setminus Z$ as will be shown in the proof.

Proof. The proof is divided into two parts. We may and do assume that $Z \neq \mathbb{P}^n$, *i.e.* one of the L_i's does not vanish on the whole $\mathbb{P}^n$.

Write $L_i(\mathbf{x}) = \sum_{0\leq j\leq n} a_{i,j}x_j$. Then $[a_{0,0} : \cdots : a_{0,n} : \cdots : a_{m,0} : \cdots : a_{m,n}]$ is a point in $\mathbb{P}^{(n+1)(m+1)-1}(\overline{\mathbb{Q}})$ and is uniquely determined by ϕ.

Part I Prove: there exists a constant $c_1(\phi)$ depending only on ϕ such that $h(\phi(\mathbf{x}))-h(\mathbf{x}) \leq c_1(\phi)$ for all $\mathbf{x} \in (\mathbb{P}^n \setminus Z)(\overline{\mathbb{Q}})$.

Let $\mathbf{x} = [x_0 : \cdots : x_n] \in (\mathbb{P}^n \setminus Z)(\overline{\mathbb{Q}})$. Fix a number field K such that $\mathbf{x} \in \mathbb{P}^n(K)$ and all $a_{i,j}$'s are in K. Then for each $v \in M_K$, we have

$$|L_i(\mathbf{x})|_v = |\sum_{0\leq j\leq n} a_{i,j}x_j|_v \leq \epsilon(v, n+1)(\max_j |a_{i,j}|_v)(\max_j |x_j|_v)$$

where $\epsilon(v,k) := \begin{cases} 1 & \text{if } v \text{ is non-archimedean} \\ k & \text{if } v \text{ is archimedean} \end{cases}$. Raising both sides to the power of $[K_v : \mathbb{Q}_p]$ (with $\mathbb{Q}_\infty = \mathbb{R}$), we get

$$\max_i \|L_i(\mathbf{x})\|_v \leq \epsilon(v, n+1)^{[K_v:\mathbb{Q}_p]}(\max_{i,j} \|a_{i,j}\|_v)(\max_j \|x_j\|_v).$$

Now we have

$$\begin{aligned}
[K : \mathbb{Q}]h(\phi(\mathbf{x})) &= \sum_{v\in M_K} \log \max_i \|L_i(\mathbf{x})\|_v \\
&\leq \sum_{v\in M_K} \log\left(\epsilon(v, n+1) \cdot \max_{i,j} \|a_{i,j}\|_v \cdot \max_j \|x_j\|_v\right) \\
&\leq \sum_{v\in M_K} (\log \max_{i,j} \|a_{i,j}\|_v + \log \max_j \|x_j\|_v) + \sum_{v|\infty} [K_v : \mathbb{R}] \log(n+1) \\
&= \sum_{v\in M_K} \log \max_{i,j} \|a_{i,j}\|_v + [K : \mathbb{Q}]h(\mathbf{x}) + [K : \mathbb{Q}] \log(n+1) \\
&= [K : \mathbb{Q}]h([a_{0,0} : \cdots : a_{0,n} : \cdots : a_{m,0} : \cdots : a_{m,n}]) + [K : \mathbb{Q}]h(\mathbf{x}) + [K : \mathbb{Q}] \log(n+1).
\end{aligned}$$

Thus $h(\phi(\mathbf{x})) - h(\mathbf{x}) \leq h([a_{0,0} : \cdots : a_{0,n} : \cdots : a_{m,0} : \cdots : a_{m,n}]) + \log(n+1)$. The first term on the right hand side depends only on ϕ. So we are done for this part.

Part II Prove: there exists a constant constant $c_2(\phi, X)$ such that $h(\phi(\mathbf{x})) - h(\mathbf{x}) \geq c_2(\phi, X)$ for all $\mathbf{x} \in X(\overline{\mathbb{Q}})$.

Write $I(X) = (F_1, \ldots, F_r)$. Since $X \cap Z = \emptyset$, we have that the polynomials $L_0, \ldots, L_m, F_1, \ldots, F_r$ have no common zeros in $\mathbb{P}^n$. By Hilbert Nullstellensatz, we then have the following equality of ideals of $\overline{\mathbb{Q}}[X_0, \ldots, X_n]$

$$\sqrt{(L_0, \ldots, L_m, F_1, \ldots, F_r)} = (X_0, \ldots, X_n).$$

In particular, for each $j \in \{0, \ldots, n\}$, we can find polynomials $G_{i,j}$ and $H_{i,j}$ and an exponent $t \geq 1$, all depending only on X and ϕ, such that

$$G_{0,j}L_0 + \cdots + G_{m,j}L_m + H_{1,j}F_1 + \cdots + H_{r,j}F_r = X_j^t.$$

Moreover $\deg G_{i,j} = t - \deg L_i = t - 1$.

Write $G_{i,j} = \sum_{|\mathbf{e}|=t-1} b_{i,j,\mathbf{e}}X^{\mathbf{e}}$, with $\mathbf{e} = (e_0, \ldots, e_n)$ a multi-index with $|\mathbf{e}| := e_0 + \cdots + e_n$ and $X^{\mathbf{e}} = X_0^{e_0} \cdots X_n^{e_n}$. Notice that $G_{i,j}$ is the sum of at most $\binom{n+t-1}{n}$ monomials.

Now let $\mathbf{x} = [x_0 : \cdots : x_n] \in X(\overline{\mathbb{Q}})$. Evaluating the equation above at $\mathbf{x}$, we get

$$G_{0,j}(\mathbf{x})L_0(\mathbf{x}) + \cdots + G_{m,j}(\mathbf{x})L_m(\mathbf{x}) = x_j^t$$

for each $j \in \{0, \ldots, n\}$.

Fix a number field K such that $\mathbf{x} \in X(K)$ and all the coefficients $b_{i,j,\mathbf{e}}$ are in K.

For each $v \in M_K$, we have

$$|x_j|_v^t = |G_{0,j}(\mathbf{x})L_0(\mathbf{x}) + \cdots + G_{m,j}(\mathbf{x})L_m(\mathbf{x})|_v \leq \epsilon(v, m+1) \max_{0\leq i\leq m} |G_{i,j}(\mathbf{x})|_v \max_{0\leq i\leq m} |L_i(\mathbf{x})|_v.$$

Thus

$$\begin{aligned}
\max_j |x_j|_v^t &\leq \epsilon(v, m+1) \max_{i,j} |G_{i,j}(\mathbf{x})|_v \max_i |L_i(\mathbf{x})|_v \\
&= \epsilon(v, m+1)(\max_{i,j} |\sum_{|\mathbf{e}|=t-1} b_{i,j,\mathbf{e}} x_0^{e_0} \cdots x_n^{e_n}|_v)(\max_{0\leq i\leq m} |L_i(\mathbf{x})|_v) \\
&\leq \epsilon(v, m+1) \left(\epsilon(v, \binom{n+t-1}{n}) \max_{i,j,\mathbf{e}} |b_{i,j,\mathbf{e}}|_v \max_j |x_j|_v^{t-1}\right) (\max_{0\leq i\leq m} |L_i(\mathbf{x})|_v).
\end{aligned}$$

Dividing both sides by $\max_j |x_j|^{t-1}$, we get

$$\max_j |x_j|_v \leq \epsilon(v, m+1)\epsilon(v, \binom{n+t-1}{n}) \max_{i,j,\mathbf{e}} |b_{i,j,\mathbf{e}}|_v \max_{0\leq i\leq m} |L_i(\mathbf{x})|_v.$$

Raising both sides to the power of $[K_v : \mathbb{Q}_p]$ (with $\mathbb{Q}_\infty = \mathbb{R}$), we get

$$\max_j \|x_j\|_v \leq \epsilon(v, m+1)^{[K_v:\mathbb{Q}_p]}\epsilon(v, \binom{n+t-1}{n})^{[K_v:\mathbb{Q}_p]} \max_{i,j} \|b_{i,j,\mathbf{e}}\|_v \max_i \|L_i(\mathbf{x})\|_v.$$

Now we have

$$\begin{aligned}
[K : \mathbb{Q}]h(\mathbf{x}) &= \sum_{v\in M_K} \max_j \|x_j\|_v \\
&\leq \sum_{v\in M_K} \log \max_{i,j,\mathbf{e}} \|b_{i,j,\mathbf{e}}\|_v + \sum_{v\in M_K} \max_i \|L_i(\mathbf{x})\|_v + \sum_{v|\infty} [K_v : \mathbb{R}] \log(m+1)\binom{n+t-1}{n} \\
&= [K : \mathbb{Q}]h(\mathbf{b}) + [K : \mathbb{Q}]h(\phi(\mathbf{x})) + [K : \mathbb{Q}] \log(m+1)\binom{n+t-1}{n}
\end{aligned}$$

where $\mathbf{b}$ is the point in an appropriate projective space whose homogeneous coordinates are $b_{i,j,\mathbf{e}}$. Notice that $\mathbf{b}$ is uniquely determined by the $G_{i,j}$'s, and hence by X and ϕ. Now we get the desired inequality $h(\phi(\mathbf{x})) - h(\mathbf{x}) \geq c_2(\phi, X)$ for all $\mathbf{x} \in X(\overline{\mathbb{Q}})$. Hence we are done. □

1.3 Height of polynomials

In this section, we study the heights of polynomials. We will use the Weil height on projective and affine spaces defined in §1.2 of this chapter.

Definition 1.3.1. *The (affine) height of a polynomial*

$$f(t_1, \ldots, t_n) = \sum_{j_1,\ldots,j_n} a_{j_1\ldots j_n} t_1^{j_1} \cdots t_n^{j_n} = \sum_{\mathbf{j}} a_{\mathbf{j}} \mathbf{t}^{\mathbf{j}}$$

with coefficients in $\overline{\mathbb{Q}}$ is the quantity $h(\mathbf{a})$ where $\mathbf{a} = (a_{\mathbf{j}})_{\mathbf{j}}$ is viewed as a point in $\overline{\mathbb{Q}}^N$ for some N.

In other words, if we assume each $a_{\mathbf{j}} \in K$ for an appropriate number field K and define the **Gauß norm**

$$\|f\|_v := \max_{\mathbf{j}} \|a_{\mathbf{j}}\|_v \tag{1.3.1}$$

for each $v \in M_K$, then we have

$$h(f) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log^+ \|f\|_v = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log \max\{\|f\|_v, 1\}. \tag{1.3.2}$$

1.3.1 Affine height vs the Projective height

In some literature, one defines the height of f as the height of the point $[a_{\mathbf{j}}]_{\mathbf{j}}$ viewed as a point in an appropriate *projective* space. This is sometimes called the *projective height of* f and denoted by $h_{\text{proj}}(f)$, and is in general smaller than the affine height we defined above.

In this course, *we always use the affine height*. An important advantage to take this convention is the following proposition, which is about the evaluation of a polynomial at a point. The proof shares some similarities with the proof of Theorem 1.2.15.

Proposition 1.3.2. *Let d be the sum the partial degrees of f. Let $\mathbf{x} = (x_1, \dots, x_n) \in \overline{\mathbb{Q}}^n$. Then*

$$h(f(\mathbf{x})) \le h(f) + dh(\mathbf{x}) + \min\{(n+1)\log(n+d+1), (n+d+1)\log 2\}.$$

As shown by the proof, this result is not correct if we use the projective height of f.

Proof. Write $f(\mathbf{t}) = \sum_{k=0}^{d} \sum_{|\mathbf{j}|=k} a_{\mathbf{j}} \mathbf{t}^{\mathbf{j}}$. Set $\psi(n,d) := \min\{(n+d+1)^{n+1}, 2^{n+d+1}\}$. Then as in the proof of Theorem 1.2.15, it is not hard to check

$$|\sum_{|\mathbf{j}|=k} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}|_v \le \epsilon\left(v, \binom{n+k}{n}\right) \max_{|\mathbf{j}|=k}\{|a_{\mathbf{j}}|_v\} \max_i |x_i|_v^k \le \epsilon\left(v, \binom{n+k}{n}\right) \max_{|\mathbf{j}|=k}\{|a_{\mathbf{j}}|_v\} \max\{1, \max_i |x_i|_v\}^d$$

with $\epsilon(v,m)$ defined to be 1 for v non-archimedean and to be m for v archimedean. Recall that $\sum_{k=0}^{d} \binom{n+k}{n} = \binom{n+d+1}{n+1} \le \psi(n,d)$. So

$$|f(\mathbf{x})|_v = |\sum_{\mathbf{j}} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}|_v \le \sum_{i=0}^{k} |\sum_{|\mathbf{j}|=k} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}|_v \le \epsilon(v, \psi(n,d)) \max_{\mathbf{j}}\{|a_{\mathbf{j}}|_v\} \max_i\{1, |x_i|_v\}^d$$

and hence

$$\max\{1, |f(\mathbf{x})|_v\} \le \epsilon(v, \psi(n,d)) \max_{\mathbf{j}}\{1, |a_{\mathbf{j}}|_v\} \max_i\{1, |x_i|_v\}^d.$$

Raising to the power of $[K_v : \mathbb{Q}_p]$ and taking the log, we get an upper bound for $\log^+ \|f(\mathbf{x})\|_v$. Hence

$$\begin{aligned}
[K:\mathbb{Q}]h(f(\mathbf{x})) &= \sum_{v \in M_K} \log^+ \|f(\mathbf{x})\|_v \\
&\le \sum_{v \in M_K} \max_{\mathbf{j}} \log^+ \|a_{\mathbf{j}}\|_v + d \sum_{v \in M_K} \max_i \log^+ \|x_i\|_v + \sum_{v|\infty} [K_v : \mathbb{R}] \log \psi(n,d) \\
&= [K:\mathbb{Q}]h(f) + [K:\mathbb{Q}]dh(\mathbf{x}) + [K:\mathbb{Q}] \log \psi(n,d).
\end{aligned}$$

We are done. □

We mention an advantage of the projective height. It is an immediate corollary of Proposition 1.2.14 (i). However, we shall not use it. Indeed, Theorem 1.3.4 is a more general and more applicable statement concerning the height of a product of two polynomials.

Lemma 1.3.3. *Let $f(t_1, \dots, t_n)$ and $g(s_1, \dots, s_m)$ be two polynomials in disjoint sets of variables. Then*

$$h_{\mathrm{proj}}(fg) = h_{\mathrm{proj}}(f) + h_{\mathrm{proj}}(g).$$

If f and g do not have disjoint sets of variables, the estimate of $h_{\mathrm{proj}}(fg)$ in terms of $h_{\mathrm{proj}}(f)$ and $h_{\mathrm{proj}}(g)$ has the same quality of Theorem 1.3.4. In most applications, unfortunately we do not have disjoint sets of variables and hence need to use the more complicated Theorem 1.3.4.

1.3.2 Height of product

The main result of this section is to study the height of a product of two polynomials. We will prove the following theorem for this estimate.

Theorem 1.3.4. *Let $f_1, \dots, f_m$ be polynomials in n variables with coefficients in $\overline{\mathbb{Q}}$. Let d be the sum of the partial degrees of $f := f_1 \cdots f_m$. Then*

$$-d \log 2 + \sum_{j=1}^{m} h(f_j) \leq h(f) \leq d \log 2 + \sum_{j=1}^{m} h(f_j).$$

Moreover in the second inequality, one can replace d by the sum of the partial degrees of the product $f_1 \cdots f_{m-1}$.

To prove this theorem, one separates the non-archimedean places and the archimedean places. For the non-archimedean places, we prove *Gauß's Lemma*. For the archimedean places, we prove *Gelfond's Lemma*. Then we combine these two lemmas to conclude.

Non-archimedean places

The contribution at the non-archimedean places is not hard to study. In this case, we have the following:

Lemma 1.3.5 (Gauß's Lemma)**.** *If v is non-archimedean, then $\|fg\|_v = \|f\|_v \|g\|_v$.*

Proof. The direction $\|fg\|_v \leq \|f\|_v \|g\|_v$ is not hard to obtain because v is non-archimedean.

Now we focus on proving the other direction $\|fg\|_v \geq \|f\|_v \|g\|_v$.

One-variable case We start with the case where both $f(t) = \sum_j a_j t^j$ and $g(t) = \sum_j b_j t^j$ are polynomials in one variable t. Up to dividing both f and g by an appropriate element in K, we may and do assume $\|f\|_v = \|g\|_v = 1$. Then $\|fg\|_v \leq 1$.

Suppose $\|fg\|_v < 1$ and we wish to get a contradiction.

For each j, set $c_j = \sum_{j=k+l} a_k b_l$. Then $fg = \sum_j c_j t^j$. Let j_0 be the smallest integer with $\|a_{j_0}\|_v = 1$. Since $\|a_k\|_v < 1$ for each $k < j_0$, we have $\|a_k b_{j_0-k}\|_v < 1$ for each $k < j_0$. If $\|b_0\|_v = 1$, then $\|a_{j_0} b_0\|_v = 1$ and hence $\|c_{j_0}\|_v = \|a_{j_0} b_0 + \sum_{k<j_0} a_k b_{j_0-k}\|_v = 1$, contradicting $\|fg\|_v < 1$. Hence $\|b_0\|_v < 1$.

Next for each l_0, we prove that $\|b_{l_0}\|_v < 1$ by induction. Suppose we have proved for $0, \dots, l_0 - 1$. Consider $c_{j_0+l_0} = \sum_{0 \leq k \leq j_0+l_0} a_k b_{j_0+l_0-k}$. For $0 \leq k \leq j_0 - 1$, we have $\|a_k\|_v < 1$ and hence $\|a_k b_{j_0+l_0-k}\|_v < 1$. For $j_0 + 1 \leq k \leq j_0 + l_0$, we have $\|b_{j_0+l_0-k}\|_v < 1$ by induction hypothesis and hence $\|a_k b_{j_0+l_0-k}\|_v < 1$. Thus $\|b_{l_0}\|_v = 1$ would yield $\|c_{j_0+l_0}\|_v = \|a_{j_0} b_{l_0}\|_v = 1$, contradicting $\|fg\|_v < 1$. Hence we can conclude $\|b_{l_0}\|_v < 1$.

But then $\|g\|_v < 1$, contradicting $\|g\|_v = 1$. So we can conclude that $\|fg\|_v = 1$ for this case.
General case Write $f(x_1, \dots, x_n) = \sum_{\mathbf{j}} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}$ and $g(x_1, \dots, x_n) = \sum_{\mathbf{j}} b_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}$. One can reduce the general case to the one-variable case by the following standard technique. Fix an integer $d > \deg(fg)$, and consider the **Kronecker substitution**

$$x_j := t^{d^{j-1}} \quad (j = 1, \dots, n). \tag{1.3.3}$$

Then

$$f(x_1, \dots, x_n) = \sum_{\mathbf{j}} a_{\mathbf{j}} (t^{d^0})^{j_1} (t^{d^1})^{j_2} \cdots (t^{d^{n-1}})^{j_n} = \sum_{0 \le j_1, \dots, j_n \le d-1} a_{j_1, \dots, j_n} t^{j_1 + dj_2 + \cdots + d^{n-1} j_n},$$
$$\text{and} \quad g(x_1, \dots, x_n) = \sum_{0 \le j_1, \dots, j_n \le d-1} b_{j_1, \dots, j_n} t^{j_1 + dj_2 + \cdots + d^{n-1} j_n}.$$

It is not hard to see that both $f_0(t) := \sum_{0 \le j_1, \dots, j_n \le d-1} a_{j_1, \dots, j_n} t^{j_1 + dj_2 + \cdots + d^{n-1} j_n}$ and $g_0(t) := \sum_{0 \le j_1, \dots, j_n \le d-1} b_{j_1, \dots, j_n} t^{j_1 + dj_2 + \cdots + d^{n-1} j_n}$ are one-variable polynomials in simplified form. So $\|f_0\|_v = \max_{j_1, \dots, j_n} \|a_{j_1, \dots, j_n}\|_v = \|f\|_v$, $\|g_0\|_v = \max_{j_1, \dots, j_n} \|b_{j_1, \dots, j_n}\|_v = \|g\|_v$, and $\|f_0 g_0\|_v = \|fg\|_v$. Hence we can conclude by the one-variable case. □

Archimedean places

It is more complicated to handle the archimedean places. The goal is to prove *Gelfond's Lemma* (Lemma 1.3.6), which plays a similar role as Gauß's Lemma for the archimedean places.

In this subsection, we consider polynomials with coefficients in $\mathbb{C}$. We use $|\cdot|$ to denote the usual euclidean absolute value on $\mathbb{C}$.

Let $f = \sum_{\mathbf{j}} a_{\mathbf{j}} \mathbf{t}^{\mathbf{j}} \in \mathbb{C}[t_1, \dots, t_n]$. Define

$$\ell_\infty(f) = |f|_\infty := \max_{\mathbf{j}} |a_{\mathbf{j}}|. \tag{1.3.4}$$

We also call $\ell_\infty(f)$ the *L^∞-norm* of f.

Now we can state the main result of this subsection.

Lemma 1.3.6 (Gelfond's Lemma)**.** *Let $f_1, \dots, f_m \in \mathbb{C}[t_1, \dots, t_n]$ and set $f := f_1 \cdots f_m$. Let d be the sum of the partial degrees of f. Then*

$$2^{-d} \prod_{j=1}^{m} \ell_\infty(f_j) \le \ell_\infty(f) \le 2^d \prod_{j=1}^{m} \ell_\infty(f_j).$$

Moreover in the second inequality, one can replace d by the sum of the partial degrees of the product $f_1 \cdots f_{m-1}$.

Before moving on, let us see how Gauß's Lemma and Gelfond's Lemma imply Theorem 1.3.4.

Proof of Theorem 1.3.4. We have

$$[K : \mathbb{Q}] h(f) = \sum_{v \in M_K} \log^+ \|f\|_v = \sum_{v \in M_K} \log \max\{\|f\|_v, 1\} = \sum_{v \in M_K} \log \max\{\|f_1 \cdots f_m\|_v, 1\}.$$

To get the upper bound, we proceed as follows

$$\begin{aligned}
[K:\mathbb{Q}]h(f) &= \sum_{v\in M_K} \max\{\log\|f_1\cdots f_m\|_v, 0\}\\
&= \sum_{v\in M_K^0} \max\left\{\sum_{j=1}^m \log\|f_j\|_v, 0\right\} + \sum_{v|\infty}[K_v:\mathbb{R}]\log^+|f_1\cdots f_m|_v \quad \text{by Gauß's Lemma (Lemma 1.3.5)}\\
&\le \sum_{v\in M_K^0}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}\max\left\{[K_v:\mathbb{R}]\log|f_1\cdots f_m|_v, 0\right\}\\
&\le \sum_{v\in M_K^0}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}\max\left\{[K_v:\mathbb{R}]\left(\sum_{j=1}^m \log|f_j|_v + d\log 2\right), 0\right\} \quad \text{by Gelfond's Lemma (Lemma 1.3.6)}\\
&= \sum_{v\in M_K^0}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}\max\left\{\sum_{j=1}^m \log\|f_j\|_v + [K_v:\mathbb{R}]d\log 2, 0\right\}\\
&\le \sum_{v\in M_K^0}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}\max\left\{\sum_{j=1}^m \log\|f_j\|_v, 0\right\} + \sum_{v|\infty}[K_v:\mathbb{R}]d\log 2\\
&\le \sum_{v\in M_K^0}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}\sum_{j=1}^m \log^+\|f_j\|_v + \sum_{v|\infty}[K_v:\mathbb{R}]d\log 2\\
&= [K:\mathbb{Q}]\sum_{j=1}^m h(f_j) + [K:\mathbb{Q}]d\log 2.
\end{aligned}$$

The "Moreover" part holds true because of the "Moreover" part of Gelfond's Lemma.

To get the lower bound, we have

$$\begin{aligned}
[K:\mathbb{Q}]h(f) &\ge \sum_{v\in M_K} \log\|f\|_v\\
&= \sum_{v\in M_K} \log\|f_1\cdots f_m\|_v\\
&= \sum_{v\in M_K^0}\sum_{j=1}^m \log\|f_j\|_v + \sum_{v|\infty}[K_v:\mathbb{R}]\log|f_1\cdots f_m|_v \quad \text{by Gauß's Lemma (Lemma 1.3.5)}\\
&\ge \sum_{j=1}^m\sum_{v\in M_K^0} \log\|f_j\|_v + \sum_{v|\infty}[K_v:\mathbb{R}](\sum_{j=1}^m \log|f_j|_v - d\log 2) \quad \text{by Gelfond's Lemma (Lemma 1.3.6)}\\
&= \sum_{j=1}^m\sum_{v\in M_K} \log\|f_j\|_v + \sum_{v|\infty}[K_v:\mathbb{R}]d\log 2\\
&= [K:\mathbb{Q}]\sum_{j=1}^m h(f_j) + [K:\mathbb{Q}]d\log 2.
\end{aligned}$$

We are done. □

So in the rest, we aim to prove Gelfond's Lemma (Lemma 1.3.6).

Definition 1.3.7. *The **Mahler measure** of f is defined to be*

$$M(f) := \exp\left(\int_{\mathbb{T}^n} \log|f(e^{i\theta_1},\dots,e^{i\theta_n})|\mathrm{d}\mu_1\cdots\mathrm{d}\mu_n\right),$$

where $\mathbb{T}$ *is the unit circle* $\{e^{i\theta} : 0 \le \theta < 2\pi\}$ *in* $\mathbb{R}$ *equipped with the standard measure* $\mathrm{d}\mu = (1/2\pi)\mathrm{d}\theta$.

The following *multiplicative* property of the Mahler measure is easy to check:

$$M(fg) = M(f)M(g). \tag{1.3.5}$$

Definition 1.3.8. *The* $\mathbf{L^2}$***-norm*** *of* f *is defined to be*

$$\ell_2(f) := \left(\int_{\mathbb{T}^n} |f(e^{i\theta_1}, \dots, e^{i\theta_n})|^2 \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_n\right)^{1/2} = \left(\sum_{\mathbf{j}} |a_{\mathbf{j}}|^2\right)^{1/2}.$$

In fact, we have given two equivalent definitions of the L^2-norm above. They coincide by Parseval's identity.

One-variable case We start by studying the one-variable case. The following lemma is an elementary tool to study the Mahler measure.

Lemma 1.3.9 (Jensen's Lemma)**.** *Let* $f(t) = a_d t^d + \cdots + a_0 \in \mathbb{C}[t]$. *Write* $\alpha_1, \dots, \alpha_d \in \mathbb{C}$ *for the roots of* f, *i.e.* $f(t) = a_d(t-\alpha_1)\cdots(t-\alpha_d)$. *Then we have*

$$\log M(f) = \log|a_d| + \sum_{j=1}^{d} \log^+|\alpha_j|,$$

with $\log^+(x) := \max\{\log x, 0\}$.

Proof. We only give a sketch here.

Because Mahler measure is multiplicative, it suffices to prove $\log M(t-\alpha) = \log^+|\alpha|$ for each $\alpha \in \mathbb{C}$.

If $|\alpha| > 1$, then the function $\log|t-\alpha|$ is harmonic in the unit disk, and hence its mean value on the unit circle is its value at the center which is $\log|\alpha| = \log^+|\alpha|$. If $|\alpha| < 1$, then the function $\log|1-\alpha\bar{t}|$ is harmonic in the unit disk and coincides with $\log|t-\alpha|$ on the unit circle, while its value at the center is $0 = \log^+|\alpha|$. Finally, the case $|\alpha| = 1$ is obtained by continuity. □

The following lemma uses the Mahler measure $M(f)$ to bound $\ell_\infty(f)$.

Lemma 1.3.10. *Let* $f(t) = a_d t^d + \cdots + a_0 \in \mathbb{C}[t]$. *Then we have*

$$\binom{d}{\lfloor d/2 \rfloor}^{-1} \ell_\infty(f) \le M(f) \le \ell_2(f) \le (d+1)^{1/2}\ell_\infty(f).$$

Proof. The last inequality is easy to see because $\ell_2(f) = (\sum_{j=0}^d |a_j|^2)^{1/2} \le (d+1)^{1/2}\max_j\{|a_j|\} = (d+1)^{1/2}\ell_\infty(f)$.

To prove the first inequality, write $f(t) = a_d(t-\alpha_1)\cdots(t-\alpha_d)$. Then for each $r \in \{0, \dots, d\}$ we have

$$|a_{d-r}| = |a_d|\left|\sum_{j_1<\cdots<j_r} \alpha_{j_1}\cdots\alpha_{j_r}\right| \le \binom{d}{r}|a_d|\prod_{j=1}^{d}\max\{1, |\alpha_j|\}.$$

Thus Jensen's Lemma above yields

$$|a_{d-r}| \le \binom{d}{r} M(f) \le \binom{d}{\lfloor d/2 \rfloor} M(f)$$

for each $r \in \{0, \dots, d\}$. So we have $\ell_\infty(f) \le \binom{d}{\lfloor d/2 \rfloor} M(f)$ and this is the first inequality.

To prove the inequality in the middle, we use *Jensen's inequality* which applies to convex functions. It says: If Ω is a space with a measure $\mathrm{d}\mu$ such that $\mathrm{d}\mu(\Omega) = \int_\Omega \mathrm{d}\mu = 1$, if g is a real-valued μ-integrable function on Ω and φ is a convex function on $\mathbb{R}$, then we have

$$\varphi\left(\int_\Omega g \mathrm{d}\mu\right) \le \int_\Omega (\varphi \circ g) \mathrm{d}\mu. \tag{1.3.6}$$

Applying this to $\Omega = \mathbb{T}$, $\mathrm{d}\mu$ as in the definition of Mahler measure and L^2-norm, $\varphi = \exp$ and $g(t) = 2\log|f(e^{i\theta})|$, we obtain

$$M(f)^2 \le \int_{\mathbb{T}} |f(e^{i\theta})|^2 \mathrm{d}\mu = \ell_2(f)^2.$$

Hence we are done for the middle inequality. □

Multi-variable case Here is the multi-variable version of the bound of $\ell_\infty(f)$ by $M(f)$.

Lemma 1.3.11. *Let $f(t_1, \dots, t_n) \in \mathbb{C}[t_1, \dots, t_n]$ with partial degrees $d_1, \dots, d_n$. Then*

$$\prod_{j=1}^n (d_j + 1)^{-1/2} M(f) \le \ell_\infty(f) \le \prod_{j=1}^n \binom{d_j}{\lfloor d_j/2 \rfloor} M(f).$$

Proof. The desired inequality is equivalent to

$$\prod_{j=1}^n \binom{d_j}{\lfloor d_j/2 \rfloor}^{-1} \ell_\infty(f) \le M(f) \le \prod_{j=1}^n (d_j + 1)^{1/2} \ell_\infty(f).$$

The proof for the second inequality follows the same line as in the one-variable case; one uses the L^2-norm as an intermediate. More precisely, one uses *Jensen's inequality* (1.3.6) to prove $M(f) \le \ell_2(f)$, and then applies the easy bound $\ell_2(f) = (\sum_{1 \le j \le d,\ 0 \le i_j \le d_j} |a_{i_1, \dots, i_d}|^2)^{1/2} \le \prod_{j=1}^n (d_j + 1)^{1/2} \ell_\infty(f)$.

Now we prove the first inequality $\prod_{j=1}^n \binom{d_j}{\lfloor d_j/2 \rfloor}^{-1} \ell_\infty(f) \le M(f)$ by induction on n. The base step $n = 1$ is proved in Lemma 1.3.10.

Assume the result is proved for $1, \dots, n-1$. We can write uniquely

$$f(t_1, \dots, t_n) = \sum_{j=0}^{d_n} f_j(t_1, \dots, t_{n-1}) t_n^j$$

for certain polynomials $f_j \in \mathbb{C}[t_1, \dots, t_{n-1}]$. Then $\ell_\infty\left(f(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, t)\right) = \max_j |f_j(e^{i\theta_1}, \dots, e^{i\theta_{n-1}})|$.

Fixing $\theta_1, \dots, \theta_{n-1}$, we have

$$\log M\left(f(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, t)\right) = \int_{\mathbb{T}} \log|f(e^{i\theta_1}, \dots, e^{i\theta_n})| \mathrm{d}\mu_n,$$

and thus

$$\begin{aligned} \log M(f) &= \int_{\mathbb{T}^n} \log|f(e^{i\theta_1}, \dots, e^{i\theta_n})| \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_n \\ &= \int_{\mathbb{T}^{n-1}} \log M\left(f(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, t)\right) \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_{n-1}. \end{aligned}$$

Fixing $\theta_1, \dots, \theta_{n-1}$, we apply the first inequality in Lemma 1.3.10 to the one-variable polynomial $f(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, t)$. We then get

$$M\left(f(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, t)\right) \geq \binom{d_n}{\lfloor d_n/2 \rfloor}^{-1} \max_j |f_j(e^{i\theta_1}, \dots, e^{i\theta_{n-1}})|.$$

Thus we have

$$\begin{aligned}
\log M(f) &\geq \int_{\mathbb{T}^{n-1}} \log \max_j |f_j(e^{i\theta_1}, \dots, e^{i\theta_{n-1}})| \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_{n-1} - \log \binom{d_n}{\lfloor d_n/2 \rfloor} \\
&\geq \max_j \int_{\mathbb{T}^{n-1}} \log |f_j(e^{i\theta_1}, \dots, e^{i\theta_{n-1}})| \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_{n-1} - \log \binom{d_n}{\lfloor d_n/2 \rfloor} \\
&= \max_j \log M(f_j) - \log \binom{d_n}{\lfloor d_n/2 \rfloor} \\
&\geq \max_j \log \ell_\infty(f_j) - \sum_{j=1}^{n} \log \binom{d_j}{\lfloor d_j/2 \rfloor} \quad \text{by induction hypothesis} \\
&= \log \ell_\infty(f) - \sum_{j=1}^{n} \log \binom{d_j}{\lfloor d_j/2 \rfloor}.
\end{aligned}$$

This is what we desire. We are done. □

Now we are ready to prove *Gelfond's Lemma*.

Proof of Lemma 1.3.6. Recall the set-up. We have $f_1, \dots, f_m \in \mathbb{C}[t_1, \dots, t_n]$ and $f := f_1 \cdots f_m$. Let d be the sum of the partial degrees of f. We wish to prove

$$2^{-d} \prod_{j=1}^{m} \ell_\infty(f_j) \leq \ell_\infty(f) \leq 2^d \prod_{j=1}^{m} \ell_\infty(f_j).$$

Write $d_1^{(j)}, \dots, d_n^{(j)}$ for the partial degrees of f_j.

We start with the lower bound for $\ell_\infty(f)$. The proof uses the relation between $M(f)$ and $\ell_\infty(f)$ established in Lemma 1.3.11. Recall that $M(f) = M(f_1) \cdots M(f_m)$. We have

$$\begin{aligned}
\prod_{j=1}^{m} \ell_\infty(f_j) &\leq \prod_{j=1}^{m} \left(\prod_{k=1}^{n} \binom{d_k^{(j)}}{\lfloor d_k^{(j)}/2 \rfloor} M(f_j) \right) \quad \text{by the second inequality in Lemma 1.3.11} \\
&= \prod_{j=1}^{m} \prod_{k=1}^{n} \binom{d_k^{(j)}}{\lfloor d_k^{(j)}/2 \rfloor} M(f) \\
&\leq \left(\prod_{j=1}^{m} \prod_{k=1}^{n} \binom{d_k^{(j)}}{\lfloor d_k^{(j)}/2 \rfloor} \right) \left(\prod_{k=1}^{n} \left(1 + \sum_{j=1}^{m} d_k^{(j)} \right)^{1/2} \right) \ell_\infty(f) \quad \text{by the first inequality in Lemma 1.3.11.}
\end{aligned}$$

Then the upper bound is obtained from the following fact: Let $a \leq A$, $b \leq B$ and d be non-negative integers. Then $\binom{A}{a}\binom{B}{b} \leq \binom{A+B}{a+b}$ and $\binom{d}{\lfloor d/2 \rfloor}(d+1)^{1/2} \leq 2^d$.[5]

[5] The first follows from $(1+t)^A(1+t)^B = (1+t)^{A+B}$, and the second follows from Stirling's formula.

Next we prove the upper bound for $\ell_\infty(f)$. For this, we will establish

$$\ell_\infty(f) \le C \prod_{j=1}^{m} \ell_\infty(f_j)$$

with

$$C = \prod_{j=1}^{m-1} \prod_{k=1}^{n} \left(1 + d_k^{(j)}\right) \le 2^d. \tag{1.3.7}$$

Let us explain how this C is chosen. First notice that *only the degrees of the first $m-1$ polynomials count*. This observation is in many applications important. It also gives the "Moreover" part of Gelfond's Lemma.

Write $f_j = \sum_{\mathbf{k}} a_{\mathbf{k}}^{(j)} \mathbf{t}^{\mathbf{k}} = \sum_{0 \le k_1 \le d_1^{(j)}, \ldots, 0 \le k_n \le d_n^{(j)}} a_{k_1,\ldots,k_n}^{(j)} t_1^{k_1} \cdots t_n^{k_n}$. Then

$$\begin{aligned} f = \prod_{j=1}^{m} f_j &= \prod_{j=1}^{m} \left(\sum_{\mathbf{k}} a_{\mathbf{k}}^{(j)} \mathbf{t}^{\mathbf{k}} \right) \\ &= \sum_{\mathbf{e}} \left(\sum_{\mathbf{k}^{(1)} + \cdots + \mathbf{k}^{(m)} = \mathbf{e}} a_{\mathbf{k}^{(1)}}^{(1)} \cdots a_{\mathbf{k}^{(m)}}^{(m)} \right) \mathbf{t}^{\mathbf{e}}. \end{aligned}$$

Here $\mathbf{e} = (e_1, \ldots, e_n)$ is a multi-index with n components, and each $\mathbf{k}^{(j)} = (k_1^{(j)}, \ldots, k_n^{(j)})$ is also a multi-index with n components. Moreover, we have $0 \le k_1^{(j)} \le d_1^{(j)}, \ldots, 0 \le k_n^{(j)} \le d_n^{(j)}$.

Now we are reduced to the following claim: For each fixed $\mathbf{e}$, we need to prove that the number of monomials in $\sum_{\mathbf{k}^{(1)} + \cdots + \mathbf{k}^{(m)}} a_{\mathbf{k}^{(1)}}^{(1)} \cdots a_{\mathbf{k}^{(m)}}^{(m)}$ is at most C. Notice that under this assumption, if $\mathbf{k}^{(1)}, \ldots, \mathbf{k}^{(m-1)}$ are all fixed, then $\mathbf{k}^{(m)}$ is also fixed. Hence we can conclude because C is the naive upper bound for the number of choices of the tuple $(\mathbf{k}^{(1)}, \ldots, \mathbf{k}^{(m-1)})$ satisfying that $0 \le k_1^{(j)} \le d_1^{(j)}, \ldots, 0 \le k_n^{(j)} \le d_n^{(j)}$. □

1.3.3 Some other operations with polynomials

We have seen how to bound the height of the product of polynomials. Now we turn to other operations.

The first is the *sum* of polynomials. For this, Proposition 1.2.8 implies the following bound rather easily.

Proposition 1.3.12. *Let $f_1, \ldots, f_r \in \overline{\mathbb{Q}}[t_1, \ldots, t_n]$. Then we have*

$$h(f_1 + \ldots + f_r) \le \sum_{j=1}^{r} h(f_j) + \log r.$$

In what follows in this subsection, let $f(\mathbf{t}) = \sum_{\mathbf{j}} \mathbf{t}^{\mathbf{j}} = \sum_{j_1,\ldots,j_n} a_{j_1 \ldots j_n} t_1^{j_1} \cdots t_n^{j_n}$.

Next, we turn to the formal partial derivatives $\partial f / \partial t_k := \sum_{j_1,\ldots,j_n,\ j_k \ge 1} j_k a_{j_1 \ldots j_n} t_1^{j_1} \cdots t_{k-1}^{j_{k-1}} t_k^{j_k - 1} t_{k+1}^{j_{k+1}} \cdots t_n^{j_n}$.

Proposition 1.3.13. *Let $d_{\max}$ be the maximum of the partial degrees of f. Then*

$$h\left(\frac{\partial f}{\partial t_k}\right) \le h(f) + \log d_{\max}.$$

Proof. Let K be a number field such that all the coefficients of f are in K.

Each $j_k \neq 0$ appearing in the monomials of f satisfies $1 \leq j_k \leq d_{\max}$. If v is non-archimedean, then $\|j_k\|_v \leq 1$. If v is archimedean, then $\|j_k\|_v \leq \|d_{\max}\|_v$. In summary, $\|j_k\|_v \leq \max\{1, \|d_{\max}\|_v\}$ for each $v \in M_K$.

Notice that each coefficient of $\partial f/\partial t_k$ is $j_k a_{j_1 \dots j_n}$. Thus

$$\|\partial f/\partial t_k\|_v \leq \max\{1, \|d_{\max}\|_v\}\|f\|_v,$$

and hence $\max\{1, \|\partial f/\partial t_k\|_v\} \leq \max\{1, \|d_{\max}\|_v\} \max\{1, \|f\|_v\}$. So

$$\begin{aligned}[K:\mathbb{Q}]h(\partial f/\partial t_k) &= \sum_{v \in M_K} \log^+ \|\partial f/\partial t_k\|_v \\ &\leq \sum_{v \in M_K} \log^+ \|d_{\max}\|_v + \sum_{v \in M_K} \log^+ \|f\|_v \\ &= [K:\mathbb{Q}]h(d_{\max}) + [K:\mathbb{Q}]h(f).\end{aligned}$$

Hence $h(\partial f/\partial t_k) \leq \log d_{\max} + h(f)$ because $h(d_{\max}) = \log d_{\max}$ as $d_{\max}$ is a positive integer. □

1.3.4 Mahler measure and algebraic number

Let us see another application of Jensen's Lemma (Lemma 1.3.9)[6], which establishes the relation between the Mahler measure and the height of an algebraic number.

Proposition 1.3.14. *Let $\alpha \in \overline{\mathbb{Q}}$ and let f be the minimal polynomial of α over $\mathbb{Z}$. Then we have*

$$\log M(f) = \deg(\alpha)h(\alpha). \tag{1.3.8}$$

In particular, we have

$$\log |N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha)| \leq \deg(\alpha)h(\alpha). \tag{1.3.9}$$

Proof. Set $d = \deg(\alpha)$ and write $f(t) = a_d t^d + \cdots + a_0 \in \mathbb{Z}[t]$. Write $\alpha_1 = \alpha, \dots, \alpha_d$ the Galois conjugates of α. Then $f(t) = a_d(t-\alpha_1)\cdots(t-\alpha_d)$. Let $K \subseteq \overline{\mathbb{Q}}$ be the Galois closure of $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$, *i.e.* K is the smallest Galois extension over $\mathbb{Q}$ which contains α and all its Galois conjugate. Write $G = \mathrm{Gal}(K/\mathbb{Q})$. Then $\{\sigma(\alpha)\}_{\sigma \in G}$ contains every conjugate of α exactly $[K:\mathbb{Q}]/d$ times.

For each (non-archimedean) $v \in M_K^0$, Gauß's Lemma (Lemma 1.3.5) yields

$$\max\{\|a_d\|_v, \dots, \|a_0\|_v\} = \|f\|_v = \|a_d\|_v \prod_{i=1}^{d} \max\{1, \|\alpha_i\|_v\}.$$

Notice that the left hand side equals 1 because each $a_i \in \mathbb{Z}$ and $\gcd(a_d, \dots, a_0) = 1$. Thus

$$\log \|a_d\|_v + \sum_{i=1}^{d} \log^+ \|\alpha_i\|_v = 0 \tag{1.3.10}$$

for each $v \in M_K^0$.

[6] $\log M(f) = \log |a_d| + \sum_{j=1}^{d} \log^+ |\alpha_j|$ for $f(t) = a_d(t-\alpha_1)\cdots(t-\alpha_d)$.

Now we have

$$\begin{aligned}
[K:\mathbb{Q}]h(\alpha) &= \frac{[K:\mathbb{Q}]}{d}\sum_{i=1}^{d} h(\alpha_i) \quad \text{by Lemma 1.2.4} \\
&= \frac{1}{d}\sum_{i=1}^{d}\sum_{v\in M_K} \log^+ \|\alpha_i\|_v \\
&= \frac{1}{d}\left(\sum_{v|\infty}\sum_{i=1}^{d}\log^+\|\alpha_i\|_v - \sum_{v\in M_K^0}\log\|a_d\|_v\right) \quad \text{by (1.3.10)} \\
&= \frac{1}{d}\sum_{v|\infty}\left(\log\|a_d\|_v + \sum_{i=1}^{d}\log^+\|\alpha_i\|_v\right) \quad \text{by Product Formula applied to } a_d.
\end{aligned}$$

If $K_v = \mathbb{R}$, then $\|\cdot\|_v = |\cdot|$. If $K_v = \mathbb{C}$, then $\|\cdot\|_v = |\cdot|^2$. Recall, from Algebraic Number Theory, the basic fact that $\#\{v : K_v = \mathbb{R}\} + 2\#\{v : K_v = \mathbb{C}\} = [K:\mathbb{Q}]$. Hence we can apply Jensen's Lemma (Lemma 1.3.9) to each term on the right hand side and obtain

$$[K:\mathbb{Q}]h(\alpha) = \frac{[K:\mathbb{Q}]}{d}\log M(f).$$

This yields (1.3.8).

To prove the "In particular" part, recall from Algebraic Number Theory that $N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha) = \prod_{i=1}^{d}\alpha_i$. Thus $\log|N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha)| = \sum_{i=1}^{d}\log|\alpha_i|$, which then $\le \sum_{i=1}^{d}\log^+|\alpha_i|$ and hence $\le \log M(f)$ by Jensen's Lemma. □

Remark 1.3.15. *Let us have a quick look at the* ***Lehmer Conjecture****. If $\alpha \neq 0$ is an algebraic number with minimal polynomial f, the Mahler measure of α is defined to be $M(\alpha) := M(f)$. Then (1.3.8) yields $M(\alpha) = H(\alpha)^{\deg(\alpha)}$. Lehmer's conjecture predicts that there exists a contant c such that $M(\alpha) \ge c > 1$ for all $\alpha \in \overline{\mathbb{Q}}^*$ not a root of unity. Alternatively, $h(\alpha) \ge c/\deg(\alpha)$ for some absolute constant $c > 0$. This conjecture is open. Currently, we have Dobrowolski's theorem which claims ($d := \deg(\alpha)$)*

$$M(\alpha) \ge 1 + c\left(\frac{\log\log d}{\log d}\right)^3.$$