Heights in diophantine geometry: an introduction

François Gatine

Contents

1 Introduction

Our references are [BG06] and [HS13]. All fields considered are number fields unless specified otherwise, all affine and projective varieties are defined over such fields. We only deal with $\overline{\mathbb{Q}}$-rational points of these varieties, where $\overline{\mathbb{Q}}$ denotes a fixed algebraic closure of $\mathbb{Q}$.

As projective space is homogeneous, all points are geometrically equal. Arithmetically speaking however, some points are more equal then others. Heights are tools helping us quantify the arithmetic complexity of points in projective space.

Let us begin humbly in the usual affine space $\mathbb{A}^n$, and fix a coordinate basis. Intuitively speaking, a point has a high complexity if some of its coordinates define massive extensions of $\mathbb{Q}$. In $\mathbb{A}^2$ for instance, one expects $(1, 3)$ to be less complex than $(\sqrt{2}, 1)$, itself being less complex than (i, α) where α is a root of some irreducible polynomial of degree 48.

Similarly, one can play the same game in $\mathbb{P}^n$ by considering homogeneous coordinates, as long as we make sure to define "complexity" to be insensitive to scalar multiples of the homogeneous coordinates. Over $\mathbb{P}^1$ for instance, the point $[1, \sqrt{2}]$ can also be representend as $[\sqrt{2}, 2]$, and both representations indeed seem to have identical complexities ; however it can also be expressed as $[\sqrt{3}, \sqrt{6}]$, which would seem more complex, even though it is not. Thus, we have to be careful in the definition of heights to account for "artificial complexity" induced by a poor choice of representative.

Defining heights by looking at the "complexity of coordinates" has an obvious flaw: changing coordinates changes the height. Although affine and projective space come equipped with a canonical set of coordinates, real issues will arise when extending the notion of heights to projective varieties, which have a collection of projective embeddings, none of which seem a priori more or less canonical. It will thus be of the utmost importance to study the behavior of heights under change of coordinates, or even finite maps between projective spaces, so as to be able to handle how two different projective embeddings of a variety may lead to two different heights on the variety.

Well-definedness questions aside, what should heights be good for? As mentionned, heights allows one to order $\overline{\mathbb{Q}}$-points by arithmetic complexity, thus allowing one to ask questions such as: what is the asymptotic behavior of the number of points of height less then H on my variety, as H grows? This is made possible by Northcott's theorem, stating that there are only finitely many points of $\mathbb{A}^1$ (thus also of $\mathbb{P}^1$) with bounded height and bounded degree over $\mathbb{Q}$. From there, one tries to achieve explicit inequalities involving heights and other helpful quantities, with such great success that heights have played a key role in numerous famous proofs, such as Mordell's theorem and Falting's theorem.

2 Heights on projective and affine spaces

We begin by fixing some notations and conventions.

Over $\mathbb{Q}$, we denote by $|\cdot|_\infty$ the standard absolute value, and for p any prime number, we consider $|\cdot|_p$ the p-adic absolute value normalized by $|p|_p = p^{-1}$.

If $K/\mathbb{Q}$ is a number field, and v is a finite place of K dividing a prime number p, we denote by $|\cdot|_v$ the absolute value on K defined as

$$\forall x \in K,\ |x|_v = |N_{K_v/\mathbb{Q}_p}(x)|_p^{1/[K:\mathbb{Q}]}.$$

Notice that as soon as $[K_v : \mathbb{Q}_p] \neq [K : \mathbb{Q}]$, $|\cdot|_v$ does not coincide with $|\cdot|_p$ over $\mathbb{Q}$.

If v is now an infinite place, $|\cdot|_v$ is defined as

$$\forall x \in K,\ |x|_v = |N_{K_v/\mathbb{R}}(x)|_\infty^{1/[K:\mathbb{Q}]}.$$

In other words, the same formula holds for $p = \infty$.

Let L/K be an extension of number fields, and v a place (finite or infinite) of K. Recall that the following equality holds:

$$\forall y \in L,\ |N_{L/K}(y)|_v = \prod_{w|v} |N_{L_w/K_v}(y)| = \prod_{w|v} |y|_w^{[L:K]} = \left(\prod_{w|v} |y|_w\right)^{[L:K]}$$

where the second equality follows from our conventions (check this).

For K a number field, we denote by Σ_k the set of all places (finite or infinite) of K. In particular,

$$\Sigma_\mathbb{Q} = \{\infty\} \cup \{p \mid p \text{ prime}\}.$$

2.1 The product formula

The product formula is an easy yet fundamental relation between all places of a number field, which is the main ingredient in defining heights. Notice the following:

Lemma 2.1: Product formula for $\mathbb{Q}$

If $x \in \mathbb{Q}^\times$, then

$$\prod_{v \in \Sigma_\mathbb{Q}} |x|_v = 1.$$

Proof. One can assume $x \in \mathbb{Z}$. It is clear that $|x|_v = 1$ for all but finitely many places. The equality is a consequence of the prime decomposition over $\mathbb{Z}$. □

We can now bootstrap from $\mathbb{Q}$ to extend the result to all number fields, as long as we use the absolute values normalized as above.

Proposition 2.2: Product formula for number fields

Let K be a number field. If $x \in K^\times$, then

$$\prod_{v \in \Sigma_K} |x|_v = 1.$$

Proof. Assume first that we know $|x|_v = 1$ for all but finitely many places of K. We can then write:

$$\prod_{v \in \Sigma_K} |x|_v = \prod_{v_0 \in \Sigma_{\mathbb{Q}}} \prod_{v | v_0} |x|_v = \left(\prod_{v_0 \in \Sigma_{\mathbb{Q}}} |x|_{v_0} \right)^{[K:\mathbb{Q}]} = 1.$$

It remains to show $|x|_v = 1$ for all but finitely many places of K, so as to ensure the product is defined, and the manipulation above is legal. As K admits only finitely many archimedean places, it suffices to consider finite places. Let $x \in K^\times$, it is algebraic over $\mathbb{Q}$ hence one can find $a_0, \ldots, a_{n-1} \in \mathbb{Q}$ for some n, such that

$$x^n + \sum_{i=0}^{n-1} a_i x^i = 0.$$

We have $|a_i|_p \in \{0, 1\}$ for all but finitely many prime numbers p. Since there are only finitely many places of K above each prime number, $|a_i|_v \leq 1$ for all but finitely many places of K. Let v be such a place; applying the ultrametric inequality to the polynomial relation yields $|x|_v \leq 1$. Thus we have shown that $|x|_v \leq 1$ for all but finitely many places of K, but as the same is true for $1/x$, we find $|x|_v = 1$ for all but finitely many places of K. □

2.2 The standard height function

We are now equiped to define the (standard) height function on $\mathbb{P}^n_{\overline{\mathbb{Q}}}$:

Definition 2.3

The height of a point $P = [x_0, \cdots, x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$ is defined as:

$$h(P) = \sum_{v \in \Sigma_K} \max_j \log |x_j|_v$$

where K is any field containing all coordinates x_j.

Fixing P as in the definition, we need to check that $h(P)$ does not depend on the choice of K, and is invariant under scaling of the coordinates. Indeed, if L/K is a finite extension, our conventions ensure that for any place $v \in \Sigma_K$:

$$\sum_{w | v} \log |x_j|_w = \log |x_j|_v$$

Moreover we have

$$\sum_{w\in\Sigma_L} \max_j \log|x_j|_w = \sum_{v\in\Sigma_K}\sum_{w|v} \max_j \log|x_j|_w$$

We claim that we can permute $\sum_{w|v}$ and $\max_j$. Indeed, let j_0 be the index maximizing $|x_{j_0}|_v$, then it is also the index maximizing $|x_{j_0}|_w = |x_{j_0}|_v^{[L_w:K_v]/[L:K]}$ for any $w|v$. Hence:

$$\sum_{w\in\Sigma_L} \max_j \log|x_j|_w = \sum_{v\in\Sigma_K} \max_j \sum_{w|v} \log|x_j|_w = \sum_{v\in\Sigma_K} \max_j \log|x_j|_v.$$

Now if K, K' are two fields containing the coordinates of P, then so is $K\cap K'$, and the height computed over any of these three fields coincide by the argument above.

We now turn to the invariance under scalar multiplication of the coordinates. Let $\lambda\in K^\times$, define $y_j = \lambda x_j$ for any j. Then:

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

and as $\prod_{v\in\Sigma_K}|\lambda|_v = 1$ by the product formula, we find the desired result.

Example 2.4. Assume $P\in\mathbb{P}^n(\mathbb{Q}) = \mathbb{P}^n(\mathbb{Z})$. Then one can choose the x_j to be integers with no common factor. In such a case, if p is a prime number, we have $\max_j \log|x_j|_p = 1$, so the height reduces to

$$h(P) = \max_j \log|x_j|_\infty.$$

More generally, if the coordinates of P are all in $\mathcal{O}_K$ for some field K, then only the infinite places will contribute to the value of $h(P)$.

Of course, because Galois acts by permutation on places of K, it is easy to see that the height is invariant under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ by precomposition.

We now restrict to the affine space $\mathbb{A}^n$, which we view as the open subset of $\mathbb{P}^n$ defined by the equation $x_0 = 1$. Because $\log|x_0|_v = 0$ for any place v, we find:

Lemma 2.5

Let $P = (x_1,\dots,x_n)\in\mathbb{A}^n(\overline{\mathbb{Q}})\subseteq\mathbb{P}^n(\overline{\mathbb{Q}})$. Then the height of P is

$$h(P) = \sum_{v\in\Sigma_K} \max_j \log^+|x_j|_v$$

where $\log^+ := \max(0,\log)$.
For $n=1$, the height of an algebraic number $\alpha\in\overline{\mathbb{Q}}$ is thus

$$h(\alpha) = \sum_{v\in\Sigma_K} \log^+|\alpha|_v.$$

Example 2.6. Let p be a prime number, $\alpha = 1/\sqrt{p}$ and $K = \mathbb{Q}(\sqrt{p})$. If $q \neq p$ is another prime number, then $|\alpha|_q = 1$. Thus

$$h(\alpha) = \log^+ |\alpha|_\infty + \log^+ |-\alpha|_\infty + \log^+ |\alpha|_p.$$

Moreover, $|\alpha|_p = p^{1/2}$ and $\log^+ |\alpha|_\infty = 0$. In conclusion, $h(\alpha) = 1/2$.

We conclude this section by proving Kronecker's theorem:

Theorem 2.7

Let $\alpha \in \overline{\mathbb{Q}}^\times$. Then $h(\alpha) = 0$ if and only if α is a root of unity.

Proof. If α is a root of unity, then all its absolute values are 1, thus $h(\alpha) = 0$.

Assume now that $h(\alpha) = 0$, i.e. $|\alpha|_v \leq 1$ for all places v of some field K containing all conjugates of α. This implies in particular that α is an integer of K, and the constant coefficient of the minimal polynomial over $\mathbb{Q}$ of α is a nonzero integer of magnitude ≤ 1, in other words it is ± 1, hence $|\alpha|_v = 1$ for all places v.

If d is a positive integer, denote by E_d the set of algebraic integers in $\overline{\mathbb{Q}}$ of degree at most d with all archimedean absolute value 1. This is a finite set, because there are only finitely many choices for the minimal polynomial of an element of E_d. Moreover, E_d is stable under raising to an integer power (as this does not raise the degree, contrary to multiplication in general), thus the elements of E_d are roots of 1.

We conclude by observing that $\alpha \in E_{\deg \alpha}$. □

3 Finiteness and estimations

We now dive deeper into the behavior of heights. Here is a simple lemma:

Lemma 3.1

Let $\alpha \in \overline{\mathbb{Q}}^\times$ and $\lambda \in \mathbb{Q}$. Then*

$$h(\alpha^\lambda) = |\lambda| h(\alpha).$$

Proof. If $\lambda > 0$, it is already true that $\log^+ x^\lambda = \lambda \log^+ x$, hence the claim. It remains to prove the case $\lambda = -1$. Let K be a field containing α, and v a place of K. Notice that

$$\log |\alpha|_v = \log^+ |\alpha|_v - \log^+ |1/\alpha|_v$$

and if we sum over all places v, the left-hand-side vanishes because of the product formula, leaving the equality

$$h(\alpha) - h(1/\alpha) = 0.$$

□

The Lemma implies the so called *fundamental inequality*:

Proposition 3.2

Let $S \subseteq \Sigma_K$ be a finite set of places of a field K. Let $\alpha \in K^\times$, then

$$-h(\alpha) \leq \sum_{v \in S} \log |\alpha|_v \leq h(\alpha).$$

Proof. Notice that $\log \leq \log^+$ and $0 \leq \log^+$, hence

$$\sum_{v \in S} \log |\alpha|_v \leq \sum_{v \in S} \log^+ |\alpha|_v \leq h(\alpha).$$

Replacing α with $-\alpha$ and using the Lemma above for $\lambda = -1$, we find

$$-\sum_{v \in S} \log |\alpha|_v \leq h(\alpha)$$

which rearranges to

$$-h(\alpha) \leq \sum_{v \in S} \log |\alpha|_v.$$

□

Transition

Definition 3.3: Mahler measure

Let $f \in \mathbb{C}[x_1, \ldots, x_n]$. The Mahler measure of f is defined to be

$$M(f) := \exp\left(\int_{\mathbb{T}^n} \log |f(\zeta_1, \ldots, \zeta_n)| d\mu(\zeta_1) \ldots d\mu(\zeta_n) \right)$$

where $\mathbb{T}$ denotes the complex unit circle with the normalized measure $d\mu = \frac{d\theta}{2\pi}$.

By definition, if f, g are two polynomials as in the definition above, we have $M(fg) = M(f)M(g)$. In one variable, it is then enough to compute the height of constant and degree 1 polynomials.

Lemma 3.4

Let $c, \alpha \in \mathbb{C}$. Then

$$M(c) = |c| \qquad M(t \mapsto t - \alpha) = \log^+(|\alpha|).$$

Proof. The first equality is clear, we focus on the second one, which amounts to show

$$\frac{1}{2\pi} \int_0^{2\pi} \log(|e^{i\theta} - \alpha|) d\theta = \begin{cases} 0 & \text{if } |\alpha| \leq 1 \\ \log |\alpha| & \text{if } \geq 1 \end{cases}.$$

Denote $f : z \mapsto \log|z-\alpha| = \frac{1}{2}\log[(z-\alpha)(\overline{z}-\overline{\alpha})]$. Notice that

$$\begin{aligned}\frac{\partial^2 f}{\partial z \partial \overline{z}} &= \frac{\partial}{\partial z}\left(\frac{(z-\alpha)}{(z-\alpha)(\overline{z}-\overline{\alpha})}\right) \\ &= \frac{\partial}{\partial z}\frac{1}{\overline{z}-\overline{\alpha}} \\ &= 0.\end{aligned}$$

which shows that f is harmonic over $\mathbb{C}\setminus\{\alpha\}$, so if $|\alpha|>1$, α is not in the unit disk, so the average of f over $\mathbb{T}$ is $f(0)=\log|\alpha|$.
If $|\alpha|<1$, define $g : z \mapsto \log|1-\alpha\overline{z}|$, it is harmonic over $\mathbb{C}\setminus\{1/\alpha\}$ so its average over $\mathbb{T}$ is $g(0)=0$. But f and g coincide over $\mathbb{T}$, which concludes this case.
If $|\alpha|=1$, we can assume $\alpha=1$, then $|e^{i\theta}-1| = \frac{1}{2}|\sin(\theta/2)|$. After changing variables in the integral and rearranging, we now need to show

$$\int_0^\pi \log|\sin(\varphi)|d\varphi = -\pi\log 2$$

which is a standard exercice, left to the reader. □

The Lemma hints at how the Mahler measure of a rational polynomial is related to the heights of its roots. Combining the Lemma with the multiplicativity of M, we find

Proposition 3.5: Archimedean Jensen's formula

Let $f(t) = a_d(t-\alpha_1)\dots(t-\alpha_d)$ with roots algebraic numbers. Then

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

We are now equipped to precisely state the link between Mahler measure and height.

Definition 3.6: Height of a polynomial

If $K/\mathbb{Q}$ is a number field and v is a place of K, we define the v-adic absolute value of a polynomial $f = \sum_{j_1,\dots,j_n} a_{j_1,\dots,j_n} t_1^{j_1}\dots t_n^{j_n} \in K[t_1,\dots,t_n]$ as

$$|f|_v := \max_{\underline{j}} |a_{\underline{j}}|_v.$$

We need this definition for the proof of the following.

Proposition 3.7: Nonarchimedean Jensen's formula

Let $\alpha \in \overline{\mathbb{Q}}$, and f its minimal polynomial *over* $\mathbb{Z}$, with leading coefficient $a_d \in \mathbb{Z}$. Let $K/\mathbb{Q}$ a Galois extension of group G, containing α, and let v be a finite place of K. Then

$$\sum_{\sigma\in G}\log^+|\sigma\alpha|_v = \frac{[K:\mathbb{Q}]}{d}\log|a_d|_v.$$

Proof. Let $d = \deg \alpha$. From field theory we have the following relation between the minimal polynomial of α over $\mathbb{Q}$, and the characteristic polynomial of multiplication by α on the $\mathbb{Q}$-vector space K:

$$(a_d^{-1} f)^{[K:\mathbb{Q}]/d} = \prod_{\sigma \in G} (t - \sigma(\alpha))$$

Consider v a finite place of K. Applying Gauss's Lemma (Lemma 3.8 below) to the above, we find

$$1 = |f|_v = |a_d|_v \prod_{\sigma \in G} |t - \sigma(\alpha)|_v^{d/[K:\mathbb{Q}]} = |a_d|_v \prod_{\sigma \in G} \max(1, |\sigma(\alpha)|_v)^{d/[K:\mathbb{Q}]}$$

hence

$$\sum_{\sigma \in G} \log^+ |\sigma\alpha|_v = \frac{[K:\mathbb{Q}]}{d} \log |a_d|_v.$$

□

Lemma 3.8: Gauss's Lemma

Let K and v be a finite place of K. If $f, g \in K[x_1, \ldots, x_n]$, then

$$|fg|_v = |f|_v |g|_v.$$

Proof. Omitted. See [BG06], Lemma 1.6.3. □

Proposition 3.9

Let $\alpha \in \overline{\mathbb{Q}}$, and f its minimal polynomial *over* $\mathbb{Z}$. Then

$$\log M(f) = \deg(\alpha) h(\alpha).$$

Proof. Let $d = \deg \alpha$, with leading coefficient $a_d \in \mathbb{Z}$. Fix $K/\mathbb{Q}$ a Galois extension of group G, containing α. Then:

$$\begin{aligned}
[K:\mathbb{Q}] h(\alpha) &= \sum_{\sigma \in G} h(\sigma\alpha) \\
&= \sum_{v \in \Sigma_K} \sum_{\sigma \in G} \log^+ |\sigma\alpha|_v \\
&= \sum_{v|\infty} \sum_{\sigma \in G} \log^+ |\sigma\alpha|_v + \frac{[K:\mathbb{Q}]}{d} \sum_{v \nmid \infty} \log |a_d|_v
\end{aligned}$$

where the last step follows from Proposition 3.7. Denoting $\alpha_1, \ldots, \alpha_d$ all conjugates

of α, we find:

$$\begin{aligned}
[K:\mathbb{Q}]h(\alpha) &= \frac{[K:\mathbb{Q}]}{d}\sum_{v|\infty}\left(\sum_{j=1}^{d}\log^+|\alpha_j|_v + \log|a_d|_v\right)\\
&= \frac{[K:\mathbb{Q}]}{d}\left(\log\prod_{v|\infty}|a_d|_v + \sum_{j=1}^{d}\log^+\prod_{v|\infty}|\alpha_j|_v\right)\\
&= \frac{[K:\mathbb{Q}]}{d}\left(\log|a_d|_\infty + \sum_{j=1}^{d}\log^+|\alpha_j|_\infty\right)\\
&= \frac{[K:\mathbb{Q}]}{d}\log M(f)
\end{aligned}$$

which proves the claim after rearrangement. □

We can now state and prove Northcott's Theorem.

Theorem 3.10: Northcott's theorem

Let d be a positive integer and $B > 0$. Then the set

$$\{\alpha \in \overline{\mathbb{Q}} \mid h(\alpha) \leq B,\ \deg\alpha \leq d\}$$

is finite.

Proof. Let α be in this set, denote by f its minimal polynomial over $\mathbb{Z}$, which has degree $\leq d$. By Proposition 3.9 above, $M(f)$ is bounded. Moreover, we will show that the norm $|f|_\infty$ (as in Definition 3.6) is bounded in terms of $M(f)$. Once this is done, there are only finitely many candidates for f, thus α lies in a finite set.
It remains to show the following claim:

$$|f|_\infty \leq \binom{d}{\lfloor d/2\rfloor} M(f).$$

Indeed, write $f = a_d t^d + \cdots + a_0$, denote $\alpha_1, \ldots, \alpha_d$ its roots, i.e. the conjugates of α. Let $0 \leq r \leq d$, we have

$$|a_{d-r}| = |a_d|\left|\sum_{j_1<\cdots<j_r}\alpha_{j_1}\ldots\alpha_{j_r}\right| \leq \binom{d}{r}|a_d|\prod_{j=1}^{d}\max(1,|\alpha_j|) \leq \binom{d}{\lfloor d/2\rfloor}M(f)$$

using Proposition 3.5 and the inequality $\binom{d}{r} \leq \binom{d}{\lfloor d/2\rfloor}$ for the last step. □

This result extends easily to projective space. If $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$, denote by $\mathbb{Q}(P)$ the field generated by the coordinates of P (after setting an arbitrary nonzero coordinate to 1). We have

Corollary 3.11

Let d be a positive integer and $B > 0$. Then the set

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

is finite.

Proof. Up to permuting coordinates, write $P = [1 : x_1 : \cdots : x_n] \in \mathbb{P}^n(K)$ for some number field K. Then for any place v of K, and any $1 \leq i \leq n$:

$$\max(1, |x_1|_v, \ldots, |x_n|_v) \geq \max(1, |x_i|_v)$$

which implies $h(P) \geq h(x_i)$. Thus the coordinates of P belong to a finite set, which concludes the proof. □

4 Weil's height machine

The height function we defined on $\mathbb{P}^n$ depends on the choice of projective coordinates. Indeed, if the height were insensitive to the transitive action of PGL_n, then it would be constant. That being said, composing the height function with a projective transformation does not have such a big impact, as we show next.

Consider a projective transformation $M \in \mathrm{PGL}_n(\overline{\mathbb{Q}})$, viewed as an $(n+1) \times (n+1)$ matrix $M = (m_{ij})$. Then if $x = [x_0, \ldots, x_n] \in \mathbb{P}^n(\overline{\mathbb{Q}})$, if K is a number field containing the coordinates of M and x, and if v is a place of K, we find:

$$\begin{aligned} \ln \max_i |(Mx)_i|_v &\leq \ln(\epsilon_v(n+1) \max_{i,j} |m_{ij} x_j|_v) \\ &\leq \ln \epsilon_v(n+1) + \ln \max_{i,j} |m_{ij}|_v + \ln \max_j |x_j| \end{aligned}$$

where $\epsilon_v(a)$ is a if v is infinite, or 1 if v is finite. Summing over all places v, we find

$$h(Mx) \leq h(x) + h(M) + \ln(n+1) \sum_{v|\infty} 1$$

where $h(M)$ denotes the height of M viewed as an element of $\mathbb{P}^{(n+1)^2-1}$. Symplifying, we have shown:

$$h \circ M \leq h + C_M$$

where C_M is a constant which only depends on M. Replacing M with M^{-1} shows

$$h \circ M \geq h - C_{M^{-1}}$$

thus we have established the following:

Proposition 4.1

Let $M \in \mathrm{PGL}_n$. Then

$$h \circ M = h + O(1).$$

This should be our cue to consider the following motto: *height functions are canonically defined up to bounded functions.*

That being said, let us move on to finite maps, *e.g.* the N-uple embedding

$$\begin{aligned} f : [x_0, \dots, x_n] &\mapsto [\text{all degree N monomials in the } x_i] \\ \mathbb{P}^n &\to \mathbb{P}^{(n+1)^N - 1} \end{aligned}$$

which is a degree N morphism onto its image. It is an easy exercice to show that $h_{\mathbb{P}^{(n+1)^N-1}} \circ f = N h_{\mathbb{P}^n}$. The same cannot be said for more general finite maps, yet the motto above fixes this issue.

Proposition 4.2

Let $f : \mathbb{P}^n \to \mathbb{P}^m$ be a morphism of degree N over its image. Then

$$h_{\mathbb{P}^m} \circ f = N h_{\mathbb{P}^n} + O(1).$$

Proof. Write $f = [f_0, \dots, f_m]$ with f_i's homogeneous polynomials of degree N which do not vanish all at the same time. Let $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$, K a field containing the coordinates of x and the f_i's, and v a place of K. The same argument as in the case of projective transformations shows that

$$h_{\mathbb{P}^m}(f(x)) \leq N h_{\mathbb{P}^n}(x) + C_f$$

except this time, f need not be invertible, so the reverse inequality is not immediate. We note in passing that this argument also holds if f is rational (that is, the f_i's may have common zeroes).

We now establish the converse inequality. By the Nullstellensatz, there is a positive integer $t > m$ and homogeneous polynomials $g_{i,j}$ of degree $t - N$ such that

$$\forall 0 \leq j \leq m, \ \sum_i g_{i,j} f_i = X_j^t.$$

Let $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$ and K a number field containing the coordinates of x, all f_i's and all $g_{i,j}$'s. Evaluating the above at x, we find that for any place v of K:

$$\max_i |x_i^t|_v \leq \epsilon_v (m+1) \max_{i,j} |g_{i,j}(x)|_v \max_i |f_i(x)|_v.$$

Taking logarithm and summing over all places yields

$$t h(x) \leq C + h(g(x)) + h(f(x))$$

where C is a constant independent of x, and g is the rational map with coordinates $g_{i,j}$. However, the first part of the proof shows

$$h(g(x)) \leq (t - d) h(x) + C_g$$

thus, rearranging, we obtain

$$d h(x) \leq C' + h(f(x))$$

with C' a constant independent of x. This concludes the proof. □

Remark 4.3. As mentionned in the proof, if $f : \mathbb{P}^n \dashrightarrow \mathbb{P}^m$ is only a rational map (*i.e.* it may not be defined everywhere), then we still have $h_{\mathbb{P}^m} \circ f \leq N h_{\mathbb{P}^n} + O(1)$, but the converse may not hold in general. Consider for instance

$$\begin{aligned} f : \mathbb{P}^2 &\to \mathbb{P}^2 \\ [x : y : z] &\mapsto [x^2 : y^2 : xz] \end{aligned}$$

which is defined except at $[0 : 0 : 1]$. Let $P = [x, y, z] \in \mathbb{P}^2(\mathbb{Q})$ with coprime integer coordinates. Then

$$h(f(P)) = \log \max\{|x^2|, |y^2|, |xz|\} + \sum_p \log \max\{|x^2|_p, |y^2|_p, |xz|_p\}.$$

Notice that the integers x^2, y^2 and xz may no longer be coprime, because of common factors between x and y^2. Let $d = \gcd(x, y^2)$, then $\max\{|x^2|_p, |y^2|_p, |xz|_p\} = |d|_p \max\{|(x/d)^2|_p, |y^2/d|_p, |(x/d)z|_p\} = |d|_p$ as x/d and y^2/d are coprime. Moreover, by the product formula,

$$\sum_p |d|_p = -\log |d| = -\log d$$

hence we have shown

$$h(f(P)) = \log \max\{|x^2|, |y^2|, |xz|\} - \log \gcd(x, y^2).$$

Consider now $P = [2^n : 2^n : 1]$ for any $n \geq 1$, thus $2h(P) = 2n \log 2$, but

$$h(f(P)) = 2n \log 2 - n \log 2 = n \log 2$$

which shows that $2h$ and $h \circ f$ differ by an unbounded function.

One can adapt the proof above to show:

Proposition 4.4

Let $f : \mathbb{P}^n \to \mathbb{P}^m$ be a rational map of degree N, defined outside a closed subset Z. Let X be a closed subvariety of $\mathbb{P}^n$ such that $X \cap Z = \varnothing$. Then:

$$h_{\mathbb{P}^m} \circ f_{|X} = N h_{\mathbb{P}^n} + O(1).$$

Proof. See [HS13], Theorem B.2.5. □

We are now equipped to show the following: given a projective variety V, closed embedding in projective spaces corresponding to equivalent ample divisors yield height identical height functions on V, up to a bounded function.

Definition 4.5

Let $\varphi : V \to \mathbb{P}^n$ be a morphism of varieties over $\overline{\mathbb{Q}}$. The height on V relative to f is defined as

$$\begin{aligned} h_\varphi : V(\overline{\mathbb{Q}}) &\to \mathbb{R}_+^\times \\ P &\mapsto h_{\mathbb{P}^n}(\varphi(P)). \end{aligned}$$

If $\varphi : \mathbb{P}^n \to \mathbb{P}^m$ is a morphissm of degree N, we have shown that $h_\varphi = N h_{\mathbb{P}^n} + O(1)$.

Theorem 4.6

Let V be a variety defined over $\overline{\mathbb{Q}}$, $\varphi : V \to \mathbb{P}^n$ and $\psi : V \to \mathbb{P}^m$ morphisms, H and H' hyperplanes in $\mathbb{P}^n$ and $\mathbb{P}^m$ respectively. If $\varphi^* H$ and $\psi^* H'$ are linearly equivalent in $\mathrm{Pic}(V)$, then

$$h_\varphi = h_\psi + \mathcal{O}(1).$$

Before we move on to the proof, here is a sanity check. If $\varphi : \mathbb{P}^n \to \mathbb{P}^m$ is a morphissm of degree N, then $\varphi^* H'$ is equivalent to NH where H, H' are hyperplanes in $\mathbb{P}^n, \mathbb{P}^m$. In particular, NH is equivalent to the pullback of a hyperplane under the N-uple embedding, call it ψ, for which we have the *equality* $h_\psi = N h_{\mathbb{P}^n}$. Thus, the theorem asserts

$$h \circ \varphi = h_\psi + O(1) = N h_{\mathbb{P}^n} + O(1)$$

which is what has already been shown above.

Proof. Let $D \in \mathrm{Div}(V)$ be an effective divisor in the linear equivalence class of the pullbacks $\varphi^* H$ and $\psi^* H'$. Then $\varphi = [f_0, \ldots, f_n]$ and $\psi = [g_0, \ldots, g_m]$ have components f_i and g_i which belong to the finite dimensional space

$$L(D) := \{f \in K(V) \mid div(f) + D \geq 0\}$$

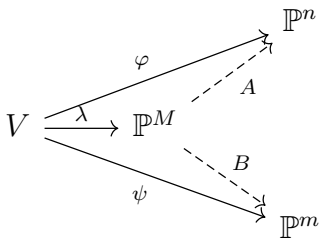
where $K(V)$ denotes the rational functions on V. Fix $h_0, \ldots, h_M$ a basis of $L(D)$, then one can write

$$f_i = \sum_{j=0}^{M} a_{i,j} h_j$$

$$g_i = \sum_{j=0}^{M} b_{i,j} h_j$$

with $a_{i,j}$ and $b_{i,j}$ constants in $\overline{\mathbb{Q}}$.
Define the morphism $\lambda = [h_0, \ldots, h_M] : V \to \mathbb{P}^M$, as well as the rational maps $A : \mathbb{P}^n \to \mathbb{P}^M$ and $B : \mathbb{P}^m \to \mathbb{P}^M$ defined by the matrices $(a_{i,j})$ and $(b_{i,j})$ respectively.

By construcction we have the following commutative diagram



with A and B well defined on the image X of λ. This image is closed because λ is proper. Thus we can apply Proposition 4.4 to see that

$$h_{\mathbb{P}^n} \circ A = h_{\mathbb{P}^M} + O(1) \qquad\qquad h_{\mathbb{P}^m} \circ B = h_{\mathbb{P}^M} + O(1)$$

and thus

$$h_{\mathbb{P}^n} \circ A - h_{\mathbb{P}^m} \circ B = O(1).$$

Precomposing by λ, we find the expected result. □

The theorem above is the first step towards constructing *Weil's height machine*. Before, considering a morphism $\varphi : V \to \mathbb{P}^n$ (a closed embedding, for instance) associate to it a height function $h_{\mathbb{P}^n} \circ \varphi$, which only depends, up to a bounded function, on the class in $\mathrm{Pic}(V)$ of the divisor associated to φ. Weil's height machine generalizes this procedure by associating a height function to *any divisor* of V, in such a way that linearly equivalent divisors produce heights differing by a bounded function.

It turns out that once stated in terms of quotient spaces (*i.e.* as a map from $\mathrm{Pic}(V)$ to functions up to bounded functions) Weil's height machine is in fact a group homomorphism, which allows one to state it cleanly as follows:

Theorem 4.7

Let V be a variety defined over $\overline{\mathbb{Q}}$. There exists a unique group homomorphism

$$h_V : \mathrm{Pic}(V) \to \frac{\{\text{functions } V(\overline{\mathbb{Q}}) \to \mathbb{R}\}}{\{\text{bounded functions } V(\overline{\mathbb{Q}}) \to \mathbb{R}\}}$$

with the property that if $\mathcal{L} \in \mathrm{Pic}(V)$ is very ample and $\varphi_{\mathcal{L}} : V \to \mathbb{P}^n$ is the corresponding embedding, then

$$h_{V,\mathcal{L}} = h_{\mathbb{P}^n} \circ \varphi + O(1)^a.$$

Additionally, the homomorphism h_V satisfies:

(a) (Functoriality) Let $f : V \to W$ be a morphism of smooth varieties, and let $D \in \mathrm{Pic}(W)$. Then

$$h_{V,f^*D} = h_{W,D} \circ \varphi + O(1)$$

(b) (Positivity) Let B be the base locus of an effective divisor class $D \in \mathrm{Pic}(V)$ and assume that $B \neq V$. Then

$$h_{V,D} \geq O(1) \qquad \text{on } V \setminus B.$$

(c) (Algebraic equivalence) Let $\mathcal{L}, D \in \mathrm{Pic}(V)$ with $\mathcal{L}$ ample and D algebraically equivalent to zero. Then

$$\lim_{\substack{P \in V(\overline{\mathbb{Q}}) \\ h_{V,\mathcal{L}}(P) \to \infty}} \frac{h_{V,D}(P)}{h_{V,\mathcal{L}}(P)} = 0$$

[a]Notice that for $V = \mathbb{P}^n$ and $\varphi = \mathrm{id}$, we recover (up to a bounded function) the classical height on $\mathbb{P}^n$

For a proof in full detail, see See [HS13], Theorem B.3.6. We give ideas to some of the steps below. The reader is encouraged to write down the details.

The statement explicitely defines $h_{V,\mathcal{L}}$ whenever $\mathcal{L}$ is very ample, and Theorem 4.6 ensures that it only depends on its divisor class, up to a bounded function. For an arbitrary divisor D, choose very ample divisors $\mathcal{M}_1$ and $\mathcal{M}_2$ such that $D = \mathcal{M}_1 - \mathcal{M}_2$, and set $h_{V,D} := h_{V,\mathcal{M}_1} - h_{V,\mathcal{M}_2}$. To see that changing $\mathcal{M}_1, \mathcal{M}_2$ adds a bounded function to the difference of heights, it suffices to check the additivity property for ample divisors

$$h_{V,\mathcal{L}_2} + h_{V,\mathcal{L}_1} = h_{V,\mathcal{L}_1+\mathcal{L}_2} + O(1),$$

which can be done by considering the Segre embedding and applying Theorem 4.6. Similarly, one shows that changing D for a linearly equivalent divisor, which we express as the difference of very ample divisors, only adds a bounded function.

Thus, the map of Weil's height machine is well defined, is a homomorphism. We omit the proof for positivity and algebraic equivalence.

Corollary 4.8

Let V be a variety defined over a number field k, and $D \in \mathrm{Div}(V)$ be ample. Let $B > 0$, and k' a finite extension of k. Then the set

$$\{P \in V(k') \mid h_{V,D}(P) \leq B\}$$

is finite, where $h_{V,D}$ is any choice of height function for the divisor class of D.

Proof. Let m be a positive integer such that mD is very ample. One has $h_{V,D} = \frac{1}{m} h_{V,mD} + O(1)$, so it suffices to prove the statement for D very ample. In that case, let $\varphi : V \to \mathbb{P}^n$ a corresponding embedding, thus

$$\forall P \in V(k'),\ h_{V,D}(P) = h_{\mathbb{P}^n}(\varphi(P)) + O(1).$$

Recall that Corollary 3.11 ensures that there are only finitely many k'-rational points of bounded height on $\mathbb{P}^n$, thus if $P \in V(k')$ has height bounded by B, $\varphi(P)$ belongs to a finite set. The injectivity of φ concludes the argument. □

Weil's height machine plays quite well with abelian varieties, due to the fact that relations between divisors turn into relations between heights. For an abelian variety A and an integer m, we denote by $[m]$ the multiplication-by-m map.

Corollary 4.9

Let k be a number field, and A/k be an abelian variety. Let $D \in \mathrm{Div}(A)$.

(a) Let m be an integer. Then

$$h_{A,D} \circ [m] = \frac{m^2+m}{2} h_{A,D} + \frac{m^2-m}{2} h_{A,D} \circ [-1] + O(1).$$

If the divisor class of D is symmetric (*i.e.* $[-1]^* D \sim D$), then

$$h_{A,D} \circ [m] = m^2 h_{A,D} + O(1).$$

If it is antisymmetric (*i.e.* $[-1]^* D \sim D$), then

$$h_{A,D} \circ [m] = m h_{A,D} + O(1).$$

(b) Let $P, Q \in A(\overline{\mathbb{Q}})$. If the divisor class of D is symmetric, then

$$h_{A,D}(P+Q) + h_{A,D}(P-Q) = 2h_{A,D}(P) + 2h_{A,D}(Q) + O(1).$$

(c) Let $P, Q \in A(\overline{\mathbb{Q}})$. If the divisor class of D is antisymmetric, then

$$h_{A,D}(P+Q) = h_{A,D}(P) + h_{A,D}(Q) + O(1).$$

Using the notations of the corollary, the statements follow from

- Mumford's formula (see [HS13], Corollary A.7.2.5)

$$[m]^* D \sim \frac{m^2+m}{2} D + \frac{m^2-m}{2} [-1]^* D$$

- if D has symmetric divisor class (see [HS13], Proposition A.7.3.3)

$$\sigma^* D + \delta^* D \sim 2\pi_1^* D + 2\pi_2^* D$$

- if D has antisymmetric divisor class (see [HS13], Proposition A.7.3.2)

$$\sigma^* D \sim \pi_1^* D + \pi_2^* D$$

where $\sigma, \delta, \pi_1, \pi_2 : A \times A \to A$ denote the maps

$$\sigma(P,Q) = P+Q, \quad \delta(P,Q) = P-Q, \quad \pi_1(P,Q) = P, \quad \pi_2(P,Q) = Q.$$

Néron and Tate have developped a theory of *canonical heights* which applies in particular in the cases $V = \mathbb{P}^n$ or $V = A$ an abelian variety. The theory states that under certain conditions, it is possible to find representatives for height functions in Weil's height machine such that most equalities up to bounded functions turn into strict equalities on the nose. Such representatives can be used to define real linear or quadratic forms on $V(\overline{\mathbb{Q}}) \otimes \mathbb{R}$. We refer the reader to [HS13], B.4, B.5.

References

[BG06] Enrico Bombieri and Walter Gubler. *Heights in Diophantine geometry*. 4. Cambridge university press, 2006.

[HS13] Marc Hindry and Joseph H Silverman. *Diophantine geometry: an introduction*. Vol. 201. Springer Science & Business Media, 2013.