



HAL
open science

Upper bounds on the heights of polynomials and rational fractions from their values

Jean Kieffer

► **To cite this version:**

Jean Kieffer. Upper bounds on the heights of polynomials and rational fractions from their values. 2021. hal-03226568v1

HAL Id: hal-03226568

<https://hal.science/hal-03226568v1>

Preprint submitted on 14 May 2021 (v1), last revised 9 Oct 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UPPER BOUNDS ON THE HEIGHTS OF POLYNOMIALS AND RATIONAL FRACTIONS FROM THEIR VALUES

JEAN KIEFFER

ABSTRACT. Let F be a univariate polynomial or rational fraction of degree d defined over a number field. We give bounds from above on the absolute logarithmic Weil height of F in terms of the heights of its values at small integers: we review well-known bounds obtained from interpolation algorithms given values at $d+1$ (resp. $2d+1$) points, and obtain tighter results when considering a larger number of evaluation points.

1. INTRODUCTION

Let F be a univariate rational fraction of degree d defined over \mathbb{Q} . The *height* of F , denoted by $h(F)$, measures the size of the coefficients of F . To define it, write $F = P/Q$ where $P, Q \in \mathbb{Z}[X]$ are coprime; then $h(F)$ is the maximum value of $\log |c|$, where c runs through the nonzero coefficients of P and Q . In particular, if $x = p/q$ is a rational number in irreducible form, then $h(x) = \log \max\{|p|, |q|\}$.

Heights can be generalized to arbitrary number fields, and are a basic tool in diophantine geometry [2, Part B]. They are also meaningful from an algorithmic point of view: the amount of memory needed to store F in a computer is in general $O(d h(F))$, and the cost of manipulating F grows with the size of its coefficients.

In this paper, we are interested in the relation between the height of F and the heights of evaluations $F(x)$, where x is an integer. One direction is easy: by [2, Prop. B.7.1], we have

$$(1) \quad h(F(x)) \leq d h(x) + h(F) + \log(d+1).$$

In the other direction, when we want to bound $h(F)$ from the heights of its values, matters are more complicated.

An easy case is when $F \in \mathbb{Z}[X]$ is a polynomial with integer coefficients of degree at most $d \geq 1$. Then, looking at the archimedean absolute value of the coefficients of F is sufficient to bound $h(F)$. Moreover, given height bounds on $d+1$ values of F , the Lagrange interpolation

Date: First version January 13, 2020; this version May 17, 2021.

formula allows us to bound $h(F)$ in a satisfactory way. For instance, assuming that

$$(2) \quad h(F(i)) \leq H \quad \text{for every } 0 \leq i \leq d,$$

we easily obtain

$$(3) \quad h(F) \leq H + d \log(2d) + \log(d + 1).$$

This result can be refined and adapted to other sets of interpolation points [1, Lem. 20], [5, Lem. 4.1]; in any case the bound on $h(F)$ is roughly H up to additional terms in $O(d \log d)$. This is consistent with inequality (1).

When F is a rational fraction or even a polynomial with rational coefficients, this result breaks down, and surprisingly little information appears in the literature despite the simplicity of the question.

1.1. Polynomials. Let us first consider the case where F is a polynomial in $\mathbb{Q}[X]$, of degree at most $d \geq 1$. Then F is determined by its values at $d + 1$ distinct points. Let x_1, \dots, x_{d+1} be distinct integers, let $H \geq 1$, and assume that $h(F(x_i)) \leq H$ for every i . This time, the Lagrange interpolation formula yields a bound on $h(F)$ which is roughly $O(dH)$ (see Proposition 3.2). This is intuitive enough: in general, computing F from its values $F(x_i)$ involves reducing the rational numbers $F(x_i)$ to the same denominator, thus multiplying the heights of the input by the number of evaluation points. But then, inequality (1) is very pessimistic at each of the evaluation points x_i : massive cancellations occur with the denominator of F , and the “expected” size of $F(x_i)$ is divided by d .

However, if we consider *more* than $d + 1$ evaluation points x_1, \dots, x_N such that $h(F(x_i)) \leq H$, we will likely find an evaluation point where inequality (1) is accurate, and hence obtain a bound on $h(F)$ of the form $O(H)$ rather than $O(dH)$. We prove the following result in this direction.

Proposition 1.1. *Let L be a number field, and let $\llbracket A, B \rrbracket$ be an interval in \mathbb{Z} . Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L[X]$ be a polynomial of degree at most $d \geq 1$, let $N \geq d + 1$, and let x_1, \dots, x_N be distinct elements of $\llbracket A, B \rrbracket$. Assume that $h(F(x_i)) \leq H$ for every $1 \leq i \leq N$. Then we have*

$$h(F) \leq \frac{N}{N-d} H + D \log(D) + d \log(2M) + \log(d + 1).$$

For instance, we obtain a bound on $h(F)$ which is linear in H when considering $N = 2d$ evaluation points.

1.2. Rational fractions. Second, consider the case where $F \in \mathbb{Q}(X)$ is a rational fraction of degree at most $d \geq 1$. Then F is determined by its values at $2d + 1$ points. If x_1, \dots, x_{2d+1} are distinct integers which are not poles of F , and if $h(F(x_i)) \leq H$ for every i , then a direct analysis of the interpolation algorithm yields a bound on $h(F)$ which is roughly $O(d^2 H)$ (see Proposition 5.2). As above, we can ask for a bound which is linear in H when more evaluation points are given.

In this case we could imagine cases where $F = P/Q$ has a very large height, but massive cancellations happen in all the quotients $P(x_i)/Q(x_i)$. This makes the result more intricate.

Proposition 1.2. *Let L be a number field, and let $\llbracket A, B \rrbracket$ be an interval in \mathbb{Z} . Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(X)$ be a univariate rational fraction of degree at most $d \geq 1$. Let S be a subset of $\llbracket A, B \rrbracket$ which contains no poles of F , let $\eta \geq 1$, and let $H \geq \max\{4, \log(2M)\}$. Assume that*

- (1) $h(F(x)) \leq H$ for every $x \in S$.
- (2) S contains at least D/η elements.
- (3) $D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.

Then we have

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d + 1),$$

where C_L is a constant depending only on L . We can take $C_{\mathbb{Q}} = 960$.

The number of evaluation points needed in this result is quite large, and depends on H . Still, Proposition 1.2 is strong enough to imply the following result.

Corollary 1.3. *Let $c \geq 1$, and let $F \in \mathbb{Q}(X)$ be a rational fraction of degree at most $d \geq 1$. Let $V \subset \mathbb{Z}$ be a finite set such that F has no poles in $\mathbb{Z} \setminus V$. Assume that for every $x \in \mathbb{Z} \setminus V$, we have*

$$h(F(x)) \leq c \max\{1, d \log d + d h(x)\}.$$

Then there exists a constant $C = C(c, \#V)$ such that

$$h(F) \leq C d \log(4d).$$

Explicitly, we can take $C = (4c + 1923)(12 + \log \max\{1, \#V\} + 2 \log(c))$.

It would be interesting to know whether we can obtain an efficient bound on $h(F)$ using only $O(d)$ evaluation points, as was the case for polynomials, instead of $O(d^3 H)$. The constants in Proposition 1.2 and Corollary 1.3 are not optimal; smaller constants can be obtained following the same proofs, at the cost of lengthier expressions.

The author has applied these results to obtain tight height bounds for modular equations on PEL Shimura varieties [3], for instance modular equations of Siegel and Hilbert type for abelian surfaces, generalizing existing works in the case of classical modular polynomials [5]. These modular equations are examples of rational fractions whose evaluations can be shown to have small height.

1.3. Organization of the paper. In Section 2, we recall the definition of heights over a number field that we use in the whole paper. In Section 3, we prove Proposition 1.1 about the heights of polynomials. To prepare for the proofs in the case of rational fractions, we study the relations between heights and norms of integers in number fields in Section 4. We prove height bounds for rational fractions using the minimal number of evaluation points in Section 5. Finally, Sections 6 and 7 are devoted to the proof of Proposition 1.2 and Corollary 1.3.

2. HEIGHTS OVER NUMBER FIELDS

Let L be a number field of degree d_L over \mathbb{Q} . Write \mathcal{V}_L^0 (resp. \mathcal{V}_L^∞) for the set of all nonarchimedean (resp. archimedean) places of L , and write $\mathcal{V}_L = \mathcal{V}_L^0 \sqcup \mathcal{V}_L^\infty$. Let $\mathcal{P}_\mathbb{Q}$ (resp. \mathcal{P}_L) be the set of primes in \mathbb{Z} (resp. prime ideals in \mathbb{Z}_L).

For each place v of L , the local degree of L/\mathbb{Q} at v is $d_v = [L_v : \mathbb{Q}_v]$, where subscripts denote completion. Denote by $|\cdot|_v$ the normalized absolute value associated with v : when $v \in \mathcal{V}_L^0$, and $p \in \mathcal{P}_\mathbb{Q}$ is the prime below v , we have $|p|_v = 1/p$.

The absolute logarithmic Weil height of projective tuples, affine tuples, polynomials and rational fractions over L is defined as follows [2, §B.2 and §B.7].

Definition 2.1. Let $n \geq 1$, and let $a_0, \dots, a_n \in L$.

- (1) The projective height of $(a_0 : \dots : a_n) \in \mathbb{P}_L^n$ is

$$h(a_0 : \dots : a_n) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log \left(\max_{0 \leq i \leq n} |a_i|_v \right).$$

- (2) The affine height of $(a_1, \dots, a_n) \in L^n$ is the projective height of the tuple $(1 : a_1 : \dots : a_n)$:

$$h(a_1, \dots, a_n) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log \left(\max \{ 1, \max_{1 \leq i \leq n} |a_i|_v \} \right).$$

In particular, for $a \in L$, we have

$$h(a) = h(1 : a) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log \left(\max \{ 1, |a|_v \} \right).$$

- (3) Let $P = \sum_{i=0}^n a_i X^i \in L[X]$. For every place $v \in \mathcal{V}_L$, we write

$$|P|_v = \max_i |a_i|_v.$$

The height of P is defined as the affine height of (a_0, \dots, a_n) . In other words

$$h(P) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log(\max\{1, |P|_v\}).$$

If $\mathfrak{p} \in \mathcal{P}_L$ is a prime ideal, we also define the \mathfrak{p} -adic valuation of P as

$$v_{\mathfrak{p}}(P) = \min_{0 \leq i \leq n} v_{\mathfrak{p}}(a_i).$$

- (4) Finally, if $F \in L(X)$ is a rational fraction, and $F = P/Q$ where $P, Q \in L[X]$ are coprime, we define $h(F)$ as the height of the projective tuple formed by all the coefficients of P and Q .

Elementary properties of heights.

- (1) The projective height of $(a_0 : \dots : a_n)$ is indeed invariant under multiplication by a common scalar factor, by the product formula [2, Lem. B.2.1(a)]. Therefore the height of a fraction is also well defined.
- (2) Heights are independent of the ambient field [2, Lem. B.2.1(c)]. This is another consequence of the product formula: in particular, we have

$$\sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} = 1.$$

- (3) If $x, y \in L$, then $h(xy) \leq h(x) + h(y)$; if $x \neq 0$, then $h(1/x) = h(x)$.
- (4) Heights satisfy the *Northcott property*: for every bound $H \in \mathbb{R}$, the number of projective tuples $(a_0 : \dots : a_n) \in \mathbb{P}_L^n$ such that $h(a_0 : \dots : a_n) \leq H$ is finite [2, Thm. B.2.3].
- (5) If $L = \mathbb{Q}$, then Definition 2.1 coincides with the naive definition of heights given in the introduction.

As Definition 2.1 suggests, in order to obtain height bounds for polynomials and rational fractions, we will try to bound their coefficients from above in the absolute values associated with all the places of L .

3. HEIGHTS OF POLYNOMIALS FROM THEIR VALUES

In this section, we estimate the height of a polynomial $F \in L[X]$ of degree at most $d \geq 1$ in terms of the heights of evaluations of F . We take our evaluation points to be integers in an interval $\llbracket A, B \rrbracket \subset \mathbb{Z}$,

and we write $D = B - A$ and $M = \max\{|A|, |B|\}$. Our tool is the Lagrange interpolation formula: if $x_1, \dots, x_{d+1} \in \llbracket A, B \rrbracket$ are distinct, then

$$(4) \quad F = \frac{1}{D!} \sum_{i=0}^{d+1} F(x_i) Q_i \quad \text{where } Q_i = D! \frac{\prod_{j \neq i} (X - x_j)}{\prod_{j \neq i} (x_i - x_j)} \in \mathbb{Z}[X].$$

Lemma 3.1. *In the notation of equality (4), we have $|Q_i| \leq D! (2M)^d$.*

Proof. Since the denominator $\prod_{j \neq i} (x_i - x_j)$ divides $D!$, we have

$$Q_i = N_i \prod_{j \neq i} (X - x_j)$$

for some $N_i \in \mathbb{Z}$ dividing $D!$. Therefore, for every $0 \leq k \leq d$, if c_k denotes the coefficient of X^{d-k} in Q_i , we have

$$|c_k| \leq |N_i| \binom{d}{k} M^k \leq D! 2^d M^d. \quad \square$$

A straightforward application of the Lagrange formula on $d+1$ evaluation points yields the following result.

Proposition 3.2. *Let $F \in L[X]$ be a univariate polynomial of degree at most $d \geq 1$, and let x_1, \dots, x_{d+1} be distinct integers in $\llbracket A, B \rrbracket$. Write $D = B - A$ and $M = \max\{A, B\}$. Assume that $h(F(x_i)) \leq H$ for every $1 \leq i \leq d+1$. Then we have*

$$h(F) \leq (d+1)H + D \log(D) + d \log(2M) + \log(d+1).$$

Proof. Let v be a place of L . If $v \in \mathcal{V}_L^0$, then the Lagrange interpolation formula (4) yields

$$\begin{aligned} \max\{1, |F|_v\} &\leq \left| \frac{1}{D!} \right|_v \max\{1, |F(x_1)|_v, \dots, |F(x_{d+1})|_v\} \\ &\leq \left| \frac{1}{D!} \right|_v \prod_{i=1}^{d+1} \max\{1, |F(x_i)|_v\}. \end{aligned}$$

If $v \in \mathcal{V}_L^\infty$, then Lemma 3.1 yields

$$\max\{1, |F|_v\} \leq \sum_{i=1}^{d+1} |F(x_i)|_v 2^d M^d \leq (d+1) 2^d M^d \prod_{i=1}^{d+1} \max\{1, |F(x_i)|_v\}.$$

Since $h(1/D!) = h(D!) \leq D \log(D)$, taking logarithms and summing gives the result. \square

It is interesting to compare Proposition 3.2 with [2, Cor. B.2.6], using the evaluation maps at x_i as linear maps from $L[X]$ to L : under the hypotheses of the proposition, the height of the *tuple* $(F(x_1), \dots, F(x_{d+1}))$ can be as large as $(d+1)H$.

Better upper bounds on $h(F)$ can be obtained given height bounds on more than $d+1$ values of F : this is the content of Proposition 1.1, which we recall here for convenience.

Proposition 3.3. *Let L be a number field, and let $\llbracket A, B \rrbracket$ be an interval in \mathbb{Z} . Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L[X]$ be a polynomial of degree at most $d \geq 1$, let $N \geq d+1$, and let x_1, \dots, x_N be distinct elements of $\llbracket A, B \rrbracket$. Assume that $h(F(x_i)) \leq H$ for every $1 \leq i \leq N$. Then we have*

$$h(F) \leq \frac{N}{N-d}H + D \log(D) + d \log(2M) + \log(d+1).$$

We need a lemma.

Lemma 3.4. *Keep the notation from Proposition 3.3, and let $v \in \mathcal{V}_L^0$ (resp. $v \in \mathcal{V}_L^\infty$). Then the number of elements $x \in \llbracket A, B \rrbracket$ satisfying the inequality*

$$|F(x)|_v < |D! F|_v \quad \left(\text{resp. } |F(x)|_v < \frac{|F|_v}{(2M)^d(d+1)} \right)$$

is at most d .

Proof of Lemma 3.4. We argue by contradiction. Let $(x_i)_{1 \leq i \leq d+1}$ be distinct elements of $\llbracket A, B \rrbracket$ satisfying the given inequality. If $v \in \mathcal{V}_L^0$, then the Lagrange interpolation formula (4) gives

$$|D! F|_v \leq \max_i |F(x_i)|_v < |D! F|_v$$

which is a contradiction. If $v \in \mathcal{V}_L^\infty$, the contradiction is

$$|F|_v \leq (2M)^d \sum_{i=1}^{d+1} |F(x_i)|_v < |F|_v$$

by Lemma 3.1. □

Proof of Proposition 3.3. Let v be a place of L . If $v \in \mathcal{V}_L^0$, then $|F(x_i)|_v \geq |D! F|_v$ for at least $N-d$ values of i by Lemma 3.4. Therefore

$$\prod_{i=1}^N \max\{1, |F(x_i)|_v\} \geq |D! F|_v^{N-d},$$

hence

$$\log \max\{1, |F|_v\} \leq \log \left| \frac{1}{D!} \right|_v + \frac{1}{N-d} \sum_{i=1}^N \log \max\{1, |F(x_i)|_v\}.$$

If $v \in \mathcal{V}_L^\infty$, then at least $N-d$ of the $F(x_i)$ satisfy the inequality $|F(x_i)|_v \geq |F|_v / (2M)^d (d+1)$, so

$$\log \max\{1, |F|_v\} \leq d \log(2M) + \log(d+1) + \frac{1}{N-d} \sum_{i=1}^N \log \max\{1, |F(x_i)|_v\}.$$

Since $h(1/D!) \leq D \log D$, summing all the contributions yields the result. \square

4. HEIGHTS AND NORMS OF INTEGERS

In this section, we study the relation between the height of integers in a number field L and their norms. Let us denote the norm of elements and fractional ideals in L by $N_{L/\mathbb{Q}}$.

Definition 4.1. Let $x \in L \setminus \{0\}$. Then we define

$$\tilde{h}(x) = \frac{1}{d_L} \log |N_{L/\mathbb{Q}}(x)| = \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log |x|_v.$$

If \mathfrak{a} is a fractional ideal in L , we also write

$$\tilde{h}(\mathfrak{a}) = \frac{1}{d_L} \log N_{L/\mathbb{Q}}(\mathfrak{a}).$$

If the reader is interested in the case $L = \mathbb{Q}$, then the remainder of this section can be safely skipped since \tilde{h} and h are equal on \mathbb{Z} . In general, they are not equal: for instance, \tilde{h} is invariant under multiplication by units. This is not the case of h as soon as L admits a fundamental unit, by the Northcott property.

Lemma 4.2. Let $x \in \mathbb{Z}_L \setminus \{0\}$. Then we have

$$0 \leq \tilde{h}(x) \leq h(x).$$

Equality holds on the right if and only if $|x|_v \geq 1$ for every $v \in \mathcal{V}_L^\infty$.

Proof. We have $N_{L/\mathbb{Q}}(c) \in \mathbb{Z} \setminus \{0\}$, hence $|N_{L/\mathbb{Q}}(c)| \geq 1$ and therefore $\tilde{h}(x) \geq 0$. The rest is obvious. \square

Proposition 4.3. There exists a constant C depending only on L such that for every $x \in \mathbb{Z}_L \setminus \{0\}$, there exists a unit $\varepsilon \in \mathbb{Z}_L^\times$ such that

$$h(\varepsilon x) \leq \max\{C, \tilde{h}(x)\}.$$

We can take $C = 2d_L \max_{i \in I} h(\varepsilon_i)$, where $(\varepsilon_i)_{i \in I}$ is any basis of units in \mathbb{Z}_L .

Proof. Let $m = \#\mathcal{V}_L^\infty$. In \mathbb{R}^m , we define the hyperplane H_s for $s \in \mathbb{R}$ as follows:

$$H_s = \{(t_1, \dots, t_m) \in \mathbb{R}^m : t_1 + \dots + t_m = s\}.$$

We also define the convex cone Δ_s as follows:

$$\Delta_s = \{(t_1, \dots, t_m) \in \mathbb{R}^m : \forall i, t_i \geq -s\}.$$

The image of \mathbb{Z}_L^\times under the logarithmic embedding

$$\text{Log} = \left(\frac{d_v}{d_L} \log |\cdot|_v \right)_{v \in \mathcal{V}_L^\infty}$$

is a full rank lattice Λ in H_0 . There exists a real number s_{\min} such that for every $s \geq s_{\min}$, the convex set $H_0 \cap \Delta_s$ contains a fundamental cell V of Λ ; we can take $s_{\min} = \max_{i \in I} h(\varepsilon_i)$ in the notation of the proposition. By translating in the direction $(1, \dots, 1)$, we also have:

- (1) For every $s \geq ms_{\min}$, the set $H_s \cap \Delta_0$ contains a translate of V .
- (2) For every $s \geq 0$, the set $H_s \cap \Delta_{s_{\min}}$ contains a translate of V .

Let $x \in \mathbb{Z}_L \setminus \{0\}$, and consider the point

$$\text{Log}(x) = \left(\frac{d_v}{d_L} \log |x|_v \right)_{v \in \mathcal{V}_L^\infty} \in \mathbb{R}^m.$$

The sum of its coordinates is $s_x = \tilde{h}(x)$. If $s_x \geq ms_{\min}$, then by (1) there exists a unit $\varepsilon \in \mathbb{Z}_L^\times$ such that $\text{Log}(x) + \text{Log}(\varepsilon)$ belongs to Δ_0 . Then $|\varepsilon x|_v \geq 1$ for every $v \in \mathcal{V}_L^\infty$, so

$$h(\varepsilon x) = \tilde{h}(\varepsilon x) = \tilde{h}(x)$$

by Lemma 4.2.

On the other hand, if $0 \leq s_x < ms_{\min}$, then by (2) we can still find a unit ε such that $\text{Log}(x) + \text{Log}(\varepsilon) \in \Delta_{s_{\min}}$, in other words

$$\frac{d_v}{d_L} \log |\varepsilon x|_v \geq -s_{\min}$$

for all $v \in \mathcal{V}_L^\infty$. Then

$$h(\varepsilon x) = \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log \max\{1, |\varepsilon x|_v\} \leq \tilde{h}(\varepsilon x) + \sum_{v \in \mathcal{V}_L^\infty} s_{\min} \leq 2ms_{\min}.$$

This proves the proposition with $C = 2ms_{\min} \leq 2d_L s_{\min}$. \square

Corollary 4.4. *Let C be as in Proposition 4.3. Then every principal ideal \mathfrak{a} of \mathbb{Z}_L admits a generator $a \in \mathbb{Z}_L$ such that*

$$h(a) \leq \max\{C, \tilde{h}(\mathfrak{a})\}.$$

Proof. Apply Proposition 4.3 with x an arbitrary generator of \mathfrak{a} . \square

This corollary allows us to bound the height of a common denominator of a given polynomial $P \in L[X]$.

Proposition 4.5. *There exists a constant C depending only on L such that for every $P \in L[X]$, there exists an element $a \in \mathbb{Z}_L$ such that $aP \in \mathbb{Z}_L[X]$ and $\max\{h(a), h(aP)\} \leq h(P) + C$.*

Proof. Let \mathfrak{C} be a set of ideals in \mathbb{Z}_L that are representatives for the class group of L , and let $P \in L[X]$, which we may assume to be nonzero. Let

$$\mathfrak{a} = \prod_{\mathfrak{p} \in \mathcal{P}_L} \mathfrak{p}^{\max\{0, -v_{\mathfrak{p}}(P)\}}$$

be the denominator ideal of P . Then

$$\tilde{h}(\mathfrak{a}) = \sum_{\mathfrak{p} \in \mathcal{P}_L} \frac{d_{\mathfrak{p}}}{d_L} \log \max\{1, |P|_{\mathfrak{p}}\} \leq h(P).$$

Let $\mathfrak{c} \in \mathfrak{C}$ be an ideal such that $\mathfrak{c}\mathfrak{a}$ is principal. By Corollary 4.4, if C denotes the constant from Proposition 4.3, we can find a generator a of $\mathfrak{c}\mathfrak{a}$ such that

$$h(a) \leq \max\{C, \tilde{h}(\mathfrak{c}\mathfrak{a})\} \leq h(P) + C'$$

where

$$C' = \max\{C, \max_{\mathfrak{c} \in \mathfrak{C}} \tilde{h}(\mathfrak{c})\}.$$

Then aP has integer coefficients, and we have

$$\begin{aligned} h(aP) &\leq \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} (\log \max\{1, |P|_v\} + \log \max\{1, |a|_v\}) \\ &= h(P) + h(a) - \sum_{v \in \mathcal{V}_L^0} \frac{d_v}{d_L} \log \max\{1, |P|_v\} \\ &= h(P) + h(a) - \tilde{h}(\mathfrak{a}) \\ &\leq h(P) + C'. \end{aligned} \quad \square$$

5. A NAIVE HEIGHT BOUND FOR FRACTIONS

Let L be a number field, and let $F \in L(X) \setminus \{0\}$ be a rational fraction of degree at most $d \geq 1$. Write $F = P/Q$ where P and Q are coprime polynomials in $L[X]$, and let d_P and d_Q be the degrees of P and Q respectively. Let x_i for $1 \leq i \leq d_P + d_Q + 1$ be distinct elements in an interval $\llbracket A, B \rrbracket \subset \mathbb{Z}$ that are not poles of F .

We recall the interpolation algorithm to reconstruct F given the pairs $(x_i, F(x_i))$ [7, §5.7]. Let $S \in L[X]$ be the polynomial of degree at most $d_P + d_Q$ interpolating the points $(x_i, F(x_i))$. Let a be a common denominator for the coefficients of S , so that $T = aS$ has coefficients in \mathbb{Z}_L . We compute the d_P -th subresultant of T and the polynomial

$$Z = \prod_{i=1}^{s+1} (X - x_i) \in \mathbb{Z}[X].$$

We obtain a Bézout relation of the form

$$UT + VZ = R$$

where $U, V, R \in \mathbb{Z}_L[X]$, and moreover $\deg(U) \leq d_Q$ and $\deg(R) \leq d_P$. Then $F = R/aU$.

In order to obtain a bound on $h(F)$, we first bound $h(S)$ using Proposition 3.2. Then, we use the following well-known fact about the size of subresultants in $\mathbb{Z}_L[X]$.

Lemma 5.1. *Let $P, Q \in \mathbb{Z}_L[X] \setminus \{0\}$ be polynomials of degrees d_P and d_Q respectively, and let $0 \leq k \leq \min\{d_P, d_Q\} - 1$. Let R be the k -th subresultant of P and Q , and let U and V be the associated Bézout coefficients. Write $s = d_P + d_Q$. Then we have*

$$\begin{aligned} h(R) &\leq (d_Q - k) h(P) + (d_P - k) h(Q) + \frac{s - 2k}{2} \log(s - 2k), \\ h(U) &\leq (d_Q - k - 1) h(P) + (d_P - k) h(Q) + \frac{s - 2k - 1}{2} \log(s - 2k - 1), \quad \text{and} \\ h(V) &\leq (d_Q - k) h(P) + (d_P - k - 1) h(Q) + \frac{s - 2k - 1}{2} \log(s - 2k - 1). \end{aligned}$$

For instance, Lemma 5.1 allows one to bound coefficient sizes in the subresultant version of the Euclidean algorithm in $\mathbb{Q}(X)$.

Proof. Let $v \in \mathcal{V}_L^\infty$. Every coefficient r of R has an expression as a determinant of size $d_P + d_Q - 2k$; its entries in the first $d_Q - k$ columns are coefficients of P , and its entries in the last $d_P - k$ columns are coefficients of Q . By Hadamard's lemma, we can bound $|r|_v$ by the product of L^2 -norms of the columns. Hence

$$|r|_v \leq \left(\sqrt{d_P + d_Q - 2k} |P|_v \right)^{d_Q - k} \left(\sqrt{d_P + d_Q - 2k} |Q|_v \right)^{d_P - k}.$$

Taking logarithms and summing over v , we obtain the desired height bound on R . Similarly, the coefficients of U (resp. V) are determinants of size $d_P + d_Q - 2k - 1$, with one column less containing coefficients of P (resp. Q). \square

Proposition 5.2. *Let L be a number field, and let $\llbracket A, B \rrbracket \subset \mathbb{Z}$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(X) \setminus \{0\}$ be a rational fraction of degree $d \geq 1$. Let d_P and d_Q be the degrees of its numerator and denominator respectively. Let x_i for $1 \leq i \leq d_P + d_Q + 1$ be distinct elements of $\llbracket A, B \rrbracket$ that are not poles of F , and assume that $h(F(x_i)) \leq H$ for every i . Then there exist polynomials $P, Q \in \mathbb{Z}_L[X]$ such that $P, Q \in \mathbb{Z}_L[Y]$ such that $F = P/Q$, $\deg P = d_P$, $\deg Q = d_Q$, and*

$$\begin{aligned} \max\{h(P), h(Q)\} &\leq (d+1)(2d+1)H + (d+1)D \log(D) + (4d^2 + 3d) \log(2M) \\ &\quad + (2d+2) \log(2d+1) + (d+1)C, \end{aligned}$$

where C is the constant from Proposition 4.5.

Proof. Let S, a, T, R, U , and V be as above; to choose a , we use Proposition 4.5, so that

$$\max\{h(a), h(T)\} \leq h(S) + C.$$

By Proposition 3.2, we have

$$(5) \quad h(S) \leq (2d+1)H + D \log(D) + 2d \log(2M) + \log(2d+1).$$

The coefficients of Z are bounded above by $(2M)^{2d+1}$, hence

$$h(Z) \leq (2d+1) \log(2M)$$

By Lemma 5.1, we have

$$\begin{aligned} h(R) &\leq (d+1)h(T) + d(2d+1) \log(2M) + \frac{2d+1}{2} \log(2d+1), \quad \text{and} \\ h(U) &\leq d h(T) + d(2d+1) \log(2M) + d \log(2d+1). \end{aligned}$$

Then $F = R/aU$, and

$$\begin{aligned} \max\{h(R), h(aU)\} &\leq \max\{h(R), h(a) + h(U)\} \\ &\leq (d+1)(h(S) + C) + d(2d+1) \log(2M) + \frac{2d+1}{2} \log(2d+1). \end{aligned}$$

Using the bound (5) on $h(S)$ ends the proof. \square

The bound obtained on $h(F)$ in Proposition 5.2 is roughly $O(d^2 H)$. This motivates a result like Proposition 1.2, where the dependency on H is only linear.

6. PREPARATIONS FOR THE PROOF OF PROPOSITION 1.2

In this section, we state preparatory lemmas for the proof of Proposition 1.2; the reader might wish to skip them until their use in the proof becomes apparent.

Lemma 6.1. *Let $\llbracket A, B \rrbracket \subset \mathbb{Z}$, let $D = B - A$, and let $\eta \geq 1$. Let S be a subset of $\llbracket A, B \rrbracket$ containing at least D/η elements, and let $1 \leq k \leq \frac{D}{2\eta}$ be an integer. Then there exists a subinterval of $\llbracket A, B \rrbracket$ of amplitude at most $\lceil 2\eta k \rceil$ containing at least $k + 1$ elements of S .*

Proof. Assume the contrary. We can partition $\llbracket A, B \rrbracket$ in at most $\lceil \frac{D}{2\eta k} \rceil$ intervals of amplitude at most $\lceil 2\eta k \rceil$, hence

$$\frac{D}{\eta} \leq \#S \leq k \left\lceil \frac{D}{2\eta k} \right\rceil < \frac{D}{2\eta} + k.$$

This is absurd because $k \leq \frac{D}{2\eta}$. □

Lemma 6.2. *Let $R \in \mathbb{Z}_L \setminus \{0\}$ be a non-unit. Then*

$$\sum_{\substack{\mathfrak{p} \in \mathcal{P}_L, \mathfrak{p} | R \\ \mathfrak{p} | p \in \mathcal{P}_\mathbb{Q}}} \frac{\log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \leq d_L(2 \log \log |N_{L/\mathbb{Q}}(R)| + 3.5).$$

Proof. First, we assume that $L = \mathbb{Q}$, so that $R \in \mathbb{Z}$ and $|R| \geq 2$. Let m be the number of prime factors in R , and let (p_i) be the sequence of prime numbers in increasing order. It is enough to prove the claim for the integer $R' = \prod_{i=1}^m p_i$, which has both a greater left hand side, since $\log(p)/(p-1)$ is a decreasing function of p , and a smaller right hand side, since $R' \leq R$. We can assume that $m \geq 2$. Then

$$\sum_{i=1}^m \frac{\log(p_i)}{p_i - 1} = \sum_{i=1}^m \frac{\log(p_i)}{p_i} + \sum_{i=1}^m \frac{\log(p_i)}{p_i(p_i - 1)} \leq \log(p_m) + 3$$

by Mertens's first theorem [4], and because the sum of the second series is less than 0.76. By [6], we have $p_m < m \log m + m \log \log m$ if $m \geq 6$; thus the rough bound $p_m \leq m^2$ holds. Since $m \leq \log(R')/\log(2)$, the result in the case $L = \mathbb{Q}$ follows.

In the general case, if $\mathfrak{p}|R$ lies above p , then p divides $N_{L/\mathbb{Q}}(R)$, and $|N_{L/\mathbb{Q}}(R)| \geq 2$. We apply Lemma 6.2 to $N_{L/\mathbb{Q}}(R) \in \mathbb{Z}$: hence

$$\begin{aligned} \sum_{\mathfrak{p}|R} \frac{\log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} &\leq \sum_{p|N_{L/\mathbb{Q}}(R)} \frac{\sum_{\mathfrak{p}|p} \log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \\ &= d_L \sum_{p|N_{L/\mathbb{Q}}(R)} \frac{\log(p)}{p-1} \leq d_L(2 \log \log |N_{L/\mathbb{Q}}(R)| + 3.5). \square \end{aligned}$$

Lemma 6.3. *Let $\mathfrak{p} \in \mathcal{P}_L$ be a prime ideal lying over $p \in \mathcal{P}_{\mathbb{Q}}$, and let $L_{\mathfrak{p}}$ be the \mathfrak{p} -adic completion of L . Let $Q \in L_{\mathfrak{p}}[X]$ be a polynomial of degree $d \geq 0$, and assume that $v_{\mathfrak{p}}(Q) = 0$. Let x_1, \dots, x_n be distinct values in $\llbracket A, B \rrbracket$, and write $D = B - A$; assume that $D \geq 1$. Let $\beta \in \mathbb{N}$. Then*

$$(6) \quad \sum_{i=1}^n \min\{\beta, v_{\mathfrak{p}}(Q(y_i))\} \leq d \left(\beta + d_L \frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})} + \frac{D}{p-1} \right).$$

Proof. We can assume that $d \geq 1$. Let λ be the leading coefficient of Q , and let $\alpha_1, \dots, \alpha_d$ be the roots of Q in an algebraic closure of $L_{\mathfrak{p}}$, where we extend $|\cdot|_{\mathfrak{p}}$ and $v_{\mathfrak{p}}$. Up to reindexation, we may assume that $|\alpha_j|_{\mathfrak{p}} \leq 1$ for $1 \leq j \leq t$, and $|\alpha_j|_{\mathfrak{p}} > 1$ for $t+1 \leq j \leq d$. For every i , we have

$$|Q(x_i)|_{\mathfrak{p}} = |\lambda|_{\mathfrak{p}} \prod_{i=1}^d |x_i - \alpha_j|_{\mathfrak{p}} = \left(|\lambda|_{\mathfrak{p}} \prod_{j=t+1}^d |\alpha_j|_{\mathfrak{p}} \right) \prod_{j=1}^t |x_i - \alpha_j|_{\mathfrak{p}}.$$

Note that

$$\left(|\lambda|_{\mathfrak{p}} \prod_{j=t+1}^d |\alpha_j|_{\mathfrak{p}} \right) \geq 1,$$

for otherwise all the coefficients of Q would belong to \mathfrak{p} . Therefore,

$$v_{\mathfrak{p}}(Q(x_i)) \leq \sum_{j=1}^t v_{\mathfrak{p}}(x_i - \alpha_j).$$

Let $k \in \mathbb{N}$ be such that $p^k \leq D < p^{k+1}$. Since the x_i are all distinct modulo p^{k+1} , there exist at most d values of i such that $v_{\mathfrak{p}}(x_i - \alpha_j) > k$ for some j . For these indices i , we bound $\min\{\beta, v_{\mathfrak{p}}(Q(x_i))\}$ from above by β . This accounts for the term $d\beta$ in inequality (6).

For all other values of i (say $i \in I$), we have $v_{\mathfrak{p}}(x_i - \alpha_j) \leq k$ for every $1 \leq j \leq t$, thus

$$v_{\mathfrak{p}}(x_i - \alpha_j) = \int_0^k \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} du.$$

Any two x_i that fall in the same disk $\{u \leq v_{\mathfrak{p}}(x - \alpha_j)\}$ coincide modulo $p^{\lceil u \rceil}$. Hence, for a given α_j , and a given $u \in]l, l+1]$, there exist at most $\lceil D/p^{l+1} \rceil$ values of i such that x_i belongs to this disk. Therefore,

$$\begin{aligned} \sum_{i \in I} v_{\mathfrak{p}}(Q(x_i)) &\leq \sum_{i \in I} \sum_{j=1}^t v_{\mathfrak{p}}(x_i - \alpha_j) \\ &= \sum_{i \in I} \sum_{j=1}^t \sum_{l=0}^{k-1} \int_l^{l+1} \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} du \\ &= \sum_{j=1}^t \sum_{l=0}^{k-1} \int_l^{l+1} \left(\sum_{i \in I} \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} \right) du \\ &\leq t \sum_{l=0}^{k-1} \left\lceil \frac{D}{p^{l+1}} \right\rceil \leq t \sum_{l=0}^{k-1} \left(\frac{D}{p^{l+1}} + 1 \right) \leq tk + \frac{tD}{p-1}. \end{aligned}$$

We have $t \leq d$, and

$$k \leq \frac{\log(D)}{\log(p)} \leq d_L \frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})}.$$

This accounts for the two remaining terms in inequality (6). \square

7. HEIGHTS OF FRACTIONS FROM THEIR VALUES

This final section is devoted to the proof of Proposition 1.2 and its corollary. We recall the main statement for the reader's convenience.

Proposition 7.1. *Let L be a number field, and let $\llbracket A, B \rrbracket$ be an interval in \mathbb{Z} . Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(X)$ be a univariate rational fraction of degree at most $d \geq 1$. Let S be a subset of $\llbracket A, B \rrbracket$ which contains no poles of F , let $\eta \geq 1$, and let $H \geq \max\{4, \log(2M)\}$. Assume that*

- (1) $h(F(x)) \leq H$ for every $x \in S$.
- (2) S contains at least D/η elements.
- (3) $D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.

Then we have

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d+1),$$

where C_L is a constant depending only on L . We can take $C_{\mathbb{Q}} = 960$.

Proof. We can assume that $F \neq 0$. We have $D \geq 4\eta d$, so by Lemma 6.1 with $k = 2d$, we can find a subinterval $\llbracket A', B' \rrbracket$ of $\llbracket A, B \rrbracket$ with amplitude at most $\lceil 4\eta d \rceil$ containing $2d+1$ elements of S , denoted by x_1, \dots, x_{2d+1} .

We use these x_i as evaluation points in Proposition 5.2: we can write $F = P/Q$ where $P, Q \in \mathbb{Z}_L[X]$ have the correct degrees and

$$\begin{aligned} \max\{h(P), h(Q)\} &\leq (d+1)(2d+1)H + 2d \lceil 4\eta d \rceil \log(\lceil 4\eta d \rceil) \\ &\quad + (4d^2 + 3d) \log(2M) + (2d+2) \log(2d+1) + (d+1)C_1 \\ &\leq (27 + C_1)\eta d^2 H, \end{aligned}$$

where C_1 is the constant from Proposition 4.3. To simplify the right hand side, we used the inequalities $1 \leq d$, $1 \leq \eta$, $\lceil 4\eta d \rceil \leq D \leq 2M$, $\lceil 4\eta d \rceil \leq 5\eta d$, and $\log(2M) \leq H$.

Let $x \in S$. We define ideals \mathfrak{s}_x , \mathfrak{n}_x and \mathfrak{d}_x of \mathbb{Z}_L as follows:

$$\mathfrak{s}_x = \gcd(P(x), Q(x)), \quad (P(x)) = \mathfrak{n}_x \mathfrak{s}_x, \quad (Q(x)) = \mathfrak{d}_x \mathfrak{s}_x.$$

Then $(F(x)) = \mathfrak{n}_x \mathfrak{d}_x^{-1}$. The ideal \mathfrak{s}_x encodes the simplifications that occur when evaluating P/Q at x . The heart of the proof is to show that \mathfrak{s}_x has small norm for at least some values of x . Let \mathfrak{r} be the greatest common divisor of all the coefficients of P and Q .

Claim 7.2. There exist at least $2dd_L + 1$ elements x of S such that

$$\tilde{h}(\mathfrak{s}_x) \leq \tilde{h}(\mathfrak{r}) + C\eta d \log(\eta d H)$$

for some constant C depending only on L .

Let us explain how to finish the proof if Claim 7.2 holds. By Lemma 3.4, we can find an $x \in S$ among these $2dd_L + 1$ values such that for every $v \in \mathcal{V}_L^\infty$, we have

$$|P(x)|_v \geq \frac{|P|_v}{(2M)^d(d+1)} \quad \text{and} \quad |Q(x)|_v \geq \frac{|Q|_v}{(2M)^d(d+1)}.$$

Then, we can calculate $h(F)$ as

$$\begin{aligned} &\sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log \max\{|P|_v, |Q|_v\} - \tilde{h}(\mathfrak{r}) \\ &\leq \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log \max\{|P(x)|_v, |Q(x)|_v\} - \tilde{h}(\mathfrak{r}) + d \log(2M) + \log(d+1) \\ &\leq \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log \max\{|P(x)|_v, |Q(x)|_v\} + \tilde{h}(\mathfrak{s}_x) - \tilde{h}(\mathfrak{r}) + d \log(2M) + \log(d+1) \\ &\leq H + C\eta d \log(\eta d H) + d \log(2M) + \log(d+1), \end{aligned}$$

as claimed.

In order to prove Claim 7.2, a crucial remark is that \mathfrak{s}_x divides the resultant R of P and Q . By Lemma 5.1, we have

$$h(R) \leq d h(P) + d h(Q) + d \log(2d) \leq (55 + 2C_1)\eta d^3 H.$$

Let $\mathfrak{p} \in \mathcal{P}_L$ be a prime factor of R with valuation $\beta_{\mathfrak{p}}$, and let I be a subset of S with n elements. We claim:

$$(7) \quad \sum_{x \in I} v_{\mathfrak{p}}(\mathfrak{s}_x) \leq n v_{\mathfrak{p}}(\mathfrak{r}) + d \left(\beta_{\mathfrak{p}} + d_L \frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})} + \frac{D}{p-1} \right).$$

To prove (7), we can work in the \mathfrak{p} -adic completion $L_{\mathfrak{p}}$ of L . Let π be a uniformizer of $L_{\mathfrak{p}}$, and let $r = \min\{v_{\mathfrak{p}}(P), v_{\mathfrak{p}}(Q)\}$ be the \mathfrak{p} -adic valuation of \mathfrak{r} . Write $P_1 = P/\pi^r$, $Q_1 = Q/\pi^r$. Then one of P_1 and Q_1 is not divisible by π ; for instance, assume that π does not divide Q_1 . Then, for every $x \in S$,

$$v_{\mathfrak{p}}(\mathfrak{s}_x) \leq \min\{\beta_{\mathfrak{p}}, v_{\mathfrak{p}}(Q(x))\} \leq v_{\mathfrak{p}}(\mathfrak{r}) + \min\{\beta_{\mathfrak{p}}, v_{\mathfrak{p}}(Q_1(x))\}.$$

Therefore inequality (7) follows from Lemma 6.3.

Summing over the prime factors \mathfrak{p} of R , we obtain an upper bound on the product of the norms of the ideals \mathfrak{s}_x , for $x \in I$. We can assume that R is not a unit, otherwise Claim 7.2 holds trivially. We have

$$\begin{aligned} \left| \prod_{x \in I} N_{L/\mathbb{Q}}(\mathfrak{s}_x) \right| &\leq N_{L/\mathbb{Q}}(\mathfrak{r})^n |N_{L/\mathbb{Q}}(R)|^d \\ &\quad \cdot \exp \left(\sum_{\substack{\mathfrak{p} \in \mathcal{P}_L, \mathfrak{p}|R \\ \mathfrak{p}|p \in \mathcal{P}_{\mathbb{Q}}}} \left(dd_L \log(D) + dD \frac{\log N_{L/\mathbb{Q}}(\mathfrak{p})}{p-1} \right) \right) \\ &\leq N_{L/\mathbb{Q}}(\mathfrak{r})^n |N_{L/\mathbb{Q}}(R)|^d \\ &\quad \cdot \exp \left(dd_L \log(D) \log |N_{L/\mathbb{Q}}(R)| / \log(2) \right. \\ &\quad \left. + dd_L D (\log \log |N_{L/\mathbb{Q}}(R)| + 3.5) \right). \end{aligned}$$

Indeed, R has at most $\log |N_{L/\mathbb{Q}}(R)| / \log(2)$ prime factors, and we can apply Lemma 6.2. Since $\tilde{h}(R) \leq (55 + 2C_1)\eta d^3 H$, we obtain

$$\begin{aligned} \sum_{x \in I} \tilde{h}(\mathfrak{s}_x) &\leq n \tilde{h}(\mathfrak{r}) + d \tilde{h}(R) + dd_L \frac{\log(D)}{\log(2)} \tilde{h}(R) \\ &\quad + dD (\log \log |N_{L/\mathbb{Q}}(R)| + 3.5) \\ &\leq n \tilde{h}(\mathfrak{r}) + C_2 (\eta d^4 H \log(D) + dD \log(\eta d H)) \end{aligned}$$

with

$$C_2 = \max \left\{ \frac{3d_L(55 + 2C_1)}{2 \log(2)}, 6.5 + \log(d_L) + \log(55 + 2C_1) \right\}.$$

Here we use that $\log(\eta d H) \geq 1$, and $\log(D) \geq 2 \log 2$.

Now we put into play our assumptions about D and S being sufficiently large. Since $D \geq \eta d^3 H \geq 4 > \exp(1)$, and the function $t/\log(t)$

is increasing for $t > \exp(1)$, we have

$$\frac{D}{\log(D)} \geq \frac{\eta d^3 H}{3 \log(\eta d H)}.$$

Moreover,

$$\#S - 2dd_L \geq \frac{D}{\eta} - \frac{D}{2\eta} = \frac{D}{2\eta}.$$

Therefore,

$$\begin{aligned} \sum_{x \in I} \tilde{h}(\mathfrak{s}_x) &\leq n \tilde{h}(\mathfrak{r}) + 4C_2 d D \log(\eta d H) \\ &\leq n \tilde{h}(\mathfrak{r}) + 8C_2 \eta d \log(\eta d H) (\#S - 2dd_L). \end{aligned}$$

This shows that in every subset of $\#S - 2dd_L$ elements of S , at least one satisfies the upper bound. $\tilde{h}(\mathfrak{s}_x) \leq \tilde{h}(\mathfrak{r}) + 8C_2 \eta d \log(\eta d H)$. Hence Claim 7.2 holds with $C = 8C_2$.

If $L = \mathbb{Q}$, we have $C_1 = 0$, so that we can take $C_2 = 120$. \square

To conclude, we give the proof of Corollary 1.3.

Corollary 7.3. *Let $c \geq 1$, and let $F \in \mathbb{Q}(X)$ be a rational fraction of degree at most $d \geq 1$. Let $V \subset \mathbb{Z}$ be a finite set such that F has no poles in $\mathbb{Z} \setminus V$. Assume that for every $x \in \mathbb{Z} \setminus V$, we have*

$$h(F(x)) \leq c \max\{1, d \log d + d h(x)\}.$$

Then there exists a constant $C = C(c, \#V)$ such that

$$h(F) \leq Cd \log(4d).$$

Explicitly, we can take $C = (4c + 1923)(12 + \log \max\{1, \#V\} + 2 \log(c))$.

Proof. We want to use Proposition 1.2 on an interval of the form $\llbracket 0, D \rrbracket$ for some integer $D \geq 4d$, with $\eta = 2$ and $S = \llbracket 0, D \rrbracket \setminus V$. The set S contains at least D/η elements as soon as $D \geq 2\#V$.

For every $x \in S$, we have $h(x) \leq \log(D)$, hence

$$h(F(x)) \leq c \max\{1, d \log d + d \log D\}.$$

Hence, if we let

$$H(D) = \max\{4, \log(2D), c(d \log d + d \log D)\}$$

we can apply Proposition 1.2 with $H = H(D)$ as soon as the condition

$$D \geq 2d^3 H(D)$$

holds. We let the reader check that we can choose

$$D = \max\{2\#V, \lceil 4cd^4 \log(4cd^4) \rceil\}.$$

Then, Proposition 1.2 yields

$$h(F) \leq H(D) + 1920d \log(2dH(D)) + d \log(2D) + \log(d+1).$$

We have $H(D) \leq 4cd \log(dD)$ and $2dH(D) \leq D$, hence

$$\begin{aligned} h(F) &\leq 4cd \log(dD) + 1920d \log(D) + d \log(2D) + \log(d+1) \\ &\leq (4c + 1923)d \log(dD) \\ &\leq (4c + 1923)d(\log(2d \max\{1, \#V\}) + \log(5cd^5 \log(4cd^4))) \end{aligned}$$

To simplify this expression further, we write

$$\log(5cd^5 \log(4cd^4)) \leq \log(20c^2d^9) \leq 3 + 2 \log(c) + 9 \log(d).$$

hence, after other simplifications,

$$h(F) \leq Cd \log(4d)$$

with

$$C = (4c + 1923)(12 + \log \max\{1, \#V\} + 2 \log(c)),$$

as claimed. \square

REFERENCES

- [1] R. Bröker and A. V. Sutherland. An explicit height bound for the classical modular polynomial. *Ramanujan J.*, 22(3):293–313, 2010.
- [2] M. Hindry and J. H. Silverman. *Diophantine geometry*. Springer, 2000.
- [3] J. Kieffer. Degree and height estimates for modular equations on PEL Shimura varieties. 2020.
- [4] F. Mertens. Ein Beitrag zur analytischen Zahlentheorie. *J. Reine Angew. Math.*, 78:46–62, 1874.
- [5] F. Pazuki. Modular invariants and isogenies. *Int. J. Number Theory*, 15(3):569–584, 2019.
- [6] B. Rosser. Explicit bounds for some functions of prime numbers. *Am. J. Math.*, 63(1):211–232, 1941.
- [7] J. von zur Gathen and J. Gerhard. *Modern computer algebra*. Cambridge University Press, third edition, 2013.