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. Acta Arithmetica, 2022, 203 (1), pp.49-68. 10.4064/aa210816-26-1 . hal-03226568v3

HAL Id: hal-03226568

https://hal.science/hal-03226568v3

Submitted on 9 Oct 2022

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 [5, 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 [5, Prop. B.7.1], we have

$$h(F(x)) \le d\,h(x) + h(F) + \log(d+1). \tag{1}$$

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 $d \ge 1$. 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 formula allows us

2020 *Mathematics Subject Classification.* 11C08, 11G50.
Key words and phrases. Heights, polynomials, rational fractions.

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

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

we easily obtain

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

This result can be refined and adapted to other sets of intepolation points [2, Lem. 20], [9, 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 $x_1, \ldots, x_{d+1}$ 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 height of $F(x_i)$ is just a fraction $1/d$ of the expected value.

However, if we consider *more* than $d+1$ evaluation points $x_1 \ldots, x_N$ 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.

Theorem 1.1. *Let L be a number field, and let $[\![A, B]\!]$ 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 $x_1, \ldots, x_N$ be distinct elements of $[\![A, B]\!]$. 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. See also Theorem 3.4 for local versions of this result.

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 $x_1, \ldots, x_{2d+1}$ 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 many quotients $P(x_i)/Q(x_i)$. This makes the result more intricate.

Theorem 1.2. *Let L be a number field of degree d_L over $\mathbb{Q}$ and discriminant Δ_L. Let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$, and 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 $[\![A, B]\!]$ 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 d_L and Δ_L. We can take $C_{\mathbb{Q}} = 960$.

We can give a general explicit expression for the constant C_L in terms of d_L and Δ_L (see §7). The number of evaluation points needed in this result is quite large, and depends on H. Still, Theorem 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 $\mathbb{Z} \backslash V$. Assume that for every $x \in \mathbb{Z} \backslash V$, 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 Theorem 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 asymptotic height bounds for modular equations on PEL Shimura varieties [6], for instance modular equations of Siegel and Hilbert type for abelian surfaces, generalizing existing works in the case of classical modular polynomials [9]. These modular equations are examples of rational fractions whose evaluations can be shown to have small height.

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 Theorem 1.1 about the heights of polynomials. To prepare for 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 Theorem 1.2.

Acknowledgements. Thanks are due to the anonymous referee for pointing out several errors in an earlier version of this paper. This work is part of the author's PhD dissertation at the University of Bordeaux (France), and he warmly thanks Damien Robert and Aurel Page for their advice and encouragement.

2. Heights over number fields

Let L be a number field of degree d_L over $\mathbb{Q}$. Write $\mathcal{V}_L^0$ (resp. $\mathcal{V}_L^\infty$) 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 the ring of integers $\mathbb{Z}_L$ of 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 $|p|_v = 1/p$. When v is archimedean, $|\cdot|_v$ is the usual real or complex absolute value.

The absolute logarithmic Weil height of projective tuples, affine tuples, polynomials and rational fractions over L is defined as follows [5, §B.2 and §B.7].

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

(1) If the a_i are not all zero, the projective height of $(a_0 : \cdots : a_n) \in \mathbb{P}^n_L$ is

$$h_{\mathrm{proj}}(a_0 : \ldots : a_n) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log\Big(\max_{0 \le i \le n} |a_i|_v\Big).$$

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

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

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

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

(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 $(a_0, \ldots, a_n)$. In other words

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

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

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

(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.

If $L = \mathbb{Q}$, then Definition 2.1 coincides with the naive definition of heights given in the introduction. By the product formula, heights are independent of the ambient field [5, Lem. B.2.1(c)]. Recall that

$$\sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} = 1, \tag{2}$$

a fact we will use many times when computing archimedean parts of heights. Moreover, if $x, y, z \in L$ with $z \neq 0$, then we have

$$h(xy) \le h(x) + h(y) \quad \text{and} \quad h(1/z) = h(z). \tag{3}$$

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 choose our evaluation points to be integers in an interval $[\![A, B]\!] \subset \mathbb{Z}$, and we write $D = B - A$ and $M = \max\{|A|, |B|\}$ (here $|\cdot| = |\cdot|_\infty$ is the archimedean absolute value). Our tool is the Lagrange interpolation formula: if $x_1, \ldots, x_{d+1} \in [\![A, B]\!]$ are distinct, then

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

Lemma 3.1. *In the notation of equality* (4)*, we have* $|Q_i|_\infty \leq D!\,(2M)^d$ *for all* $1 \leq i \leq d+1$.

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

$$|c_k|_\infty \leq |N_i|_\infty \binom{d}{k} M^k \leq D!\, 2^d M^d. \qquad \square$$

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* $x_1, \ldots, x_{d+1}$ *be distinct integers in* $[\![A, B]\!]$*. Write* $D = B - A$ *and* $M = \max\{|A|, |B|\}$.

(1) *For every* $v \in \mathcal{V}_L^0$*, we have*

$$|F|_v \leq \left|\frac{1}{D!}\right|_v \max\{|F(x_1)|_v, \ldots, |F(x_{d+1})|_v\},$$

and for every $v \in \mathcal{V}_L^\infty$*, we have*

$$|F|_v \leq (d+1)(2M)^d \max\{|F(x_1)|_v, \ldots, |F(x_{d+1})|_v\}.$$

(2) *Assume that* $h(F(x_i)) \leq H$ *for every* $1 \leq i \leq d+1$*. Then*

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

Proof. Part 1 is an immediate consequence of the interpolation formula (4), and Lemma 3.1 for archimedean places. For part 2, let v be a place of L. By part 1, we have

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

where $C_v = |1/D!|_v$ if v is nonarchimedean, and $C_v = (d+1)(2M)^d$ if v is archimedean. Taking logarithms and summing, we obtain

$$h(F) \leq h(1/D!) + \big(d\log(2M) + \log(d+1)\big) \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} + \sum_{i=1}^{d+1} h(F(x_i)).$$

By eq. (3), we have $h(1/D!) = h(D!) = \log(D!) \leq D\log(D)$. The result follows then from eq. (2). □

It is interesting to compare Proposition 3.2 with [5, 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* $(F(x_1), \ldots, F(x_{d+1}))$ can be as large as $(d+1)H$.

Remark 3.3. The result of Proposition 3.2 takes a particularly nice form because the evaluation points x_i are integers taken in a fixed interval. If we only assume the x_i to be distinct algebraic integers of bounded height, then providing an upper bound on the height of a common multiple of all products of the form $\prod_{j \neq i}(x_i - x_j)$ seems more complicated. A similar issue arises when the x_i are only assumed to be distinct points in $\mathbb{Q}$ of bounded height. However, if the evaluation points x_i are chosen to be rational numbers with the same denominator, then one can still apply Proposition 3.2 to a rescaled polynomial. In the rest of this paper, we will continue to consider (almost) consecutive integers as evaluation points.

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 Theorem 1.1, which we recall here with additional local statements.

Theorem 3.4. *Let L be a number field, and let $[\![A, B]\!]$ 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 $x_1, \ldots, x_N$ be distinct elements of $[\![A, B]\!]$. Assume that $h(F(x_i)) \leq H$ for every $1 \leq i \leq N$. Then we have*

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

More precisely, for every $v \in \mathcal{V}_L$, we have

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

where $C_v = \log|1/D!|_v$ if $v \in \mathcal{V}_L^0$, and $C_v = d\log(2M) + \log(d+1)$ if $v \in \mathcal{V}_L^\infty$.

We will need the following lemma.

Lemma 3.5. *Keep the notation from Theorem 3.4, and let $v \in \mathcal{V}_L^0$ (resp. $v \in \mathcal{V}_L^\infty$). Then the number of elements $x \in [\![A, B]\!]$ satisfying the inequality*

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

is at most d.

Proof of Lemma 3.5. We argue by contradiction, using part 1 of Proposition 3.2. □

Proof of Theorem 3.4. It is enough to prove the local statements: after that, the global statement results from summing all the local contributions. Let v be a place of L. If $v \in \mathcal{V}_L^0$, then by Lemma 3.5, we have $|F(x_i)|_v \geq |D!\, F|_v$ for at least $N - d$ values of i. Therefore,

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

and

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

Similarly, 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

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

□

4. Heights and norms of integers

Let L be a number field, let $\mathbb{Z}_L$ be its ring of integers, and let Δ_L be its discriminant. In this section, we study the relation between the height of elements of $\mathbb{Z}_L$ and their norms. We denote the norm of elements and fractional ideals in L by $N_{L/\mathbb{Q}}$.

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

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

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

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

If the reader is interested in the case $L = \mathbb{Q}$, then the remainder of this section can be safely skipped since $\widetilde{h}$ and h are equal on $\mathbb{Z}$. In general, they are not equal: for instance, $\widetilde{h}$ is invariant under multiplication by units. This is not the case for h as soon as L admits a fundamental unit, by the Northcott property [5, Thm. B.2.3].

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

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

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

Proof. We have $N_{L/\mathbb{Q}}(c) \in \mathbb{Z}\backslash\{0\}$, hence $\left|N_{L/\mathbb{Q}}(c)\right| \geq 1$ and $\widetilde{h}(x) \geq 0$. The rest is obvious. □

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

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

We can take $C = d_L \sum_{i \in I} h(\varepsilon_i)$, 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:

$$H_s = \{(t_1, \ldots, t_m) \in \mathbb{R}^m \colon t_1 + \cdots + t_m = s\}.$$

We also define the convex cone Δ_s as follows:

$$\Delta_s = \left\{(t_1, \ldots, t_m) \in \mathbb{R}^m \colon \forall i,\ t_i \geq -s\right\}.$$

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

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

is a full rank lattice Λ in H_0. Let $(\varepsilon_i)_{1 \leq i \leq m-1}$ be a basis of units in $\mathbb{Z}_L$, and let V be the following fundamental cell of Λ:

$$V = \left\{\sum_{i=1}^{m-1} \lambda_i \,\mathrm{Log}(\varepsilon_i) \colon \lambda_i \in [-\tfrac{1}{2}, \tfrac{1}{2}] \text{ for all } i\right\}.$$

For each $v \in \mathcal{V}_L^\infty$ and each $1 \leq i \leq m-1$, we have

$$\frac{d_v}{d_L} \log |\varepsilon_i|_v \geq -\frac{d_v}{d_L} \log \max\{1, |1/\varepsilon_i|_v\} \geq -\,h(1/\varepsilon_i) = -h(\varepsilon_i).$$

Therefore V is included in $H_0 \cap \Delta_s$ for every $s \geq s_{\min} = \frac{1}{2} \sum_{i=1}^{m-1} h(\varepsilon_i)$. From this, we deduce:

(1) For every $s \geq m s_{\min}$, the set $H_s \cap \Delta_0$ contains a translate of V; indeed its translate by $-s/m \cdot (1, \ldots, 1)$ is $H_0 \cap \Delta_{s/m}$.

(2) For every $s \geq 0$, the set $H_s \cap \Delta_{s_{\min}}$ contains a translate of V; indeed its translate by $-s/m \cdot (1, \ldots, 1)$ is $H_0 \cap \Delta_{s_{\min}+s/m}$.

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

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

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

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

by Lemma 4.2.

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

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

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

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

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

Remark 4.4. We can give an explicit upper bound for an acceptable constant C in Proposition 4.3 in terms of the degree and discriminant of L only. Let $\mathfrak{R}_L$ be the regulator of L. By [3, Lem. 1], L admits a basis of units $(\varepsilon_i)_{1 \leq i \leq m-1}$ (where $m = \#\mathcal{V}_L^\infty$) such that

$$h(\varepsilon_i) \leq \frac{((m-1)!)^2}{2^{m-1} d_L^{m-1}} \Big(\frac{\delta(L)}{d_L}\Big)^{2-m} \mathfrak{R}_L$$

for each $1 \leq i \leq m-1$; here $\delta(L) > 0$ satisfies the property that all non-roots of unity in L have height at least $\delta(L)/d_L$. It is known that we can take $\delta(L) = \log(2)/d_L$ if $d_L \leq 2$, and

$$\delta(d_L) = \max\Big\{\frac{1}{53 d_L \log(6 d_L)}, \frac{1}{4}\Big(\frac{\log\log d_L}{\log d_L}\Big)^3\Big\}$$

otherwise [3, §3]. (Lehmer's conjecture asserts that $\delta(L)$ can be chosen uniformly for all number fields L). Moreover, the regulator of L is bounded above in terms of d_L and Δ_L. To see this, we use the main theorem of [11] and we note that

(1) the class number of $\mathbb{Z}_L$ is at least one,
(2) L contains at most $d_L\big(2 + \log(d_L)/\log(2)\big)$ roots of unity.

Therefore

$$\mathfrak{R}_L < d_L\Big(2+\frac{\log(d_L)}{\log(2)}\Big)\Big(\frac{4}{d_L-1}\Big)^{d_L-1}|\Delta_L|^{1/2}(\log|\Delta_L|)^{d_L-1}.$$

The final upper bound we obtain for the constant C in Proposition 4.3 grows at least linearly in $|\Delta_L|^{1/2}$ and exponentially in d_L.

Corollary 4.5. *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*

$$h(a) \le \max\{C, \widetilde{h}(\mathfrak{a})\}.$$

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

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

Proposition 4.6. *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 $\max\{h(a), h(aP)\} \le h(P) + C'$. We can take*

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

where $\mathfrak{C}$ is a set of ideals in $\mathbb{Z}_L$ that are representatives for the class group of L, and C is the constant from Proposition 4.3.

Proof. Let $\mathfrak{C}$ and C be as above, 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

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

Let $\mathfrak{c} \in \mathfrak{C}$ be an ideal such that $\mathfrak{c}\mathfrak{a}$ is principal. By Corollary 4.5, if C denotes the constant from Proposition 4.3, we can find a generator a of $\mathfrak{c}\mathfrak{a}$ such that

$$h(a) \le \max\{C, \widetilde{h}(\mathfrak{c}\mathfrak{a})\} \le \widetilde{h}(\mathfrak{a}) + C' \le h(P) + C'.$$

Then aP has integer coefficients, and we have

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

□

Remark 4.7. Minkowski's bound [7, §V.4] implies that we can always choose $\mathfrak{C}$ in such a way that

$$\max_{\mathfrak{c}\in\mathfrak{C}} N_{L/\mathbb{Q}}(\mathfrak{c}) \leq |\Delta_L|^{1/2} \left(\frac{4}{\pi}\right)^{d_L/2} \frac{d_L!}{d_L^{d_L}}.$$

Combined with Remark 4.4, this gives an upper bound on an acceptable C' in Proposition 4.6 depending only on d_L and Δ_L. Under the generalized Riemann hypothesis, a much sharper upper bound is available: we can choose $\mathfrak{C}$ in such a way that

$$\max_{\mathfrak{c}\in\mathfrak{C}} N_{L/\mathbb{Q}}(\mathfrak{c}) \leq 12 \log(|\Delta_L|)^2$$

by [1, Thm. 3].

5. A naive height bound for fractions

Let L be a number field, and let $F \in L(X)\backslash\{0\}$ 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 $[\![A, B]\!] \subset \mathbb{Z}$ that are not poles of F.

We recall the interpolation algorithm to reconstruct F given the pairs $(x_i, F(x_i))$ [12, §5.7]. Define $S \in L[X]$ as the polynomial of degree at most $d_P + d_Q$ interpolating the points $(x_i, F(x_i))$. Let $a \in \mathbb{Z}_L$ 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 [4, §3] of T and the polynomial

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

which is a polynomial $R \in \mathbb{Z}_L[X]$ of degree at most d_P; the usual resultant is the 0-th subresultant. We obtain a Bézout relation [4, §3.2] 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 $P, Q \in \mathbb{Z}_L[X]\backslash\{0\}$ 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

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

For instance, Lemma 5.1 allows one to bound coefficient sizes in the subresultant version of the Euclidean algorithm in $\mathbb{Q}(X)$ [12, §6.11].

Proof. Let $v \in \mathcal{V}_L^\infty$. By definition, 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 [12, Thm. 16.6], we can bound $|r|_v$ by the product of L^2-norms of the columns of this determinant in the absolute value v. Hence

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

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 $d_P + d_Q - 2k - 1$, where one column less contains coefficients of P (resp. Q). □

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

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

where C is the constant from Proposition 4.6.

Proof. Let S, a, T, Z, R, U, and V be as above; to choose a, we use Proposition 4.6, so that

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

By Proposition 3.2, we have

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

The archimedian absolute values of 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

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

Using the upper bound (5) on $h(S)$ ends the proof. □

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

6. Preparations for the proof of Theorem 1.2

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

We keep the notation introduced at the beginning of §2, to which we add the following. If $\mathfrak{p} \in \mathcal{P}_L$, we denote by $v_\mathfrak{p}$ the $\mathfrak{p}$-adic valuation on L, with the convention that $v_\mathfrak{p}(0) = +\infty$. When considering $\mathfrak{p}$ as a finite place of L, we write $|\cdot|_\mathfrak{p}$ for the associated absolute value. We denote by $d_\mathfrak{p}$ and $e_\mathfrak{p}$ the local degree and ramification index of $\mathfrak{p}$ in the extension $L/\mathbb{Q}$. With our normalizations, the following formula holds for every $x \in L$ and $\mathfrak{p} \in \mathcal{P}_L$:

$$|x|_\mathfrak{p} = N_{L/\mathbb{Q}}(\mathfrak{p})^{-v_\mathfrak{p}(x)/d_\mathfrak{p}}.$$

Finally, for $r \in \mathbb{R}$, we denote the upper integral part of r by $\lceil r \rceil$.

Lemma 6.1. *Let $[\![A, B]\!] \subset \mathbb{Z}$, let $D = B - A$, and let $\eta \geq 1$; assume that $D \geq 2\eta$. Let S be a subset of $[\![A, B]\!]$ 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 $[\![A, B]\!]$ of length at most $\lceil 2\eta k \rceil$ containing at least $k+1$ elements of S.*

Proof. Let $m \in \mathbb{Z}$ such that $m \geq 1$. Then for each $n \geq 1$, the following intervals of $\mathbb{Z}$:

$$[\![0, m]\!]\,,\ [\![m+1, 2m+1]\!]\,,\ \ldots,\ [\![(n-1)(m+1), n(m+1)-1]\!]$$

form a partition of $[\![0, n(m+1)-1]\!]$ in n intervals of length m. Taking $m = \lceil 2\eta k \rceil$ and $n = \lceil D/(2\eta k) \rceil$, the right endpoint of the latter interval is at least D. Therefore, by translating the above partition and intersecting it with $[\![A, B]\!]$, we obtain a partition of $[\![A, B]\!]$ in at most $\lceil D/(2\eta k) \rceil$ intervals of length at most $\lceil 2\eta k \rceil$. In the case that each of these intervals contains at most k elements of S, we deduce that

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

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

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

$$\sum_{\substack{\mathfrak{p} \in \mathcal{P}_L,\, \mathfrak{p} | R \\ \mathfrak{p} | p \in \mathcal{P}_{\mathbb{Q}}}} \frac{e_{\mathfrak{p}} \log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \leq d_L(2 \log\log \left| N_{L/\mathbb{Q}}(R) \right| + 4).$$

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 [8], and because the sum of the second series is less than 0.76. By [10], 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 $\left| N_{L/\mathbb{Q}}(R) \right| \geq 2$. We apply Lemma 6.2 to $N_{L/\mathbb{Q}}(R) \in \mathbb{Z}$: hence

$$\begin{aligned}
\sum_{\mathfrak{p}|R} \frac{e_{\mathfrak{p}} \log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} &\leq \sum_{p | N_{L/\mathbb{Q}}(R)} \frac{\sum_{\mathfrak{p}|p} e_{\mathfrak{p}} \log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \\
&= d_L \sum_{p | N_{L/\mathbb{Q}}(R)} \frac{\log(p)}{p-1} \\
&\leq d_L(2 \log\log \left| N_{L/\mathbb{Q}}(R) \right| + 4).
\end{aligned}$$

□

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 $x_1, \ldots, x_n$ be distinct values in $[\![A, B]\!]$, and write $D = B - A$; assume that $D \geq 1$. Let $\beta \in \mathbb{N}$. Then*

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

Proof. We can assume that $d \geq 1$. Let λ be the leading coefficient of Q, and let $\alpha_1, \ldots, \alpha_d$ 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

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

Since $v_{\mathfrak{p}}(Q) = 0$, we have

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

Therefore, for each $1 \leq i \leq n$,

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

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 $v_{\mathfrak{p}}(x_i - \alpha_j) > ke_{\mathfrak{p}}$ 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 $v_{\mathfrak{p}}(x_i - \alpha_j) \leq ke_{\mathfrak{p}}$ for every $1 \leq j \leq t$. For each $1 \leq w \leq ke_{\mathfrak{p}}$ and $1 \leq j \leq t$, define

$$S_{j,w} = \{i \in I \colon v_{\mathfrak{p}}(x_i - \alpha_j) \geq w\}.$$

For fixed j and w, all the values x_i for $i \in S_{j,w}$ coincide modulo $p^{\lceil w/e_{\mathfrak{p}} \rceil}$, so

$$\#S_{j,w} \leq \left\lceil \frac{D}{p^{\lceil w/e_{\mathfrak{p}} \rceil}} \right\rceil.$$

Note that for all $i \in I$ and $1 \leq j \leq t$, the number of values of $w \in [\![1, ke_{\mathfrak{p}}]\!]$ such that $i \in S_{j,w}$ is precisely $v_{\mathfrak{p}}(x_i - \alpha_j)$. Therefore,

$$\begin{aligned}\sum_{i\in I} v_{\mathfrak{p}}(Q(x_i)) &= \sum_{i\in I}\sum_{j=1}^{t} v_{\mathfrak{p}}(x_i - \alpha_j)\\ &= \sum_{j=1}^{t}\sum_{w=1}^{ke_{\mathfrak{p}}} \#S_{j,w}\\ &\leq d\sum_{w=1}^{ke_{\mathfrak{p}}}\Big(\frac{D}{p^{\lceil w/e_{\mathfrak{p}}\rceil}}+1\Big)\\ &= de_{\mathfrak{p}}\sum_{w=1}^{k}\Big(\frac{D}{p^{w}}+1\Big)\\ &\leq de_{\mathfrak{p}}k + \frac{de_{\mathfrak{p}}D}{p-1}.\end{aligned}$$

Since

$$k \leq \frac{\log(D)}{\log(p)} = \frac{d_{\mathfrak{p}}}{e_{\mathfrak{p}}}\cdot\frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})},$$

this accounts for the two remaining terms in inequality (6). □

7. Heights of fractions from their values

This final section is devoted to the proof of Theorem 1.2 and its corollary. We keep the notation from §6, and recall the main statement for the reader's convenience.

Theorem 7.1. *Let L be a number field of degree d_L over $\mathbb{Q}$ and discriminant Δ_L. Let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$, and 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 $[\![A, B]\!]$ 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 d_L and Δ_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 of $[\![A, B]\!]$ of length at most $\lceil 4\eta d\rceil$ containing $2d+1$ elements of S, denoted by $x_1, \ldots, x_{2d+1}$. We use these x_i

as evaluation points to apply Proposition 5.2: we can write $F = P/Q$ where $P, Q \in \mathbb{Z}_L[X]$ are coprime in $L[X]$ and satisfy

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

where C_1 is the constant from Proposition 4.3. To simplify the right hand side, we use 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:

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

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

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

for some constant C depending only on L.

Let us explain how to finish the proof assuming that Claim 7.2 holds. By Lemma 3.5, 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, by Definition 2.1, we have

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

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

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

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:

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

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 $x \in S$,

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

Therefore inequality (7) follows from Lemma 6.3.

Inequality (7) gives an upper bound on the $\mathfrak{p}$-adic valuation of the ideal $\prod_{x\in I}\mathfrak{s}_x$. Taking the product over the prime factors $\mathfrak{p}$ of R, we obtain an upper bound on the norm of that ideal. We can assume that R is not a unit, otherwise Claim 7.2 holds trivially. We obtain

$$\begin{aligned}\left|\prod_{x\in I} N_{L/\mathbb{Q}}(\mathfrak{s}_x)\right| &\le N_{L/\mathbb{Q}}(\mathfrak{r})^n \left|N_{L/\mathbb{Q}}(R)\right|^d \\ &\qquad \cdot \exp\Biggl(\sum_{\substack{\mathfrak{p}\in\mathcal{P}_L,\,\mathfrak{p}\mid R\\ \mathfrak{p}\mid p\in\mathcal{P}_{\mathbb{Q}}}}\Bigl(dd_{\mathfrak{p}}\log(D) + dD\frac{e_{\mathfrak{p}}\log N_{L/\mathbb{Q}}(\mathfrak{p})}{p-1}\Bigr)\Biggr) \\ &\le N_{L/\mathbb{Q}}(\mathfrak{r})^n \left|N_{L/\mathbb{Q}}(R)\right|^d \\ &\qquad \cdot \exp\bigl(dd_L\log(D)\log\left|N_{L/\mathbb{Q}}(R)\right|/\log(2) \\ &\qquad\qquad + dd_L D(2\log\log\left|N_{L/\mathbb{Q}}(R)\right| + 4)\bigr).\end{aligned}$$

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

$$\begin{aligned}\sum_{x\in I}\widetilde{h}(\mathfrak{s}_x) &\le n\,\widetilde{h}(\mathfrak{r}) + d\,\widetilde{h}(R) + dd_L\frac{\log(D)}{\log(2)}\,\widetilde{h}(R) \\ &\qquad + dD(2\log\log\left|N_{L/\mathbb{Q}}(R)\right| + 4) \\ &\le n\,\widetilde{h}(\mathfrak{r}) + C_2\bigl(\eta d^4 H\log(D) + dD\log(\eta dH)\bigr)\end{aligned}$$

with

$$C_2 = \max\left\{\frac{3d_L(55+2C_1)}{2\log(2)}, 10 + 2\log(d_L) + 2\log(55+2C_1)\right\}. \tag{8}$$

Here we use that $\log(\eta dH) \ge 1$, and $\log(D) \ge 2\log 2$.

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,

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

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

In general, C_2 is defined in (8); in this equation, C_1 is a constant such that Proposition 4.6 holds. By Remarks 4.4 and 4.7, C_1 can be bounded above explicitly in terms of d_L and Δ_L only, so the same property holds for C_L. If $L = \mathbb{Q}$, we have $C_1 = 0$, so we can take $C_2 = 120$. □

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 $\mathbb{Z}\backslash V$. Assume that for every $x \in \mathbb{Z}\backslash V$, 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 apply Theorem 1.2 on an interval of the form $[\![0, D]\!]$ for some integer $D \geq 4d$, with $\eta = 2$ and $S = [\![0, D]\!]\backslash V$. 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 Theorem 1.2 with $H = H(D)$ as soon as the condition

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

holds. We check that we can choose

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

Then, Theorem 1.2 yields

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

We have $H(D) \leq 4cd \log(dD)$ and $2d\, H(D) \leq D$, 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. □

References

[1] E. Bach. Explicit bounds for primality testing and related problems. *Math. Comp.*, 55(191):355–380, 1990.

[2] R. Bröker and A. V. Sutherland. An explicit height bound for the classical modular polynomial. *Ramanujan J.*, 22(3):293–313, 2010.

[3] Y. Bugeaud and K. Győry. Bounds for the solutions of unit equations. *Acta Arith.*, 74(1):67–80, 1996.

[4] M. El Kahoui. An elementary approach to subresultants theory. *J. Symb. Comp.*, 35(3):281–292, 2003.

[5] M. Hindry and J. H. Silverman. *Diophantine Geometry*. Springer, 2000.

[6] J. Kieffer. Degree and height estimates for modular equations on PEL Shimura varieties. *Journal of the LMS*, to appear.

[7] S. Lang. *Algebraic Number Theory*. Springer, second edition, 1994.

[8] F. Mertens. Ein Beitrag zur analytischen Zahlentheorie. *J. Reine Angew. Math.*, 78:46–62, 1874.

[9] F. Pazuki. Modular invariants and isogenies. *Int. J. Number Theory*, 15(3):569–584, 2019.

[10] B. Rosser. Explicit bounds for some functions of prime numbers. *Amer. J. Math.*, 63(1):211–232, 1941.

[11] J. Sands. Generalization of a theorem of Siegel. *Acta Arith.*, 58(1):47–57, 1991.

[12] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, third edition, 2013.

Harvard University, Mathematics Department, Cambdidge, MA 02138, United States

Email address: kieffer@math.harvard.edu