Periodica Mathematica Hungarica Vol. 46 (2), 2003, pp. 135–155

ON METRIC HEIGHTS

A. Dubickas (Vilnius) and C. J. Smyth (Edinburgh)

[Communicated by: Attila Pethő]

Abstract

Metric heights are modified height functions on the non-zero algebraic numbers $\overline{\mathbb{Q}}^*$ which can be used to define a metric on certain cosets of $\overline{\mathbb{Q}}^*$. They have been defined with a view to eventually applying geometric methods to the study of $\overline{\mathbb{Q}}^*$. In this paper we discuss the construction of metric heights in general. More specifically, we study in some detail the metric height obtained from the naïve height of an algebraic number (the maximum modulus of the coefficients of its minimal polynomial). In particular, we compute this metric height for some classes of surds.

1. Introduction

For the study of number fields, the geometry of numbers is a very useful tool. To use it to study their rings of algebraic integers or unit groups, for instance, it is well-known that one embeds these structures as lattices in a Euclidean space, so that one can apply geometric methods to them. When one considers how one might do something similar for the field $\overline{\mathbb{Q}}$ of *all* algebraic numbers, however, it is not easy to see how best to proceed. In this paper, we study a class of height functions which we call *metric heights.* These functions can be used to put a metric structure on certain cosets of the multiplicative group $\overline{\mathbb{Q}}^*$, and so give possible tools for applying geometric methods to its study. This paper represents only a preliminary stage in this possible programme, as it is not yet clear what the most satisfactory metric structure for the study of $\overline{\mathbb{Q}}$ will be. In a recent paper [3], we studied one such height, the metric Mahler measure. These kind of heights are not always easy to compute exactly. Indeed, as we show below, even the computation of the metric heights of surds is a non-trivial business.

First we give a general definition of metric heights, and look at some of their properties. In particular, we study the *metric naïve height*, and show (Theorem 2) that its value for surds coincides with that of the metric Mahler measure. Next,

Mathematics subject classification number: Primary 11R04, 11R09, Secondary 11D68, 11Y16.

Key words and phrases: algebraic numbers, heights.

0031-5303/03/$20.00

Akadémiai Kiadó, Budapest
Kluwer Academic Publishers, Dordrecht

we compute these common values for the special surds $m\sqrt{n}$ (Theorem 3) and, for primes p and q, for the surds $(p/q^2)^{1/d}$ (Theorem 4), $(pq^2)^{1/d}$ (Theorem 5), and $(pq)^{2/d}$ (Theorem 6). The complication of the last two results indicates that there seems little prospect of very general formulas for the metric heights, even for surds. At the end of the paper we do, however, give an algorithm for finding the metric height of a given surd. In some cases, we compute the metric naïve height for algebraic numbers whose minimal polynomial has a large leading or constant coefficient (Corollary 2).

In deriving some of our results, we have been led to the study of a new arithmetical function $g(n)$, defined in Section 4, which seems of independent interest. It is a generalization of the 'smallest prime factor of n' function.

2. Heights and metric heights

Generally speaking, a real non-negative function defined on the group of non-zero algebraic numbers $\overline{\mathbb{Q}}^*$ is called a *height function* if, for a given bound, there are only finitely many algebraic numbers whose height and whose degree does not exceed that bound (see, for instance, [4]). Without too much loss of generality (namely, on replacing h by $\exp|h-h(1)|$, if necessary) we can assume that $h \geqslant 1$ and $h(1)=1$. Also, on replacing $h(\alpha)$ by $\max\{h(\alpha), h(1/\alpha)\}$, if necessary, we can assume that $h(\alpha)=h(1/\alpha)$ for every $\alpha \in \overline{\mathbb{Q}}^*$. In what follows we shall take it for granted that a height (function) also has these extra properties.

Let $P(z)=a_dz^d+\cdots+a_1z+a_0=a_d(z-\alpha^{(1)})\dots(z-\alpha^{(d)})$ be the minimal polynomial of α. Both the *naïve height* of α, $H(\alpha)=\max_{0\leqslant k\leqslant d}|a_k|$, and the *Mahler measure* of α,

$$M(\alpha)=a_d\prod_{1\leqslant j\leqslant d}\max\{|\alpha^{(j)}|,1\},$$

are heights, according to our definition. So also are each of $\exp\{L(\alpha)-2\}$, $L(\alpha)-1$, where $L(\alpha)=\sum_{k=0}^{d}|a_k|$ is the *length* of α, $\max\{\overline{|\alpha|},\overline{|1/\alpha|}\}$, the *'symmetric house'* of α, where $\overline{|\alpha|}=\max_{1\leqslant j\leqslant d}|\alpha^{(j)}|$, and $M(\alpha)^{1/d}$, where $(1/d)\log M(\alpha)$ is called the *Weil height.* More generally $\exp\{L_p(\alpha)-2\}$, $L_p(\alpha)-1$, where $p>0$ and $L_p(\alpha)=\sum_{k=0}^{d}|a_k|^p$ is the sum of pth powers of coefficients of P, are also heights. In a recent paper [2] we defined another new height, the *Remak height.* Note that all of these heights have the property that $h(\alpha')=h(\alpha)$ for any conjugate α' of α.

Given any height h, we can define another height $\widehat{h}$ by the formula

$$\widehat{h}(\alpha)=\inf h(\alpha_1)h(\alpha_2)\dots h(\alpha_s).$$

Here, the infimum is taken over every positive integer s and over all algebraic numbers $\alpha_1,\alpha_2,\dots,\alpha_s$ whose product is equal to α. In [3] we applied this definition for $h=M$, the height $\widehat{M}$ obtained by the procedure described being called there the *metric Mahler measure.*

Let h be a height. For arbitrary non-zero algebraic numbers α and β we have the following properties of $\widehat{h}$:

(i) $1 \leqslant \widehat{h}(\alpha) \leqslant h(\alpha)$,
(ii) $\widehat{h}(\alpha) = \widehat{h}(1/\alpha)$,
(iii) $\widehat{h}(\alpha\beta) \leqslant \widehat{h}(\alpha)\widehat{h}(\beta)$.

We denote by $\Omega = \Omega(\widehat{h})$ the non-zero algebraic numbers α such that $\widehat{h}(\alpha) = 1$. By (i)–(iii), Ω is a group. Clearly, the set $\Omega(h) = \{\alpha \mid h(\alpha) = 1\}$ is a subset of $\Omega(\widehat{h})$. In particular, $\Omega(M) = \Omega(\widehat{M})$ is the group of all roots of unity, which we denote by $\mathbb{U}$ (see [3]).

For any $\widehat{h}$, the function $\mathcal{D}$ given by

$$\mathcal{D}(\alpha\Omega, \beta\Omega) = \log \widehat{h}(\alpha/\beta)$$

is well-defined, and is a metric on the factor group $\overline{\mathbb{Q}}^*/\Omega$. Indeed, the triangle inequality for $\mathcal{D}$ follows immediately, by the definition of $\mathcal{D}$ and by (iii). Also, $\mathcal{D}(\alpha\Omega, \beta\Omega) = \mathcal{D}(\beta\Omega, \alpha\Omega)$, by (ii). Finally, $\mathcal{D}(\alpha\Omega, \beta\Omega) \geqslant 0$ with $\mathcal{D}(\alpha\Omega, \beta\Omega) = 0$ if and only if $\widehat{h}(\alpha/\beta) = 1$, that is, $\alpha/\beta \in \Omega$ which is equivalent to $\alpha\Omega = \beta\Omega$. Note that

$$\widehat{\widehat{h}} = \widehat{h},$$

so that we do not get another new height by repeating this procedure. In some cases, for instance for $h^*(\alpha) = M(\alpha)^{1/d}$, we may have $\widehat{h^*} = h^*$ (see [4, Property 3.3 and Lemma 3.10] for the inequality $h^*(\alpha\beta) \leqslant h^*(\alpha)h^*(\beta)$ with $\alpha, \beta \in \overline{\mathbb{Q}}^*$).

We say that α is a *surd* if some positive integer power of α is rational. Let us denote by $\mathbb{S}$ the set of all surds. Also, let $\mathbb{E}$ be the set of all algebraic numbers α whose conjugates (including α itself) are either all in the unit circle $|z| \leqslant 1$ or are all strictly outside the unit circle. The set $\mathbb{E}$ was considered in [1], where it was shown that $M(\alpha) \in \mathbb{N}$ if and only if $\alpha \in \mathbb{E}$. (Throughout, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Z}^*$ and $\mathbb{Q}$ stand for the sets of positive integers, integers, non-zero integers and rational numbers, respectively.) It is clear that

$$\mathbb{U} \subset \mathbb{S} \subset \mathbb{E}.$$

3. The metric naïve height

The naïve height H is discrete, its values all being in $\mathbb{N}$. It follows that the infimum in the definition of $\widehat{H}$ is in fact the minimum

$$\widehat{H}(\alpha) = \min H(\alpha_1)H(\alpha_2)\dots H(\alpha_s)$$

taken over every $s \in \mathbb{N}$ and all $\alpha_1, \alpha_2, \dots, \alpha_s \in \overline{\mathbb{Q}}^*$ such that $\alpha_1 \dots \alpha_s = \alpha$. Since the values of H are positive integers, so are the values of $\widehat{H}$. We call $\widehat{H}$ the *metric naïve height*. (We will drop the 'naïve' when the context is clear.) Without loss of generality we can assume that if

$$\widehat{H}(\alpha) = H(\alpha_1)H(\alpha_2)\dots H(\alpha_s),$$

then $\widehat{H}(\alpha_k) = H(\alpha_k)$ for every $k = 1, 2, \dots, s$. Indeed, if, say $\widehat{H}(\alpha_1) < H(\alpha_1)$, then there exist $\beta_1, \dots \beta_l \in \overline{\mathbb{Q}}^*$ such that $\alpha_1 = \beta_1 \dots \beta_l$ and $H(\beta_1) \dots H(\beta_l) < H(\alpha_1)$. We thus can replace the numbers $\alpha_1, \alpha_2, \dots, \alpha_s$ by the numbers $\beta_1, \dots, \beta_l, \alpha_2, \dots, \alpha_s$ having the same product, but a smaller product of heights, contradicting the definition of $\widehat{H}$.

We now state some properties of the metric height $\widehat{H}$. Clearly, it satisfies the conditions (i)–(iii) of Section 2. Furthermore, we have that

$$\widehat{H}(\alpha) = \widehat{H}(\alpha') \tag{1}$$

if α' is a conjugate to α.

Some of the simplest inequalities for the metric height are similar to those for the metric Mahler measure. Given $\alpha \in \overline{\mathbb{Q}}^*$ of degree d, let $a_d \in \mathbb{N}$ and $a_0 \in \mathbb{Z}^*$ be the extreme coefficients of its minimal polynomial $\sum_{k=0}^{d} a_k x^k$, and $\mathcal{N}(\alpha) = |a_0|/a_d$ be the absolute value of the norm of α. Also, given $n \in \mathbb{N}$, define $\mathcal{R}(n)$, the *radical* of n, by $\mathcal{R}(1) = 1$ and $\mathcal{R}(n) = \prod_{p|n} p$, where the product is taken over every prime divisor of n.

THEOREM 1. *Let $\alpha \in \overline{\mathbb{Q}}^*$ be of degree d. Then*
(a) $\widehat{H}(\alpha) \geqslant \widehat{H}\big(\mathcal{N}(\alpha)^{1/d}\big)$,
(b) $\widehat{H}(\alpha) \geqslant \max\{\mathcal{R}(a_d), \mathcal{R}(|a_0|)\}$.

COROLLARY 1. *Suppose that $m, n \in \mathbb{N}$, $m > n$, with mn square-free and $d \in \mathbb{Z}^*$. Then*

$$\widehat{H}((m/n)^{1/d}) = m.$$

Theorem 1 allows us to find the metric height of some numbers which belong to $\mathbb{E}$. The next result, combined with Theorem 4 below, shows that the metric naïve height $\widehat{H}(\alpha)$ of α having minimal polynomial $z^{35} - 7z - 175$ is equal to 175. (The polynomial is irreducible, by Eisenstein's criterion. Also, $\widehat{H}(175^{1/35}) = 175$, by Theorem 5.)

COROLLARY 2. *Let $\alpha \in \mathbb{E}$ be of degree d and of height m. Suppose that m and n are the moduli of the extreme coefficients of the minimal polynomial for α. If $\widehat{H}((m/n)^{1/d}) = m$, then*

$$\widehat{H}(\alpha) = m.$$

From (i) it follows that $\Omega(h) \subset \Omega(\widehat{h})$. Recall that the set $\Omega(\widehat{M})$ is the set of all roots of unity (see [2]). So is also $\Omega(M)$, hence $\Omega(\widehat{M}) = \Omega(M)$. However, the sets $\Omega(H)$ and $\Omega(\widehat{H})$ are distinct. For instance, setting $\varphi = (3+\sqrt{5})/2 = \big((1+\sqrt{5})/2\big)^2$, we have that $H(\varphi) = 3$, but $\widehat{H}(\varphi) = 1$.

There exists an $\alpha \in \mathbb{S}$ such that $H(\alpha) \neq M(\alpha)$. Consider, for instance, the number $\alpha = \exp\{2\pi i/105\} \in \mathbb{U}$. Clearly, $\Phi_{105}(\alpha) = 0$, where $\Phi_{105}(z) = \sum_{k=0}^{48} a_k z^k$ is the 105th cyclotomic polynomial, having coefficients $a_0, a_1, \ldots, a_{24}$ equal to

$$1, 1, 1, 0, 0, -1, -1, -2, -1, -1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, -1, 0, -1, 0, -1,$$

respectively. The remaining coefficients of $\Phi_{105}(z)$ can be obtained by the formula $a_{48-k} = a_k$, where $k = 0, 1, \ldots, 23$. Consequently, $H(\alpha) = 2$, but $M(\alpha) = 1$, since $\alpha \in \mathbb{U}$. The next theorem shows that the metric height is equal to the metric Mahler measure on the multiplicative group $\mathbb{S}$.

THEOREM 2. *If $\alpha \in \mathbb{S}$, then $\widehat{H}(\alpha) = \widehat{M}(\alpha)$.*

We proved in [3] that $\widehat{M}(r) = M(r) = H(r)$ for every non-zero $r \in \mathbb{Q}$. In particular, the following corollary holds.

COROLLARY 3. *For every $n \in \mathbb{N}$ we have $\widehat{H}(n) = n$ and thus $\widehat{H} : \overline{\mathbb{Q}}^* \to \mathbb{N}$ is a surjection.*

4. The metric height of surds

Every element of the set $\mathbb{S}$ can be written in the form $\omega(m/n)^{1/d}$. Throughout, $(m/n)^{1/d}$ is the real dth root, $\omega \in \mathbb{U}$ and $m, n, d \in \mathbb{N}$. In particular, every irrational surd α which is a real number such that $\alpha^2 \in \mathbb{Q}$ and is an algebraic integer can be written as $m\sqrt{n}$ with $m \in \mathbb{Z}^*$ and a square-free integer $n > 1$. For such numbers, the metric naïve height is given by the following simple formula (see also Corollary 5 below).

THEOREM 3. *If $m \in \mathbb{N}$ and $n > 1$ is a square-free integer, then $\widehat{H}(m\sqrt{n}) = mn$.*

It seems that the only irrational surds for which such a formula can be obtained are the quadratic irrationalities. Even the cubic case, namely, determining a similar formula for *all* algebraic integers $\alpha \in \mathbb{S}$ such that $\alpha^3 \in \mathbb{Q}$, but $\alpha, \alpha^2 \notin \mathbb{Q}$, is almost certainly a very difficult question. See Theorems 4 and 5 (the case $d = 3$) for some results in special cases. However, in some very particular cases, for instance, if $\mathcal{R}(mn)$ is the product of two distinct primes, it is possible to find the metric naïve height even for surds of an arbitrary degree.

Let $p \neq q$ be two primes. In the next three theorems we investigate the metric naïve height of the numbers $(p/q^2)^{1/d}$, $(pq^2)^{1/d}$ and $(pq)^{2/d}$ respectively. These are the simplest types of 'non-square-free' surds. For mn square-free, we know from Corollary 1 that $(m/n)^{1/d}$ has metric naïve height $\max\{m, n\}$. Our results have discouraged us from looking at the metric naïve height of more complicated surds!

THEOREM 4. *We have*

$$\widehat{H}((p/q^2)^{1/d}) = \begin{cases} pq \text{ if } p < q \text{ and } d \text{ is even} \\ \max\{p, q^2\} \text{ otherwise.} \end{cases}$$

Set $e = e(p,q) = \lfloor \log\max\{p,q\}/\log\min\{p,q\} \rfloor$, so that, for instance, if $p > q$ then

$$q^e < p < q^{e+1}.$$

Here $\lfloor x \rfloor$ is the greatest integer $\leqslant x$. Also, let $f(d)$ be the smallest prime divisor of $d > 1$.

THEOREM 5. *If d is even, then $\widehat{H}((pq^2)^{1/d}) = pq$. If $d > 1$ is odd, then*

$$\widehat{H}((pq^2)^{1/d}) = \begin{cases} q^2 & \text{if } p < q \text{ and } f(d) \leqslant 2e+1 \\ p^{e+1}q & \text{if } p < q \text{ and } f(d) = 2e+3 \\ pq & \text{if } p > q \text{ and } f(d) \leqslant e+2 \\ q^{e+2} & \text{if } p > q \text{ and } f(d) = e+3 \\ pq^2 & \text{otherwise.} \end{cases}$$

In order to state our next theorem, we need to define a new function $g^*(d)$. To do this, we first define the function $g(d)$ to be the smallest integer L such that there exist positive or negative divisors k_1, k_2, k_3 of d with the k_i not all of the same modulus and $\ell_1, \ell_2, \ell_3 \in \{1, 2, \ldots, L\}$ such that $\ell_1 k_1 + \ell_2 k_2 + \ell_3 k_3 = 0$. The function $g(d)$ can be regarded as a kind of generalization of $f(d)$, as $f(d)$ is the smallest L with k_1, k_2, ℓ_1, ℓ_2 as above and $\ell_1 k_1 + \ell_2 k_2 = 0$. The function $g^*(d)$ is then defined as for $g(d)$, but with the additional condition that $2\ell_1$ must divide $k_2 + k_3$. So certainly $g^*(d) \geqslant g(d)$, though for many values of d the two functions are equal – see the end of Section 6.

THEOREM 6. *Let primes $p > q$ be given, and $d > 1$ be odd. Then*

$$\widehat{H}((pq)^{2/d}) = \begin{cases} p^2 & \text{if } e \geqslant f(d) - 1 \\ pq^{e+1} & \text{if } e = f(d) - 2 \\ p^2 q & \text{if } g^*(d) - 1 \leqslant e \leqslant f(d) - 3 \\ & \text{or } e = g^*(d) - 2,\ f(d) = 2g^*(d) - 1 \text{ and } p > q^{e+1/2} \\ q^{2e+2} & \text{if } e = g^*(d) - 2,\ f(d) = 2g^*(d) - 1 \text{ and } p < q^{e+1/2} \\ pq^{e+2} & \text{if } e = g^*(d) - 2 \text{ and } f(d) > 2g^*(d) - 1 \\ (pq)^2 & \text{if } e \leqslant g^*(d) - 3. \end{cases}$$

Note that for d even, $\widehat{H}((pq)^{2/d}) = pq$, by [2, Corollary 6(i)] combined with Theorem 2. Also, $f(d) \geqslant 2g^*(d) - 1$, by Lemma 3(ii), so that all possibilities are covered by the theorem.

5. Alternative form of Theorem 2

The following is an alternative form of Theorem 2. This is because, in [2, Theorem 2], we obtained the same result, but with the metric Mahler measure $\widehat{M}$ in place of the metric naïve height $\widehat{H}$. Throughout, (m, n) denotes the greatest common divisor of m and n.

PROPOSITION 1. *Suppose that $m, n, d \in \mathbb{N}$, where $(m, n) = 1$, and $\zeta \in \mathbb{U}$. Then*

$$\widehat{H}(\zeta(m/n)^{1/d}) = \prod_i \max\{m_i, n_i\},$$

where $m_i, n_i \in \mathbb{N}$ are such that the representation of

$$(m/n)^{1/d} = \prod_i \big(m_i/n_i\big)^{1/d_i}, \tag{2}$$

for some $d_i \in \mathbb{Z}^*$, is chosen so that $\prod_i \max\{m_i, n_i\}$ is minimal. In particular, $\widehat{H}((m/n)^{1/d}) \leqslant \max\{m, n\}$.

From the minimality it follows that the prime divisors of $\widehat{H}((m/n)^{1/d})$ are also prime divisors of mn, so that $\mathcal{R}(\widehat{H}((m/n)^{1/d}))$ divides $\mathcal{R}(mn)$. Also, there is no loss of generality in assuming that $m_i n_i > 1$ for every i. Furthermore, there is a representation (2) with $\widehat{H}((m/n)^{1/d}) = \prod_i \max\{m_i, n_i\}$ such that the prime divisors of $\prod_i m_i n_i$ are also prime divisors of mn:

$$\mathcal{R}\Big(\prod_i m_i n_i\Big) = \mathcal{R}(mn).$$

We also have [2, Theorem 2(d)] the lower bound

$$\widehat{H}((m/n)^{1/d}) \geqslant \max_{p^t \| mn} p^{u(t/d)} = q_0 \tag{3}$$

say. Here, for every $r \in \mathbb{Q}$, $u(r)$ denotes the smallest $u \in \mathbb{N}$ for which r can be represented as

$$r = 1/s_1 + 1/s_2 + \cdots + 1/s_u$$

with some non-zero integers $s_1, \ldots, s_u$. In particular, if p is a prime and t is an integer, then $\widehat{H}(p^{t/d}) = p^{u(t/d)}$. See also [2, Theorem 2] for some other inequalities for the metric Mahler measure which for surds is equal to the metric naïve height.

Finally, by Corollary 1, if $n, d \in \mathbb{N}$ and n is square-free, then $\widehat{H}(n^{1/d}) = n$. If however there is a prime p such that p^3 divides n, then $\widehat{H}(n^{1/d}) < n$ (see also [2, Corollary 6]).

Recall that $\Omega(\widehat{H})$ is the group of algebraic numbers which can be written as products of algebraic numbers each having height 1. Certainly, every α of metric height 1 must be a unit. From Proposition 1 we will deduce that all roots of unity are of metric height 1. In fact, the proof of this corollary forms the first part of the proof of the proposition.

COROLLARY 4. *The group $\mathbb{U}$ of all roots of unity is a subgroup of $\Omega(\widehat{H})$.*

Thus every root of unity has metric naïve height 1, and so the metric height is invariant under multiplication by elements of $\Omega(\widehat{H})$.

COROLLARY 5. *If $\zeta \in \mathbb{U}$ and $\alpha \in \overline{\mathbb{Q}}^*$, then $\widehat{H}(\alpha) = \widehat{H}(\alpha\zeta)$.*

6. Lemmas and their proofs

The first lemma is the main difference between the proofs of Theorem 2 for the metric Mahler measure and that below for the metric naïve height. We will also use it for the proof of Proposition 1.

LEMMA 1. *If $m, n, d \in \mathbb{N}$ and $\omega \in \mathbb{U}$, then there is an $\zeta \in \mathbb{U}$ such that*

$$H(\zeta\omega(m/n)^{1/d}) \leqslant \max\{m, n\}.$$

PROOF. If either $m = n$ or if $d = 1$, then the lemma follows, with $\zeta = 1/\omega$. The statement is therefore true for $m+n+d = 3$. Let us fix a positive integer $S \geqslant 4$ and assume that the lemma is true for every triplet $\langle m, n, d\rangle$, where $m+n+d < S$, and for arbitrary $\omega \in \mathbb{U}$. We shall prove the statement for arbitrary $\omega \in \mathbb{U}$ and $m+n+d = S$, where m and n (at least one of which is greater than 1) are coprime and $d > 1$. We write m/n in the form

$$m/n = \prod_{p\in\mathcal{P}} p^{\nu_p(m/n)},$$

where $\mathcal{P}$ denotes the set of all prime divisors of mn and $\nu_p(m/n)$ denotes the integer power of p in the factorization of the rational number m/n.

If the polynomial $nz^d - m$ is irreducible, then once again we can take $\zeta = 1/\omega$. Let $P_\ell(z)$ of degree ℓ, where $1 \leqslant \ell \leqslant d-1$, with integer coefficients be a divisor of $nz^d - m$. The roots of $P_\ell(z)$ are all of modulus $(m/n)^{1/d}$, hence their product is of modulus $(m/n)^{\ell/d}$. We deduce that

$$\big(m/n\big)^{\ell/d} = \prod_{p\in\mathcal{P}} p^{\nu_p(m/n)\ell/d}.$$

Here, the left-hand side is a rational number, hence so is the right-hand side. The set $\mathcal{P}$ is not empty, because $m > 1$ or $n > 1$. Therefore, by an easy irrationality argument, d divides $r\ell$, where r is the largest common divisor of the numbers $\nu_p(m/n)$ as p runs over every element of $\mathcal{P}$. Since $\ell < d$, we deduce that $(r, d) > 1$. There exist positive integers m_0, n_0, r_0, d_0 such that $m = m_0^r$, $n = n_0^r$, $r_0 = r/(r, d)$, $d_0 = d/(r, d)$. We have that

$$\big(m/n\big)^{1/d} = \big(m_0/n_0\big)^{r/d} = \big(m_0/n_0\big)^{r_0/d_0} = \big(m_0^{r_0}/n_0^{r_0}\big)^{1/d_0}.$$

Because $m_0^{r_0} \leqslant m$, $n_0^{r_0} \leqslant n$ and $d_0 < d$, the inequality

$$m_0^{r_0} + n_0^{r_0} + d_0 < m + n + d = S$$

holds. There exists a ζ such that the height of $\zeta\omega(m_0^{r_0}/n_0^{r_0})^{1/d_0}$, is, by induction, at most

$$\max\{m_0^{r_0}, n_0^{r_0}\} \leqslant \max\{m, n\}.$$

So the height of $\zeta\omega(m/n)^{1/d}$ is at most $\max\{m, n\}$, as claimed. The proof of the lemma is completed.

The diophantine system considered in the next lemma arises in the proof of Theorem 6.

LEMMA 2. *Let $d > 1$ be an odd integer. Then the equations*

$$1/d_2 + 1/d_3 = 2/d, \qquad 1/d_1 + (1 - \ell)/d_2 + (1 - \ell')/d_3 = 2/d,$$

have at least one integer solution $\langle d_1, d_2, d_3, \ell, \ell'\rangle$ with $\ell, \ell' \in \{1, 2, \dots, L\}$ iff $L \geqslant g^*(d)$. In fact, in all such solutions, d_1, d_2, d_3 are parameterized by

$$d_1 = -d(k_2 + k_3)/2\ell_1 k_1, \quad d_2 = d(k_2 + k_3)/2k_2, \quad d_3 = d(k_2 + k_3)/2k_3,$$

for some divisors k_1, k_2, k_3 of d with $k_2 \neq -k_3$ and some $\ell_1, \ell, \ell' \in \{1, 2, \dots, L\}$, for which $\ell_1 k_1 + \ell k_2 + \ell' k_3 = 0$ and ℓ_1 divides $(k_2 + k_3)/2$.

PROOF. The expressions for d_2 and d_3 follow immediately by writing the first equation as $(2d_2 - d)(2d_3 - d) = d^2$ and then $2d_3 - d$ as a product $k_2 d/k_3$ of the divisors k_2 and d/k_3 of d. Note that $k_2 \neq -k_3$. It is an easy exercise, using the expressions for d_2 and d_3, to get the expression for d_1. To show that this expression is an integer, note first that we may assume that $(\ell_1, d/k_1) = 1$, for otherwise ℓ_1 could be decreased, without changing d_1. Thus for $d_1 \in \mathbb{Z}^*$ we need, as well as the condition $k_1|d$, the condition $\ell_1|(k_2+k_3)/2$. Note that then $\ell_1|(k_2+k_3, \ell k_2+\ell' k_3)$, so ℓ_1 divides both $(\ell-\ell')k_2$ and $(\ell-\ell')k_3$. Now if $(k_2, k_3) > 1$ we can reduce k_2, k_3 and either k_1 or ℓ_1, without changing $\langle d_1, d_2, d_3\rangle$. So we can assume that $(k_2, k_3) = 1$. Hence $\ell_1|(\ell - \ell')$, so that $\ell_1 < \max\{\ell, \ell'\}$ and $\max\{\ell_1, \ell, \ell'\} \leqslant L$.

Finally, if $k_2 = k_3$, then $k_2 = k_3 = (k_2 + k_3)/2 = 1$, giving $\ell_1 = 1$, $|k_1| = \ell + \ell' > 1$. Thus the k_i do not all have the same modulus, and so, for $\ell = \ell_2, \ell' = \ell_3$, the conditions of the k_i and ℓ_i in the definition of $g^*(d)$ are satisfied. This shows

that $L \geqslant g^*(d)$. Conversely, if $L \geqslant g^*(d)$ then the k_i and ℓ_i from the definition of $g^*(d)$ parameterize a solution to the given system of equations.

We now present some properties of the function $g^*(d)$ used in the statement of Theorem 6.

LEMMA 3. *Let $d \in \mathbb{N}$, $d > 1$. The function $g^*(d)$ has the following properties.*

(i) *We have $g^*(d) = 1$ iff d is even.*
(ii) *If d is odd then $2 \leqslant g^*(d) \leqslant (f(d)+1)/2$.*
(iii) *If d is odd then $g^*(d) \leqslant \min\max\{\ell, |d' - \ell d''|\}$, where the minimum is taken over every $\ell \in \mathbb{N}$ and over every pair $\langle d', d''\rangle$ of divisors of d such that $1 < d'' < d'$.*
(iv) *If d is odd, then $g^*(d) \leqslant \min\max\{\ell_1, \ell_2\}$, where the minimum is taken over all pairs $\langle \ell_1, \ell_2\rangle$ and over all positive or negative divisors k_1, k_2 of d such that $\ell_1 k_1 + \ell_2 k_2 = 1$ and $1 < |k_1| < |k_2|$.*
(v) *If d is odd, then $g^*(d) = 2$ iff d has three distinct positive or negative divisors in arithmetic progression. In particular, $g^*(d) = 2$ when d has a factor of the form $r(r+2)$ or $r(2r+1)$ or $(r+1)(2r+1)$, where $r \in \mathbb{N}$.*
(vi) *If d is odd and a multiple of $d' > 1$ then $g^*(d) \leqslant g^*(d')$.*
(vii) *If $d = p^r$, where $r \in \mathbb{N}$, is an odd prime power, then $g^*(d) = (p+1)/2$.*

Furthermore, the above properties (i)–(vii) also hold for the function $g(d)$.

PROOF. To evaluate or bound $g^*(d)$, we need to consider solutions of $\ell_1 k_1 + \ell_2 k_2 + \ell_3 k_3 = 0$ used in the definition of g and g^*.

(i) If $g^*(d) = 1$, then $\ell_1 = \ell_2 = \ell_3 = 1$, so that not all the k_i can be odd. Hence d is even. Conversely, if d is even, then the identity $1 \cdot (-2) + 1 \cdot 1 + 1 \cdot 1 = 0$ shows that $g^*(d) = 1$.

To prove (ii), (iii), (iv) and (v), it is enough to write down the following identities:

$$\text{(ii)} \qquad ((f(d)-1)/2) \cdot 1 + ((f(d)+1)/2) \cdot 1 + 1 \cdot (-f(d)) = 0,$$
$$\text{(iii)} \qquad 1 \cdot (-d') + \ell \cdot d'' + |d' - \ell d''| \cdot \operatorname{sgn}(d' - \ell d'') = 0,$$
$$\text{(iv)} \qquad 1 \cdot (-1) + \ell_1 \cdot k_1 + \ell_2 \cdot k_2 = 0,$$
$$\text{(v)} \qquad 1 \cdot k_1 + 2 \cdot (-k_2) + 1 \cdot k_3 = 0.$$

Note that for $g^*(d) = 2$ only one of ℓ_1, ℓ_2, ℓ_3 can be 2. The special cases of (v) then follow from the arithmetic progressions $\{-r, 1, r+2\}$, $\{-1, r, 2r+1\}$, and $\{1, r+1, 2r+1\}$ respectively. Note too that in all the above identities $\ell_1 = 1$, except for (ii), so that it certainly divides $(k_2 + k_3)/2$. (In (ii), we have $2\ell_1 = f(d) - 1$ and $k_2 + k_3 = 1 - f(d)$.)

The proof of (vi) is immediate. For (vii), note that if say k_2 and k_3 are divisible by p, then we can divide k_2, k_3 and either k_1 or ℓ_1 by p. Thus we may assume that $k_2 = k_3 = \pm 1$. Hence, as $k_1 \neq -k_2, |k_1| \geqslant p$, so that $|\ell_2 + \ell_3| \geqslant p$. The inequality $\max\{\ell_2, \ell_3\} \geqslant (p+1)/2$ combined with (ii) gives the result.

Finally, we remark that the proofs are easily modified, using the inequality $g \leqslant g^*$ where necessary, to show that the results all hold for g as well as for g^*.

A little Maple computation shows that the bound for $g^*(d)$ and $g(d)$ in (iii) above is sharp for all $d < 187$ ($g^*(187) = g(187) = 3$, while the bound in (iii) is 5). The minimum of the bounds in (iii) and (iv) is sharp for all odd $d < 589$ (with $g^*(589) = g(589) = 5$, where the bounds of (iii) and (iv) are 7 and 10 respectively). The functions $g^*(d)$ and $g(d)$ are equal for all $d < 1073$, with $g^*(1073) = 8$ and $g(1073) = 5$, using the identities $1\cdot(-37)+8\cdot1+1\cdot29 = 0$ and $4\cdot(-37)+5\cdot29+3\cdot1 = 0$, respectively.

7. Proofs of the results of Sections 3 and 5

The main ingredient in the proofs of Theorems 1 and 2 is inequality (4) below. We then prove Corollary 4, and, via Corollary 5 (which is its immediate consequence), deduce Theorem 1(a). After proving Proposition 1, which, as we remarked earlier, implies Theorem 2, we will conclude this section by proving Theorem 1(b) and Corollary 2. (Corollary 1 follows from Theorem 1(b), whereas Corollary 3 is an immediate consequence of Theorem 2.)

Given $\alpha \in \overline{\mathbb{Q}}^*$, let $\alpha_1, \dots, \alpha_s$ be such that $\alpha = \alpha_1 \dots \alpha_s$ and

$$\widehat{H}(\alpha) = H(\alpha_1) \dots H(\alpha_s).$$

Assume that α is of degree d and α_i is of degree d_i for every $i = 1, \dots, s$. Let also $n_i, \pm m_i$, where $m_i, n_i \in \mathbb{N}$, be the pair of extreme coefficients of the minimal polynomial of α_i so that $\mathcal{N}(\alpha_i) = m_i/n_i$. From Lemma 1 we deduce that there exist $\omega_1, \dots, \omega_s \in \mathbb{U}$ such that

$$H\big(\omega_i(m_i/n_i)^{1/d_i}\big) \leqslant \max\{m_i, n_i\} \leqslant H(\alpha_i)$$

with every $i = 1, \dots, s$. By [2, Lemma 3(i)], we have that

$$\mathcal{N}(\alpha)^{1/d} = \mathcal{N}(\alpha_1)^{1/d_1} \dots \mathcal{N}(\alpha_s)^{1/d_s} = (m_1/n_1)^{1/d_1} \dots (m_s/n_s)^{1/d_s}.$$

Set $\omega = \omega_1 \dots \omega_s$. It follows that

$$\omega\mathcal{N}(\alpha)^{1/d} = \omega_1(m_1/n_1)^{1/d_1} \dots \omega_s(m_s/n_s)^{1/d_s}.$$

Consequently,

$$\widehat{H}(\omega\mathcal{N}(\alpha)^{1/d}) \leqslant H\big(\omega_1(m_1/n_1)^{1/d_1}\big) \dots H\big(\omega_s(m_s/n_s)^{1/d_s}\big),$$

and so

$$\widehat{H}(\omega\mathcal{N}(\alpha)^{1/d}) \leqslant \max\{m_1, n_1\} \dots \max\{m_s, n_s\} \leqslant \widehat{H}(\alpha). \tag{4}$$

PROOF OF COROLLARY 4. Suppose that r is the smallest positive integer such that the primitive rth root of unity $\zeta_r = \exp\{2\pi i/r\}$ is not yet written as a product of roots of unity whose naïve height is 1. Clearly, r and $\mathcal{R}(r)$ both must be composite. Write r in the form $r_1 r_2$ for some positive coprime integers $r_1, r_2 \geqslant 2$. Since the numbers ζ_{r_1} and ζ_{r_2} can be represented as products of roots of unity all of naïve height 1, so can their product $\exp\{2\pi(r_1+r_2)i/r\}$. As $(r, r_1+r_2) = 1$, this number is conjugate to ζ_r. So, by (1), ζ_r is also the product of roots of unity of naïve height 1.

PROOF OF THEOREM 1(a). By Corollary 5, $\widehat{H}(\mathcal{N}(\alpha)^{1/d}) = \widehat{H}(\omega\mathcal{N}(\alpha)^{1/d})$, from which, with (4), we obtain the inequality of Theorem 1(a).

PROOF OF PROPOSITION 1. Clearly, $(m/n)^{1/d} = \mathcal{N}((m/n)^{1/d})^{1/D}$, where $(m/n)^{1/d}$ is, say, of degree D (which may be smaller than d). Now, by Corollary 5, inequality (4), with α and d being replaced by $(m/n)^{1/d}$ and D, respectively, is throughout an equality. It follows that in the representation $(m/n)^{1/d} = \alpha_1 \dots \alpha_s$ with $\widehat{H}((m/n)^{1/d}) = H(\alpha_1) \dots H(\alpha_s)$ every α_i can be replaced by $\omega_i (m_i/n_i)^{1/d_i}$, giving the new surd

$$\omega_1 \dots \omega_s (m/n)^{1/d} = \omega_1 (m_1/n_1)^{1/d_1} \dots \omega_s (m_s/n_s)^{1/d_s}$$

of the same metric height as that of $(m/n)^{1/d}$ (by Corollary 5). Hence

$$(m/n)^{1/d} = (m_1/n_1)^{1/d_1} \dots (m_s/n_s)^{1/d_s}$$

and $\widehat{H}(\zeta(m/n)^{1/d}) = \widehat{H}((m/n)^{1/d}) = \prod_i \max\{m_i, n_i\}$ with arbitrary $\zeta \in \mathbb{U}$. If (2) happened to be chosen so that $\prod_i \max\{m_i, n_i\}$ was not minimal, then, by taking another (2) with this product being minimal, and on applying Lemma 1, we would get a contradiction.

PROOF OF THEOREM 1(b). This part follows from [2, Lemma 3(ii)] combined with Theorem 2. Since a_d divides the product $m_1^{D/d_1} \dots m_s^{D/d_s}$, where $D = d_1 \dots d_s$, we have that $\mathcal{R}(a_d) \leqslant m_1 \dots m_s$. Consequently,

$$\widehat{H}(\alpha) = H(\alpha_1) \dots H(\alpha_s) \geqslant m_1 \dots m_s \geqslant \mathcal{R}(a_d).$$

Finally, $\widehat{H}(\alpha) \geqslant \mathcal{R}(|a_0|)$, because $\widehat{H}(\alpha) = \widehat{H}(1/\alpha)$.

PROOF OF COROLLARY 2. Using Theorem 1(a) and $\mathcal{N}(\alpha) = m/n$ or $= n/m$, we deduce that

$$m = H(\alpha) \geqslant \widehat{H}(\alpha) \geqslant \widehat{H}\big((m/n)^{\pm 1/d}\big) = \widehat{H}\big((m/n)^{1/d}\big) = m.$$

We thus have equality throughout, giving $\widehat{H}(\alpha) = m$, as claimed in Corollary 2.

8. Proofs of Theorems 3, 4, 5, 6

PROOF OF THEOREM 3. The trivial representation $m\sqrt{n} = m \cdot n^{1/2}$ gives the inequality $\widehat{H}(m\sqrt{n}) \leqslant mn$. It suffices to prove the opposite inequality for $m \in \mathbb{N}$, $n > 1$. In fact, it is more convenient to prove it for $\widehat{M}$, and then apply Theorem 2.

Let $\alpha = m\sqrt{n} = \alpha_1 \dots \alpha_s$ be such that

$$\widehat{M}(\alpha) = M(\alpha_1) \dots M(\alpha_s),$$

where s is minimal for which such a representation is possible. We shall prove that $s = 2$. Clearly, $\widehat{M}(\alpha) \leqslant mn < m^2 n = M(\alpha)$, thus $s > 1$. We now deduce, by [3, Property 3.3 and Lemma 3.10], the inequality

$$M(\alpha)^{1/2} \leqslant M(\alpha_1)^{1/d_1} \dots M(\alpha_s)^{1/d_s} \leqslant (M(\alpha_1) \dots M(\alpha_s))^{1/d_1} < M(\alpha)^{1/d_1}.$$

Here, d_i stands for the degree of α_i and, without loss of generality, $d_1 \leqslant d_2 \leqslant \dots \leqslant d_s$. Consequently, $d_1 < 2$, so α_1 is rational. If $s > 2$, from the representation

$$\alpha/\alpha_1 = \alpha_2 \dots \alpha_s$$

and from the minimality of s, it is easy to see that

$$M(\alpha/\alpha_1) > M(\alpha_2) \dots M(\alpha_s).$$

Now, by the above argument, we have that $d_2 = 1$ and so $\alpha_2 \in \mathbb{Q}$. However, $M(\alpha_1\alpha_2) \leqslant M(\alpha_1)M(\alpha_2)$ for $\alpha_1, \alpha_2 \in \mathbb{Q}$, contradicting the minimality of s.

It follows that $\alpha = r\alpha'$ with $r \in \mathbb{Q}$. If $r = u/v$, where $u, v \in \mathbb{N}$, $(u, v) = 1$, then

$$\alpha = (u/v)(v^2 m^2 n/u^2)^{1/2}.$$

By Proposition 1,

$$\widehat{M}(\alpha) = \frac{\max\{u, v\} \max\{v^2 m^2 n, u^2\}}{(v^2 m^2 n, u^2)}.$$

On replacing v by 1, we certainly do not increase the right-hand side. Then, on estimating the maximum by $m^2 n$, we get

$$\widehat{M}(\alpha) \geqslant um^2 n/(m^2 n, u^2).$$

In order to complete the proof it suffices to show that $um \geqslant (m^2 n, u^2)$ for n square-free.

Let us fix a prime p, and assume that $p^g \| u$, $p^h \| m$, $p^t \| n$. Here, $t \in \{0, 1\}$, since n is square-free. By considering two cases $h \geqslant g$ and $g \geqslant h + 1$, we see that

$$g + h \geqslant \min\{2h + t, 2g\}.$$

Thus $\nu_p(um) \geqslant \nu_p((m^2 n, u^2))$ for every prime p. The inequality $um \geqslant (m^2 n, u^2)$ follows, and so the proof of Theorem 3 is completed.

PROOF OF THEOREM 4. First note that, by Eisenstein's criterion, the polynomial $q^2z^d - p$ is irreducible, so that

$$\widehat{H}((p/q^2)^{1/d}) \leqslant H\big((p/q^2)^{1/d}\big) = \max\{p, q^2\}.$$

We consider the various possibilities.

1. If $p > q^2$, then, as $\mathcal{R}(\prod_i m_i n_i) = \mathcal{R}(mn) = pq$, in (2) at least one m_i or n_i is divisible by p. Thus $\max\{m_i, n_i\} \geqslant p$, and so $\widehat{H}((p/q^2)^{1/d}) \geqslant p$. It follows that $\widehat{H}((p/q^2)^{1/d}) = p$.

2. We can therefore assume that $p < q^2$. If $\widehat{H}((p/q^2)^{1/d}) < q^2$, then q can appear in only one fraction of (2), say in m_j/n_j. Furthermore, q must be to the first power. If further d is odd, then the equality $-2/d = \pm 1/d_j$ with $d_j \in \mathbb{Z}^*$ is impossible. This shows that in this case, for d odd, we have the equality $\widehat{H}((p/q^2)^{1/d}) = q^2$.

3. We can assume that $p < q^2$ and d is even. If also $p > q$, then $pq > q^2$. If in (2), namely in $\prod_{i=1}^{s}(m_i/n_i)^{1/d_i}$, we have $s \geqslant 2$ then the product $\prod_i \max\{m_i, n_i\}$ is at least pq which is not minimal. If however $s = 1$, then (2) is the trivial representation $(p/q^2)^{1/d} = (p/q^2)^{1/d}$ which gives $\widehat{H}((p/q^2)^{1/d}) = q^2$.

4. We can assume that $p < q$ and d even. The representation $(p/q^2)^{1/d} = p^{1/d}q^{-2/d}$ shows that $\widehat{H}((p/q^2)^{1/d}) \leqslant pq$. Note that $s \geqslant 2$, for otherwise we have $\widehat{H}((p/q^2)^{1/d}) = q^2 > pq$. In the product $\prod_i \max\{m_i, n_i\}$ the prime q appears only once. But the product is strictly greater than q, so it is at least pq. This shows that $\widehat{H}((p/q^2)^{1/d}) = pq$ for even d and $p < q$.

PROOF OF THEOREM 5. The polynomial $z^d - pq^2$ is irreducible (again Eisenstein), so with Theorem 1(b) we have

$$pq^2 \geqslant \widehat{H}((pq^2)^{1/d}) \geqslant \mathcal{R}(pq^2) = pq.$$

On the other hand, the representation $(pq^2)^{1/d} = p^{1/d}q^{2/d}$ shows that $\widehat{H}((pq^2)^{1/d}) \leqslant pq$ for even d.

Let $d > 1$ be odd for the rest of the proof. Clearly, in $P = \prod_i^s \max\{m_i, n_i\}$ either $s \geqslant 2$ or $s = 1$. In the latter case $P = \widehat{H}((pq^2)^{1/d}) = pq^2$. So assume that $s \geqslant 2$ and $P < pq^2$.

1. We begin with the case $p < q$. The power of q in P is at most 2, for otherwise $P \geqslant q^3 > pq^2$, a contradiction. Thus q must appear in exactly two distinct places of the representation (2) or in one place as a square at most. The latter case is impossible, because then we either have $2/d = \pm 1/d_1$ or, using $s \geqslant 2$, we would have that $P \geqslant pq^2$. If q appears in two places of (2) and $s \geqslant 3$, then again we obtain the inequality $P \geqslant pq^2$. Therefore $s = 2$. On replacing d_i by $-d_i$, if necessary, we see that (2) must be nothing else, but

$$(pq^2)^{1/d} = (q/p^{\ell-1})^{1/d_1}(q/p^{\ell'-1})^{1/d_2}$$

with integer ℓ, ℓ', d_1 and d_2. If either $\ell - 1$ or $\ell' - 1$ is negative, then $P \geqslant pq^2$, a contradiction. Recall that $p^e < q < p^{e+1}$. If ℓ or ℓ' is greater than or equal to $e + 3$, then $P \geqslant p^{e+2}q > pq^2$, a contradiction. Thus $1 \leqslant \ell, \ell' \leqslant e + 2$.

The system

$$2/d = 1/d_1 + 1/d_2, \quad -1/d = (\ell - 1)/d_1 + (\ell' - 1)/d_2$$

has no solutions if $\ell = \ell' \geqslant 1$. Let $\ell \neq \ell'$. Without loss of generality we assume that $\ell < \ell'$. The solution is

$$d_1 = (\ell' - \ell)d/(2\ell' - 1), \quad d_2 = -(\ell' - \ell)d/(2\ell - 1).$$

Let $g = (2\ell' - 1, d)$. If $g = 1$, then $\ell' - \ell$ is divisible by $2\ell' - 1$, which is impossible, because $\ell' - \ell < 2\ell' - 1$. Hence $g > 1$ and g is an odd number. If $f(d) > 2e + 3$, then $2\ell' - 1 > 2e + 3$ and so $\ell' > e + 2$, a contradiction (with $P < pq^2$).

Assume that $f(d) < 2e+3$, so that, as $f(d)$ is odd, $f(d) \leqslant 2e+1$. Consequently, there is an $\ell' \leqslant e+1$ such that $g = (2\ell' - 1, d) > 1$. Now, setting $\ell = (2\ell' - 1 + g)/2g$, which is a non-negative integer smaller than ℓ', we deduce that $d_1 = (g - 1)d/2g$, $d_2 = -(g - 1)d/2$ is a solution of the system, thus $P \leqslant q^2$. Since d is odd, we have that, for the function u of Section 5, $u(2/d) = 2$. From the inequality

$$\widehat{H}((pq^2)^{1/d}) = P \geqslant \max\{p^{u(1/d)}, q^{u(2/d)}\} = q^2,$$

we deduce that $P = q^2$ for $p < q$ and $f(d) \leqslant 2e + 1$. We have the first alternative of the theorem. If however $f(d) > 2e + 1$, then $P > q^2$.

2. Assume now that $f(d) = 2e+3$. Then $g = (2\ell' - 1, d) > 1$ only for $\ell' = e+2$, so that $g = 2\ell' - 1 = 2e + 3$. Thus $P \geqslant p^{e+1}q$. On the other hand, setting $u = 0$, we see again that $d_1 = (g - 1)d/2g$, $d_2 = -(g - 1)d/2$ is a solution of the system, thus $P \leqslant p^{e+1}q$. We deduce that $P = p^{e+1}q$ for $p < q$ and $f(d) = 2e + 3$, so we have the second alternative.

3. Let $p > q$. The power of p in P is at most 2, for otherwise $P \geqslant p^3 > pq^2$, a contradiction. Thus p must appear in at most two distinct places of the representation (2). If p appears in one place of (2), then its power is 1, for otherwise $P \geqslant p^2q > pq^2$. We have that (2) is

$$(pq^2)^{1/d} = (p/q^{\ell-1})^{1/d} q^{1/d_2}.$$

It follows that $d_2 = d/(\ell+1)$. As $q^e < p < q^{e+1}$, $\ell+1 \leqslant e+2$ so that, if $f(d) \leqslant e+2$, then there is an $\ell \in \mathbb{N}$ such that d_2 is an integer, giving $P \leqslant pq$. By Theorem 1(b), we have the inequality $P \geqslant \mathcal{R}(pq^2) = pq$. Thus $P = pq$, and we have the third alternative.

4. If $f(d) > e + 3$, then $\ell + 1 \geqslant f(d) \geqslant e + 4$, $\ell - 1 \geqslant e + 2$ and so once again $P \geqslant q^{e+3} > pq^2$. If $f(d) = e + 3$, then, setting $\ell = e + 2$, we see that d_2 is an integer, thus $P \leqslant q^{e+2}$. If p appears in two places of (2) and $s \geqslant 3$, then again we obtain the inequality $P \geqslant p^2q > pq^2$. Therefore $s = 2$. We still have that $P \geqslant p^2 > pq^2$, unless $e = 1$. However, the equalities $e = 1$ and $f(d) = e + 3$ cannot both hold at the same time, since $f(d)$ is prime. Thus p appears in one place of (2). Hence, if $P > pq$, then $P = q^{e+2}$ iff $f(d) = e + 3$ (fourth alternative). If $p > q$ and $f(d) > e + 3$, then we have the fifth alternative. The proof of Theorem 5 is now completed.

Proof of Theorem 6. Let $\beta = (pq)^{2/d}$. First note that, by (3), $\widehat{H}(\beta) \geqslant p^2$, where (without loss of generality) $p > q$. Also of course $\widehat{H}(\beta) \leqslant M(\beta) \leqslant (pq)^2$.

Hence with $q^e < p < q^{e+1}$, we see from Proposition 1 that the possible values of $\widehat{H}(\beta)$ are all integers in the range $\lfloor p^2, (pq)^2 \rfloor$ whose prime factors belong to $\{p, q\}$. It is straightforward to find out which these are. One way is the following: put $E = \log p/\log q$, $e = \lfloor E \rfloor \geqslant 1$. Then the numbers we want are $\{p^i q^j | i, j \in \{0\} \cup \mathbb{N}\}$, where $2E \leqslant Ei + j \leqslant 2E + 2$. Furthermore, since E is irrational, equality holds only if $i = 2$ and $j = 0$ or $j = 2$. The numbers we obtain are

$$p^2 < pq^{e+1} < q^{2e+2} < p^2 q < pq^{e+2} < q^{2e+3} < (pq)^2$$

if $p > q^{e+1/2}$, while if $p < q^{e+1/2}$ they are

$$p^2 < q^{2e+1} < pq^{e+1} < p^2 q < q^{2e+2} < pq^{e+2} < (pq)^2$$

Also, if $e = 1$, $\widehat{H}(\beta) = p^3$ is an extra possibility.

We now look at the possible α_i such that $\beta = \prod_k \alpha_k$ and $\widehat{H}(\beta) = \prod_k M(\alpha_k)$, for the possible values of $\widehat{H}(\beta)$ we have just found. We know from Proposition 1 that each α_k is of the form $(p^{i_k} q^{j_k})^{1/d_k}$, where $i = i_k \in \mathbb{Z}, j = j_k \in \mathbb{Z}$ and $d_k \in \mathbb{Z}^*$. Since d_k can be negative we can assume that $i \geqslant 0$. Next, as

$$M((p^i q^j)^{1/d_k}) \geqslant M((p^i)^{1/d_k}) M((q^j)^{1/d_k})$$

for $i, j \geqslant 0$, we can assume that there are no $\alpha_k = p^i q^j$ with both $i > 0, j > 0$ among the α_k. Thus every α_k is of the form $(p^i/q^j)^{1/d_k}$ with $i, j \geqslant 0$. We now list all such $(p^i/q^j)^{1/d_k}$ with $M((p^i/q^j)^{1/d_k}) \leqslant (pq)^2$. We write $\max\{p^i, q^j\} = p^I q^J$ and then record the vector $\langle i/d_k, -j/d_k, I, J \rangle$.

Type #	Range of j	$\langle i/d_k, -j/d_k, I, J \rangle$
1	$0 < j$	$\langle 0, -j/d_k, 0, j \rangle$
2	$0 \leqslant j \leqslant e$	$\langle 1/d_k, -j/d_k, 1, 0 \rangle$
3	$e \leqslant j$	$\langle 1/d_k, -j/d_k, 0, j \rangle$
4a	$0 \leqslant j \leqslant 2e + 1$	$\langle 2/d_k, -j/d_k, 2, 0 \rangle$
4b	$0 \leqslant j \leqslant 2e$	$\langle 2/d_k, -j/d_k, 2, 0 \rangle$
5a	$2e + 2 \leqslant j$	$\langle 2/d_k, -j/d_k, 0, j \rangle$
5b	$2e + 1 \leqslant j$	$\langle 2/d_k, -j/d_k, 0, j \rangle$
6	$j = 0$	$\langle 3/d_k, 0, 3, 0 \rangle$.

Here $d_k \in \mathbb{Z}^*$ and types 4a, 5a refer to the case $p > q^{e+1/2}$, while types 4b, 5b refer to the case $p < q^{e+1/2}$. In order to write $(pq)^{2/d}$ as a product of terms $(p^i/q^j)^{1/d_k}$ whose Mahler measures have product $p^a q^b$, we must find a (multi)set of vectors from the table above which sum to $\langle 2/d, 2/d, a, b \rangle$ for some choices of the integers $d_k \in \mathbb{Z}^*$.

We now consider whether or not the various possible values of $\widehat{H}(\beta)$ we have listed do really occur as a value of $\widehat{H}(\beta)$. Note that it is never possible for the

multiset of vectors to consist of just one element. So in what follows only the possible sets with at least two elements will be considered.

1. $\widehat{H}(\beta) = p^2$. In this case the only possible set of vectors is a set of two vectors of type 2, where the values of j for these vectors are $\ell - 1$, $\ell' - 1$ say, with $1 \leqslant \ell,\ \ell' \leqslant e + 1$. Requiring that these add to $\langle 2/d, 2/d, 2, 0\rangle$ gives the system of two equations

$$1/d_1 + 1/d_2 = 2/d, \quad (\ell - 1)/d_1 + (\ell' - 1)/d_2 = -2/d. \tag{5}$$

It follows that

$$d_1 = (\ell' - \ell)d/2\ell', \quad d_2 = (\ell - \ell')d/2\ell.$$

From this we have that, if $\ell' > \ell$ say, then $(d, \ell') > 1$, because $\ell' - \ell$ is not divisible by $2\ell'$. Hence $f(d) \leqslant e + 1$. Conversely, if $f(d) \leqslant e + 1$ we can take $\ell = 1$, $\ell' = f(d)$ which gives the identity

$$(pq)^{2/d} = p^{1/d_1}(p/q^{f(d)-1})^{1/d_2} \tag{6}$$

with $d_1, d_2 \in \mathbb{Z}^*$.

2. $\widehat{H}(\beta) = q^{2e+1}$ for $p < q^{e+1/2}$. Clearly only vectors of types 1, 3 can be used. We need at least two vectors of type 3 to get the first components adding to $2/d$. But then the sum of j_1 and j_2 would be at least $2e + 2$, which is of no use. Thus $\widehat{H}(\beta) = q^{2e+1}$ is impossible.

3. $\widehat{H}(\beta) = pq^{e+1}$. The only possibility for the set of vectors is one vector of type 2 and one of type 3 with $j = e+1$. The resulting equations are the same as in 1. The only difference is that here we must have $j = \ell' - 1 = e + 1$, giving $f(d) \leqslant e+2$ and hence $f(d) = e+2$, for otherwise case 1 would apply. Conversely, for $f(d) = e+2$ system (5) and its associated identity (6) apply, giving $\widehat{H}(\beta) = pq^{e+1}$.

4. $\widehat{H}(\beta) = q^{2e+2}$ for $p > q^{e+1/2}$. The only possibility is two vectors of type 3, each with $j = e + 1$. This gives (5), where $\ell = \ell' = e + 2$, with no solution at all.

5. $\widehat{H}(\beta) = p^2q$. First consider the case $p < q^{e+1/2}$. There are two possibilities for the set of vectors.

We can have one vector of type 1 with $j = 1$, and two of type 2. This gives the equations of Lemma 2, and we have a solution in this case iff $e + 1 \geqslant g^*(d)$. Then the solution of the equations gives the identity

$$(pq)^{2/d} = q^{1/d_1}(p/q^{\ell-1})^{1/d_2}(p/q^{\ell'-1})^{1/d_3},$$

showing that indeed $\widehat{H}(\beta) \leqslant p^2q$.

Alternatively, we can have one vector of type 1 with $j = 1$, and one vector of type 4b, giving

$$(pq)^{2/d} = q^{1/d_1}(p^2/q^{\ell-1})^{1/d}$$

with $\ell \leqslant 2e + 1$, so that $d_1 = d/(\ell + 1)$. Thus we have a solution with $\ell = f(d) - 1$ so long as $f(d) - 1 \leqslant 2e + 1$, or, as d is odd, $f(d) \leqslant 2e + 1$. But then, by Lemma 3(ii), $g^*(d) \leqslant (f(d) + 1)/2 \leqslant e + 1$.

Hence, if $g^*(d) \leqslant e+1$, in both cases we have a solution of the diophantine system always giving an example with $\widehat{H}(\beta) \leqslant p^2q$. Furthermore, we must have $e \leqslant f(d)-3$ so that $\widehat{H}(\beta) \neq p^2$ or pq^{e+1} or q^{2e+2} (see case 6 below).

Now consider the case $p > q^{e+1/2}$. The case with one vector of type 1 and two of type 2 is the same as for $p < q^{e+1/2}$. If we have one vector of type 1 and one of type 4a, then the only difference is that $\ell \leqslant 2e+2$ replaces $\ell \leqslant 2e+1$. This gives $f(d) \leqslant 2e+3$ and so, by Lemma 3(ii), $g^*(d)-1 \leqslant (f(d)-1)/2 \leqslant e+1$. If $g^*(d) \leqslant e+1$ then, as above, we have a solution. In addition to that, if $f(d) = 2e+3$, we have an additional solution with $\ell = 2e+2$ provided that $g^*(d) = e+2 = (f(d)+1)/2$.

6. $\widehat{H}(\beta) = q^{2e+2}$ for $p < q^{e+1/2}$. The possibility of two type 3 vectors can be ruled out as above. The only remaining possibility is with vectors of types 1 ($j_1 = 1$) and 5b ($j_2 = 2e+1$) giving $(pq)^{2/d} = q^{(2e+3)/d}(p^2/q^{2e+1})^{1/d}$. It follows that $(2e+3)|d$, so that $f(d) \leqslant 2e+3$. But if $f(d) < 2e+3$ then $f(d) \leqslant 2e+1$, $e+1 \geqslant (f(d)+1)/2 \geqslant g^*(d)$, and so $\widehat{H}(\beta) \leqslant p^2q$. Or, if $g^*(d) < (f(d)+1)/2$ then $g^*(d) \leqslant e+1$, so that again $\widehat{H}(\beta) \leqslant p^2q$. So we must have $f(d) = 2e+3 \geqslant 5$ and $g^*(d) = (f(d)+1)/2 \geqslant 3$ in this case, giving also $e = g^*(d)-2$.

7. $\widehat{H}(\beta) = pq^{e+2}$. There are two possibilities. We can have one type 2 vector and one type 3 vector with $j = e+2$, giving system (5). As in case 1, there is an integer solution if $f(d) \leqslant e+3$, whereas our case is $(f(d)+1)/2 > g^*(d) = e+2 \geqslant f(d)-1$, a contradiction with $f(d) \geqslant 3$. The second possibility is one type 1 vector with $j_1 = 1$, one type 2 vector and one type 3 vector with $j_2 = e+1$ the resulting equations being those of Lemma 2 with $\ell \leqslant e+1$ and $\ell' = e+2$, giving $e+2 \geqslant g^*(d)$. As $\widehat{H}(\beta) > q^{2e+2}$, we must therefore have $e = g^*(d)-2$ and $g^*(d) < (f(d)+1)/2$ in this case.

8. $\widehat{H}(\beta) = q^{2e+3}$ for $p > q^{e+1/2}$. We have three possibilities. Two type 3 vectors with $j_1 = e+1$ and $j_2 = e+2$ are impossible, because d is odd. One type 1 vector with $j_1 = 1$ and one type 5a vector with $j_2 = 2e+2$ is also impossible, as it gives $2e+1 = 2$. Finally, one type 1 vector with $j_1 = 1$ and two type 3 vectors with $j_2 = j_3 = e+1$, giving equations of Lemma 2 with $\ell = \ell' = e+2$, which is impossible.

9. $\widehat{H}(\beta) = p^3$ and $e = 1$. There are just two possibilities. Firstly we can have three type 2 vectors, with resulting equations

$$1/d_1 + 1/d_2 + 1/d_3 = 2/d, \quad (1-\ell)/d_1 + (1-\ell')/d_2 + (1-\ell'')/d_3 = 2/d$$

for $1 \leqslant \ell, \ell', \ell'' \leqslant e+1 = 2$ and, as usual, the $d_i \in \mathbb{Z}^*$. One easily shows that it is not possible to have none, or one, or two, or three of ℓ, ℓ', ℓ'' equal to 2.

The alternative is one type 2 vector and one type 4a or 4b vector. This gives the equations

$$1/d_1 + 2/d_2 = 2/d, \quad (1-\ell)/d_1 + (1-\ell')/d_2 = 2/d$$

for $1 \leqslant \ell \leqslant e+1 = 2$ and $1 \leqslant \ell' \leqslant 2e+2 = 4$and the $d_i \in \mathbb{Z}^*$. Then, if $\ell = 1$, we have $\ell' = 3$ and $d_2 = -d$, in which case $1/d_1 = 4/d$, which is impossible. If $\ell' = 2$, then, adding the equations we have $(3-\ell')/d_2 = 2/d$, again impossible as $\ell' \leqslant 4$.

If none of the conditions for 1-9 to hold are valid, then $e \leqslant g^*(d) - 3$, in which case we have $\widehat{H}(\beta) = (pq)^2$ as a result of the trivial identity $(pq)^{2/d} = p^{1/d}p^{1/d}q^{1/d}q^{1/d}$. This completes the proof of Theorem 6.

9. Finding the metric naïve height of a surd

In this section we describe an algorithm for finding $\widehat{H}\big((m/n)^{1/d}\big)$. To do this, we first need an algorithm for deciding whether a system of linear equations has a solution in positive or negative unit fractions.

Let $A = \|a_{ij}\|$ be an $r \times s$ matrix with rational entries, and $\mathbf{b} = (b_i) \in \mathbb{Q}^r$. Denote by $\mathbb{Z}_0^{-1}$ the set

$$\mathbb{Z}_0^{-1} = \{0\} \cup \{1/q \mid q \in \mathbb{Z}^*\}.$$

of positive and negative unit fractions, augmented by 0.

PROPOSITION 2. *There is a finite algorithm which, given an equation $A\mathbf{x} = \mathbf{b}$, either finds a solution $\mathbf{x} \in (\mathbb{Z}_0^{-1})^s$ or shows that the equation has no such solution.*

PROOF. The algorithm is trivial if $s = 1$. So we can assume that $s > 1$, and proceed by recursion. If $\mathbf{b} = \mathbf{0}$ we can take $\mathbf{x} = \mathbf{0}$. Thus we can assume that some $b_i \neq 0$. If for any i with $b_i \neq 0$ we have $a_{ij} = 0$ for all j, then clearly there is no solution. So we can assume that for each i with $b_i \neq 0$ there is a j with $a_{ij} \neq 0$. Then from $\sum_j a_{ij}x_j = b_i$ we get $\max_j |x_j| \geqslant B := \max_{i:b_i\neq 0} \quad |b_i|/\sum_j |a_{ij}|$. Hence there is some $j \in \{1, \ldots, s\}$ for which $x_j = \pm 1/\ell$, where $\ell \in \{1, 2, \cdots, \lfloor 1/B \rfloor\}$. For each such particular j and x_j, we have an $(s-1)$-variable equation of the same type, which, by recursion, has an algorithm of the kind we are seeking. Then we obtain a solution to the s-variable equation from any of the $(s-1)$-variable equations having a solution. If none of these $(s-1)$-variable equations has a solution, then neither does the s-variable equation. This completes the proof.

Of course we do not claim that the algorithm given here is particularly efficient. It could clearly be refined in several ways to be made into a more practical algorithm.

We can now describe an algorithm for finding $\widehat{H}\big((m/n)^{1/d}\big)$. By allowing d to also be negative, we can assume that $m > n \geqslant 1$.

1. The first step in the algorithm is to find the list $\mathcal{F}$ defined as

$$\mathcal{F} = \langle q \in \mathbb{N} \mid q_0 \leqslant q \leqslant m \quad \text{and all prime factors of } q \text{ divide } mn \rangle,$$

with elements in ascending order. Here $\mathcal{F}$ is the list of all possible values taken by $\widehat{H}\big((m/n)^{1/d}\big)$, with q_0 the lower bound for $\widehat{H}\big((m/n)^{1/d}\big)$ of equation (3).

This is essentially a question of finding all integer lattice points in a bounded region, so it can clearly be done in finite time. Specifically, if the p_ℓ with say $\ell \in L$

are the prime factors of mn, then we need to find all non-negative integers e_ℓ with $\sum_{\ell\in L} e_\ell \log p_\ell \leqslant \log(mn)$. Then the values of q are the numbers $\prod_{\ell\in L} p_\ell^{e_\ell}$. Since we know (see Section 5) that the metric naïve height of our surd $(m/n)^{1/d}$ is an integer $\leqslant m$ all of whose prime factors divide mn, we have $\widehat{H}\big((m/n)^{1/d}\big) \in \mathcal{F}$.

2. Now step through each $q \in \mathcal{F}$ in turn, until SUCCESS (described below) is achieved. For the given q, find all factorizations $q = m_1 \dots m_s$ with say $1 < m_1 \leqslant \dots \leqslant m_s$.

3. Given such a factorization, find all $n_1, \dots, n_s$ such that $n_j < m_j$ and

$$(m_j, n_j) = 1, \quad j = 1, \dots, s,$$

while also $\mathcal{R}(\prod_j m_j n_j) = \mathcal{R}(mn)$.

4. For given q, m_j and n_j try to solve the equation

$$(m/n)^{1/d} = (m_1/n_1)^{1/d_1} \dots (m_s/n_s)^{1/d_s}$$

in non-zero integers $d_1, \dots, d_s$. If $m/n = \prod_{\ell\in L} p_\ell{}^{f_\ell}$ and $m_j/n_j = \prod_{\ell\in L} p_\ell{}^{f_{\ell j}}$, then this equation is equivalent to the system of equations

$$f_\ell/d = \sum_{j=1}^{s} f_{\ell j}/d_j, \quad \ell \in L.$$

By Proposition 2, we can either find a solution to this system, or show that there is no solution. Note that although that the proposition allows solutions with some zero components, this cannot happen here, as it would imply that there was a solution with all components non-zero, but for a smaller value of q, whereupon the algorithm would already have terminated.

(Note that, as a small refinement, we can assume that no m_j/n_j is a perfect power. This is because any solution with $m_j/n_j = (m'_j/n'_j)^t$, $t > 1$, would have $(m_j/n_j)^{1/d_j} = ((m'_j/n'_j)^{1/d_j})^t$, so that in the factorization $q = \prod_j m_j$ we can replace m_j by t copies of m'_j, with associated n'_j.)

5. If a solution is obtained in 4, then SUCCESS! Stop, with $\widehat{H}\big((m/n)^{1/d}\big) = q$. Otherwise, repeat with different $\{n_j\}$, then with a different factorization $\{m_j\}$ of q, then for the next value of $q \in \mathcal{F}$.

Note that SUCCESS is eventually guaranteed thanks to the trivial factorization $m = m$, giving the largest possible value $H\big((m/n)^{1/d}\big) = m$.

ACKNOWLEDGEMENT. The research of the first named author was partially supported by the Lithuanian State Science and Studies Foundation.

REFERENCES

[1] A. Dubickas, Mahler measures close to an integer, *Canadian Math. Bull.* **45** (2002), 196–203.

[2] A. Dubickas and C. J. Smyth, On the Remak height, the Mahler measure, and conjugate sets of algebraic numbers lying on two circles, *Proc. Edinburgh Math. Soc.* **44** (2001), 1–17.

[3] A. Dubickas and C. J. Smyth, On the metric Mahler measure, *J. Number Th.* **86** (2001), 368–387.

[4] M. Waldschmidt, *Diophantine Approximation on Linear Algebraic Groups*, Springer-Verlag, Berlin, New York, 2000.

(Received: April 6, 2002)

A. Dubickas
Department of Mathematics and Informatics
Vilnius University
Naugarduko 24, Vilnius 2600
Lithuania
E-mail: arturas.dubickas@maf.vu.lt

C. J. Smyth
School of Mathematics,
Edinburgh University, J.C.M.B.
King's Buildings
Mayfield Road, Edinburgh EH9 3JZ
Scotland, UK
E-mail: chris@maths.ed.ac.uk