Illinois Journal of Mathematics
Volume 46, Number 1, Spring 2002, Pages 63–80
S 0019-2082

IDEAL CONSTRUCTIONS AND IRRATIONALITY MEASURES OF ROOTS OF ALGEBRAIC NUMBERS

PAULA B. COHEN AND ALFRED J. VAN DER POORTEN

ABSTRACT. This paper addresses the problem of determining the best results one can expect using the Thue-Siegel method as developed by Bombieri in his equivariant approach to effective irrationality measures to roots of high order of algebraic numbers, in the non-archimedean setting. As an application, we show that this method, under a non-vanishing assumption for the auxiliary polynomial which replaces the appeal to Dyson's Lemma type arguments and together with a version of Siegel's Lemma due to Struppeck and Vaaler, yields a result comparable to the best results obtained to date by transcendence methods.

1. Introduction

This paper is motivated by recent work of van der Poorten [18] on some conjectures of Bombieri, Hunt and van der Poorten [8]. It addresses the question of the best result one can hope for with the Thue-Siegel method as developed by Bombieri [4] in his equivariant approach to effective approximations to roots of high order of algebraic numbers. To simplify the computations, we shall in fact work in the non-archimedean situation as in [6], but the auxiliary construction will be a polynomial and not an interpolation determinant. One should be able to treat the archimedean case in a similar manner.

We obtain in our Theorem 4.2 an analogue of Theorem 1 of [6] by applying the Thue-Siegel principle using $(\alpha, 1)$ as anchor pair, where α is an r-th root of a non-zero number a in an algebraic number field K. The new feature is that we assume the non-vanishing of the auxiliary polynomial in two variables at a point $(\alpha, \alpha\gamma^{-1})$, $\gamma \in K$, which is well-approximated in an appropriate non-archimedean valuation by $(1, 1)$. We thereby forgo any appeal to a Dyson's Lemma type argument. We do not introduce powers of the anchor pair as in [6] as the gain in Dyson's Lemma by having more points no longer applies.

Received December 18, 2000; received in final form November 23, 2001.
2000 *Mathematics Subject Classification.* 11J99.
The first author acknowledges support from a Macquarie University Research Grant and from the Ellentuck Fund at the Institute for Advanced Study, Princeton. The second author acknowledges support from a grant from the Australian Research Council.

Our Theorem 4.2 represents in the above sense the limit of the method of [6]. Our auxiliary construction is universal in the sense that it depends only on r and not on α, so that the non-vanishing assumption at $(\alpha, \alpha\gamma^{-1})$ seems not unreasonable to attain. Nonetheless, experience has shown it to be elusive and to represent one of the main technical difficulties of the Thue-Siegel method.

Combining Theorem 4.2 with a version of Siegel's Lemma due to Struppeck and Vaaler [16], we obtain in Theorem 6.2 a result showing that under this non-vanishing assumption our application of the Thue-Siegel method can yield effective irrationality measures for roots of high order of algebraic numbers comparable to the best results obtainable to date by transcendence techniques.

We also make some comparisons to recent work of Bennett [3] and of Bombieri-Cohen [7]. The method of [3] uses a so-called "almost-perfect" construction derived from Padé approximation techniques for which the non-vanishing assumption is immediate. We show that even Bennett's conjectured bounds for the height of the resulting auxiliary polynomial are too weak to allow this method to be applied to our situation. This justifies the less than almost-perfect method applied in [7] whereby requiring less vanishing of the auxiliary construction at $(1, 1)$ enables one to reduce the height of the auxiliary polynomial.

2. The main results

We use the same notation as in [6]. Therefore, if K is a number field, then the absolute values $|\ |_v$ in K are normalised by requiring that, for $x \in K$,

$$|x|_v = \|x\|_v^{d_v/d},$$

where $[K_v : \mathbb{Q}_v] = d_v$, $[K : \mathbb{Q}\,] = d$ and where $\|x\|_v$ is the unique extension to the completion K_v of the ordinary real or p-adic absolute value in $\mathbb{Q}_v$. For a vector $\mathbf{x} = (x_1, \ldots, x_m)$ in K^m and a place $v \in M_K$, we define

$$|\mathbf{x}|_v = \max(|x_1|_v, \ldots, |x_m|_v).$$

The (homogeneous) absolute height of $\mathbf{x}$ is defined as

$$H(\mathbf{x}) = \prod_{v \in M_K} |\mathbf{x}|_v\,.$$

The logarithmic absolute height of $\mathbf{x}$ is then defined as

$$h(\mathbf{x}) = \log H(\mathbf{x}).$$

For $x \in K$ we denote by $H(x)$ the height of the vector $(1, x) \in K^2$, so that

$$H(x) = \prod_{v \in M_K} \max(1,\, |x|_v).$$

The logarithmic absolute height of x is then given by $h(x) = \log H(x)$.

This height definition may be further extended to polynomials P in several variables with coefficients in K by taking $H(P)$ to be the absolute height of the vector of all the coefficients of P, with the corresponding logarithmic absolute height $h(P) = \log H(P)$. For $v \in M_K$, we let $|P|_v$ be the maximum of the v-adic valuations of the coefficients of P.

Let $v \in M_K$ and $v \big| p$ where p is a rational prime. Denote by f_v the residue class degree and by e_v the ramification index of the extension $K_v/\mathbb{Q}_v$. Let a be a non-zero element of K which is not a root of unity. Suppose that $|a-1|_v < 1$. Let r be a positive integer coprime with p. Then a has an r-th root $\alpha \in K_v$ satisfying $0 < |\alpha - 1|_v < 1$. We wish to obtain an effective irrationality measure $\mu > 0$ for α of the form

$$|\alpha\gamma^{-1} - 1|_v \geq c(\alpha) H(\gamma)^{-\mu}, \qquad \text{for all } \gamma \in K \setminus \{0\}. \tag{2.1}$$

Here, the positive constant $c(\alpha)$ is effectively computable and may depend on α but not on γ.

In making the auxiliary construction in §3, we assume the existence of a non-zero polynomial $P = P(x,y)$ with rational integer coefficients, which vanishes to high order at all $(\varepsilon, \varepsilon)$ with $\varepsilon^r = 1$.

Let $\gamma \in K, \gamma \neq 0$. We make the further strong assumption that we can ensure $P(\alpha, \alpha') \neq 0$, where $\alpha' = \alpha\gamma^{-1}$. This replaces the Dyson's Lemma type arguments of [4, p. 70] and [6, p. 209]. We then examine the dependence of the above irrationality measure μ on the logarithmic absolute height $h(P)$ of P. In particular, if N_1 is the degree in x and N_2 the degree in y of P, we know by existing versions of Siegel's Lemma (see §6) that we can choose P in such a way that we have an upper bound for its height of the form

$$h(P) \leq N_1 l_1 + N_2 l_2\,, \tag{2.2}$$

for N_1 and N_2 sufficiently large. Here l_1 and l_2 are positive, finite and independent of N_1 and N_2, although of course they in general may depend on the other parameters of the problem, and in particular on α and α'. We assume from now on an upper bound for $h(P)$ of the form (2.2).

We show in Theorem 1 of §4 that, under the above assumptions, we can obtain using the equivariant Thue-Siegel method an effective irrationality measure of the form

$$|\alpha' - 1|_v \geq \{\exp(l_2) H(\alpha')\}^{-\mu}, \tag{2.3}$$

where

$$\mu = \frac{4}{2-\delta} \cdot \frac{1}{\Lambda} \cdot (h(a) + r l_1), \tag{2.4}$$

and

$$\Lambda = \log |\alpha - 1|_v^{-1}. \tag{2.5}$$

The parameter $0 < \delta < 2$ measures the amount of vanishing imposed on $P(x,y)$ at $(x,y) = (\varepsilon, \varepsilon), \varepsilon^r = 1$ (see §3). The higher the vanishing, the smaller is δ, so that the overdetermined situation would correspond to $\delta = 0$. In general, the quantity $(2-\delta)^{-1}$ can be easily controlled. Intrinsic to the expression for the irrationality measure is an "anchor condition" as in previous papers (see [8]), whereby the quality of the effective irrationality measure μ depends on how close we can take α to 1 with respect to the valuation v, that is, on how large is Λ. This is of course the essence of the Thue-Siegel Principle. Indeed, regardless of the quality of the bound for $h(P)$ (that is, even supposing that up to constants l_1 is bounded above by $h(\alpha)$ and l_2 is bounded above by $h(\alpha'))$, the expression (2.4) shows that the best irrationality measure we can hope for with these methods would be of the form, for some absolute constant $c > 0$,

$$|\alpha\gamma^{-1} - 1|_v \geq c(\alpha)H(\gamma)^{-ch(a)/\Lambda}. \tag{2.6}$$

In §6 we use results of Struppeck and Vaaler [16] to obtain estimates for $l_1(P)$ and $l_2(P)$ which show that the assumption $P(\alpha, \alpha') \neq 0$ allows one to obtain, with existing versions of Siegel's Lemma, an irrationality measure in (2.1) of the form

$$\mu = c' h(a)(D_v^*)^3 \log\left(\frac{r}{h(a)} + 1\right), \tag{2.7}$$

for some absolute constant $c' > 0$ and $D_v^* = \max(1, d/f_v \log p)$. This same irrationality measure (up to the constants depending on K and v) is the best obtainable from current methods using linear forms in logarithms. The archimedean analogue of the fact that this follows using logarithmic forms was announced in a 1994 lecture of Baker [1], and both the archimedean and non-archimedean logarithmic forms proof is worked out in more recent work of Bugeaud [11]. For a general treatment of archimedean forms in logarithms see [2], and of non-archimedean forms in logarithms see [19], [20].

In §5 we derive a version of the Thue-Siegel Lemma as applied to our situation using a construction derived from [3]. By design, this yields an auxiliary polynomial satisfying the non-vanishing assumption at (α, α') but at the cost of having a height which seems unreasonably large for applications. Finally, in §7 we end with some remarks about the motivating work [18] of van der Poorten.

3. The auxiliary construction

For $I = (i_1, i_2) \in \mathbb{Z}_{\geq 0}$, set D^I for the partial derivative

$$D^I = \frac{1}{i_1!}\frac{1}{i_2!}\frac{\partial^{i_1}}{\partial x^{i_1}}\frac{\partial^{i_2}}{\partial y^{i_2}},$$

which is to act on polynomials in $\mathbb{C}[x, y]$. Let L be a field of characteristic 0, and suppose that $(\beta_1, \beta_2) \in L^2$. For real $M_1,\ M_2 > 0$ and a polynomial $P(x, y) \in L[x, y]$, the index of P at (β_1, β_2) relative to (M_1, M_2) is defined as

$$\mathrm{ind}_{(\beta_1,\beta_2)}(P; M_1, M_2) = \min\left\{\frac{i_1}{M_1} + \frac{i_2}{M_2} \ \middle|\ D^I P(\beta_1, \beta_2) \neq 0\right\}.$$

Let N_1 and N_2 be positive integers. In what follows, our estimates are true for N_1 and N_2 sufficiently large. We suppose further that for any fixed choice of a finite positive real number z we can ensure that

$$\lim_{N_1, N_2 \mapsto \infty} N_1/N_2 = z. \tag{3.1}$$

This should pose no problem, as our discussion should go through quite generally; in particular it is to be expected that our postulated upper bound (2.2) for $h(P)$ should be derivable for general large N_1 and N_2, as are the other results of this paper.

THEOREM 3.1 (Box Principle Lemma). *Let* $0 < \theta_i \leq 1$, $i = 1, 2$, *and* $0 < \delta < 2$. *Let* $T = \frac{1}{2}\theta_1\theta_2$ *satisfy* $rT = (1 - \frac{1}{2}\delta)$. *Then there is a polynomial* $P \in \mathbb{Z}[x, y]$, $P \neq 0$, *with* $deg_x P \leq N_1$ *and* $deg_y P \leq N_2$, *and with index*

$$\mathrm{ind}_{(\varepsilon,\varepsilon)}(P; \theta_1 N_1, \theta_2 N_2) \geq 1$$

at every point $(\varepsilon, \varepsilon)$, $\varepsilon^r = 1$, *for* N_1 *and* N_2 *sufficiently large.*

Proof. The vanishing requirement gives rise to a system of

$$\frac{1}{2} r\theta_1\theta_2 N_1 N_2 + O\big(\max(N_1, N_2)\big)$$

homogeneous linear equations over the rationals in $(N_1+1)(N_2+1)$ unknowns, namely the coefficients of P. As $rT < 1$, the lemma follows by basic linear algebra. □

We let

$$rT = \left(1 - \frac{1}{2}\delta\right), \qquad 0 < \delta < 2 \tag{A.1}$$

denote the assumption of the above lemma, and suppose that there exists a polynomial P as in the Box Principle Lemma with

$$P(\alpha, \alpha') \neq 0. \tag{A.2}$$

As remarked in §2, the assumption (A.2) is of course very strong, replacing in one fell swoop a Dyson's Lemma type argument, as in [4, p. 70] and [6, p. 209]. It also considerably simplifies the computations.

4. The Thue-Siegel Principle

Continuing with the situation of §3, let P be a polynomial as in the Box Principle Lemma, with logarithmic absolute height $h(P)$ bounded as in (2.2), and which satisfies (A.2). Let $L = K(\alpha, \zeta)$ where ζ is a primitive r-th root of unity. We recall the product formula in the form

$$\sum_{w \nmid v} \log |P(\alpha, \alpha')|_w = -\sum_{w | v} \log |P(\alpha, \alpha')|_w,$$

the sum being over the valuations w of L. Following standard practice, we shall estimate the left hand side trivially and the right hand side using the vanishing to high order of P at the points $(\varepsilon, \varepsilon)$ with $\varepsilon^r = 1$.

If $w \nmid v$ and $w \nmid \infty$ then we have

$$\log |P(\alpha, \alpha')|_w \leq \log |P|_w + N_1 \max(1, |\alpha|_w) + N_2 \max(1, |\alpha'|_w). \tag{4.1}$$

If $w \nmid v$ and $w | \infty$ then we have

$$\log |P(\alpha, \alpha')|_w \leq \log |P|_w + N_1 \max(1, |\alpha|_w) + N_2 \max(1, |\alpha'|_w) + O(\log N_1 N_2). \tag{4.2}$$

If $w | v$, then we expand P in a Taylor series around $(\varepsilon, \varepsilon)$ where $\varepsilon^r = 1$ and ε will be chosen suitably. We have, with $J = (j_1, j_2)$,

$$P(\alpha, \alpha') = \sum_{J \notin G} D^J P(\varepsilon, \varepsilon)(\alpha - \varepsilon)^{j_1} (\alpha' - \varepsilon)^{j_2}$$

where

$$G = \left\{ (i_1, i_2) : \frac{i_1}{\theta_1 N_1} + \frac{i_2}{\theta_2 N_2} < 1 \right\}.$$

As $w \nmid \infty$, the binomial coefficients caused by the differentiation in Taylor's formula do not contribute and we obtain, for an appropriate choice of ε (see [6]),

$$\begin{aligned} \log |P(\alpha, \alpha')|_w &\leq \log |P|_w + \max_{(j_1, j_2) \notin G} \{ j_1 \log |\alpha - \varepsilon|_w + j_2 \log |\alpha' - \varepsilon|_w \} \\ &\leq \log |P|_w - \min \left(N_1 \theta_1 \log |\alpha - \varepsilon|_w^{-1}, N_2 \theta_2 \log |\alpha' - \varepsilon|_w^{-1} \right) \\ &= \log |P|_w - \widetilde{\delta}_w \min \left(N_1 \theta_1 \log |\alpha - 1|_v^{-1}, N_2 \theta_2 \log |\alpha' - 1|_v^{-1} \right) \end{aligned}$$

where $\widetilde{\delta}_w = [L_w : K_v]/[L : K]$. As $\sum_{w|v} \widetilde{\delta}_w = 1$, we have from the product formula that

$$\begin{aligned} &\min(N_1 \theta_1 \log |\alpha - 1|_v^{-1}, N_2 \theta_2 \log |\alpha' - 1|_v^{-1}) \\ &\qquad \leq h(P) + N_1 h(\alpha) + N_2 h(\alpha') + O(\log N_1 N_2). \end{aligned} \tag{4.3}$$

Let

$$\Lambda = \log |\alpha - 1|_v^{-1}, \qquad \Lambda' = \log |\alpha' - 1|_v^{-1}. \tag{4.4}$$

We choose

$$z = \frac{\theta_2 \Lambda'}{\theta_1 \Lambda}. \tag{4.5}$$

Dividing by N_2, taking the limit when N_1, N_2 go to infinity and multiplying by $\theta_1 \Lambda$ in (4.3), we have

$$\theta_1 \theta_2 \Lambda \Lambda' \leq \theta_1 \Lambda \lim_{N_2 \to \infty} \frac{h(P)}{N_2} + \theta_2 \Lambda' h(\alpha) + \theta_1 \Lambda h(\alpha').$$

Then, using (2.2) we deduce

$$\theta_1\theta_2\Lambda\Lambda' \leq \theta_2\Lambda'\big(l_1 + h(\alpha)\big) + \theta_1\Lambda\big(l_2 + h(\alpha')\big). \tag{4.6}$$

Therefore, if

$$\theta_1\Lambda \geq 2\{l_1 + h(\alpha)\}, \tag{4.7}$$

we deduce from (4.6) that,

$$\theta_2\Lambda' \leq 2\{l_2 + h(\alpha')\}. \tag{4.8}$$

The above computations may be summarised as follows.

THEOREM 4.1 (Thue-Siegel Lemma). *Let $0 < \theta_i \leq 1, i = 1, 2$, and $0 < \delta < 2$ with $r\theta_1\theta_2 = 2 - \delta$. Suppose that there exists a polynomial P as in the Box Principle Lemma, with $h(P)$ bounded above as in* (2.2)*, and which satisfies* (A.2)*. Then, under the anchor condition*

$$|\alpha - 1|_v \leq \{\exp(l_1)H(\alpha)\}^{-2/\theta_1}, \tag{4.9}$$

we have

$$|\alpha' - 1|_v \geq \{\exp(l_2)H(\alpha')\}^{-2/\theta_2}.$$

We may rewrite (A.1) as

$$2/\theta_2 = \big(2/(2-\delta)\big)r\theta_1. \tag{4.10}$$

We can rewrite condition (4.9) of the Thue-Siegel Lemma as

$$r\theta_1\Lambda \geq 2\{rl_1 + h(a)\}. \tag{4.11}$$

Setting

$$\mu = 2/\theta_2,$$

from (4.10) and (4.11) we deduce that we can take

$$\mu = \frac{4}{2-\delta} \cdot \frac{1}{\Lambda} \cdot \max(h(a) + rl_1).$$

We therefore have the following result.

THEOREM 4.2 (Theorem 1). *Suppose that there exists a polynomial P as in the Box Principle Lemma, with $h(P)$ bounded above as in* (2.2)*, and which satisfies* (A.2)*. Then we have*

$$|\alpha' - 1|_v \geq \{\exp(l_2)H(\alpha')\}^{-\mu},$$

where

$$\mu = \frac{4}{2-\delta} \cdot \frac{1}{\Lambda} \cdot (h(a) + rl_1).$$

Notice that if for a fixed $0 < \delta < 2$ we have an ideal height estimate for our auxiliary construction of the form

$$h(P) \leq c_3 N_1 h(\alpha) + c_4 N_2 h(\alpha'),$$

then the Thue-Siegel Lemma implies a resulting irrationality measure for α of the form

$$\log \frac{1}{|\alpha - 1|_v} \cdot \log \frac{1}{|\alpha' - 1|_v} \leq c_5 r h(\alpha) h(\alpha'). \tag{4.12}$$

5. Some comparisons between Padé techniques and the Thue-Siegel Principle

In this section we make some comparative remarks about an old construction of Mahler, recently reapplied in [3] to obtain irrationality measures to roots of rational numbers, and the approach of [7] which exploits an equivariant Thue-Siegel principle to obtain general irrationality measures to roots of algebraic numbers. Both Mahler's construction and a limiting case of the construction of [7] can be seen as special cases of the auxiliary construction of the Box Principle Lemma of §3 of the present paper.

In [6], an auxiliary polynomial $P(x, y) \in \mathbb{Z}[x, y]$ is constructed which vanishes to high order at all $(\varepsilon, \varepsilon)$ with $\varepsilon^r = 1$ as in the Box Principle Lemma. The irrationality measure for α is then obtained by working with the product formula applied to the algebraic number $P^*(\alpha, \alpha')$, where P^* is an appropriate derivative of P chosen using Dyson's Lemma. In [7], by only working with derivatives with respect to x, we were able to replace the two variable Dyson's Lemma argument by a Wronskian argument, thereby rendering the method completely elementary. This in fact leads to a better irrationality measure than that of [6]. Of course in these approaches we do not assume (A.2) of §3.

It is not difficult to see that the auxiliary construction of [7] can be derived from that of [6], and for convenience we now explain why. We take a polynomial $P \in \mathbb{Z}[x, y]$ and write it as

$$P(x, y) = \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} a(j_1, j_2) x^{j_1} y^{j_2}. \tag{5.1}$$

We now apply the vanishing condition at $(\varepsilon, \varepsilon)$ of the Box Principle Lemma for the case $\theta_2 = 1/N_2$ and $\theta_1 = k/N_1$ where k is an integer and $rk < N_1 N_2$, so that we require

$$D^{(l,0)} P(\varepsilon, \varepsilon) = 0, \qquad l = 0, \dots, k-1;\ \varepsilon^r = 1. \tag{5.2}$$

From (5.1), we can write this as

$$\sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} \binom{j_1}{l} a(j_1, j_2) \varepsilon^{j_1 - l} \varepsilon^{j_2} = 0, \qquad l = 0, \dots, k-1;\ \varepsilon^r = 1. \tag{5.3}$$

Multiplying, for each ε with $\varepsilon^r = 1$ and each $t = 0, \ldots, r-1$, the above equation by ε^{l-t} and averaging over all ε we derive the kr new equations,

$$\sum_{j_1+j_2 \equiv t \bmod r} \binom{j_1}{l} a(j_1, j_2) = 0, \qquad l = 0, \ldots, k-1;\ t = 0, \ldots, r-1, \tag{5.4}$$

where in the above sum $0 \le j_1 \le N_1$, $0 \le j_2 \le N_2$. Let $N_2 = s < r$, $N_1 = nr + s$ and consider the equation with $t = s$, namely,

$$\sum_{j_1+j_2 \equiv s \bmod r} \binom{j_1}{l} a(j_1, j_2) = 0, \qquad l = 0, \ldots, k-1, \tag{5.5}$$

where in the above sum $0 \le j_1 \le nr + s$, $0 \le j_2 \le s$. In the range of this sum, we have therefore that j_1 is of the form $ri + s - j$ for $j = j_2 = 0, \ldots, s$ and $i = 0, \ldots, n$. This is equivalent to constructing an auxiliary polynomial of the form

$$Q(x, y) = \sum_{j=0}^{s} A_j(x^r) x^{s-j} y^j, \tag{5.6}$$

where the $A_j(x)$ are polynomials of degree at most n, such that $Q(x, 1)$ vanishes to order k at $x = 1$.

The auxiliary construction of [7] (with the parameter $l = 1$ of that paper) has the form (5.6) with $k < (s+1)(n+1)$, and with the coefficients of the polynomials A_j rational numbers (which will in general have controlled denominators in what follows). The condition that $Q(x, 1)$ vanish to order $k < (n+1)(s+1)$ at $x = 1$ is equivalent to the construction of a Padé approximation to the algebraic function $(1-z)^{1/r}$. Indeed, there is a polynomial $H(x)$ such that

$$Q(x, 1) = (1-x)^k H(x). \tag{5.7}$$

Consider the multivalued function $u(z) = (1-z)^{1/r}$. We have (after an appropriate choice of branch) a formal power series expansion of $u(z)$ around $z = 0$, convergent in the disc $|z| < 1$, as follows:

$$u(z) = 1 + \sum_{i=1}^{\infty} (-1)^i \binom{1/r}{i} z^i. \tag{5.8}$$

Substituting $x = u(z)$ into (5.6), we have

$$Q(u(z), 1) = \left(\sum_{i=1}^{\infty} (-1)^{i+1} \binom{1/r}{i} z^i \right)^k H\big(u(z)\big),$$

which is divisible by z^k in $\mathbb{C}[[z]]$. That is, the formal power series $Q(u(z), 1)$ converges in the disc $|z| < 1$ and has a zero of order k at $z = 0$. Setting

$B_j(z) = A_j(1-z)$, we have a Padé approximation for $(1-z)^{1/r}$,

$$R(z) := Q(u(z), 1) = \sum_{j=0}^{s} B_j(z) u(z)^{s-j}. \tag{5.9}$$

Conversely, suppose that we have a solution to the Padé approximation problem as in (5.9). Then,

$$\frac{1}{l!}\frac{d^l}{dz^l}\left(\sum_{j=0}^{s} B_j(z)u(z)^{s-j}\right)\Bigg|_{z=0} = 0, \quad l = 0, \ldots, k-1,$$

which implies, on changing variables to $x = u(z)$,

$$\frac{1}{l!}\left(\frac{1}{rx^{r-1}}\frac{d}{dx}\right)^l \left(\sum_{j=0}^{s} A_j(x^r)x^{s-j}\right)\Bigg|_{x=1} = 0, \qquad l = 0, \ldots, k-1.$$

This gives back inductively the system,

$$\frac{1}{l!}\frac{d^l}{dx^l}\left(\sum_{j=0}^{s} A_j(x^r)x^{s-j}\right)\Bigg|_{x=1} = 0, \qquad l = 0, \ldots, k-1.$$

We understand the almost perfect situation, corresponding to the case $k = (n+1)(s+1)-1$, for the above Padé approximation problem. The functions $(1-z)^{j/r}$ for $j = 0, \ldots, s$ are normal at $(n, n, \ldots, n) \in \mathbb{Z}^{s+1}$ and, up to a constant, the remainder function $R(z)$ is uniquely determined as

(5.10)

$$R(z) = \sum_{j=0}^{s} B_j(z)(1-z)^{\frac{s-j}{r}} = \frac{1}{2\pi i}\int_C \prod_{j=0}^{s}\prod_{l=0}^{n}\left(\zeta - \left(\frac{s-j}{r}\right) - l\right)^{-1}(1-z)^{\zeta}d\zeta,$$

where C is a closed contour containing all the $(\frac{s-j}{r}) + l$, $j = 0, \ldots, s$. By multiplication by a suitable constant, we can take the above expression to have the form (see [3, §3]) used by Mahler

$$R(z) = R(z, n) = \sum_{j=0}^{s} r_j(z, n)(1-z)^{(s-j)/r},$$

where

$$r_j(z, n) = (-1)^{(n+1)(s+1)-1}(n!)^s \sum_{l=0}^{n}(1-z)^l \prod_{h\neq j}^{s}\prod_{h'\neq l}^{n}\left(\frac{j-h}{r} + (l-h')\right)^{-1}.$$

Now with $k = (s+1)(n+1)-1$, let

$$R_h(z, n) = \sum_{j=0}^{s} B_{hj}(z)(1-z)^{\frac{s-j}{r}}$$

be the solution of the almost perfect Padé approximation problem, $\deg(B_{hj}) \leq n$ for $h \neq j$ and $\deg(B_{hh}) \leq (n+1)$. We use the same normalisation of the

B_{hj} as in [3, §3], but instead of the notation $A_{ij}(z,r)$ we use $B_{hj}(z)$, and instead of the parameters m, n, r we use, respectively, $s+1$, r, $n+1$. Mahler [15] showed that there is an explicit non-zero constant $\lambda_{r,n,s}$ such that

$$\det\big(B_{hj}(z)\big)_{h,j=0,\ldots,s} = \lambda_{r,n,s} z^{(n+1)(s+1)}. \tag{5.11}$$

Let $A_{hj}(z) = B_{hj}(1-z)$, $h,j = 0,\ldots,s$. Then we see from (5.11) that, if $a \neq 1$, then

$$\det\big(B_{hj}(1-a)\big)_{h,j=0,\ldots,s} = \det\big(A_{hj}(a)\big)_{h,j=0,\ldots,s} \neq 0. \tag{5.12}$$

For $h = 0, \ldots, s$, let

$$Q_h(x,y) = \sum_{j=0}^{s} A_{hj}(x^r) x^{s-j} y^j.$$

From (5.12) we deduce that, for some $h \in \{0,\ldots,s\}$, we have

$$\beta = \beta_h = \alpha^{-s} Q_h(\alpha,\alpha') = \sum_{j=0}^{s} A_{hj}(a)\gamma^{-j} \neq 0.$$

We then apply the product formula to $\beta \in K$, that is

$$\sum_{w\in M_K} \log|\beta|_w = 0,$$

estimating $\log|\beta|_w$ in a trivial way when $w \neq v$, and using a two-variable Taylor expansion when $w = v$. This represents a departure from the method of [3].

For every $w \neq v$ we have

$$\log|\beta|_w \leq (n+1)\log^+|a|_w + s\log^+|1/\gamma|_w + \max_j \log|A_{hj}|_w,$$

where $|A_{hj}|_w$ is the maximum of the w-adic valuations of the coefficients of A_{hj}. If instead $w = v$, we have $|\alpha - 1|_v < 1$ and we may assume that also $|\alpha' - 1|_v < 1$. The Taylor series of $Q_h(x,y)$ with center $(1,1)$ has rational coefficients because $Q_h(x,y) \in \mathbb{Q}[x,y]$. The divided differentiation occasioned by the Taylor expansion introduces no new denominators into the coefficients of the A_{hj}. Moreover, by construction, the polynomial $Q_h(x,1)$ has a zero of order at least $(s+1)(n+1)$ at $x = 1$. Therefore,

$$\begin{aligned}\log|\beta|_v &= |\alpha^{-s}Q_h(\alpha,\alpha')|_v \\ &\leq \max\Big(|\alpha-1|_v^{(s+1)(n+1)}, |\alpha'-1|_v\Big) + \max_j \log|A_{hj}|_v.\end{aligned}$$

Combining these estimates with the product formula we find

$$\min((s+1)(n+1)\Lambda, \Lambda') \leq (n+1)h(a) + sh(\gamma) + \max_j h(A_{jh}),$$

where $\Lambda \le \log(1/|\alpha - 1|_v)$ and $\Lambda' \le \log(1/|\alpha' - 1|_v)$. Now, as $A_{hj}(z) = B_{hj}(1-z)$, we have

$$h(A_{hj}) \le h(B_{hj}) + (n+1)\log 2 + \log(n+1).$$

Therefore,

$$\begin{aligned}&\min((s+1)(n+1)\Lambda, \Lambda')\\ &\quad < (n+1)h(a) + (s+1)h(\gamma) + \max_j h(B_{hj}) + (n+1)\log 2 + \log(n+1).\end{aligned}$$

So finally we have the following result.

THEOREM 5.1 (Almost-Perfect Thue-Siegel Lemma). *Suppose that, for* $\Lambda \le \log(1/|\alpha - 1|_v)$, *we have*

$$\Lambda \ge \frac{h(a)}{s+1} + \frac{h(\gamma)}{n+1} + \frac{1}{(s+1)(n+1)}\left(\max_j h(B_{hj}) + (n+1)\log 2 + \log(n+1)\right).$$

Then

$$\log(1/|\alpha' - 1|_v) \le (s+1)(n+1)\Lambda\,.$$

In order to apply this lemma, it is therefore crucial to have a good bound for $\max_j h(B_{hj})$, which is precisely one of the major preoccupations of [3]. Inspection of the estimates of that paper show, in particular, that the applicability of the Almost-Perfect Thue-Siegel Lemma above is governed by the contribution to $\max_j h(B_{hj})$ of the denominators of the coefficients of the B_{hj}. These are bounded in turn by numbers $\Delta_{s+1,r,n+1}$, studied in [3] where, following [12], estimates are obtained for the limit

$$\mathrm{Chr}_r^{s+1} := \limsup_{n\to\infty} \frac{1}{n}\log \Delta_{s+1,r,n+1}.$$

Unfortunately, the upper bounds for the Chr_r^{s+1} calculated in [3] and indeed even the conjectured bounds of §3 of that paper are not of a quality which seems readily exploitable for an application of the Almost-Perfect Thue-Siegel Lemma. This problem can be avoided by requiring vanishing to smaller order than the nominal $(n+1)(s+1) - 1$ in (5.2), however, at the expense of dealing with a vector space of equivariant auxiliary functions (5.1) of dimension greater than 1. This is the less than almost-perfect situation of the method applied in [7].

6. An estimate for the height of the auxiliary construction

In this section we use an estimate for the height $h(P)$ of the auxiliary construction P to derive from the Thue-Siegel Lemma of §4 an effective irrationality measure for α under the assumption (A.2). This estimate for $h(P)$ is certainly not the best possible and indeed it is hoped that the methods being developed in [18] will indicate how to obtain better results. Nonetheless, it leads to a result (subject to (A.2)) which compares well with the best known results to date using linear forms in logarithms. The vanishing condition for P at the points $(\varepsilon, \varepsilon)$, $\varepsilon^r = 1$, required by the Box Principle Lemma of §3 leads to a linear system whose solution space over $\mathbb{Q}$ is the space of solutions of the matrix equation

$$\mathcal{A}\vec{x} = \vec{0},$$

where $\vec{x}$ is in $\mathbb{Q}^{(N_1+1)(N_2+1)}$ and

$$\mathcal{A} = \left(\tbinom{u}{i_1} \tbinom{v}{i_2} \varepsilon_h^{u-i_1} \varepsilon_h^{v-i_2} \right).$$

Here, the columns of $\mathcal{A}$ are indexed by (u, v) with $0 \le u \le N_1$ and $0 \le v \le N_2$ and so are $C = (N_1+1)(N_2+1)$ in number. The R rows of $\mathcal{A}$ are indexed by (i_1, i_2, h), where $h = 1, \dots, r$ indexes the r-th roots ε_h of unity and where (i_1, i_2) are the solutions to

$$\frac{i_1}{\theta_1 N_1} + \frac{i_2}{\theta_2 N_2} < 1.$$

We have $R = rTN_1N_2 + O\big(\max(N_1, N_2)\big)$ and by the hypothesis of the Box Principle Lemma we have $R < C$. The linear system defined by $\mathcal{A}$ is equivalent to a linear system defined over $\mathbb{Q}$ (see [4], proof of Lemma 1). Let $\mathcal{V}$ be the vector subspace of $\mathbb{Q}^{(N_1+1)(N_2+1)}$ generated over $\mathbb{Q}$ by the solution space to this linear system. It has dimension $C - S$ where $S = \operatorname{rank}(\mathcal{A})$. Equation 1.11 of Theorem 2 and Corollary 6 in [16] give directly the following estimate for the height $H(\mathcal{V})$ of $\mathcal{V}$ (as defined in [10]):

$$\text{(6.1)} \quad \log H(\mathcal{V}) \le rTN_1N_2 \left\{ N_1 \cdot \frac{1}{3}\theta_1 \left(\log(1/4\theta_1) + \frac{11}{18} \right) + N_2 \cdot \frac{1}{3}\theta_2 \left(\log(1/4\theta_2) + \frac{11}{18} \right) \right\}.$$

By [10, Th. 9] we know that there is a basis $\mathcal{B}$ of $\mathcal{V}$ in $\mathbb{Z}^{C-S}$ with

$$\text{(6.2)} \quad \prod_{\vec{x} \in \mathcal{B}} H(\vec{x}) \le H(\mathcal{V}),$$

where $H(\vec{x})$ is just the maximum of the absolute values of the components of $\vec{x}$. Hence, there is a polynomial P satisfying the requirements of the Box Principle Lemma with

(6.3)
$$h(P) \le \frac{rT}{1 - rT} \left\{ N_1 \cdot \frac{1}{3}\theta_1 \left(\log(1/4\theta_1) + \frac{11}{18} \right) + N_2 \cdot \frac{1}{3}\theta_2 \left(\log(1/4\theta_2) + \frac{11}{18} \right) \right\}$$

We have shown the following theorem.

Theorem 6.1 (Siegel's Lemma). *There is a polynomial $P = P(x, y)$ satisfying the requirements of the Box Principle Lemma and with $h(P)$ bounded above as in* (2.2) *for*

$$l_i = \frac{rT}{1-rT}\left\{\frac{1}{3}\theta_i\left(\log(1/4\theta_i)+\frac{11}{18}\right)\right\}, \qquad i = 1,\ 2. \tag{6.4}$$

We now put this estimate for $h(P)$ into the Thue-Siegel Lemma of §4. Let $X = rT = 1 - \delta/2$, and suppose that $0 < X < 1/2$. We then set $\mu = 2/\theta_2$, so that $\theta_1 = 2X/r\theta_2 = X\mu/r$. For (4.9) of the Thue-Siegel Lemma to be satisfied we must have

$$\Lambda \geq \frac{2X}{1-X}\left(\frac{1}{3}\log\left(\frac{r}{4X\mu}\right)+\frac{11}{54}\right)+\frac{2h(a)}{X\mu}. \tag{6.5}$$

By inspection of the above inequality, as we may suppose that $r > \mu$, we see that (6.5) follows from the two conditions

$$\frac{X}{1-X}\left(\frac{1}{3}\log\left(\frac{r}{4X\mu}\right)+\frac{11}{54}\right) \leq \frac{1}{4}\Lambda\,, \qquad h(a) \leq \frac{1}{4}\Lambda X\mu\,. \tag{6.6}$$

There is an appropriate constant $c_6 > 0$ such that the first inequality in (6.6), is fulfilled with

$$(2-\delta) = 2X = c_6(D_v^*)^{-2}(\log(\frac{r}{\mu})+1)^{-1},$$

where $1/D_v^* = \min(1, f_v \log p/d)$, and d is the degree of K over $\mathbb{Q}$. We have used Lemma 1 of [6] which remarks that we always have $\Lambda \geq f_v \log p/d$. There is an absolute constant $c_7 > 0$ such that the second inequality in (6.6) is satisfied once

$$\mu \geq c_7 h(a)(D_v^*)^3\left(\log\left(\frac{r}{\mu}\right)+1\right).$$

For l_2 we have the expression

$$l_2 = \frac{X}{1-X}\left(\frac{1}{3}\,\frac{2}{\mu}\log\left(\frac{\mu}{8}\right)+\frac{11}{54}\right),$$

which is bounded by an absolute constant once μ is itself larger than some absolute constant.

We have proved the following theorem.

Theorem 6.2 (Theorem 2). *Let $P \in \mathbb{Z}[x, y]$ be as in Siegel's Lemma and suppose it satisfies* (A.2)*. Then, there are effective positive absolute constants c_1 and c_2 such that if $0 < \kappa < 1$ and*

$$r \geq c_1(D_v^*)^3\kappa^{-1}(\log(\kappa^{-1})+1)h(a) \tag{6.7}$$

and

$$h(\alpha') \geq c_2, \tag{6.8}$$

we have

$$|\alpha' - 1|_v \geq H(\alpha')^{-\kappa r}.$$

7. Determinants and ideal height estimates

The paper [8] discusses attempts to predict best ideal estimates for the asymptotic quantities $l_1(P)$ and $l_2(P)$ as defined in (2.2) for P a polynomial with vanishing much as in the Box Principle Lemma and satisfying (A.2). Specifically, let α_1, α_2 be generators of some algebraic number field of degree r over $\mathbb{Q}$. In the first instance, the vanishing demanded is at the r conjugates of the point (α_1, α_2). However, the authors note (see [5] for general principles, and [9] for details of the cubic case) that in the case $r = 3$, and in the case $\alpha_1 = \sqrt[r]{a_1}$, $\alpha_2 = \sqrt[r]{a_2}$, each an r-th root of some rational, it suffices to construct an 'invariant' P independent of the particular generators. In the cubic case, $r = 3$, the polynomial P must vanish at the three points $(0, 0)$, $(1, 1)$ and (∞, ∞); in the r-th root case, P is to vanish at the r points $(\varepsilon, \varepsilon)$ with $\varepsilon^r = 1$.

It seems best to illustrate the state of play as regards the construction of those invariant polynomials P by sketching some examples.

Consider the simultaneous approximation problem of constructing polynomials $A_0, \ldots, A_m$ satisfying $\deg A_j < \rho_j$, with $\rho_0 + \cdots + \rho_m = \sigma$, so that

$$R(z) = A_0(z)(1-z)^{\alpha_0} + A_1(z)(1-z)^{\alpha_1} + \cdots + A_m(z)(1-z)^{\alpha_m} = O(z^{\sigma-1}).$$

These are $\sigma - 1$ equations in σ unknowns. On adding the normalisation $R^{(\sigma-1)}(z) = 1$, say, one could endeavour to solve this problem professionally by the Bombieri–Vaaler Siegel Lemma [10], or naïvely by Cramer's rule for systems of linear equations. One discovers, of course not just fortuitously, that the determinants one is led to study are the same. Nonetheless, the first principles viewpoint encourages one to rewrite $R(z)$ as $\sum_{i=0}^{m}\sum_{h=1}^{\rho_i} a_{ih}(1-z)^{\alpha_i+h-1}$ and to determine the coefficients a_{ih}. Cramer's rule now tells us that each coefficient a_{ih} is the quotient Δ_{ih}/Δ of two determinants. Here the $\sigma \times \sigma$ 'master' determinant Δ has entry $\binom{\alpha_i+h-1}{j-1}$ in its j-th row and (i, h)-th column. Δ_{ih} is Δ with its (i, h)-th column replaced by the column $[0, 0, \ldots, 0, 1]$.

There is modern literature on evaluating recalcitrant determinants, very usefully summarised in [13]. Among the many valuable principles [13] recommends to the reader is the advice that 'the more parameters the better'. That's why our example, which is no more than a generalisation of the Padé approximation problem (5.10), has its present frills.

Indeed, it is plain that Δ vanishes whenever two of the quantities $\alpha_i + h - 1$ and $\alpha_i' + h' - 1$ — with the pair (i, h) different from (i', h') — happen to coincide. It follows that, with a lexicographic ordering on the pairs, the difference product

$$\prod_{(i,h)<(i',h')} \big((\alpha_i + h - 1) - (\alpha'_i + h' - 1)\big) \tag{7.1}$$

divides Δ. But it is easy to check that both this product, and the determinant Δ, has degree $\rho_i(\sigma - \rho_i)$ in each α_i, and is of total degree $\frac{1}{2}(\sigma^2 - \sum_{i=0}^{m} \rho_i^2)$ in the α's. Plainly therefore, (7.1) gives Δ up to an easily evaluated constant multiplier.

It is now relatively straightforward to see which difference factors are respectively missing from each Δ_{ih} and to rediscover the formulas given by the integrals of §5.

The above evaluation is trivial. Δ is just a Vandermonde determinant in very slight disguise. However, let

$$R(x,y) = A_0(x)(1-x)^{\alpha_0} + A_1(x)(1-x)^{\alpha_1} y + \cdots + A_m(x)(1-x)^{\alpha_m} y^m .$$

Then the simultaneous Padé approximation problem

$$R(x,1) = O(x^{\tau_0}) ,$$
$$\vdots$$
$$R_{y^l}(x,1) = \frac{\partial^l}{\partial y^l} R(x,y)\Big|_{y=1} = O(x^{\tau_l}),$$
$$\vdots$$

say, with $\sum \tau_l = \sigma - 1$, is rather less penetrable.

It happens that the case $\alpha_k = k\alpha$, for each k, is an appropriate specialisation. Denote the $\sigma \times \sigma$ determinant of this system by $M(\alpha)$, and its determinant by $\Delta(\alpha)$. With $\alpha = a$, a positive integer, $R(x,y)$ is a polynomial in x and y.

Specifically, $\deg R_{y^l}(x,1) = \max_{l \le k \le m} ka + \rho_k$. If this is less than $\tau_l - 1$ for one or more $l = 0, 1, \ldots,$ the approximation problem has redundant constraints and its master matrix $M(\alpha)$ is of rank less than σ for $\alpha = a$. Say its rank is $\sigma - t$. Then the master determinant $\Delta(\alpha) = \det\big(M(\alpha)\big)$ has a factor $(\alpha - a)^t$. One can similarly discover factors $\alpha + a$ by putting $\alpha = -a$ and considering the degree $\max_{l \le k \le m}(\rho_k + (m-k)a)$ of $(1-x)^{ma} R_{y^l}(x,1)$.

Perhaps surprisingly, the cubic case also gives rise to a Padé approximation problem of the present shape (see [17, §7] or [18]); that is, just as that of the r-th root case. This is applied in [18], where both the conjectured evaluations of [8, §5] are established. Krattenthaler and Zeilberger [14] had already proved the second of those conjectures. They introduce a parameter into the determinant and then carefully manipulate the determinant to discover its factors. In contrast, [18], though strongly guided by [14], views the determinants as belonging to Padé approximation problems.

That motivates the introduction of apparently natural parameters, which, once specialised to functions of a single parameter α, allows factors of the determinant to be discovered by way of blatant reduction in its row or column rank.

The particular specialisation $\alpha_k = k\alpha$ seems to be 'appropriate' for the following reason. For each l the conditions $R_{y^l}(x,1) = O(x^{\tau_l})$ are the same as

$$(1-x)^{-l\alpha} R_{y^l}(x,1) = O(x^{\tau_l}).$$

Now differentiate these conditions with respect to α and then divide by $\log(1-x)$. Remarkably, we obtain conditions equivalent to

$$R_{y^{l+1}}(x,1) = O(x^{\tau_l - 1}).$$

In this way one discovers yet additional multiplicity of already discovered factors of $\Delta(\alpha)$, allowing it to be completely evaluated, at any rate in the two cases studied in detail in [18]. Careful consideration of the phenomenon suggests always choosing $\alpha_k = k\alpha + c_k$, for some constants c_k. A conjecture, mildly supported by experiment, to the effect that the determinants take values 'of a combinatorial nature' further suggests that one can then succeed in evaluating the master determinant $\Delta(\alpha)$ by the techniques applied in [18].

However, a successful application of these ideas seems at the least to require that the parameter N_2 of §5, thus m of the present remarks, be allowed to be arbitrarily large. The work [18] struggles to detail the general case $m = 2$. It does indeed deal with a special case of general m, but so special that that case is actually that of arbitrary powers of the case $m = 1$.

References

[1] A. Baker, *Logarithmic forms*, Notes of a lecture delivered at the Symposium on Diophantine Problems, Boulder, Colorado, June 1994; private communication.

[2] A. Baker and G. Wüstholz, *Logarithmic forms and group varieties*, J. Reine Angew. Math. **442** (1993), 19–62.

[3] M.A. Bennett, *Rational approximation to algebraic numbers of small height: the diophantine equation* $|ax^n - by^n| = 1$, J. Reine Angew. Math. **535** (2001), 1–49.

[4] E. Bombieri, *Effective diophantine approximation on* $\mathbf{G}_m$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **20** (1993), 61–89.

[5] ———, *The equivariant Thue–Siegel Principle*, Arithmetic geometry (F. Catanese, ed.), Symposia Mathematica, vol. 37, Cambridge University Press, Cambridge, 1997, pp. 70–86.

[6] E. Bombieri and P. B. Cohen, *Effective diophantine approximation on* $\mathbf{G}_m$*, II*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **24** (1997), 205–225.

[7] ———, *Effective diophantine approximation on* $\mathbf{G}_m$*, III*, in progress.

[8] E. Bombieri, D.C. Hunt, and A.J. van der Poorten, *Determinants in the study of Thue's method and curves with prescribed singularities*, Experimental Math. **4** (1995), 87–96.

[9] E. Bombieri, A. J. van der Poorten, and J. D. Vaaler, *Effective measures of irrationality for cubic extensions of number fields*, Ann. Scuola Norm. Super. Pisa Cl. Sci. (4) **23** (1996), 211–248.

[10] E. Bombieri and J. D. Vaaler, *On Siegel's Lemma*, Invent. Math. **73** (1983), 11-32; Addendum, ibid. **75** (1984), 377.

[11] Y. Bugeaud, *Bornes effectives pour les solutions des équations en S-unités et des équations de Thue-Mahler*, J. Number Theory, **71** (1998), 227–244.

[12] G.V. Chudnovsky, *On the method of Thue-Siegel*, Ann. of Math. (2) **117** (1983), 325–382.

[13] C. Krattenthaler, *Advanced determinant calculus*, The Andrews Festschrift (Maratea, 1998), Sém. Lothar. Combin. **42** (1999), Art. B42q, 67pp; see `http://cartan.u-strasbg.fr:80/~slc/`.

[14] C. Krattenthaler and D. Zeilberger, *Proof of a determinant evaluation conjectured by Bombieri, Hunt and van der Poorten*, New York J. Math. **3** (1997), 54–102; see `http://nyjm.albany.edu:8000/nyjm.html`.

[15] K. Mahler, *Ein Beweis des Thue-Siegelschen Satzes über die Approximation algebraischer Zahlen für binomische Gleichungen*, Math. Ann. **105** (1931), 267–276.

[16] T. Struppeck and J. D. Vaaler, *Inequalities for heights of subspaces and the Thue-Siegel Principle*, Analytic Number Theory, Proceedings of a Conference in Honor of Paul T. Bateman (B. C. Berndt et al., eds.), Progr. Math., vol. 85, Birkhäuser, Boston, 1990, pp. 493–528.

[17] A. J. van der Poorten, *Generalised simultaneous approximation of functions*, J. Austral. Math. Soc. Ser. A **51** (1991), 50–61.

[18] ———, *A powerful determinant*, Experimental Math. **10** (2001), 307–320.

[19] K. Yu, *p-adic logarithmic forms and group varieties I*, J. Reine Angew. Math. **502** (1998), 29–92.

[20] ———, *p-adic logarithmic forms and group varieties II*, Acta Arith. **89** (1999), 337–378.

Paula B. Cohen, Department of Mathematics, Texas A&M University, College Station, TX 77843, USA and ceNTRe for Number Theory Research, Macquarie University, Sydney, NSW 2109, Australia

E-mail address: `pcohen@math.tamu.edu`

Alfred J. van der Poorten, ceNTRe for Number Theory Research, Macquarie University, Sydney, NSW 2109, Australia

E-mail address: `alf@math.mq.edu.au`