

## SOME QUANTITATIVE RESULTS RELATED TO ROTH'S THEOREM

E. BOMBIERI and A. J. VAN DER POORTEN

(Received 27 January 1987)

Communicated by J. H. Loxton

### Abstract

We employ the Dyson's Lemma of Esnault and Viehweg to obtain a new and sharp formulation of Roth's Theorem on the approximation of algebraic numbers by algebraic numbers and apply our arguments to yield a refinement of the Davenport-Roth result on the number of exceptions to Roth's inequality and a sharpening of the Cugiani-Mahler theorem. We improve on the order of magnitude of the results rather than just on the constants involved.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 11 J 68.

The new proof [5] by Esnault and Viehweg of Roth's Theorem relies on replacing the celebrated Roth Lemma by a new result of a purely algebraic nature. This result is called "Dyson's Lemma" because Dyson was the first to prove it in the case of two variables, exactly for this kind of application. It is significant to note that there are quantitative consequences of Dyson's Lemma which go beyond those previously accessible and which show that the new lemma lies essentially deeper than the original result.

We begin by once again proving Roth's Theorem. We use the invariant approach (see [1]) which with equal ease yields the result over general number fields and for arbitrary valuations. Moreover our formulation is a statement that for  $n$  approximations  $\beta_j$  to  $\alpha$  satisfying given gap conditions we have

$$|\alpha - \beta_j|_v \geq c(\alpha, n)h(\beta_j)^{-2-\varepsilon_n}$$

---

This work and the visit of the first author to Macquarie University was supported by ARGS grants B8515957 and B15893.

© 1988 Australian Mathematical Society 0263-6115/88 \$A2.00 + 0.00

for at least one  $j$ ,  $1 \leq j \leq n$ ; the constants  $c(\alpha, n)$  and  $\epsilon_n$  are stated explicitly. This appears to be new. In fact we prove a more general inequality for the maximum of  $n$  expressions  $|\alpha_j - \beta_j|_v$  with the numbers  $\alpha_j$  in a fixed number field and not necessarily the same. We then obtain a refinement of the Davenport-Roth result [4] on the number of exceptions to Roth's inequality and a sharpening of the Cugiani-Mahler Theorem [3], [6]. We improve on the order of magnitude of the results involved rather than just on the constants. We have been told of unpublished work of J. Luckhardt in late 1984 which employed the Esnault-Viehweg Lemma to yield a refinement of the Davenport-Roth result of similar quality to the one given here.

Of course the present note does not exhaust the potential applications of Dyson's inequality.

The height  $h(\alpha)$  of an algebraic number  $\alpha$  is its absolute height as defined and discussed in, for example, [1], pages 259–261 and our absolute values are normalised as in [1]. We use a compact notation appropriate to situations in which one deals with functions of several variables. This should be self-explanatory except perhaps for such constructs as

$$\frac{1}{I!} \Delta^I = \frac{1}{i_1! i_2! \dots i_n!} \frac{\partial^{i_1 + \dots + i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}}.$$

We wish to thank E. Calabi and J. Vaaler for providing the arguments in the proof of Lemma 4.

### 1. Dyson's Lemma

Let  $\alpha_1, \alpha_2, \dots, \alpha_n$  be elements of a number field  $K$  of degree  $r$  over the field  $k$ . Set

$$\mathcal{T}(t) = \{ \mathbf{x} : x_1 + x_2 + \dots + x_n \leq t, 0 \leq x_j \leq 1 \}$$

and write  $V(t) = \text{Vol}(\mathcal{T}(t))$ .

**LEMMA 1.** *Suppose  $t > 0$  and  $rV(t) < 1$ . Then for sufficiently large positive integers  $d_1, d_2, \dots, d_n$  there is a polynomial  $P(X_1, \dots, X_n) \in k[X_1, \dots, X_n]$  of multi-degree at most  $D = (d_1, d_2, \dots, d_n)$  and not identically zero such that*

$$\frac{1}{I!} \Delta^I P(\alpha_1, \alpha_2, \dots, \alpha_n) = 0$$

for all  $I = (i_1, i_2, \dots, i_n)$  with  $I/D = (i_1/d_1, \dots, i_n/d_n) \in \mathcal{T}(t)$ , and

$$\log h(P) \lesssim \frac{rV(t)}{1 - rV(t)} \sum_{i=1}^n \log(2h(\alpha_i)) d_i \text{ as } d_i \rightarrow \infty, \quad i = 1, 2, \dots, n.$$

PROOF. Write  $\alpha^\sigma = (\alpha_1^\sigma, \dots, \alpha_n^\sigma)$  with  $\sigma = \sigma_1, \dots, \sigma_r$  running over a full set of distinct embeddings of  $K$  over  $k$  into  $\mathbf{C}$  (that is, over all conjugates of  $\alpha = (\alpha_1, \dots, \alpha_n)$  over  $k$ ). Set  $P(X) = \sum p_J X^J$  and consider the set of equations

$$\frac{1}{I!} \Delta^I P(\alpha^\sigma) = 0, \quad I/D \in \mathcal{T}(t); \quad \sigma = \sigma_1, \dots, \sigma_r.$$

This is a linear system in the coefficients  $p_J$  of  $P$  which we are to solve in  $k$ . The number of unknown is  $(d_1 + 1) \cdots (d_n + 1) \sim d_1 \cdots d_n$  as  $D \rightarrow \infty$  (meaning: all of  $d_1 \rightarrow \infty, \dots, d_n \rightarrow \infty$ ), while the number of equations is asymptotic to  $rV(t)d_1 \cdots d_n$  as  $D \rightarrow \infty$ .

The matrix of coefficients  $\mathcal{A}$  has entries

$$\mathcal{A} = \left( \binom{J}{I} (\alpha^\sigma)^{j-1} \right)$$

with rows indexed by  $(I, \sigma)$  and columns by  $J$ .

The height  $H(\mathcal{A})$  of  $\mathcal{A}$  may be estimated by Bombieri and Vaaler [2], Corollary 13, page 30 and we find

$$H(\mathcal{A}) \leq \prod_{I, \sigma} H(A_{I, \sigma})$$

where  $A_{I, \sigma}$  is the row vector with entries  $\binom{J}{I} (\alpha^\sigma)^{j-I}$ . (Recall that the height  $H(\zeta)$  of a row vector  $(\zeta_1, \dots, \zeta_n)$  does not exceed  $\sqrt{nh}(\zeta)$  where  $\log h(\zeta) = \sum_v \max_i \log |\zeta_i|_v$  is the usual height in projective space.) Hence with  $N = (d_1 + 1) \cdots (d_n + 1)$ ,

$$\begin{aligned} \log(H(A_{I, \sigma})) &\leq \frac{1}{2} \log N + \sum_v \max_J \log \left| \binom{J}{I} (\alpha^\sigma)^J \right|_v \\ &\leq \log N + \sum_v \max_J \log \left| \binom{J}{I} \right|_v \\ &\quad + \sum_v (d_1 \log^+ |\alpha_1^\sigma|_v + \cdots + d_n \log^+ |\alpha_n^\sigma|_v) \\ &\leq \log N + d_1 \log 2 + \cdots + d_n \log 2 + d_1 \log h(\alpha_1) + \cdots + d_n \log h(\alpha_n) \end{aligned}$$

which yields

$$H(\mathcal{A}) \leq \left( \prod (d_i + 1)(2h(\alpha_i))^{d_i} \right)^{rM}$$

where  $M$  is the cardinality of the set  $\{I: I/D \in \mathcal{T}(t)\}$ . Recalling that  $M \sim V(t)d_1 \cdots d_n$  as  $D \rightarrow \infty$ , Lemma 1 now follows from Siegel's Lemma. (For a formulation consistent with our present approach see [2].)

LEMMA 2 (Esnault and Viehweg). *Let  $\alpha, t, P$  be as in Lemma 1 and assume further that each  $\alpha_i$  in  $\alpha$  has exact degree  $r$  over  $k$  and that  $d_1, d_2, \dots, d_n$  is rapidly decreasing in the sense that, for some  $\eta < 1$ ,*

$$\frac{d_{i+1}}{d_i} \leq \frac{\eta}{4rn} \quad \text{for } i = 1, \dots, n - 1.$$

Let  $\tau$  be such that

$$rV(t) + V(\tau) \geq 1 + \eta$$

(or, if there is no such  $\tau$ , set  $\tau = 1$ ). Then for any point  $\beta = (\beta_1, \dots, \beta_n) \in k^n$  we can find  $I^* = (i_1^*, \dots, i_n^*)$  such that

$$\frac{1}{I^*!} \Delta^{I^*} P(\beta_1, \dots, \beta_n) \neq 0$$

and  $I^*/D \in \mathcal{T}(\tau)$ .

The proof of the lemma is immediate from Theorem 0.4 of [5]. We apply that result with  $M = M' = r + 1$ ,

$$\xi_\mu = (\alpha_1^{\sigma_\mu}, \dots, \alpha_n^{\sigma_\mu}), \mu = 1, \dots, r \text{ and } \xi_M = (\beta_1, \dots, \beta_n);$$

$$\mathbf{a} = (d_1^{-1}, d_2^{-1}, \dots, d_n^{-1}) \text{ and } t_1 = \dots = t_r = t, t_{r+1} = \tau.$$

The admissibility condition  $\xi_{\mu,1} \neq \xi_{\mu',i}$  for  $\mu \neq \mu'$  and  $i = 1, \dots, n$  is assured by each  $\alpha_i$  having exact degree  $r$  over  $k$  and the  $\beta_i$  belonging to  $k$ .

If Lemma 2 were false we would obtain

$$\sum V(t_\mu) = rV(t) + V(\tau) \leq \prod_{j=1}^n \left( 1 + (M - 2) \sum_{i=j+1}^n \frac{d_i}{d_j} \right),$$

and noting that the product is majorised by

$$\prod_{j=1}^n \left( 1 + r \frac{\eta}{2rn} \left( \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots \right) \right) = \left( 1 + \frac{\eta}{2n} \right)^n < e^{\eta/2} \leq 1 + \eta$$

we would contradict the conditions of the lemma.

It is convenient to combine Lemmas 1 and 2 so that the polynomial constructed in Lemma 1 does not vanish at  $\beta$ . The required polynomial is just an appropriate derivative of  $P(X)$  and a simple estimate of the increase in height occasioned by the differentiation yields

LEMMA 3. Let  $\alpha, \beta, D, t, \tau$  be as in the hypotheses of Lemmas 1 and 2. Then there is a polynomial  $Q(\mathbf{X}) \in k[\mathbf{X}]$  of multi-degree  $\leq D$ , and not identically zero, such that

$$\frac{1}{I!} \Delta^I Q(\alpha) = 0 \quad \text{for } \frac{I}{D} \in \mathcal{T}(t - \tau)$$

and

$$Q(\beta) \neq 0,$$

with

$$\log h(Q) < \frac{rV(t)}{1 - rV(t)} \sum_{i=1}^n \log(2h(\alpha_i)) d_i + \sum_{i=1}^n (\log 2) d_i.$$

NOTE. If  $\tau > t$  the condition of vanishing at  $\alpha$  is vacuous and the lemma is trivial.

The following comments may help the reader in comparing Lemma 2 (the main result of Esnault and Viehweg) with the celebrated Roth Lemma. In applying Roth's Lemma one imposes a growth condition of the type

$$\frac{\log h(\beta_{i+1})}{\log h(\beta_i)} > \frac{1}{\eta}$$

on the  $\beta$ 's, and this ultimately yields a  $\tau$  of the size  $\tau = (c\eta)^{2^{-n+1}}$  with  $\eta$  comparable to that in Lemma 2 and some  $c > 0$  [8]. We apply Lemma 3 with a similar condition on the  $\beta$ s and with  $rV(t)$  extremely close to 1; hence with  $\tau$  determined by  $V(\tau) \sim \eta$ . But, it turns out,  $V(\tau) = \tau^n/n!$  for  $\tau \leq 1$  and this gives  $\tau \sim (n/e)\eta^{1/n}$ . As  $\eta \rightarrow 0$  the relative advantage of this estimate becomes obvious.

### 2. Cube slices

To use the results of the previous section one needs accurate information about the volume  $V(t)$  of the set  $\mathcal{T}(t)$ :

$$\mathcal{T}(t) = \{x: x_1 + \dots + x_n \leq t \leq x_1 \leq 1\}.$$

The following analytical treatment gives very precise results and is quite different from the estimate from below occurring in Mignotte [7], Lemma 3. We have not found an explicit account of the estimates given here in the literature.

Let  $C_n$  be the  $n$ -dimensional unit cube  $\{x_i: 0 \leq x_i \leq 1, i = 1, \dots, n\}$  and let  $H_z$  be the hyperplane slice

$$H_z = C_n \cap \left\{ \frac{x_1 + \dots + x_n}{n} = \frac{1}{2} - z \right\}.$$

LEMMA 4. Set  $U(x) = \log((\sinh x/2)/(x/2))$  and let  $v(x)$  be the inverse function of  $U'(x)$ . Then we have, uniformly in  $z: -\frac{1}{2} \leq z \leq \frac{1}{2}$ , as  $n \rightarrow \infty$ :

$$\text{Area}(H_z) \sim \frac{1}{\sqrt{2\pi U''(v(z))}} e^{-(zv(z) - U(v(z)))n}.$$

PROOF. Denote by  $\phi(x)$  the characteristic function of the interval  $-1/2 \leq x \leq 1/2$  and let  $\delta(x)$  be Dirac's  $\delta$ -function. Then it is clear that

$$\text{Area}(H_z) = \sqrt{n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \phi(x_1) \dots \phi(x_n) \delta(nz - x_1 - \dots - x_n) dx_1 \dots dx_n.$$

The right hand side is the convolution of  $\phi$  ( $n$  times) and  $\delta$ .

To evaluate it we use Fourier transforms, recalling that the Fourier transform of a convolution is the product of the Fourier transforms. We have

$$\hat{\phi}(u) = \frac{\sin u/2}{u/2}, \quad \hat{\delta}(u) = 1,$$

hence

$$\text{Area}(H_z) = \frac{\sqrt{n}}{2\pi} \int_{-\infty}^{\infty} \left( \frac{\sin u/2}{u/2} e^{izu} \right)^n du.$$

Recall  $U'(v(z)) = z$ , and shift the line of integration  $u \in \mathbf{R}$  into the complex plane so that it becomes the line  $u \in iv(z) + \mathbf{R}$ , noting that the choice of the shift  $iv(z)$  is motivated by the fact that it is the stationary phase point of the integrand.

With  $u = iv(z) + y$ , thus near the stationary phase point, we have

$$\frac{\sin u/2}{u/2} e^{izu} = \exp \left( -zv(z) + U(zv(z)) - \frac{1}{2}U''(v(z))y^2 + O(y^3) \right).$$

Moreover

$$\left| \frac{\sin u/2}{u/2} e^{izu} \right| = \sqrt{\frac{(\sinh v(z)/2)^2 + (\sin y/2)^2}{(v(z)/2)^2 + (y/2)^2}} e^{-zv(z)}$$

takes its maximum at  $y = 0$ .

The result of the lemma now follows from a standard asymptotic expansion.

LEMMA 5. *We have*

(i)  $V \left( \frac{1}{2} - s\sqrt{n} \right) \sin \frac{1}{\sqrt{\pi}} \int_{s\sqrt{6}}^{\infty} e^{-x^2} dx, \quad \text{uniformly for } s = o(n^{1/6});$

(ii)  $V(\tau) = \frac{\tau^n}{n!} \quad \text{if } \tau \leq 1.$

PROOF. By Lemma 4

$$V \left( n \left( \frac{1}{2} - z \right) \right) = \sqrt{n} \int_z^{1/2} \text{Area}(H_y) dy \sim \sqrt{\frac{n}{2\pi}} \int_z^{1/2} \frac{e^{-(yv(y) - U(v(y)))n}}{\sqrt{U''(v(y))}} dy.$$

Set  $W(x) = xv(x) - U(v(x))$ . Then  $W(x)$  is increasing with  $x$  and its Taylor expansion at 0 is  $W(x) = 6x^2 + O(x^3)$ ; also  $U''(v(x)) = \frac{1}{12} + O(x^2)$ .

We break the integral into two parts, namely with  $\varepsilon > 0$

$$\int_z^{1/2} = \int_z^{\varepsilon n^{-1/3}} + \int_{\varepsilon n^{-1/3}}^{1/2} = \int_{I_1} + \int_{I_2}, \quad \text{say.}$$

The integral over  $I_2$  is majorised by  $e^{-b\epsilon^2 n^{1/3}(1+o(1))}$ . Since  $ny^3 = O(\epsilon^3)$  and  $U''(v(y)) \sim \frac{1}{12}$  on  $I_1$  we have

$$\begin{aligned} \int_{I_1} &= (1 + O(\epsilon^3)) \int_z^{\epsilon n^{-1/3}} \frac{e^{-6ny^2}}{\sqrt{1/12}} dy = ((1 + O(\epsilon^3)) \sqrt{\frac{2}{n}} \int_{z\sqrt{6n}}^{\epsilon\sqrt{6n}^{1/6}} e^{-x^2} dx \\ &\sim \sqrt{\frac{2}{n}} \int_{z\sqrt{6n}}^{\infty} e^{-x^2} dx, \end{aligned}$$

provided, recalling  $s = z\sqrt{n}$ , that  $z = o(n^{1/3})$  and  $\epsilon \rightarrow 0$  with  $\epsilon/zn^{1/3} \rightarrow \infty$ , which we suppose. Indeed the last condition ensures that we can neglect the integral over  $I_2$  relative to the integral over  $I_1$  and completes the proof of part (i) of the lemma.

Part (ii) is a standard elementary exercise easily done by induction on  $n$ . We remark, omitting the proof, that

$$V(\tau) \sim \frac{\tau^n}{n!} \quad \text{for } \tau < \frac{cn}{\log n} \text{ with } c < 1,$$

as  $n \rightarrow \infty$ .

For numerical calculations, the following bound is useful.

LEMMA 6. For  $0 < \lambda < 1$  we have

$$V(n/2 - nz) \leq e^{-(z\lambda - U(\lambda))n}.$$

Also  $U(\lambda) \leq \lambda^2/6$ .

PROOF (Mignotte).

$$\begin{aligned} V\left(\frac{n}{2} - nz\right) &= \int \cdots \int_{(x_1 + \cdots + x_n > n/2 - nz) \cap C_n} \cdots \int dx_1 \cdots dx_n \\ &\leq \int \cdots \int_{C_n} \cdots \int e^{\lambda(\sum(x_i - 1/2) - zn)} dx_1 \cdots dx_n \\ &= e^{-z\lambda n} \left( \int_0^1 e^{\lambda(x - 1/2)} dx \right)^n, \end{aligned}$$

which is the first inequality. The second inequality is easy.

### 3. Roth's Theorem

The following result clearly includes Roth's Theorem for general number fields and arbitrary absolute values.

**THEOREM 1.** *Let  $\alpha_1, \dots, \alpha_n$  be elements of a number field  $K$  of degree  $r$  over the field  $k$  with each  $\alpha_i$  of exact degree  $r$  over  $k$ . Suppose  $n \geq c_0 \log r$  (where  $c_0$  is a sufficiently large constant) and set  $\eta: 0 < \eta \leq 1/2n!$ . Let  $\beta_1 \in k$  be approximations to  $\alpha_i, i = 1, \dots, n$ , such that we have the gap condition*

$$\frac{1}{\eta} \log(4h(\alpha_{i+1})) + \log(4h(\beta_{i+1})) \geq \frac{4rn}{\eta} \left( \frac{1}{\eta} \log(4h(\alpha_i)) + \log(4h(\beta_i)) \right).$$

Then

$$\max_{i=1, \dots, n} |\alpha_i - \beta_i|_v \geq ((4h(\alpha_n))^{1/\eta} 4h(\beta_n))^{-2-3} \sqrt{\log r} / \sqrt{n}.$$

**PROOF.** We first choose  $t = t(n, \eta)$  so that  $rV(t) = 1 - \eta$  and  $\tau = \tau(n, \eta)$  so that  $V(\tau) = 2\eta$  and apply Lemma 3 to obtain a polynomial  $P \in k[X]$  of multi-degree  $\leq D$  and not identically zero such that

$$\frac{1}{I!} \Delta^I P(\alpha) = 0 \quad \text{for } \frac{I}{D} \in \mathcal{T}(t - \tau)$$

and  $P(\beta) \neq 0$  with

$$\log h(P) < \sum_{i=1}^n \left( \frac{1}{\eta} \log(2h(\alpha_i)) + \log 2 \right) d_i \quad \text{as } D \rightarrow \infty.$$

The choice of the  $d_i$  is at our disposal subject to  $d_{i+1}/d_i \leq \eta/4rn$  as  $D \rightarrow \infty$ . For each absolute value  $v$  of  $k$  we estimate  $\log |P(\beta)|_v$  as follows. If  $v \neq v_0$  we use the trivial estimate (which is an exact inequality at all finite places):

$$\log |P(\beta)|_v \lesssim \log |P|_v + \sum_{i=1}^n d_i \log^+ |\beta_i|_v.$$

If, instead,  $v = v_0$  we expand  $P$  in a Taylor series about the point  $\alpha$  to obtain

$$P(\beta) = \sum_{I/D \notin \mathcal{T}(t-\tau)} \frac{1}{I!} P(\alpha) (\beta - \alpha)^I$$

because the Taylor coefficients vanish for  $I/D \in \mathcal{T}(t - \tau)$ . An easy calculation yields

$$\begin{aligned} \log |P(\beta)|_{v_0} &\lesssim \varepsilon_{v_0} \sum_{i=1}^n (\log 2) d_i + \log |P|_{v_0} + \sum_{i=1}^n d_i \log^+ |\alpha_i|_{v_0} \\ &\quad + \max_{J/D \notin \mathcal{T}(t-\tau)} \sum_{i=1}^n j_i \log |\beta_i - \alpha_i|_{v_0}, \end{aligned}$$

where

$$\varepsilon_{v_0} = \begin{cases} [K_{v_0} : k_{v_0}]/r & \text{if } v_0 | \infty, \\ 0 & \text{if } v_0 \nmid \infty. \end{cases}$$



The first term on the right of the asymptotic inequality is a trivial estimate of the contribution of the derivatives. We have  $|\beta_i - \alpha_i|_{v_0} < 1$ , because this is what is meant by saying that  $\beta_i$  is “an approximation to  $\alpha_i$ ”, and this implies

$$\log^+ |\alpha_i|_{v_0} \leq \varepsilon_{v_0} \log 2 + \log^+ |\beta_i|_{v_0}.$$

Substituting this bound above, and summing the estimates for  $\log |P(\beta)|_v$  over all absolute values of  $k$ , the product formula yields

$$\begin{aligned} 0 &= \sum_v \log |P(\beta)|_v \lesssim 2\varepsilon_{v_0} \sum_{i=1}^n (\log 2) d_i + \sum_v \log |P|_v + \sum_v \sum_{i=1}^n \log^+ |\beta_i|_v d_i \\ &\quad + \max_{J/D \notin \mathcal{T}(t-\tau)} \sum_{i=1}^n j_i \log |\beta_i - \alpha_i|_{v_0} \\ &\leq 2\varepsilon_{v_0} \sum_i (\log 2) d_i + \log h(P) + \sum_i (\log h(\beta_i)) d_i \\ &\quad + \max_{J/D \notin \mathcal{T}(t-\tau)} \sum_i j_i \log |\beta_i - \alpha_i|_{v_0}. \end{aligned}$$

The estimate for  $\log h(P)$  yields a fortiori

$$0 \lesssim \sum_{i=1}^n \left( \frac{1}{\eta} \right) \log(4h(\alpha_i)) + \log(4h(\beta_i)) d_i + \max_{J/D \notin \mathcal{T}(t-\tau)} \sum_{i=1}^n j_i \log |\beta_i - \alpha_i|_{v_0}.$$

Now choose

$$d_i = \left( \frac{d}{\frac{1}{\eta} \log(4h(\alpha_i)) + \log(4h(\beta_i))} \right),$$

as we may since the gap condition implies an appropriate inequality for the  $d_{i+1}/d_i$ . After dividing the inequality by  $d$  we may let  $d \rightarrow \infty$  to obtain

$$0 \leq n + \max_{i=1}^n x_i \log |\alpha_i - \beta_i|_{v_0} \frac{d_i}{d}$$

with the max taken over real  $x_1, \dots, x_n$  with  $\sum x_i \geq t - \tau$ ,  $0 \leq x_i \leq 1$ . The maximum can only increase if we replace the conditions  $0 \leq x_i \leq 1$  by just  $0 \leq x_i$ . Since  $\log |\alpha_i - \beta_i|_{v_0} < 0$  for all  $i$ , it follows by linearity that the maximum is attained at a point at which  $n - 1$  of the co-ordinates equal 0 and the non-zero co-ordinate equals  $t - \tau$ .

We have shown

$$0 \leq n + (t - \tau) \max_i \log |\alpha_i - \beta_i|_{v_0} \frac{d_i}{d}.$$

The theorem now follows from the bound

$$\frac{n}{t - \tau} \leq 2 + 3 \frac{\sqrt{\log \tau}}{\sqrt{n}}$$

which holds for  $n \geq c_0 \log r$ . To see this we use Lemma 5.

Set  $t = n/2 - s\sqrt{n}$ . Then

$$V(t) \sim \frac{1}{\sqrt{\pi}} \int_{s\sqrt{6}}^{\infty} e^{-x^2} dx.$$

Define  $s_r$  to be the value of  $s$  making this integral equal to  $1/r$ . But  $V(\tau) = 2\eta \leq 1/n!$  so that we are in the range in which  $V(\tau) = \tau^n/n!$  and  $\tau \leq 1$ . Thus

$$\frac{n}{t - \tau} \leq \frac{n}{n/2 - (1 + \varepsilon_n)s_r\sqrt{n} - 1} \quad \text{with } \varepsilon_n \rightarrow 0 \text{ as } n \rightarrow \infty.$$

To estimate  $s_r$  note that

$$\frac{1}{\sqrt{\pi}} \int_z^{\infty} e^{-x^2} dx \leq \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2z} \int_z^{\infty} e^{-x^2} dx^2 = \frac{1}{\sqrt{\pi}} \frac{e^{-z^2}}{2z},$$

so  $s_r \leq z_r/\sqrt{6}$  whenever  $z = z_r$  satisfies  $e^{-z^2}/\sqrt{\pi}2z \leq 1/r$ .

It is now clear that we can choose  $z_r = \sqrt{\log r}$ . This makes it plain that

$$\frac{n}{t - \tau} \leq \frac{n}{\frac{n}{2} - (1 + \varepsilon_n)s_r\sqrt{n} - 1} \leq 2 + 3\frac{\sqrt{\log r}}{\sqrt{n}} \quad \text{for } n \geq c_0 \log r,$$

completing the proof. A more refined calculation using Lemma 6, which we omit here, shows that  $c_0 = 28$  is admissible.

#### 4. The number of exceptions to Roth’s inequality

We next prove a generalisation and sharpening of a result of Davenport and Roth [4]:

**THEOREM 2.** *Let  $\alpha$  be an element of a number field  $K$  of degree  $r$  over the field  $k$ . Let  $\zeta_0 = \min(1, 6/\sqrt{c_0})$ . If  $0 < \zeta \leq \zeta_0$ , the number of solutions  $\beta \in k$  of the inequality*

$$|\alpha - \beta|_v < \frac{1}{64h(\beta)^{2+\zeta}}$$

*does not exceed*

$$\frac{2}{\zeta} \log \log(4h(\alpha)) + 3000 \frac{(\log r)^2}{\zeta^5} \log \left( \frac{50 \log r}{\zeta^2} \right).$$

**PROOF.** Since the bound increases with  $r$  there is no loss of generality in supposing that  $k(\alpha) = K$ . Choose  $\eta = 1/2n!$ . Choose  $n$  minimal such that  $\zeta_0 \geq \zeta \geq 6\sqrt{\log r}/\sqrt{n}$ ; hence  $n \geq c_0 \log r$ . Then if

$$4h(\beta) > (4h(\alpha))^{5/\eta\zeta}$$

we have

$$\frac{1}{64}h(\beta)^{-2-\varsigma} \leq (4h(\beta))^{-2-\varsigma} \leq ((4h(\alpha))^{1/\eta}4h(\beta))^{-2-\varsigma/2}.$$

Therefore our set of solutions is contained in the union of the subset  $S_0$  of solutions for which

$$\log(4h(\beta)) \leq \frac{5}{\eta\varsigma} \log(4h(\alpha))$$

and the set of solutions to

$$|\alpha - \beta|_v \leq ((4h(\alpha))^{1/\eta}4h(\beta))^{-2-3\sqrt{\log \tau}/\sqrt{n}}.$$

We classify the solutions to this latter inequality into sets  $S_i$  with

$$\log(4h(\beta)) \in \left[ \log(4h(\beta_i)), \frac{4rn}{\eta} \left( \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta_i)) \right) \right]$$

where the  $\beta_i$  are chosen inductively by taking  $\beta_1$  a solution with smallest height with  $\log(4h(\beta_1)) > 5/\eta\varsigma \log(4h(\alpha))$  and  $\beta_{i+1}$  a solution with smallest height so that

$$\log(4h(\beta_{i+1})) > \frac{4rn}{\eta} \left( \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta_i)) \right).$$

By Theorem 1 the sequence  $\beta_1, \dots, \beta_i, \dots$  is finite with at most  $n - 1$  elements.

Let  $\beta, \beta'$  with  $h(\beta) \leq h(\beta')$  be distinct solutions in  $S_\nu$  of the original inequality  $|\alpha - \beta|_v < 1/64h(\beta)^{-2-\varsigma}$ . By the Liouville bound

$$\frac{1}{2h(\beta)h(\beta')} \leq |\beta - \beta'|_v \leq |\alpha - \beta|_v + |\alpha - \beta'|_v \leq 2 \cdot \frac{1}{64}h(\beta)^{-2-\varsigma},$$

so  $4h(\beta') > 4(4h(\beta))^{1+\varsigma}$ . It follows that if there are  $m_\nu$  solutions in  $S_\nu$  then

$$\begin{aligned} (4h(\beta_\nu))^{(1+\varsigma)m_\nu-1} &\leq ((4h(\alpha))^{1/\eta}4h(\beta_\nu))^{4rn/\eta} \\ &\leq ((4h(\beta_i))^{s/5}(4h(\beta_\nu))^{4rn/\eta}) \\ &\leq (4h(\beta_\nu))^{5rn/\eta}. \end{aligned}$$

Hence  $(1 + \varsigma)^{m_\nu-1} \leq 5rn/\eta$  so

$$m_\nu \leq 1 + \frac{\log(5rn/\eta)}{\log(1 + \varsigma)}.$$

By our remark on the number of sets  $S_\nu$  we have proved that the number of solutions to  $|\alpha - \beta|_v < (1/64)h(\beta)^{-2-\varsigma}$  with  $\log(4h(\beta)) > (5/\eta\varsigma) \log(4h(\alpha))$  is at most

$$(n - 1) + (n - 1) \frac{\log(5rn/\eta)}{\log(1 + \varsigma)}.$$

Note that this bound is independent of  $\alpha$ .

For the set  $S_0$  we use the same procedure. As before, the gap principle  $4h(\beta') > 4(4h(\beta))^{1+\varsigma}$  holds, and as before the number  $m'_0$  of solutions

$$\log(4h(\beta)) \in \left[ \delta \log 4h(\alpha), \frac{5}{\eta\varsigma} \log(4h(\alpha)) \right]$$

does not exceed

$$1 + \frac{\log(5/\eta\varsigma\delta)}{\log(1 + \varsigma)};$$

here  $\delta$  satisfying  $0 < \delta < 5/\eta\varsigma$  is a parameter to be chosen below.

For solutions  $\beta$  so that  $\log(4h(\beta)) \in [0, \delta \log(4h(\alpha))]$  we use  $4h(\beta) \geq 4$  to obtain the estimate

$$4^{m''_0-1} \leq (4h(\alpha))^\delta$$

where  $m''_0$  is the number of solutions with  $\log(4h(\beta))$  in that interval. Hence

$$m''_0 < 1 + \delta \frac{\log(4h(\alpha))}{\log(1 + \varsigma)}.$$

We collect all these estimates and obtain the bound

$$|S| \leq n + 1 + (n - 1) \frac{\log(5rn/\eta)}{\log(1 + \varsigma)} + \frac{\log(5/\eta\varsigma\delta)}{\log(1 + \varsigma)} + \delta \frac{\log(4h(\alpha))}{\log(1 + \varsigma)}$$

for the totality of solutions. Choosing  $\delta = 1/\log(4h(\alpha))$ , which is optimal, converts the bound to

$$|S| \leq n + 1 + \frac{1}{\log(1 + \varsigma)} \left( \log \log(4h(\alpha)) + 1 + (n - 1) \log \frac{5rn}{\eta} + \log \frac{5}{\eta\varsigma} \right)$$

whilst  $1/\log(1 + \varsigma) < 2/\varsigma$  and  $1/rn < 1/\varsigma$  allows a simplification to

$$|S| \leq n + 1 + \frac{2}{\varsigma} \left( \log \log(4h(\alpha)) + 1 + n \log \frac{5rn}{\eta} \right).$$

Now

$$6 \frac{\sqrt{\log r}}{\sqrt{n}} \leq \varsigma \leq 6 \frac{\sqrt{\log r}}{\sqrt{n-1}} \quad \text{implies} \quad n \leq 1 + \frac{36 \log r}{\varsigma^2}$$

and since  $n > 2(\log r + 1)$  we get

$$\log \frac{5rn}{\eta} = \log r + \log(10n.n!) < \log r + n \log n - \frac{1}{2}n < n \log n - 1.$$

Hence

$$\begin{aligned} |S| &\leq n + 1 + \frac{2}{\varsigma} \log \log(4h(\alpha)) + \frac{2}{\varsigma} n^2 \log n - \frac{2}{\varsigma} (n - 1) \\ &< \frac{2}{\varsigma} \log \log(4h(\alpha)) + \frac{2}{\varsigma} n^2 \log n. \end{aligned}$$

Finally we eliminate  $n$  by noting that  $(2/\varsigma)n^2 < 3000\varsigma^{-5}(\log r)^2$  obtaining, say,

$$|S| < \frac{2}{\varsigma} \log \log(4h(\alpha)) + \frac{3000}{\varsigma^5} (\log r)^2 \log \left( \frac{50}{\varsigma^2} \log r \right)$$

as claimed.

NOTE. As in [4], Corollary 1, page 693 the bound of Theorem 2 can be used to estimate the number of solutions of the more general problem

$$|\alpha - \beta|_v < \frac{C}{h(\beta)^{2+\zeta}}.$$

We remark that the term  $3000\zeta^{-5}(\log r)^2 \log(50\zeta^{-2} \log r)$  in our result replaces  $\exp(70r^2\zeta^{-2})$  in [4]. It is also clear that the numerical constants can be substantially reduced with a little extra work.

Note that (though we use  $\zeta \leq 1$  in the course of our argument) the example of a zero of  $a_0(x - 1)(x - 2) \cdots (x - r) + 1$  with  $a_0$  sufficiently large and  $\beta$  an integer  $1 \leq \beta \leq r$  shows that  $\zeta \ll \log r$  is essential for the validity of our bound.

### 5. The Cugiani-Mahler Theorem

**THEOREM 3.** *Let  $\alpha$  be an element of a number field  $K$  of degree  $r$  over the field  $k$ . Set*

$$f(t) = 6\sqrt{\log r} \sqrt[4]{\frac{\log \log \log 4t}{\log \log 4t}}$$

and let  $\{\beta_i\}$  be a sequence of elements of  $k$ , arranged according to increasing height and satisfying for  $i = 1, 2, 3, \dots$

$$|\alpha - \beta_i|_v < (4h(\beta_i))^{-2-f(h(\beta_i))}.$$

Then

$$\limsup \frac{\log h(\beta_{i+1})}{\log h(\beta_i)} = \infty.$$

**PROOF.** We prove Theorem 3 by contradiction. As before we may suppose that  $K = k(\alpha)$ . Accordingly, let  $\{\beta^{(\nu)}\}$  be an infinite sequence of elements of  $k$  arranged according to increasing height so that

$$|\alpha - \beta^{(\nu)}|_v < (4h(\beta^{(\nu)}))^{-2-f(h(\beta^{(\nu)}))},$$

and

$$\frac{\log 4h(\beta^{(\nu+1)})}{\log 4h(\beta^{(\nu)})} < C.$$

As before, let  $n$  be a sufficiently large natural number, and set  $\eta = 1/2n!$ . Now select indices  $\nu_0$  and  $\nu_1$  so that

$$(4h(\beta^{(\nu_0-1)}))\sqrt{\log r/\sqrt{n}} < (4h(\alpha))^{1/\eta} \leq (4h(\beta^{(\nu_0)}))\sqrt{\log r/\sqrt{n}}$$

and

$$f(h(\beta^{(\nu_1-1)})) \geq 6\frac{\sqrt{\log r}}{\sqrt{n}} > f(h(\beta^{(\nu_1)})).$$

With  $n$  sufficiently large  $\nu_0$  and  $\nu_1$  may be chosen so that  $\nu_0 < \nu_1$ . But for  $\nu \geq \nu_0$

$$(4h(\beta^{(\nu)}))^{-2-6\sqrt{\log r}/\sqrt{n}} < ((4h(\alpha))^{1/\eta}(4h(\beta^{(\nu)})))^{-2-3\sqrt{\log r}/\sqrt{n}},$$

and for  $\nu < \nu_1$

$$(4h(\beta^{(\nu)}))^{-2-f(h(\beta^{(\nu)}))} < (4h(\beta^{(\nu)}))^{-2-6\sqrt{\log r}/\sqrt{n}}.$$

Thus for  $\nu$  with  $\nu_0 \leq \nu < \nu_1$  we have

$$|\alpha - \beta^{(\nu)}| < (4h(\beta^{(\nu)}))^{-2-f(h(\beta^{(\nu)}))} \leq ((4h(\alpha))^{1/\eta}4h(\beta^{(\nu)}))^{-2-3\sqrt{\log r}/\sqrt{n}}.$$

Our next step is to construct a subsequence  $\{\beta_i\}$  satisfying the gap condition of Theorem 1. We choose  $\beta_1 = \beta^{(\nu_0)}$  and inductively set  $\beta_{i+1} = \beta^{(\nu)}$  with  $\nu$  defined by

$$\begin{aligned} \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta^{(\nu)})) &\geq \frac{4nr}{\eta} \left( \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta_i)) \right) \\ &\geq \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta^{(\nu-1)})). \end{aligned}$$

Now by definition

$$\begin{aligned} \log(4h(\beta_{i+1})) &= \log(4h(\beta^{(\nu)})) \\ &\leq C \log(4h(\beta^{(\nu-1)})) \quad \text{by the slow-growth assumption,} \\ &\leq 4 \frac{Cnr}{\eta} \left( \frac{1}{\eta} \log(4h(\alpha)) + \log(4h(\beta_i)) \right) \\ &\leq 4 \frac{Cnr}{\eta} \left( 1 + \frac{\sqrt{\log r}}{\sqrt{n}} \right) \log(4h(\beta_i)) \quad \text{since } \nu \geq \nu_0, \\ &\leq Cr^2n^n \log(4h(\beta_i)) \end{aligned}$$

once  $n$  is sufficiently large, say  $n > n_1$ .

Thus, subject to  $n > n_1$ , say, there is a subsequence  $\{\beta_i\}$  of  $\{\beta^{(\nu)}\}$  which satisfies both the gap condition of Theorem 1 and the growth condition

$$\log(4h(\beta_i)) < (Cr^2n^n)^{i-1} \log(4h(\beta^{(\nu_0)})).$$

Suppose  $\nu_2$  is the index defined by  $\beta_n = \beta^{(\nu_2)}$ . Then we must have  $\nu_2 > \nu_1$ , otherwise the inequalities

$$|\alpha - \beta_i|_v < (4h(\beta_i))^{-2-f(h(\beta_i))} \leq ((4h(\alpha))^{1/\eta}4h(\beta_i))^{-2-3\sqrt{\log r}/\sqrt{n}}$$

for  $i = 1, 2, \dots, n$  contradict Theorem 1. It follows that

$$(Cr^2n^n)^{n-1} \log(4h(\beta_1)) > \log(4h(\beta_n)) \geq \log(4h(\beta^{(\nu_1+1)})),$$

where

$$f(h(\beta^{(\nu_1+1)})) < 6 \frac{\sqrt{\log r}}{\sqrt{n}}.$$

Moreover

$$\begin{aligned} \log(4h(\beta_1)) = \log(4h(\beta^{(\nu_0)})) &\leq C \log(4h(\beta^{(\nu_0-1)})), \quad \text{by the growth hypothesis,} \\ &\leq C \frac{\sqrt{n}}{\sqrt{\log r}} \frac{1}{\eta} \log(4h(\alpha)), \quad \text{by definition of } \nu_0, \\ &< Cr^2 n^n \log(4h(\alpha)), \quad \text{as a crude estimate.} \end{aligned}$$

Thus we have shown that

$$(Cr^2 n^n)^n \log(4h(\alpha)) > \log 4h(\beta^{(\nu_1+1)})$$

where

$$f(h(\beta^{(\nu_1+1)})) < 6 \frac{\sqrt{\log r}}{\sqrt{n}} \leq f(h(\beta^{(\nu_1)}))$$

and

$$\log(4h(\beta^{(\nu_1+1)})) \leq C \log(4h(\beta^{(\nu_1)})).$$

Now choose

$$f(t) = B \sqrt[4]{\frac{\log \log \log t}{\log \log t}}.$$

Then  $f(h(\beta^{(\nu_1+1)})) \sim 6\sqrt{\log r}/\sqrt{n}$  is equivalent to

$$\log(4h(\beta^{(\nu_1+1)})) = \exp\left((1 + o(1)) \frac{B^4 n^2 \log n}{648 (\log r)^2}\right)$$

where the left-hand side does not exceed

$$(Cr^2 n^n)^n \log(4h(\alpha)).$$

This is absurd for  $n > n_1$  and  $B = 6\sqrt{\log r}$ , as we wished to show.

REMARKS. Cugiani [3] obtains such a result but (in effect) with  $f(t)$  replaced by

$$9r(\log \log \log t)^{-1/2}.$$

Mahler [6], Appendix B, proves a  $g$ -adic generalisation of similar quality and it is this result that is sharpened by Mignotte [7] so that the conclusion

$$\limsup \frac{\log h(\beta_{i+1})}{(\log h(\beta_i))^\rho} = \infty \quad \text{if } f(t) = 4 \left( \frac{\log(r+2) \log 2}{a} \right)^{1/2} (\log \log \log t)^{-1/2}$$

with  $0 < a < b$  and  $1 < \rho < 2^{(b-a)/a}$  is obtained.

If we attempt to obtain a gap condition on the  $\log h(\beta_i)$  as in [7] we obtain a result only of the quality of that note. However, by changing the constant factor  $B$  in  $f(t)$  our gap condition may be refined to

$$\log h(\beta_{i+1}) > \log(h(\beta_i) \exp(c\sqrt{\log \log(h(\beta_i)) \log \log \log(h(\beta_i))})).$$

## References

- [1] E. Bombieri, 'On the Thue-Siegel-Dyson theorem', *Acta Math.* **148** (1982), 255–296.
- [2] E. Bombieri and J. Vaaler, 'On Siegel's lemma', *Invent. Math.* **73** (1983), 11–32.
- [3] Marco Cugiani, 'Sull'approssimabilità di un numero algebrico mediante numeri algebrici di un corpo assegnato', *Boll. Un. Mat. Ital Serie III* **14** (1959), 1–12.
- [4] H. Davenport and K. F. Roth, 'Rational approximations to algebraic numbers', *Mathematika* **2** (1955), 160–167.
- [5] H. Esnault and E. Viehweg, 'Dyson's Lemma for polynomials in several variables (and the theorem of Roth)', *Invent. Math.* **78** (1984), 445–490.
- [6] K. Mahler, *Lectures on diophantine approximation*, Part 1: *g-adic numbers and Roth's theorem* (University of Notre Dame, 1961).
- [7] M. Mignotte, 'Une généralisation d'un théorème de Cugiani-Mahler', *Acta Arith.* **22** (1972), 57–67.
- [8] K. F. Roth, 'Rational approximations to algebraic numbers', *Mathematika* **2** (1955), 1–20; Corrigendum, 168.

School of Mathematics  
The Institute for Advanced Study  
Princeton, New Jersey 08540  
U.S.A.

School of Mathematics and Physics  
Macquarie University  
N.S.W. 2109  
Australia