SIMULTANEOUS APPROXIMATION TO ALGEBRAIC NUMBERS BY RATIONALS

BY

WOLFGANG M. SCHMIDT

University of Colorado, Boulder, Colorado, USA

1. Introduction

We shall prove theorems on simultaneous approximation which generalize Roth's well-known theorem [3] on rational approximation to a single algebraic irrational α.

Throughout the paper, $\|\xi\|$ will denote the distance from a real number ξ to the nearest integer.

THEOREM 1. *Let $\alpha_1, \ldots, \alpha_n$ be real algebraic numbers such that $1, \alpha_1, \ldots, \alpha_n$ are linearly independent over the field Q of rationals. Then for every $\varepsilon > 0$ there are only finitely many positive integers q with*

$$\|q\alpha_1\| \cdot \|q\alpha_2\| \ldots \|q\alpha_n\| \cdot q^{1+\varepsilon} < 1. \tag{1}$$

COROLLARY. *Suppose $\alpha_1, \ldots, \alpha_n, \varepsilon$ are as above. There are only finitely many n-tuples $(p_1/q, \ldots, p_n/q)$ of rationals satisfying*

$$|\alpha_i - (p_i/q)| < q^{-1-(1/n)-\varepsilon} \qquad (i = 1, 2, \ldots, n). \tag{2}$$

A dual to Theorem 1 is as follows.

THEOREM 2. *Let $\alpha_1, \ldots, \alpha_n, \varepsilon$ be as in Theorem 1. There are only finitely many n-tuples of nonzero integers $q_1, \ldots, q_n$ with*

$$\|q_1\alpha_1 + \ldots + q_n\alpha_n\| \cdot |q_1 q_2 \ldots q_n|^{1+\varepsilon} < 1. \tag{3}$$

COROLLARY. *Again let $\alpha_1, \ldots, \alpha_n, \varepsilon$ be as in Theorem 1. There are only finitely many $(n+1)$-tuples of integers $q_1, q_2, \ldots, q_n, p$ with $q = \max(|q_1|, \ldots, |q_n|) > 0$ and with*

$$|q_1\alpha_1 + \ldots + q_n\alpha_n + p| > q^{-n-\varepsilon}. \tag{4}$$

When $n=1$, these two theorems are the same, and are in fact Roth's theorem mentioned above. A few years ago [4] I had proved these theorems in the case $n=2$. Our proofs will depend on a result of this earlier paper. What is new now is the use of Mahler's theory [2] of compound convex bodies.

2. Approximation by algebraic numbers of bounded degree

By *algebraic number* we shall understand a real algebraic number. Let ω be algebraic of degree at most k. There is a polynomial $f(t)=a_k t^k+\dots+a_1 t+a_0 \neq 0$, unique up to a factor ± 1, whose coefficients $a_k, \dots, a_1, a_0$ are coprime rational integers and which is irreducible over the rationals, such that $f(\omega)=0$. This polynomial is usually called the *defining polynomial* of ω. Define the *height* $H(\omega)$ of ω by

$$H(\omega)=\max(|a_k|, \dots, |a_1|, |a_0|). \tag{5}$$

THEOREM 3. *Let α be algebraic, k a positive integer, and $\varepsilon>0$. There are only finitely many algebraic numbers ω of degree at most k such that*

$$|\alpha-\omega|<H(\omega)^{-k-1-\varepsilon}. \tag{6}$$

When $k=1$, this result reduces again to Roth's theorem, and when $k=2$ it had been proved in [4]. Wirsing had proved(¹) a weaker version of Theorem 3, with $-k-1-\varepsilon$ in the exponent in (6) replaced by $-2k-\varepsilon$.

Theorem 3 may be deduced from Theorem 2 as follows. Let $f(t)$ be the defining polynomial of ω. Then $f(\alpha)=f(\omega)+(\alpha-\omega)f'(\tau)=(\alpha-\omega)f'(\tau)$ where τ lies between α and ω. Now since α is fixed, and by (6), τ lies in a bounded interval. Hence $|f'(\tau)|\leqslant c_1(k,\omega)H(\omega)$, and (6) yields

$$|a_k\alpha^k+\dots+a_1\alpha+a_0|<c_1(k,\omega)\,H(\omega)^{-k-\varepsilon}. \tag{7}$$

Now if α is not algebraic of degree at most k, then $1, \alpha, \dots, \alpha^k$ are linearly independent over Q, and the corollary to Theorem 2 implies that (7) has only finitely many solutions in integers $a_k, \dots, a_1, a_0$.

Suppose now that α is algebraic of degree m where $1\leqslant m\leqslant k$. There are rational integers d and b_{ij} $(0\leqslant i\leqslant k,\ 0\leqslant j\leqslant m-1)$ such that

$$d\alpha^i=b_{i0}+b_{i1}\alpha+\dots+b_{im-1}\alpha^{m-1} \quad (0\leqslant i\leqslant k).$$

Putting $y_j=\sum_{i=0}^{k}a_ib_{ij}$ $(0\leqslant j\leqslant m-1)$, we obtain

(¹) See his paper "Approximation to algebraic numbers by algebraic numbers of bounded degree", to appear in the report on the number theory institute at Stony Brook, July 1969.

$$|y_j| \leqslant c_2(k, \alpha)\, H(\omega) \quad (0 \leqslant j \leqslant m-1) \tag{8}$$

and
$$|\alpha^{m-1} y_{m-1} + \ldots + \alpha y_1 + y_0| < c_3(k, \alpha)\, H(\omega)^{-k-\varepsilon}. \tag{9}$$

By the corollary to Theorem 2, the inequalities (8), (9) have only the trivial solution $y_0 = \ldots = y_{m-1} = 0$ if $H(\omega)$ is large. But $a_k\alpha^k + \ldots + a_1\alpha + a_0 = d^{-1}(\alpha^{m-1}y_{m-1} + \ldots + y_0)$, and hence (7) implies that $f(\alpha) = a_k\alpha^k + \ldots + a_0 = 0$ if $H(\omega)$ is large. But $f(\alpha) = 0$ is possible only if ω is a conjugate of α, and there are only finitely many such conjugates.

3. Quoting a theorem

Let l be a positive integer greater than 1 and let

$$M_i = \beta_{i1}x_1 + \ldots + \beta_{il}x_l \quad (1 \leqslant i \leqslant l)$$

be l linear forms in $\mathbf{x} = (x_1, \ldots, x_l)$ with algebraic coefficients β_{ij} of determinant 1. Also let S be a subset of $\{1, 2, \ldots, l\}$. We say the system $\{M_1, \ldots, M_l; S\}$ is *regular* if

(i) for every $i \in S$, the *nonzero* elements among $\beta_{i1}, \ldots, \beta_{il}$ are linearly independent over Q.

(ii) for every k in $1 \leqslant k \leqslant l$, there is an $i \in S$ with $\beta_{ik} \neq 0$.

Now let

$$L_i = \alpha_{i1}x_1 + \ldots + \alpha_{il}x_l \quad (1 \leqslant i \leqslant l)$$

again be l linear forms with algebraic coefficients of determinant 1. There exist unique linear forms $M_1, \ldots, M_l$, the *adjoint* forms to $L_1, \ldots, L_l$, such that

$$L_1(\mathbf{x})\,M_1(\mathbf{y}) + \ldots + L_l(\mathbf{x})\,M_l(\mathbf{y}) = x_1y_1 + \ldots + x_ly_l$$

for any two vectors $\mathbf{x} = (x_1, \ldots, x_l)$, $\mathbf{y} = (y_1, \ldots, y_l)$. The forms $M_1, \ldots, M_l$ again have algebraic coefficients of determinant 1. Let S be a subset of $\{1, 2, \ldots, l\}$. We say the system $\{L_1, \ldots, L_l; S\}$ is *proper* if $\{M_1, \ldots, M_l; S\}$ is regular. It is clear that this definition is the same as the one given in § 1.4 of [4].

We now state Theorem 6 of [4].

THEOREM A. ("Theorem on the next to last minimum"). *Suppose $L_1, \ldots, L_l; S$ are proper, and $A_1, \ldots, A_l$ are positive reals satisfying*

$$A_1 A_2 \ldots A_l = 1 \tag{10}$$

and
$$A_i \geqslant 1 \quad \text{if } i \in S. \tag{11}$$

The set defined by
$$|L_i(\mathbf{x})| \leqslant A_i \quad (1 \leqslant i \leqslant l) \tag{12}$$
is a parallelopiped of volume 2^l; *denote its successive minima (in the sense of the Geometry of Numbers) by* $\lambda_1, \ldots, \lambda_{l-1}, \lambda_l$.

For every $\delta > 0$ *there is then a* $Q_0 = Q_0(\delta; L_1, \ldots, L_l; S)$ *such that*
$$\lambda_{l-1} > Q^{-\delta} \tag{13}$$
if
$$Q \geqslant \max(A_1, \ldots, A_l, Q_0). \tag{14}$$

4. A corollary to the quoted theorem

COROLLARY. *Let* $L_1, \ldots, L_l$; S *and* $A_1, \ldots, A_l$ *be as in the theorem. Again let* $\lambda_1, \ldots, \lambda_{l-1}, \lambda_l$ *be the successive minima of the parallelopiped defined by* (12). *For every* δ *in* $0 < \delta < 1$ *there is a* $Q_1 = Q_1(\delta; L_1, \ldots, L_l; S)$ *such that*
$$\lambda_{l-1} > \lambda_l Q^{-\delta} \tag{15}$$
provided
$$\lambda_1 A_i > Q^{-\delta/(2l)} \quad (i \in S) \tag{16}$$
and
$$Q \geqslant \max(A_1, \ldots, A_l, Q_1). \tag{17}$$

To prove this corollary we need to recall Lemma 7 of [4]:

LEMMA 1. (Davenport). *Let* $L_1, \ldots, L_l$ *be linear forms of determinant* 1, *and let* $\lambda_1, \ldots, \lambda_l$ *be the successive minima of the parallelopiped given by*
$$|L_i(\mathbf{x})| \leqslant 1 \quad (i = 1, \ldots, l). \tag{18}$$
Suppose $\varrho_1, \ldots, \varrho_l$ *are positive real numbers having*
$$\varrho_1 \varrho_2 \cdots \varrho_l = 1, \tag{19}$$
$$\varrho_1 \geqslant \varrho_2 \geqslant \ldots \geqslant \varrho_l > 0, \tag{20}$$
$$\varrho_1 \lambda_1 \leqslant \varrho_2 \lambda_2 \leqslant \ldots \leqslant \varrho_l \lambda_l. \tag{21}$$
Then, after a suitable permutation of $L_1, \ldots, L_l$, *the successive minima* $\lambda_1', \ldots, \lambda_l'$ *of the new parallelopiped*
$$\varrho_i |L_i(\mathbf{x})| \leqslant 1 \quad (i = 1, \ldots, l) \tag{22}$$
satisfy
$$\varrho_i \lambda_i << \lambda_i' << \varrho_i \lambda_i \quad (i = 1, \ldots, l). \tag{23}$$
Here the constants in (23) *depend only on* l.

The corollary is now proved as follows. Let $\lambda_1, \dots, \lambda_l$ be the successive minima of the parallelopiped (12). This parallelopiped may also be defined by $|L_i^*(\mathbf{x})| \leqslant 1$ $(i = 1, \dots, l)$ where $L_i^*(\mathbf{x}) = L_i(\mathbf{x}) A_i^{-1}$ $(i = 1, \dots, l)$. Put

$$\varrho_0 = (\lambda_1 \lambda_2 \dots \lambda_{l-2} \lambda_{l-1}^2)^{1/l}, \tag{24}$$

$$\varrho_1 = \varrho_0/\lambda_1, \varrho_2 = \varrho_0/\lambda_2, \dots, \varrho_{l-1} = \varrho_0/\lambda_{l-1}, \varrho_l = \varrho_0/\lambda_{l-1}. \tag{25}$$

Then (19), (20) and (21) hold. Applying Lemma 1 to $L_1^*, \dots, L_l^*$ we see that there is a permutation $(j_1, \dots, j_l)$ *of* $(1, \dots, l)$ such that the successive minima $\lambda_1', \dots, \lambda_l'$ of the parallelopiped

$$|L_i(\mathbf{x})| \leqslant A_i \varrho_{j_i}^{-1} (= A_i', \text{ say}) \quad (1 \leqslant i \leqslant l), \tag{26}$$

satisfy (23).

Suppose first that $A_i' \leqslant 1$ for some $i \in S$. Since for $i \in S$,

$$A_i' = A_i \varrho_{j_i}^{-1} \geqslant A_i \varrho_1^{-1} = \lambda_1 A_i \varrho_0^{-1} > Q^{-\delta/(2l)} \varrho_0^{-1}$$

by (16), we have $\varrho_0 > Q^{-\delta/(2l)}$. On the other hand, $\lambda_1 \lambda_2 \dots \lambda_l << 1$, whence $\varrho_0 << (\lambda_{l-1}/\lambda_l)^{1/l}$. Thus $\lambda_{l-1}/\lambda_l >> Q^{-\delta/2}$, and (15) holds provided Q is large.

The other possibility is that $A_i' > 1$ for every $i \in S$. We may then apply the theorem on the next to last minimum to the parallelopiped (26). Thus $\lambda_{l-1}' > Q^{-\delta/(8l^2)}$ provided $Q \geqslant \max (Q_2, A_1', \dots, A_l')$. Or, put differently, we have

$$\lambda_{l-1}' > Q^{-\delta/(2l)} \tag{27}$$

if

$$Q \geqslant \max (Q_3, A'^{1/(4l)}) \tag{28}$$

with $A' = \max (A_1', \dots, A_l')$. On the other, hand, by (23), we have $\lambda_{l-1}' << \varrho_{l-1} \lambda_{l-1} = \varrho_0 << (\lambda_{l-1}/\lambda_l)^{1/l}$. In conjunction with (27) this implies that $\lambda_{l-1}/\lambda_l >> Q^{-\delta/2}$, hence that $\lambda_{l-1} > \lambda_l Q^{-\delta}$ if Q is large.

It remains to be shown that (16) and (17) imply (28). Put $A = \max (A_1, \dots, A_l)$. We have $A' \leqslant A/\varrho_{l-1} = A\lambda_{l-1}/\varrho_0 << A\lambda_{l-1}/\lambda_1 << A\lambda_1^{-l}$, since $\lambda_1^{l-1}\lambda_{l-1} << 1$. Further by (16) we have $A\lambda_1 > Q^{-\delta/(2l)}$, whence

$$A' << A\lambda_1^{-l} << A^{1+l} Q^{\delta/2}.$$

Thus (17) implies that

$$Q > A^{1/2} Q^{\delta/2} > (A^{1+l} Q^{\delta/2})^{1/(4l)} Q_1^{\delta/8} > A'^{1/(4l)}$$

provided Q_1 is large.

5. The compounds of linear forms

Suppose $k > 1$ and let $\sigma, \tau, \dots$ denote subsets of $\{1, 2, \dots, k\}$. Write σ' for the complement of σ in $\{1, 2, \dots, k\}$. Define $(-1)^\sigma$ by

$$(-1)^{\sigma}=\prod_{j\in\sigma}(-1)^{j}. \tag{29}$$

For any integer p with $1\leqslant p<k$, let $C(k,p)$ consist of all sets σ with exactly p elements. Then $C(k,p)$ consists of $l(p)=\binom{k}{p}$ sets σ.

Let
$$L_i=\alpha_{i1}x_1+\ldots+\alpha_{ik}x_k \quad (i=1,\ldots,k) \tag{30}$$

be k linear forms of determinant 1 in $\mathbf{x}=(x_1,\ldots,x_k)$. Let p with $1\leqslant p<k$ be fixed at the moment. For every $\sigma\in C(k,p)$, $\tau\in C(k,p)$, write $\alpha_{\sigma\tau}$ for the $(p\times p)$-determinant formed from all ith rows with $i\in\sigma$ and all jth columns with $j\in\tau$ of the matrix (α_{ij}). We shall construct linear forms $L^{(p)}$ in vectors $\mathbf{x}^{(p)}$ with $l(p)$ components which are denoted by x_τ where $\tau\in C(k,p)$. Namely, for every $\sigma\in C(k,p)$, we put

$$L_{\sigma}^{(p)}(\mathbf{x}^{(p)})=\sum_{\tau\in C(k,p)}\alpha_{\sigma\tau}x_\tau. \tag{31}$$

We call these linear forms the pth *compounds* of $L_1,\ldots,L_k$. There are exactly $l(p)$ such pth compounds.

Again, for every σ in $C(k,p)$, put

$$\hat{L}_{\sigma}^{(p)}(\mathbf{x}^{(p)})=\sum_{\tau\in C(k,p)}(-1)^{\sigma}(-1)^{\tau}\alpha_{\sigma'\tau'}x_\tau. \tag{32}$$

Let $\mathbf{e}_{\tau}^{(p)}$ be the basis vector whose component $x_\tau=1$, and all of whose other components are zero. Then for any τ_1,τ_2 in $C(k,p)$, one has

$$\sum_{\sigma\in C(k,p)}L_{\sigma}^{(p)}(\mathbf{e}_{\tau_1}^{(p)})\,\hat{L}_{\sigma}^{(p)}(\mathbf{e}_{\tau_2}^{(p)})=\begin{cases}1 & \text{if} \quad \tau_1=\tau_2\\ 0 & \text{otherwise.}\end{cases}$$

This follows from Laplace's rule on the expansion of determinants, applied to the determinant $[\alpha_{ij}]$ $(1\leqslant i,j\leqslant k)$. It follows immediately that

$$\sum_{\sigma\in C(k,p)}L_{\sigma}^{(p)}(\mathbf{x}^{(p)})\,\hat{L}_{\sigma}^{(p)}(\mathbf{y}^{(p)})\equiv\sum_{\sigma\in C(k,p)}x_\sigma y_\sigma.$$

We have therefore shown the following result, which is essentially equivalent with Mahler's remark in [2, § 18].

LEMMA 2. *The system of linear forms $L_{\sigma}^{(p)}$ where $\sigma\in C(k,p)$ and the system of forms $\hat{L}_{\sigma}^{(p)}$ where $\sigma\in C(k,p)$ are adjoint to each other.*

Throughout the rest of this section let p in $1\leqslant p<k$ and $l=l(p)$ be fixed. The inequalities

$$|L_i(\mathbf{x})|\leqslant 1 \quad (i=1,\ldots,k) \tag{33}$$

define a parallelopiped Π in E^k. Since $L_1, \ldots, L_k$ have determinant 1, it follows from determinant theory that the l forms $L_\sigma^{(p)}(\mathbf{x}^{(p)})$ with $\sigma \in C(k, p)$ again have determinant 1. In particular these l linear forms are linearly independent. Hence the inequalities

$$|L_\sigma^{(p)}(\mathbf{x}^{(p)})| \leqslant 1 \quad (\sigma \in C(k, p)) \tag{34}$$

define a certain parallelopiped $\Pi^{(p)}$ in E^l. This parallelopiped is in general not exactly the same as Mahler's pth *compound* of Π, but as Mahler points out in [2, § 21], it is closely related to it.

Denote the successive minima of Π by $\lambda_1, \ldots, \lambda_k$, and for every σ write

$$\lambda_\sigma = \prod_{i \in \sigma} \lambda_i. \tag{35}$$

There is an ordering $\sigma_1, \sigma_2, \ldots, \sigma_l$ of the $l = l(p)$ elements σ of $C(k, p)$ such that

$$\lambda_{\sigma_1} \leqslant \lambda_{\sigma_2} \leqslant \ldots \leqslant \lambda_{\sigma_l}.$$

Denote the successive minima of $\Pi^{(p)}$ by $\nu_1, \nu_2, \ldots, \nu_l$.

THEOREM B. (Mahler.) *One has*

$$\nu_j << \lambda_{\sigma_j} << \nu_j \quad (1 \leqslant j \leqslant l(p)), \tag{36}$$

with the constants in $<<$ *only depending on* k.

Proof. This follows from Theorem 3 in [2] together with Mahler's remarks at the beginning of [2, § 21] which show that the successive minima of $\Pi^{(p)}$ and of the pth compound of Π differ only by bounded factors.

Now let $A_1, \ldots, A_k$ be positive reals with

$$A_1 A_2 \ldots A_k = 1. \tag{37}$$

Then if we put

$$A_\sigma = \prod_{i \in \sigma} A_i, \tag{38}$$

we have

$$\prod_{\sigma \in C(k, p)} A_\sigma = 1. \tag{39}$$

The inequalities

$$|L_i(\mathbf{x})| \leqslant A_i \quad (i = 1, \ldots, k) \tag{40}$$

define a parallelopiped Π_A in E^k, and the inequalities

$$|L_\sigma^{(p)}(\mathbf{x}^{(p)})| \leqslant A_\sigma \quad (\sigma \in C(k, p)) \tag{41}$$

define a parallelopiped $\Pi_A^{(p)}$ in E^l.

COROLLARY TO THEOREM B. *Define* λ_i $(1 \leqslant i \leqslant k)$, λ_σ $(\sigma \in C(k,p))$, ν_i $(1 \leqslant i \leqslant l)$ *as above, but with reference to* Π_A *and* $\Pi_A^{(p)}$ *instead of to* Π *and* $\Pi^{(p)}$. *Then one has again*

$$\nu_j << \lambda_{\sigma_j} << \nu_j \quad (1 \leqslant j \leqslant l(p)). \tag{42}$$

Proof. This follows from an application of Theorem B to the forms $L_i^* = A^{-1} L_i$ $(i = 1, \dots, k)$.

6. Special linear forms

Suppose now that $\alpha_1, \dots, \alpha_n$ are algebraic, and $1, \alpha_1, \dots, \alpha_n$ linearly independent over the rationals. Put

$$k = n + 1 \tag{43}$$

and
$$L_1(\mathbf{x}) = x_1 - \alpha_1 x_k,\ L_2(\mathbf{x}) = x_2 - \alpha_2 x_k, \dots,\ L_n(\mathbf{x}) = x_n - \alpha_n x_k,\ L_k(\mathbf{x}) = x_k. \tag{44}$$

For every p in $1 \leqslant p \leqslant n = k-1$, there are $l(p)$ compound forms $L_\sigma^{(p)}(\mathbf{x}^{(p)})$ with $\sigma \in C(k,p)$. Let $S^{(p)}$ consist of those $\sigma \in C(k,p)$ which contain the integer k.

LEMMA 3. *The forms* $L_\sigma^{(p)}(\mathbf{x}^{(p)})$ *with* $\sigma \in C(k,p)$ *together with* $S^{(p)}$ *form a proper system.*

Proof. By the definition of proper systems we have to show that the adjoint forms of $L_\sigma^{(p)}$ form a regular system with $S^{(p)}$. Hence in view of Lemma 2 we have to show that the forms $\hat{L}_\sigma^{(p)}$ where $\sigma \in C(k,p)$ together with $S^{(p)}$ form a regular system. Now except for the signs of the coefficients and the notation for the variables, the forms $\hat{L}_\sigma^{(p)}$ are the same as the forms $L_{\sigma'}^{(k-p)}$. We have to show that $L_{\sigma'}^{(k-p)}$ with $\sigma \in C(k,p)$ together with $S^{(p)}$ form a regular system. Let $\hat{S}^{(k-p)}$ consist of all sets σ' with $\sigma \in S^{(p)}$. Replacing p by $k-p$ we thus have to show that for every p in $1 \leqslant p \leqslant k-1 = n$,

$$L_\sigma^{(p)} \quad \text{with } \sigma \in C(k,p),\ \hat{S}^{(p)}$$

form a regular system. Note that $\hat{S}^{(p)}$ consists precisely of all $\sigma \in C(k,p)$ which do not contain the integer k.

Suppose now that $\sigma \in \hat{S}^{(p)}$. Then with the special forms given by (44) we have

$$L_\sigma^{(p)}(\mathbf{x}^{(p)}) = x_\sigma + \sum_{i \in \sigma} \pm \alpha_i x_{\sigma - i + k}. \tag{45}$$

Here $\sigma - i + k$ denotes the set obtained from σ by removing its element i and adding the integer k. The summands here have signs $+$ or $-$, but there is no need to evaluate these signs. From (45) it follows that except for their signs, the nonzero coefficients of $L_\sigma^{(p)}$ are 1 and the numbers α_i with $i \in \sigma$. These numbers form a subset of $1, \alpha_1, \dots, \alpha_n$, and hence they are linearly independent over the rationals. Thus condition (i) in the definition of regular systems is satisfied. It also is clear that for every τ in $C(k, p)$ there is a $\sigma \in \hat{S}^{(p)}$ such that the coefficient of x_τ in $L_\sigma^{(p)}$ is not zero. Hence (ii) holds.

7. Special parallelopipeds

LEMMA 4. *Assume that $\alpha_1, \ldots, \alpha_n$ are algebraic, and $1, \alpha_1, \ldots, \alpha_n$ linearly independent over the rationals. Put $k = n+1$ and define $L_1(\mathbf{x}), \ldots, L_k(\mathbf{x})$ by (44). Suppose $A_1, \ldots, A_k$ are positive and have*

$$A_1 A_2 \ldots A_k = 1 \tag{46}$$

and

$$A_1 < 1, \ldots, A_n < 1; \quad A_k > 1. \tag{47}$$

Let $\lambda_1, \ldots, \lambda_k$ be the successive minima of the parallelopiped Π_A given by

$$|L_i(\mathbf{x})| \leqslant A_i \quad (i = 1, \ldots, k). \tag{48}$$

Then for every $\delta > 0$ there is a $Q_2 = Q_2(\delta, \alpha_1, \ldots, \alpha_n)$ such that

$$\lambda_1 > Q^{-\delta} \tag{49}$$

provided

$$Q \geqslant \max (A_k, Q_2). \tag{50}$$

Proof. Our proof will be by induction on n. When $n = 1$ we may apply Theorem A with $l = 2, L_1, L_2$ and $S = \{2\}$. It follows that $\lambda_1 = \lambda_{l-1} > Q^{-\delta}$ provided $Q \geqslant \max (A_2, Q_0)$.

Now assume the truth of the lemma for integers less than n. It will suffice to prove for every p in $1 \leqslant p \leqslant k-1 = n$ and every $\delta > 0$ that

$$\lambda_{k-p} > \lambda_{k-p+1} Q^{-\delta} \tag{51}$$

provided $Q \geqslant \max (A_k, Q_3)$ where $Q_3 = Q_3(\delta, \alpha_1, \ldots, \alpha_n)$. Namely, repeated application of (51) yields $\lambda_1 > \lambda_k Q^{-n\delta} >> Q^{-n\delta}$. Since $\delta > 0$ was arbitrary, the lemma follows.

It remains to show (51). Let σ be the set in $C(k, p)$ consisting of $1, 2, \ldots, p-1, k$. (Hence σ consists of k only if $p = 1$). Our first aim is to show that with A_σ defined by (38), we have

$$\lambda_1 A_\sigma^{1/p} > Q^{-\delta} \tag{52}$$

if $Q \geqslant \max (A_k, Q_4)$. Take at first the case when $p = 1$. Then since there is an integer point $\mathbf{x}_0 \neq 0$ with $|L_i(\mathbf{x}_0)| \leqslant \lambda_1 A_i$ $(i = 1, \ldots, k)$, it follows that

$$1 \leqslant \max (\lambda_1 A_1, \ldots, \lambda_1 A_k) = \lambda_1 A_k = \lambda_1 A_\sigma^{1/p},$$

and (52) is true. Now assume that $1 < p \leqslant n = k-1$. Put

$$B_i = A_i/A_\sigma^{1/p} \quad (i \in \sigma). \tag{53}$$

Then by (46) and (47) we have

$$\prod_{i \in \sigma} B_i = B_1 B_2 \dots B_{p-1} B_k = 1 \tag{54}$$

and
$$B_i < 1 \ (1 \leqslant i \leqslant p-1), \ B_k > 1. \tag{55}$$

By definition of λ_1 there is an integer point $\mathbf{x}_0 \neq \mathbf{0}$ with $|L_i(\mathbf{x}_0)| \leqslant \lambda_1 A_i \ (i = 1, \dots, k)$. Now since Π_A has volume 2^k, the first minimum λ_1 is at most 1 by Minkowski's theorem. Hence $\lambda_1 A_i < 1 \ (i = 1, 2, \dots, n)$ by (47). Hence in $\mathbf{x}_0 = (x_1, \dots, x_n, x_k)$, the last coordinate x_k cannot be zero. Hence the vector $\mathbf{y}_0 = (x_1, \dots, x_{p-1}, x_k)$ in E^p is not $\mathbf{0}$. The linear forms L_i with $i \in \sigma$ may be interpreted as forms in $\mathbf{y} = (x_1, \dots, x_{p-1}, x_k)$. We have

$$|L_i(\mathbf{y}_0)| \leqslant \lambda_1 A_i = \lambda_1 A_\sigma^{1/p} B_i \quad (i \in \sigma).$$

Thus the parallelopiped in E^p defined by

$$|L_i(\mathbf{y})| \leqslant B_i \quad (i \in \sigma)$$

has a first minimum μ_1 with $\mu_1 \leqslant \lambda_1 A_\sigma^{1/p}$. In view of (54) and (55) it follows from our induction hypothesis that

$$\lambda_1 A_\sigma^{1/p} \geqslant \mu_1 > Q^{-\delta}$$

provided $Q \geqslant \max (B_k, Q_5)$. Since $B_k = A_k/A_\sigma^{1/p} \leqslant A_k$, the inequality (52) is true provided $Q \geqslant \max (A_k, Q_4)$.

Recall that $S_\sigma^{(p)}$ consists of all $\sigma \in C(k, p)$ which contain k. It is clear that (52) is in fact true for every $\sigma \in S^{(p)}$ provided $Q \geqslant \max (A_k, Q_4)$.

Let $L_\sigma^{(p)}(\mathbf{x}^{(p)})$ with $\sigma \in C(k, p)$ be the pth compound forms of $L_1, \dots, L_k$, and define the parallelopiped $\Pi_A^{(p)}$ by (41). The first minimum ν_1 of $\Pi_A^{(p)}$ satisfies $\nu_1 >> \lambda_1 \lambda_2 \dots \lambda_p >> \lambda_1^p$ by (42), and hence we have

$$\nu_1 A_\sigma >> \lambda_1^p A_\sigma >> Q^{-p\delta} \quad (\sigma \in S^{(p)})$$

by (52) provided Q is large. Since $\delta > 0$ in (52) was arbitrary, we have in fact

$$\nu_1 A_\sigma > Q^{-\delta/(2l)} \quad (\sigma \in S^{(p)}) \tag{56}$$

if $Q \geqslant \max (A_k, Q_6)$. Here $Q_6 = Q_6(\delta, \alpha_1, \dots, \alpha_n)$ and $l = l(p) = \binom{k}{p}$.

We now apply the corollary proved in section 4 to the proper system $L_\sigma^{(p)}$ $(\sigma \in C(k, p))$, $S^{(p)}$. The inequality (16) now becomes (56), and hence it is true if Q is large. It follows that

$$\nu_{l-1} > \nu_l Q^{-\delta} \tag{57}$$

provided (17) holds, i.e. provided $Q \geqslant \max(A_\sigma(\sigma \in C(k,p)), Q_7)$. Since $A_\sigma \leqslant A_k$ by (47), the last condition is fulfilled if $Q \geqslant \max(A_k, Q_7)$. Now by (42) again we have

$$\nu_l << \lambda_{k-p+1}\lambda_{k-p+2}\dots\lambda_k << \nu_l$$

and
$$\nu_{l-1} << \lambda_{k-p}\lambda_{k-p+2}\lambda_{k-p+3}\dots\lambda_k << \nu_{l-1}.$$

Thus (57) yields
$$\lambda_{k-p} >> \lambda_{k-p+1}Q^{-\delta}$$

if $Q \geqslant \max(A_k, Q_7)$. Since $\delta > 0$ was arbitrary, we therefore have (51) if $Q \geqslant \max(A_k, Q_3)$. This proves the lemma.

LEMMA 5. *Suppose $\alpha_1, \dots, \alpha_n$ are as in Lemma 4, and put $k = n+1$. Define linear forms $M_1, \dots, M_k$ by*

$$M_1(\mathbf{x}) = x_1, \quad M_2(\mathbf{x}) = x_2, \dots, M_n(\mathbf{x}) = x_n, \quad M_k(\mathbf{x}) = \alpha_1 x_1 + \dots + \alpha_n x_n + x_k. \tag{58}$$

Let $B_1, \dots, B_k$ be positive numbers with

$$B_1 B_2 \dots B_k = 1, \tag{59}$$

$$B_1 > 1, \dots, B_n > 1, \; B_k < 1. \tag{60}$$

Write $\mu_1, \dots, \mu_k$ for the successive minima of the parallelopiped Π_B defined by

$$|M_i(\mathbf{x})| \leqslant B_i \quad (i = 1, \dots, k). \tag{61}$$

For every $\delta > 0$ there is a $Q_8 = Q_8(\delta, \alpha_1, \dots, \alpha_n)$ such that

$$\mu_1 > Q^{-\delta} \tag{62}$$

provided
$$Q \geqslant \max(B_k^{-1}, Q_8). \tag{63}$$

Proof. This lemma is dual to Lemma 4. Write $A_i = B_i^{-1}$ $(i = 1, \dots, k)$. Then (46), (47) hold. The forms $M_1, \dots, M_k$ are adjoint to $L_1, \dots, L_k$ given by (44), and hence the forms $M_1/B_1, \dots, M_k/B_k$ are adjoint to $L_1/A_1, \dots, L_k/A_k$. Thus if $\lambda_1, \dots, \lambda_k$ are the successive minima of Π_A defined in Lemma 4, then it is well known that

$$1 << \lambda_i \mu_{k+1-i} << 1 \quad (i = 1, \dots, k). \tag{64}$$

(See, e.g., [1]. Another way to prove this is to use the corollary of Theorem B together with the fact, established in Lemma 2, that $M_1, \dots, M_k$ are essentially the $(k-1)$-st compounds of $L_1, \dots, L_k$. Namely, it follows that μ_{k+1-i} is of the same order of magnitude as $\lambda_1 \dots \lambda_{i-1}\lambda_{i+1}\dots\lambda_k$, hence as λ_i^{-1}.)

By Lemma 4 we have $\lambda_{k-1} \geqslant \ldots \geqslant \lambda_2 \geqslant \lambda_1 > Q^{-\delta}$, and hence $\lambda_k << (\lambda_1 \ldots \lambda_{k-1})^{-1} << Q^{k\delta}$. Thus by (64), $\mu_1 >> Q^{-k\delta}$. Since $\delta > 0$ was arbitrary, we have in fact (62) provided (63) holds with a suitably large Q_8.

8. Proof of the main theorems

The proof of Theorem 1 will be by induction on n. The case $n=1$ is Roth's theorem. Suppose that $n>1$ and q is a positive integer with

$$\|q\alpha_1\| \ldots \|q\alpha_n\| \cdot q^{1+\varepsilon} < 1. \tag{65}$$

Put
$$k = n+1, \quad \eta = \varepsilon/k, \tag{66}$$

$$A_i = \|q\alpha_i\|\, q^{\eta} \quad (i=1, \ldots, n), \qquad A_k = (A_1 A_2 \ldots A_n)^{-1}. \tag{67}$$

Now if one of the numbers $A_1, \ldots, A_n$ were at least 1, say if $A_1 \geqslant$ ~~1, then~~

$$\|q\alpha_2\| \ldots \|q\alpha_n\|\, q^{1+\varepsilon-\eta} < 1,$$

and by induction hypothesis this holds for only finitely many integers q. We may therefore assume that the numbers $A_1, \ldots, A_n$ are less than 1, and that (46), (47) hold. From (65), (66) and (67) we have

$$A_k = q^{-n\eta}(\|q\alpha_1\| \ldots \|q\alpha_n\|)^{-1} > q^{1+\varepsilon-n\eta} = q^{1+\eta}, \tag{68}$$

and (67) together with Roth's theorem yields

$$A_k \leqslant (\|q\alpha_n\| \ldots \|q\alpha_n\|)^{-1} < q^{2n} \tag{69}$$

for large q.

Let $p_1, \ldots, p_n$ be integers with $\|q\alpha_i\| = |q\alpha_i - p_i|$ $(i=1, \ldots, n)$, and let $\mathbf{x}_0$ be the point $(p_1, \ldots, p_n, q)$ in E^k. Then (67) and (68) imply that

$$|L_i(\mathbf{x}_0)| \leqslant A_i q^{-\eta} \quad (i=1, \ldots, k), \tag{70}$$

where $L_1, \ldots, L_k$ are the forms given by (44). Thus the parallelopiped Π_A defined by $|L_i(\mathbf{x})| \leqslant A_i$ $(i=1, \ldots, k)$ has a first minimum λ_1 with $\lambda_1 \leqslant q^{-\eta}$. The number $Q = q^{2n}$ satisfies $Q > A_k$ by (69), and we still have $\lambda_1 \leqslant Q^{-\eta/(2n)}$. By Lemma 4 this is im~~possible if q and~~ hence Q is large.

Now let us turn to Theorem 2. Suppose that $q_1, \ldots, q_n$ are nonzero integers with

$$\|q_1\alpha_1 + \ldots + q_n\alpha_n\| \cdot |q_1 \ldots q_n|^{1+\varepsilon} < 1. \tag{71}$$

We may assume that $0 < \varepsilon < 1$. Put

$$k = n + 1, \quad \eta = \varepsilon / k, \quad q = |q_1 q_2 \dots q_n|, \tag{72}$$

$$B_i = |q_i| \, q^{\eta} \quad (i = 1, \dots, n), \qquad B_k = (B_1 B_2 \dots B_n)^{-1}. \tag{73}$$

Then (59) and (60) hold if $q > 1$. We have

$$B_k = q^{-n\eta} |q_1 q_2 \dots q_n|^{-1} > \|q_1 \alpha_1 + \dots + q_n \alpha_n\| \, q^{-n\eta} |q_1 q_2 \dots q_n|^{\varepsilon} = \|q_1 \alpha_1 + \dots + q_n \alpha_n\| \, q^{\eta} \tag{74}$$

by (71), (72), (73), and $$B_k^{-1} = q^{n\eta} |q_1 q_2 \dots q_n| \leqslant q^{2n} \tag{75}$$

by (72), (73).

Let p be the integer with $\|q_1 \alpha_1 + \dots + q_n \alpha_n\| = |q_1 \alpha_1 + \dots + q_n \alpha_n + p|$, and let $\mathbf{x}_0$ be the point $(q_1, \dots, q_n, p)$ in E^k. Then in view of (73), (74) we have

$$|M_i(\mathbf{x}_0)| \leqslant B_i q^{-\eta} \quad (i = 1, \dots, k), \tag{76}$$

where $M_1, \dots, M_k$ are the forms defined in (58). Thus the parallelopiped Π_B given by $|M_i(\mathbf{x})| \leqslant B_i$ $(i = 1, \dots, k)$ has a first minimum μ_1 with $\mu_1 \leqslant q^{-\eta}$. The number $Q = q^{2n}$ satisfies $Q \geqslant B_k^{-1}$ by (75), and we still have $\mu_1 \leqslant Q^{-\eta/(2n)}$. By Lemma 5 this is impossible unless Q and hence q are small.

References

[1]. Mahler, K., Ein Übertragungsprinzip für konvexe Körper. *Čas. Pěst. Mat.*, 68 (1939), 93–102.

[2]. Mahler, K., On compound convex bodies (I). *Proc. London Math. Soc.* (3), 5 (1955), 358–379.

[3]. Roth, K. F., Rational approximations to algebraic numbers. *Mathematika*, 2 (1955), 1–20.

[4]. Schmidt, W. M., On simultaneous approximations of two algebraic numbers by rationals. *Acta Math.*, 119 (1967), 27–50.

Received February 27, 1970