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 \( \alpha \).

Throughout the paper, \( \| \xi \| \) will denote the distance from a real number \( \xi \) 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 \( \mathbb{Q} \) of rationals. Then for every \( \varepsilon > 0 \) there are only finitely many posi-
tive integers \( q \) with

\[
\| q \alpha_1 \| \cdot \| q \alpha_2 \| \cdots \| q \alpha_n \| \cdot q^{1+\varepsilon} < 1.
\]  

(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

\[
|\alpha_i - (p_i/q)| < q^{-1-(1/n)-\varepsilon} \quad (i = 1, 2, \ldots, n).
\]

(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

\[
\| q_1 \alpha_1 + \cdots + q_n \alpha_n \| \cdot | q_1 q_2 \cdots q_n |^{1+\varepsilon} < 1.
\]

(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

\[
| q_1 \alpha_1 + \cdots + q_n \alpha_n + p | > q^{-n-\varepsilon}.
\]

(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 \( \omega \) be algebraic of degree at most \( k \). There is a polynomial \( f(t) = a_k t^k + \cdots + a_1 t + a_0 \neq 0 \), unique up to a factor \( \pm 1 \), whose coefficients \( a_k, \ldots, a_1, a_0 \) 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 \( \omega \). Define the height \( H(\omega) \) of \( \omega \) by

\[
H(\omega) = \max (|a_k|, \ldots, |a_1|, |a_0|).
\]  

Theorem 3. Let \( \alpha \) be algebraic, \( k \) a positive integer, and \( \varepsilon > 0 \). There are only finitely many algebraic numbers \( \omega \) of degree at most \( k \) such that

\[
|\alpha - \omega| < H(\omega)^{-k-1-\varepsilon}.
\]  

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 \( \omega \). Then \( f(\tau) = f(\alpha) + (\alpha - \omega) f'(\tau) - (\alpha - \omega) f'(\alpha) \) where \( \tau \) lies between \( \alpha \) and \( \omega \). Now since \( \alpha \) is fixed, and by (6), \( \tau \) lies in a bounded interval. Hence \( |f'(\tau)| < c_1(k, \omega) H(\omega) \), and (6) yields

\[
|a_k \alpha^k + \cdots + a_1 \alpha + a_0| < c_1(k, \omega) H(\omega)^{-k-\varepsilon}.
\]  

When \( \alpha \) is not algebraic of degree at most \( k \), then \( 1, \alpha, \ldots, \alpha^k \) are linearly independent over \( \mathbb{Q} \), and the corollary to Theorem 2 implies that (7) has only finitely many solutions in integers \( a_0, \ldots, a_1, a_0 \).

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

\[
d \alpha^i = b_{0i} + b_{1i} \alpha + \cdots + b_{ni} \alpha^{m-1} \quad (0 \leq i \leq k).
\]

Putting \( y_j = \sum_{i=0}^m a_i b_{ij} \) \((0 \leq j \leq m-1)\), 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.
SIMULTANEOUS APPROXIMATION TO ALGEBRAIC NUMBERS BY RATIONALS

191

\[ |y_j| \leq c_2(k, \omega) H(\omega) \quad (0 \leq j \leq m - 1) \]  

and

\[ |\alpha^{m-1} y_{m-1} + \ldots + \alpha y_1 + y_0| \leq c_2(k, \omega) H(\omega)^{-k-s}. \]

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 \( a_k \alpha^k + \ldots + a_0 = d^{-1}(\alpha^{m-1} y_{m-1} + \ldots + y_0) \), 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 \( \omega \) is a conjugate of \( \alpha \), and there are only finitely many such conjugates.

3. Quoting a theorem

Let \( l \) be a positive integer greater than 1 and let

\[ M_i = \beta_{il} x_1 + \ldots + \beta_{ii} x_i \quad (1 \leq i \leq l) \]

be \( l \) linear forms in \( x = (x_1, \ldots, x_l) \) with algebraic coefficients \( \beta_{il} \) 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 \( \beta_{il}, \ldots, \beta_{ii} \) are linearly independent over \( Q \).

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

Now let

\[ L_i = \alpha_{il} x_1 + \ldots + \alpha_{ii} x_i \quad (1 \leq i \leq l) \]

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

\[ L_1(x) M_1(y) + \ldots + L_l(x) M_l(y) = x_1 y_1 + \ldots + x_l y_l \]

for any two vectors \( x = (x_1, \ldots, x_l), y = (y_1, \ldots, y_l) \). 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

\[ A_1 A_2 \ldots A_l = 1 \]  

and

\[ A_l \geq 1 \quad \text{if} \ i \in S. \]
The set defined by 

\[ |L_i(x)| \leq A_i \quad (1 \leq i \leq l) \]  

(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

\[ \lambda_{l-1} > Q^{-\delta} \]  

(13)

if

\[ Q > \max (A_l, \ldots, A_1, Q_0). \]  

(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 \(\delta\) in \(0 < \delta < 1\) there is a \(Q_1 = Q_1(\delta; L_1, \ldots, L_l; S)\) such that

\[ \lambda_{l-1} > \lambda_l Q^{-\delta} \]  

(15)

provided

\[ \lambda_i A_i > Q^{1-\delta} \quad (i \in S) \]  

(16)

and

\[ Q > \max (A_1, \ldots, A_l, Q_1). \]  

(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

\[ |L_i(x)| \leq 1 \quad (i = 1, \ldots, l). \]  

(18)

Suppose \(\varrho_1, \ldots, \varrho_l\) are positive real numbers having

\[ \varrho_1 \varrho_2 \cdots \varrho_l = 1, \]  

(19)

\[ \varrho_1 \geq \varrho_2 \geq \ldots \geq \varrho_l > 0, \]  

(20)

\[ \varrho_1 \lambda_1 \leq \varrho_2 \lambda_2 \leq \ldots \leq \varrho_l \lambda_l. \]  

(21)

Then, after a suitable permutation of \(L_1, \ldots, L_l\), the successive minima \(\lambda'_1, \ldots, \lambda'_l\) of the new parallelopiped

\[ \varrho_i |L_i(x)| \leq 1 \quad (i = 1, \ldots, l) \]  

(22)

satisfy

\[ \varrho_1 \lambda'_1 < \varrho_2 \lambda'_2 < \ldots < \varrho_l \lambda'_l \quad (i = 1, \ldots, l). \]  

(23)

Here the constants in (23) depend only on \(l\).
The corollary is now proved as follows. Let $\lambda_1, \ldots, \lambda_l$ be the successive minima of the parallelepiped (12). This parallelepiped may also be defined by $|L_i^*(x)| \leq 1$ ($i = 1, \ldots, l$) where $L_i^*(x) = L_i(x) A_i^{-1}$ ($i = 1, \ldots, l$). Put

$$\varrho_0 = (\lambda_1 \lambda_2 \ldots \lambda_{i-2} \lambda_{i-1})^{1/2},$$

$$\varrho_1 = \varrho_0/\lambda_1, \varrho_2 = \varrho_0/\lambda_2, \ldots, \varrho_{l-1} = \varrho_0/\lambda_{l-1}, \varrho_l = \varrho_0/\lambda_l.$$  

Then (19), (20) and (21) hold. Applying Lemma 1 to $L_i^*, \ldots, L_i^*$ we see that there is a permutation $(j_1, \ldots, j_l)$ of $(1, \ldots, l)$ such that the successive minima $\lambda_{j_1}, \ldots, \lambda_{j_l}$ of the parallelepiped

$$|L_i^*(x)| \leq A_i \varrho_i^{-1} (= A_i', \text{ say}) \quad (1 \leq i \leq l),$$

satisfy (23).

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

$$A_i' = A_i \varrho_i^{-1} = A_i \varrho_i^{-1} > \lambda_i A_i \varrho_i^{-1} > Q^{-\delta_0} \varrho_0^{-1}$$

by (16), we have $\varrho_0 > Q^{-\delta_0} \varrho_0.$ On the other hand, $\lambda_1 \lambda_2 \ldots \lambda_i < 1,$ whence $\varrho_0 < < (\lambda_{i-1}/\lambda_i)^{1/2}.$ Thus $\lambda_{i-1}/\lambda_i > Q^{-\delta_0},$ 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 parallelepiped (26). Thus $\lambda_{l-1}' > Q^{-\delta_1(4i)}$ provided $Q > \max \{Q_3, A_i', \ldots, A_i'\}.$ Or, put differently, we have

$$\lambda_{l-1}' > Q^{-\delta_1(4i)}$$

if

$$Q > \max \{Q_3, A_i', A_i', \ldots, A_i'\}$$

with $A' = \max \{A_1', \ldots, A_i'\}.$ On the other hand, by (23), we have $\lambda_{i-1}' < < \varrho_{i-1}/\lambda_{i-1} = \varrho_0 < < (\lambda_{i-1}/\lambda_i)^{1/2}.$ In conjunction with (27) this implies that $\lambda_{i-1}/\lambda_i > Q^{-\delta_2},$ hence that $\lambda_{i-1} > \lambda_i Q^{-\delta_2}$ if $Q$ is large.

It remains to be shown that (16) and (17) imply (28). Put $A = \max \{A_1, \ldots, A_i\}.$ We have $A' < A|\varrho_{i-1} = A\lambda_{i-1}/\varrho_0 < A\lambda_{i-1}/\lambda_i < A\lambda_i^{-1},$ since $\lambda_i^{-1} \lambda_{i-1} < < 1.$ Further by (16) we have $A\lambda_i > Q^{-\delta_0(2i)}$, whence

$$A' < A\lambda_i^{-1} < A^{1+i} Q^{\delta_2}.$$

Thus (17) implies that

$$Q > A^{1+i} Q^{\delta_2} (A^{1+i} Q^{\delta_2})^{1+i} Q_i^{i+i} > A^{1+i(4)}$$

provided $Q_i$ is large.

5. The compounds of linear forms

Suppose $k > 1$ and let $\sigma, \tau, \ldots$ denote subsets of \{1, 2, ..., $k\}$. Write $\sigma'$ for the complement of $\sigma$ in \{1, 2, ..., $k\}. Define (−1)$^n$ by
For any integer $p$ with $1 \leq p < k$, let $C(k, p)$ consist of all sets $\sigma$ with exactly $p$ elements. Then $C(k, p)$ consists of $l(p) = \binom{k}{p}$ sets $\sigma$.

Let
\[
L_i = \alpha_{i1} x_1 + \ldots + \alpha_{ik} x_k \quad (i = 1, \ldots, k)
\]
be $k$ linear forms of determinant $1$ in $x = (x_1, \ldots, x_k)$. Let $\rho$ with $1 \leq \rho < k$ be fixed at the moment. For every $\sigma \in C(k, p)$, $\tau \in C(k, p)$, write $x_{\sigma \tau}$ for the $(p \times p)$-determinant formed from all $i$th rows with $i \in \sigma$ and all $j$th columns with $j \in \tau$ of the matrix $(x_{ij})$. We shall construct linear forms $L^{(p)}_\sigma$ in vectors $x^{(p)}$ with $l(p)$ components which are denoted by $x_\tau$ where $\tau \in C(k, p)$. Namely, for every $\sigma \in C(k, p)$, we put
\[
L^{(p)}_\sigma(x^{(p)}) = \sum_{\tau \in C(k, p)} x_{\sigma \tau} x_\tau.
\]
We call these linear forms the $p$th compounds of $L_1, \ldots, L_k$. There are exactly $l(p)$ such $p$th compounds.

Again, for every $\sigma$ in $C(k, p)$, put
\[
L^{(p)}_\sigma(x^{(p)}) - \sum_{\gamma \in C(k, p)} (-1)^{\rho} (-1)^{\gamma} x_{\sigma \gamma},
\]
Let $e^{(p)}_{\tau}$ be the basis vector whose component $x_\tau = 1$, and all of whose other components are zero. Then for any $\tau_1, \tau_2$ in $C(k, p)$, one has
\[
\sum_{\alpha \in C(k, p)} L^{(p)}_\alpha(e^{(p)}_{\tau_1}) L^{(p)}_\alpha(e^{(p)}_{\tau_2}) = \begin{cases} 1 & \text{if } \tau_1 = \tau_2 \\ 0 & \text{otherwise}. \end{cases}
\]
This follows from Laplace's rule on the expansion of determinants, applied to the determinant $[x_{ij}]$ ($1 \leq i, j \leq k$). It follows immediately that
\[
\sum_{\alpha \in C(k, p)} L^{(p)}_\alpha(x^{(p)}) L^{(p)}_\alpha(y^{(p)}) \equiv \sum_{\alpha \in C(k, p)} x_\sigma y_\sigma.
\]
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 $L^{(p)}_\sigma$ where $\sigma \in C(k, p)$ and the system of forms $L^{(p)}_\sigma$ where $\sigma \in C(k, p)$ are adjoint to each other.

Throughout the rest of this section let $p$ in $1 \leq p < k$ and $l = l(p)$ be fixed. The inequalities
\[
|L_i(x)| \leq 1 \quad (i = 1, \ldots, k)
\]
define a parallelopiped II in $E^k$. Since $L_1, ..., L_k$ have determinant 1, it follows from determinant theory that the $l$ forms $L_\sigma^{(p)}(x^{(p)})$ with $\sigma \in C(k, p)$ again have determinant 1. In particular these $l$ linear forms are linearly independent. Hence the inequalities

$$|L_\sigma^{(p)}(x^{(p)})| \leq 1 \quad (\sigma \in C(k, p))$$

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

Denote the successive minima of II by $\lambda_1, ..., \lambda_k$, and for every $\sigma$ write

$$\lambda_\sigma = \prod_{i=1}^{k} \lambda_i.$$  \tag{35}

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

$$\lambda_{\sigma_1} \leq \lambda_{\sigma_2} \leq ... \leq \lambda_{\sigma_l}.$$  

Denote the successive minima of II$^{(p)}$ by $v_1, v_2, ..., v_l$.

**Theorem B. (Mahler.)** One has

$$v_j < \lambda_{\sigma_j} << v_i \quad (1 \leq j \leq l(p)),$$

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 II$^{(p)}$ and of the $p$th compound of II differ only by bounded factors.

Now let $A_1, ..., A_k$ be positive reals with

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

Then if we put

$$A_\sigma = \prod_{i=1}^{k} A_{\sigma_i},$$

we have

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

The inequalities

$$|L_i(x)| \leq A_i \quad (i = 1, ..., k)$$

define a parallelopiped II$^*$ in $E^k$, and the inequalities

$$|L_\sigma^{(p)}(x^{(p)})| \leq A_\sigma \quad (\sigma \in C(k, p))$$

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

Corollary to Theorem B. Define \( \lambda_i \) (\( 1 \leq i \leq k \)), \( \lambda_\sigma \) (\( \sigma \in C(k, p) \)), \( v_i \) (\( 1 \leq i \leq l \)) as above, but with reference to \( \Pi_A \) and \( \Pi_A^{(p)} \) instead of to \( \Pi \) and \( \Pi^{(p)} \). Then one has again

\[
v_i < \lambda_\sigma < < v_j \quad (1 \leq j \leq l(p)).
\]

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

6. Special linear forms

Suppose now 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

\[
L_1(x) = x_1 - \alpha_1 x_n, \quad L_2(x) = x_2 - \alpha_2 x_n, \ldots, \quad L_n(x) = x_n - \alpha_n x_n, \quad L_k(x) = x_k.
\]

For every \( p \) in \( 1 \leq p \leq n = k - 1 \), there are \( l(p) \) compound forms \( L_\sigma^{(p)}(x^{(p)}) \) 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 \( L_\sigma^{(p)}(x^{(p)}) \) 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 \( L_\sigma^{(p)} \) 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 \( L_\sigma^{(p)} \) 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 \( S^{(k-p)} \) consist of all sets \( \sigma' \) with \( \sigma \in S^{(p)} \). Replacing \( p \) by \( k - p \) we thus have to show that for every \( p \) in \( 1 \leq p \leq k - 1 = n \),

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

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

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

\[
L_\sigma^{(p)}(x^{(p)}) = x_\sigma + \sum_{i \notin \sigma} \pm \alpha_i x_{i+k}.
\]

Here \( \sigma - i + k \) denotes the set obtained from \( \sigma \) 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 \( \alpha_i \) with \( i \in \sigma \). These numbers form a subset of \( 1, \alpha_1, \ldots, \alpha_n \), 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 \( \tau \) in \( C(k, p) \) there is a \( \sigma \in S^{(p)} \) such that the coefficient of \( x_\tau \) in \( L_\sigma \) 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 \( L_1(x), \ldots, L_k(x) \) by (44). Suppose \( A_1, \ldots, A_k \) are positive and have

\[
A_1 A_2 \ldots A_k = 1
\]

and

\[
A_1 < 1, \ldots, A_n < 1; \quad A_k > 1.
\]

Let \( \lambda_1, \ldots, \lambda_n \) be the successive minima of the parallelopiped \( \Pi_\lambda \) given by

\[
|L_i(x)| \leq A_i \quad (i = 1, \ldots, k).
\]

Then for every \( \delta > 0 \) there is an \( \delta_1 = \delta_1(\delta, \alpha_1, \ldots, \alpha_n) \) such that

\[
\lambda_1 > \delta_1
\]

provided

\[
\delta > \max (A_k, \delta_3).
\]

**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 \( \lambda_1 \geq \delta_1 \geq \delta \) provided \( \delta > \max (A_k, \delta_3) \).

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

\[
\lambda_{k-p} \geq \lambda_{k-p+1} \delta
\]

provided \( \delta > \max (A_k, \delta_3) \) where \( \delta_3 = \delta_3(\delta, \alpha_1, \ldots, \alpha_n) \). Namely, repeated application of (51) yields \( \lambda_1 > \delta_1 \delta > \delta \). Since \( \delta > 0 \) was arbitrary, the lemma follows.

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

\[
\lambda_1 A_1^{(p)} > \delta
\]

if \( \delta > \max (A_k, \delta_4) \). Take at first the case when \( p = 1 \). Then since there is an integer point \( x_0 \neq 0 \) with \( |L_i(x_0)| < \lambda_1 A_i \) \( (i = 1, \ldots, k) \), it follows that

\[
1 < \max (\lambda_1 A_1, \ldots, \lambda_1 A_k) = \lambda_1 A_k = \lambda_1 A_1^{(p)},
\]

and (52) is true. Now assume that \( 1 < p < n = k - 1 \). Put
Then by (46) and (47) we have
\[ B_i = \frac{A_j}{A_j^{1/p}} \quad (i \in \sigma). \]

By definition of \( \lambda_1 \) there is an integer point \( x_0 \neq 0 \) with \( |L_i(x_0)| \leq \lambda_1 A_i \quad (i = 1, \ldots, k) \).

Now since \( \Pi_d \) has volume \( 2^d \), the first minimum \( \lambda_1 \) is at most 1 by Minkowski's theorem. Hence \( \lambda_1 A_i < 1 \) \((i = 1, 2, \ldots, n)\) by (47). Hence in \( x_0 = (x_1, \ldots, x_n) \), the last coordinate \( x_k \) cannot be zero. Hence the vector \( y_0 = (x_1, \ldots, x_{p-1}, x_k) \) in \( E^p \) is not 0. The linear forms \( L_i \) with \( i \in \sigma \) may be interpreted as forms in \( y = (x_1, \ldots, x_{p-1}, x_k) \). We have
\[ |L_i(y_0)| \leq \lambda_1 A_i - \lambda_1 A_j^{1/p} B_i \quad (i \in \sigma). \]

Thus the parallelepiped in \( E^p \) defined by
\[ |L_i(y)| < B_i \quad (i \in \sigma) \]
has a first minimum \( \mu_1 \) with \( \mu_1 \leq \lambda_1 A_j^{1/p} \). In view of (54) and (55) it follows from our induction hypothesis that
\[ \lambda_1 A_j^{1/p} \geq \mu_1 > Q^{-d} \]
provided \( Q \geq \max (B_k, Q_4) \). Since \( B_k = A_k/A_1^{1/p} < A_k \), the inequality (52) is true provided \( Q \geq \max (A_k, Q_4) \).

Recall that \( S^{(p)} \) 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 \( Q \geq \max (A_k, Q_4) \).

Let \( L^{(p)}(x^{(p)}) \) with \( \sigma \in C(k, p) \) be the \( \sigma \)th compound forms of \( L_1, \ldots, L_k \), and define the parallelepiped \( \Pi_\sigma \) by (41). The first minimum \( v_1 \) of \( \Pi_\sigma \) satisfies \( v_1 > \lambda_1 \lambda_2 \ldots \lambda_p > \lambda_\sigma \) by (42), and hence we have
\[ v_1 A_\sigma > \lambda_\sigma A_\sigma > Q^{-p} \quad (\sigma \in S^{(p)}) \]
by (52) provided \( Q \) is large. Since \( \delta > 0 \) in (52) was arbitrary, we have in fact
\[ v_1 A_\sigma > Q^{-\delta(21)} \quad (\sigma \in S^{(p)}) \]
if \( Q \geq \max (A_k, Q_4) \). Here \( Q_4 = Q_4(\delta, x_1, \ldots, x_n) \) and \( l - l(p) = \binom{k}{p} \).

We now apply the corollary proved in section 4 to the proper system \( L^{(p)}(\sigma \in C(k, p)) \), \( S^{(p)} \). The inequality (16) now becomes (56), and hence it is true if \( Q \) is large. It follows that
\[ v_{l-1} > v_1 Q^{-d} \]
(57)
provided (17) holds, i.e. provided $Q > \max (A_{\sigma} (\sigma \in C(k, p)), Q_{t})$. Since $A_{\sigma} < A_{k}$ by (47), the last condition is fulfilled if $Q > \max (A_{k}, Q_{t})$. Now by (42) again we have

$$v_{t} < \frac{\lambda_{k-p+1}}{\lambda_{k-p+2}} \cdots \frac{\lambda_{k}}{\lambda_{k-1}}$$

and

$$v_{t-1} < \frac{\lambda_{k-p} \lambda_{k-p+2} \cdots \lambda_{k}}{\lambda_{k-p+1}}$$

Thus (57) yields

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

if $Q > \max (A_{k}, Q_{t})$. Since $\delta > 0$ was arbitrary, we therefore have (51) if $Q > \max (A_{k}, Q_{t})$. This proves the lemma.

**Lemma 5.** Suppose $x_{1}, ..., x_{n}$ are as in Lemma 4, and put $k = n + 1$. Define linear forms $M_{1}, ..., M_{k}$ by

$$M_{1}(x) = x_{1}, \quad M_{2}(x) = x_{2}, \quad ..., \quad M_{n}(x) = x_{n}, \quad M_{k}(x) = x_{1} x_{2} + ... + x_{n} x_{n-1} + x_{n}. \quad (58)$$

Let $B_{1}, ..., B_{k}$ be positive numbers with

$$B_{1} B_{2} \cdots B_{k} = 1, \quad (59)$$

$$B_{1} > 1, ..., B_{k} > 1, B_{k} < 1. \quad (60)$$

Write $\mu_{1}, ..., \mu_{k}$ for the successive minima of the parallelopiped $\Pi_{\delta}$ defined by

$$|M_{i}(x)| \leq B_{i} \quad (i = 1, ..., k). \quad (61)$$

For every $\delta > 0$ there is a $Q_{b} = Q_{b}(\delta, x_{1}, ..., x_{n})$ such that

$$\mu_{1} > Q^{-\delta} \quad (62)$$

provided

$$Q > \max (B_{b}^{-1}, Q_{b}). \quad (63)$$

Proof. This lemma is dual to Lemma 4. Write $A_{i} = B_{i}^{-1} (i = 1, ..., k)$. Then (46), (47) hold. The forms $M_{1}, ..., M_{k}$ are adjoint to $L_{1}, ..., L_{k}$ given by (44), and hence the forms $M_{1}/B_{1}, ..., M_{k}/B_{k}$ are adjoint to $L_{1}/A_{1}, ..., L_{k}/A_{k}$. Thus if $\lambda_{1}, ..., \lambda_{k}$ are the successive minima of $\Pi_{\delta}$ defined in Lemma 4, then it is well known that

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

(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 $M_{1}, ..., M_{k}$ are essentially the $(k-1)$-st compounds of $L_{1}, ..., L_{k}$. Namely, it follows that $\mu_{k+1-i}$ is of the same order of magnitude as $\lambda_{1} \lambda_{i-1} \lambda_{i+1} ... \lambda_{k}$, hence as $\lambda_{i}^{-1}$.)
By Lemma 4 we have \( \lambda_{k-1} \geq \cdots \geq \lambda_2 \geq \lambda_1 > Q^{-\delta} \), and hence \( \lambda_k << (\lambda_1 \cdots \lambda_{k-1})^{-1} << Q^{\delta} \). Thus by (64), \( \mu_1 > Q^{-\delta} \). Since \( \delta > 0 \) was arbitrary, we have in fact (62) provided (63) holds with a suitably large \( Q_0 \).

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

\[
\|q\alpha_1\| \cdots \|q\alpha_n\| \cdot q^{1+\varepsilon} < 1. \tag{65}
\]

Put

\[
k = n + 1, \quad \eta = \varepsilon/k, \tag{66}
\]

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

Now if one of the numbers \( A_1, \ldots, A_n \) were at least 1, say if \( A_1 > 1 \), then

\[
\|q\alpha_1\| \cdots \|q\alpha_n\| q^{1+\varepsilon-\eta} < 1,
\]

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

\[
A_k = q^{-\eta} (\|q\alpha_1\| \cdots \|q\alpha_n\|)^{-1} > q^{1+\varepsilon-\eta} = q^{1+\varepsilon}, \tag{68}
\]

and (67) together with Roth's theorem yields

\[
A_k < (\|q\alpha_1\| \cdots \|q\alpha_n\|)^{-1} < q^{2n}, \tag{69}
\]

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 \( x_0 \) be the point \((p_1, \ldots, p_n, q)\) in \( E^2 \). Then (67) and (68) imply that

\[
|L_i(x_0)| \leq A_i q^{-\eta} \quad (i = 1, \ldots, k), \tag{70}
\]

where \( L_1, \ldots, L_k \) are the forms given by (44). Thus the parallelepiped \( \Pi_A \) defined by \( |L_i(x)| \leq A_i \) (\( i = 1, \ldots, k \)) has a first minimum \( \lambda_i \leq q^{-\eta} \). The number \( Q = q^{2n} \) satisfies \( Q > A_k \) by (69), and we still have \( \lambda_i \leq Q^{-\eta(k+\eta)} \). By Lemma 4 this is impossible 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

\[
\|q_1 \alpha_1 + \cdots + q_n \alpha_n\| : |q_1 \cdots q_n|^{1+\varepsilon} < 1. \tag{71}
\]

We may assume that \( 0 < \varepsilon < 1 \). Put
SIMULTANEOUS APPROXIMATION TO ALGEBRAIC NUMBERS BY RATIONALS

\[ k = n + 1, \quad \eta = \epsilon/k, \quad q = |q_1 q_2 \ldots q_n|. \]  (72)

\[ B_i = |q_i| q^n \quad (i = 1, \ldots, n), \quad B_k = (B_1 B_2 \ldots B_n)^{-1}. \]  (73)

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

\[ B_k = q^{-\eta n} |q_1 q_2 \ldots q_n|^{-1} = \left\| q_1 \alpha_1 + \ldots + q_n \alpha_n \right\| q^{-\eta} \]  (74)

by (71), (72), (73), and

\[ B_k^{-1} = q^{\eta n} |q_1 q_2 \ldots q_n| \leq q^{2n} \]  (75)

by (72), (73).

Let \( p \) be the integer with \( \left\| q_1 \alpha_1 + \ldots + q_n \alpha_n \right\| = \left\| q_1 \alpha_1 + \ldots + q_n \alpha_n + p \right\| \), and let \( x_0 \) be the point \( (q_1, \ldots, q_n, p) \) in \( E^k \). Then in view of (73), (74) we have

\[ |M_i(x_0)| \leq B_i q^{-\eta} \quad (i = 1, \ldots, k), \]  (76)

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

References


Received February 27, 1970