ENRICO BOMBIERI

# Simultaneous approximations of algebraic numbers

*Séminaire N. Bourbaki*, 1973, exp. nº 400, p. 1-20

<http://www.numdam.org/item?id=SB_1971-1972__14__1_0>

Séminaire BOURBAKI 400-01
24e année, 1971/72, n° 400 Novembre 1971

# SIMULTANEOUS APPROXIMATIONS OF ALGEBRAIC NUMBERS

[following W. M. SCHMIDT]

by Enrico BOMBIERI

I. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be real numbers. Dirichlet's theorem in Diophantine Approximation states that

THEOREM (Dirichlet).- For every $N \geq 1$ there is $q$, $1 \leq q \leq N$, such that

$$\|q\alpha_1\| \leq N^{-1/n}, \ldots, \|q\alpha_n\| \leq N^{-1/n},$$

where $\|\ \|$ denotes the distance from the nearest integer.

COROLLARY.- Let $1, \alpha_1, \ldots, \alpha_n$ be real numbers, linearly independent over $\mathbb{Q}$. Then there are infinitely many integers $q$ such that

$$\|q\alpha_1\| \leq q^{-1/n}, \ldots, \|q\alpha_n\| \leq q^{-1/n}.$$

In 1955, after previous work by Thue, Siegel, Dyson, Gel'fond and Schneider it was proved by Roth that

ROTH'S THEOREM.- Let $\alpha$ be irrational algebraic and let $\varepsilon > 0$. There are only finitely many integers $q$ such that

$$\|q\alpha\| \leq q^{-1-\varepsilon}.$$

Now Roth's theorem has been generalized by W. M. Schmidt to the case of simultaneous approximations.

SCHMIDT'S THEOREM 1.- Let $1, \alpha_1, \ldots, \alpha_n$ be algebraic real numbers, linearly independent over $\mathbb{Q}$, and let $\varepsilon > 0$. There are only finitely many integers $q$ such that

$$\|q\alpha_1\| \ldots \|q\alpha_n\| \leq q^{-1-\varepsilon}.$$

COROLLARY.- There are only finitely many integers $q$ such that

$$\|q\alpha_1\| \leq q^{-1/n-\varepsilon}, \dots, \|q\alpha_n\| \leq q^{-1/n-\varepsilon}.$$

Schmidt also proves a dual version of this result :

THEOREM 2.- Let $\alpha_1, \dots, \alpha_n$ be as in Theorem 1, and let $\varepsilon > 0$. There are only finitely many n-ples of non-zero integers $q_1, \dots, q_n$ such that

$$\|q_1\alpha_1 + \dots + q_n\alpha_n\| \leq |q_1 \dots q_n|^{-1-\varepsilon}.$$

COROLLARY.- Let $\alpha$ be algebraic, $k$ a positive integer and $\varepsilon > 0$. There are only finitely many algebraic numbers $\omega$ of degree $\leq k$ such that

$$|\alpha - \omega| \leq H(\omega)^{-k-1-\varepsilon}$$

where $H(\omega)$ is the height of $\omega$ (maximum coefficient of an irreducible integral defining polynomial of $\omega$).

If $k = 1$ this reduces to Roth's theorem ; a weaker result, with an exponent $2k + \varepsilon$ instead of $k + 1 + \varepsilon$ has been proved by Wirsing [3] with a different method.

Schmidt's proof of these results uses Roth's method, but the extension is not straightforward and many original ideas are needed. In order to present Schmidt's arguments, it is therefore worthwhile to sketch Roth's proof.

II. Roth's Proof. For a neat exposition of Roth's proof we refer to Cassels [1]. Roth's theorem is obtained combining the following two results :

PROPOSITION 1.- Let $\alpha$ be algebraic, let $\varepsilon > 0$ and let $r_1, \dots, r_m$ be positive integers.

For $m \geq m_o(\alpha, \varepsilon)$ there is a polynomial

$$P \in \mathbb{Z}[x_1, \dots, x_m]$$

not identically $0$ of degree $\leq r_h$ in $x_h$, such that

(i) $\quad |P| \leq C_1^{r_1 + \cdots + r_m}$ ;

(ii) $\quad D^J P(\alpha, \alpha, \ldots, \alpha) = 0$

if $J = (j_1, \ldots, j_m)$ and

(2.1) $$\sum_{h=1}^{m} j_h/r_h \leq (\tfrac{1}{2} - \varepsilon)m .$$

Here $|P|$ is the sum of the moduli of the coefficients of $P$ and $D^J$ is the usual differential operator $(\partial/\partial x_1)^{j_1} \ldots (\partial/\partial x_m)^{j_m}$. The constant $C_1$ depends only on $\alpha$.

The proof is simple. Considering the $(r_1 + 1) \ldots (r_m + 1)$ coefficients of $P$ as unknowns one has a system of homogeneous linear equations $D^J P(\alpha) = 0$. Now if $\alpha$ is algebraic of degree $s$, the equation

$$\frac{1}{J!} D^J P(\alpha, \ldots, \alpha) = 0$$

splits in a system of $s$ linear equations in the coefficients of the polynomial $P$, with <u>integral</u> coefficients $\leq C_2^{r_1 + \cdots + r_m}$ where $C_2 = C_2(\alpha)$. Since equation (2.1) has at most $\frac{1}{\varepsilon\sqrt{m}} (r_1 + 1) \ldots (r_m + 1)$ solutions, we get a system of $\leq \frac{s}{\varepsilon\sqrt{m}} (r_1 + 1) \ldots (r_m + 1)$ equations in $(r_1 + 1) \ldots (r_m + 1)$ unknowns with integral coefficients $\leq C_2^{r_1 + \cdots + r_m}$. This is easily solved using Dirichlet's box principle, provided $\frac{s}{\varepsilon\sqrt{m}} \leq \frac{1}{2}$, that is $m \geq m_0(\alpha, \varepsilon)$, obtaining a non-zero solution satisfying (i).

Now let $\beta_h = p_h/q_h$ be $m$ approximations to $\alpha$ such that

(2.2) $$|\alpha - p_h/q_h| < q_h^{-k} ,$$

let $P$ be the polynomial of Proposition 1 and let $\nu = (\nu_1, \ldots, \nu_m)$ be such that, if we write

$$Q = \frac{1}{\nu!} D^{\nu} P$$

we have

$$Q(\beta_1,\dots,\beta_m) \neq 0 .$$

Then $Q(\beta) = Q(\beta_1,\dots,\beta_m)$ is a rational number with denominator $\leq q_1^{r_1} \dots q_m^{r_m}$ therefore

$$|Q(\beta)| \geq q_1^{-r_1} \dots q_m^{-r_m} .$$

Now assume that

$$\sum_{h=1}^{m} \nu_h / r_h < \varepsilon m .$$

Then $Q$ is not identically $0$ and $Q(\alpha) = 0$, therefore

$$|Q(\beta)| = |Q(\beta) - Q(\alpha)| \leq \Sigma \frac{1}{J!} |D^{J+\nu} P(\alpha)| \; |\alpha - \beta|^J$$

$$\leq C_3^{r_1 + \dots + r_m} \max |\alpha - \beta|^J$$

where the max is over the n-ples $J$ such that

$$\sum_{h=1}^{m} (j_h + \nu_h)/r_h \geq (\tfrac{1}{2} - \varepsilon)m .$$

If there are infinitely many approximations satisfying (2.2) one can take $q_1^{r_1} \sim q_2^{r_2} \sim \dots \sim q_m^{r_m}$ and more precisely

$r_1$ very large

$$r_h = [\, r_1 \frac{\log q_1}{\log q_h} ] + 1 \qquad\qquad h = 2,\dots,m$$

$q_1$ very large

and now

$$|\alpha - \beta|^J \leq q_1^{-kj_1} \dots q_m^{-kj_m}$$

$$\leq q_1^{-kr_1 \, \Sigma j_h / r_h} .$$

Since
$$\sum_{h=1}^{m} j_h/r_h \geq (\tfrac{1}{2} - \varepsilon)m - \sum_{h=1}^{m} \nu_h/r_h \geq (\tfrac{1}{2} - 2\varepsilon)m$$
we deduce
$$|Q(\beta)| \leq C_4^{r_1 + \dots + r_m} q_1^{-k(\frac{1}{2} - 2\varepsilon)mr_1} .$$
On the other hand,
$$|Q(\beta)| \geq q_1^{-r_1} \dots q_m^{-r_m} \geq C_5^{-r_1 - \dots - r_m} q_1^{-mr_1} .$$

If we choose $q_1, q_2, \dots$ rapidly increasing then $r_1, r_2, \dots$ are rapidly decreasing and we may ensure that $r_1 + \dots + r_m \leq 2r_1$. Hence, letting $q_1 \to \infty$ we find
$$m \geq k(\tfrac{1}{2} - 2\varepsilon)m$$
and
$$k(\tfrac{1}{2} - 2\varepsilon) \leq 1 .$$

Since $\varepsilon$ is arbitrary, $k \leq 2$ and Roth's theorem follows.

The difficulty consists in showing that $\Sigma \nu_h/r_h$ is small without putting conditions of the sort " $q_1$ is not too large compared with $q_m$ ". Now using an ingenious inductive method, Roth obtains

PROPOSITION 2.- Let $0 < \delta < 16^{-m}$, let $P \in \mathbb{Z}[x_1, \dots, x_m]$ of degree $\leq r_h$ in $x_h$ and not identically $0$, let
$$\delta r_h \geq r_{h+1} \quad , \quad h = 1, \dots, m-1 \quad , \quad \delta r_m \geq 10$$
and let $\beta_h = p_h/q_h$ be such that

(i) $\delta r_1 \log q_1 \gg \log|P|$

(ii) $\delta \log q_h \gg m$ , $r_h \log q_h \geq r_1 \log q_1$ .

Then there is $\nu = (\nu_1, \dots, \nu_m)$ with
$$D^{\nu} P(\beta_1, \dots, \beta_m) \neq 0$$

and

$$\sum_{h=1}^{m} \nu_h/r_h \leq 10^m \delta^{2^{-m}} .$$

It is clear that, taking $\delta$ sufficiently small, Proposition 2 is sufficient to complete the proof of Roth's theorem along the lines mentioned before.

The proof of Proposition 2 is rather intricate, and because of lack of space and time, we cannot give an indication of the ideas involved in it.

III. Schmidt's Proof. The index. In the previous argument, instead of working with polynomials of degree $\leq r_h$ in $x_h$ we could work with polynomials in pairs of variables $x_h$, $y_h$, $h = 1,\dots,m$ and homogeneous of degree $r_h$ in the pair $x_h$, $y_h$. Instead of asking that a derivative $D^J P$ should vanish at a point $(\beta_1,\dots,\beta_m)$ we could introduce the linear forms

$$L_h = x_h - \beta_h y_h$$

and ask that $P$ belong to the ideal in $\mathbb{R}[x_1, y_1,\dots,x_m, y_m]$ generated by polynomials

$$L_1^{i_1} \dots L_m^{i_m}$$

with $i_h > j_h$ for $h = 1,\dots,m$. This remark leads Schmidt to the following definitions.

Let $R = \mathbb{R}[x_{11},\dots,x_{1\ell};\dots;x_{m1},\dots,x_{m\ell}]$ be the ring of polynomials in $m\ell$ variables and let $L_1,\dots,L_m$ be linear forms (not $0$) of the type

$$L_h = L_h(x_{h1},\dots,x_{h\ell}) .$$

For $c \geq 0$ let $I(c)$ be the ideal in $R$ generated by all $L^J$ where $J = (j_1,\dots,j_m)$ satisfies

$$\sum_{h=1}^{m} j_h/r_h \geq c ,$$

where $r_1,\dots,r_m$ are positive integers.

DEFINITION.- The index of $P$ with respect to $(L_1,\dots,L_m\,;\,r_1,\dots,r_m)$ is the largest $c$ with $P \in I(c)$ and $c = +\infty$ if $P$ is identically $0$.

We have

$$\mathrm{ind}(P+Q) \geq \min(\mathrm{ind}\,P\,,\,\mathrm{ind}\,Q)$$

$$\mathrm{ind}\,PQ = \mathrm{ind}\,P + \mathrm{ind}\,Q\,.$$

If $J$ is a $\ell m$-uple

$$J = (j_{11},\dots,j_{1\ell}\,;\,\dots\,;\,j_{m+1},\dots,j_{m\ell})$$

one puts

$$(J/r) = \sum_{h=1}^{m} (j_{h1} + \dots + j_{h\ell})/r_h$$

and

$$P^{(J)} = \frac{1}{J!} D^J P\,.$$

One gets easily

$$\mathrm{ind}\,P^{(J)} \geq \mathrm{ind}\,P - (J/r)\,.$$

The first step in Schmidt's proof is to obtain the analogue of Propositions 1 and 2. We have

PROPOSITION A.- Let $L_j = \alpha_{j1}X_1 + \dots + \alpha_{j\ell}X_\ell$, $j = 1,\dots,\ell$, be independent linear forms, with algebraic integers as coefficients. Let

$$L_{hj} = L_j(x_{h1},\dots,x_{h\ell})$$

and let $\varepsilon > 0$.

For $m \geq m_o(\alpha\,,\,\varepsilon)$ there is a polynomial

$$P \in \mathbb{Z}[x_{11},\dots,x_{m\ell}]$$

not identically $0$, homogeneous of degree $r_h$ in $x_{h1},\dots,x_{h\ell}$ such that

(i) $|P| \leq C_5^{\,r_1+\dots+r_m}$ ;

(ii) $\mathrm{ind}\,P \geq (\ell^{-1} - \varepsilon)m$ ,

with respect to $(L_{1j},\dots,L_{mj}\,;\,r_1,\dots,r_m)$ for $j = 1,\dots,\ell$. Moreover, if we write

$$\frac{1}{J!} D^J P = \Sigma\, d^J(j)\, L_{11}^{j_{11}} \dots L_{m\ell}^{j_{m\ell}}$$

we have

$$|d^J(j)| \leq C_6^{r_1 + \dots + r_m}$$

for all $J$ , $j$ and $d^J(j) = 0$ unless for $k = 1,\dots,\ell$

(iii) $$\Big| \sum_{h=1}^{m} j_{hk}/r_h - \ell^{-1} m \Big| \leq \ell m \varepsilon + \ell (J/r) \ .$$

The proof of Proposition A is rather similar to that of Proposition 1.

Proposition 2 can also be extended, and one gets

PROPOSITION B.- Let $0 < \delta < C_7^{-2^m}$ , $0 < \tau \leq 1$ , let $P \in Z[x_{11},\dots,x_{m\ell}]$ be not identically $0$ , homogeneous of degree $r_h$ in $x_{h1},\dots,x_{h\ell}$ , let

$$\delta r_h \geq r_{h+1} \ , \qquad h = 1,\dots,m-1$$

and let $$M_h = m_{h1} x_{h1} + \dots + m_{h\ell} x_{h\ell}$$

be non-zero linear forms whose coefficients are integral and have no common factor. Let also

$$|M_h| = \max_j |m_{hj}|$$

and assume

(i) $\delta \tau r_1 \log|M_1| \gg \log|P|$ ;

(ii) $\delta \tau \log|M_h| \gg m$ , $r_h \log|M_h| \geq \tau r_1 \log|M_1|$ for $h = 1,\dots,m$ .

Then the index of $P$ with respect to $(M_1,\dots,M_m ; r_1,\dots,r_m)$ satisfies

$$\operatorname{ind} P \leq 10^m \delta^{2^{-m}} \ .$$

The ideas in the proof are the same as Roth's, but the technical difficulties are of course much greater.

The conclusion that may be drawn from Propositions A and B is, except in case $\ell = 2$ substantially weaker than Schmidt's theorems. In Roth's case, one takes

$$\ell = 2 \quad , \quad L_1 = X_1 - \alpha X_2 \quad , \quad L_2 = X_2$$

and in Schmidt's case one would take

$$L_j = X_j - \alpha_j X_\ell \quad , \quad j = 1,\dots,\ell - 1 \quad , \quad L_\ell = X_\ell \ .$$

However, in order to conclude the proof, one eventually has to consider many other sets of linear forms.

## IV. Schmidt's Proof. The theorem of the next to last minimum.

Let $K$ be a symmetrical convex body in $\mathbf{R}^n$ centered at the origin and let $V(K)$ be its volume. For $\lambda > 0$ let $\lambda K$ be the corresponding homothetic convex body. The successive minima $\lambda_1,\dots,\lambda_n$ are defined as follows :

$$\lambda_i = \inf\{\ \lambda \mid \lambda K \text{ contains } i \text{ linearly independent points of } \mathbf{Z}^n\} \ .$$

A basic theorem of Minkowski states

SECOND THEOREM OF MINKOWSKI.- We have

$$\frac{2^n}{n!} \leq \lambda_1 \dots \lambda_n V(K) \leq 2^n \ .$$

We need another definition. Let

$$M_i = \beta_{i1}X_1 + \dots + \beta_{i\ell}X_\ell$$

be independent linear forms with algebraic coefficients. Let $S$ be a subset of $\{1,2,\dots,\ell\}$ .

DEFINITION.- $\{M_1,\dots,M_\ell \,;\, S\}$ is regular if

(i) for $j \in S$ the non-zero elements among $\beta_{j1},\dots,\beta_{j\ell}$ are linearly independent over $\mathbf{Q}$ ;

(ii) for every $k \leq \ell$ there is $j \in S$ with $\beta_{jk} \neq 0$ .

Now let $L_1,\dots,L_\ell$ be again linear forms with algebraic coefficients and let $S \subset \{1,2,\dots,\ell\}$ .

DEFINITION.- $\{L_1,\dots,L_\ell ; S\}$ is proper if $\{M_1,\dots,M_\ell ; S\}$ is regular, where the $M_i$ are the adjoint forms of $L_j$ .

Now Schmidt proves

THEOREM <u>of the next to last minimum</u>.- Let $\{L_1,\dots,L_\ell ; S\}$ be proper and let $A_1,\dots,A_\ell$ be positive reals such that

$$A_1 \dots A_\ell = 1 \quad , \quad A_j \geq 1 \text{ if } j \in S .$$

The set in $\mathbb{R}^\ell$

$$|L_j(x)| \leq A_j \quad , \qquad j = 1,\dots,\ell$$

is a symmetric convex body centered at $0$ ; let $\lambda_1,\dots,\lambda_\ell$ denote its successive minima.

For every $\delta > 0$ there is

$$Q_o = Q_o(\delta ; L_1,\dots,L_\ell ; S)$$

such that

$$\lambda_{\ell-1} > Q^{-\delta}$$

provided

$$Q \geq \max(A_1,\dots,A_\ell ; Q_o) .$$

This is a consequence of Propositions A and B. The proof of the theorem is obtained through various reduction steps.

a) It is sufficient to prove the result when $A_j = Q^{c_j}$ and $c_1,\dots,c_\ell$ are fixed constants such that

$$c_1 + \dots + c_\ell = 0 \quad , \quad |c_j| \leq 1 \text{ for all } j \quad , \quad c_j \geq 0 \text{ for } j \in S .$$

This is easy, because one can show that if one modifies slightly the $A_j$ (say by a factor $Q^{b_j}$ , with $|b_j| < \delta/2$ ) then the minimum $\lambda_{\ell-1}$ is modified by a factor of that order of magnitude. Thus one may suppose that $A_j = Q^{c_j}$ where the $c_j$ belong to a finite set depending only on $\delta$ .

b) We may suppose that the coefficients $\alpha_{ij}$ are algebraic integers. In fact if $q$ is a common denominator for the $\alpha_{ij}$, the successive minima of $|qL_j| \leq A_j$ are $q^{-1}$ times the successive minima of $|L_j| \leq A_j$.

Now assume the theorem is false. There is $b > 0$ and an increasing sequence $Q_1, Q_2, \ldots$ going to infinity and $\ell$-uples $y_{h1}, \ldots, y_{h\ell}$ of linearly independent points of $\mathbb{Z}^\ell$ such that

$$|L_j(y_{hk})| \leq Q_h^{c_j - b}$$

for $j = 1, \ldots, \ell$, $k = 1, \ldots, \ell$ and $h = 1, 2, \ldots$.

We let $M_h$, $h = 1, 2, \ldots$ be the (unique up to sign) linear form with integer coefficients without common factor, such that

$$M_h(y_{hk}) = 0 \qquad \text{for } k = 1, \ldots, \ell .$$

Let us assume that $Q_1$ is large, take (as in Roth's proof)

$$r_h = \left[ r_1 \frac{\log Q_1}{\log Q_h} \right] + 1$$

where $r_1$ is very large and let $P$ be the polynomial of Proposition A. Then, using property (ii) of $P$ (the lower bound for the index) Schmidt shows that $P$ has index

$$\operatorname{ind} P \geq C_8 \, b \, m$$

with respect to $(M_1, \ldots, M_m ; r_1, \ldots, r_m)$, for some constant

$$C_8 = C_8(\ell) > 0 .$$

The proof goes as follows.

Let

$$y_h = \sum_{k=1}^{\ell} a_k \, y_{hk}$$

be a linear combination of $y_{h1}, \ldots, y_{h\ell}$ with integral coefficients $a_k$, with $|a_k| \leq Q_1^{\varepsilon}$. If we use Proposition A and $|L_j(y_{hk})| \leq Q_h^{c_j - b}$ we get

$$\frac{1}{J!}|P^J(y_1,\dots,y_m)| \le (C_9 Q_1^{\varepsilon})^{r_1+\dots+r_m} \max \prod_{h=1}^{m} Q_h^{j_{h1}(c_1-b)+\dots+j_{h\ell}(c_\ell-b)}$$

and, by (ii) and (iii) the max is over the $j$'s such that (iii) holds. By the choice of $r_h$ the product is

$$\le C_{10}^{mr_1} Q_1^{-b\ell(\ell^{-1}-\ell\varepsilon)mr_1+b\ell(J/r)r_1+Kr_1}$$

where

$$K = c_1 \sum_{h=1}^{m} j_{h1}/r_h + \dots + c_\ell \sum_{h=1}^{m} j_{h\ell}/r_h .$$

Now using (iii) and $c_1 + \dots + c_\ell = 0$ , $|c_i| \le 1$ we find

$$K \le C_{11} m\varepsilon + \ell^2 (J/r)$$

therefore

$$\frac{1}{J!}|P^J(y_1,\dots,y_m)| \le (C_{12}Q_1^{\varepsilon})^{mr_1} Q_1^{-bmr_1+C_{13}[\varepsilon m+(J/r)]r_1} < 1$$

if $(J/r) < C_{14}bm$ , $Q_1$ is large enough, for $\varepsilon$ sufficiently small.

Now the left hand side of this inequality is an integer, therefore

$$P^J(y_1,\dots,y_m) = 0$$

for

$$y_h = \sum_{k=1}^{\ell} a_k y_{hk} , \qquad |a_k| < Q_1^{\varepsilon} ,$$

$a_k$ integral, and all $J$ with

$$(J/r) < C_{14}bm .$$

It is not difficult to show that this implies that the restriction of $P^J(x_{11},\dots,x_{m\ell})$ to the linear space $M_1(x_{11},\dots,x_{1\ell}) = 0,\dots, M_m(x_{m1},\dots,x_{m\ell}) = 0$ vanishes identically, since it vanishes on sufficiently many well-distributed integral points of this linear space ; the required statement about the index of $P$ with respect to $M_1,\dots,M_m$ follows easily.

Now one would like to apply Proposition B and show that if the $r_h$ are rapidly decreasing then for every $\varepsilon > 0$

$$\operatorname{ind} P \leq \varepsilon m$$

with respect to $(M_1, \ldots, M_m ; r_1, \ldots, r_m)$ thus getting a contradiction. In order to be able to do this one needs first that the $r_h$ be rapidly decreasing, which means the $Q_h$ rapidly increasing and this can be done by taking a subsequence of the $Q_h$. But one also needs inequalities for the $r_h$ and the $\log|M_h|$ and one should show that

$$\log Q_h \ll \log|M_h| \ll \log Q_h .$$

It turns out without much difficulty that this follows from the condition that $\{L_1, \ldots, L_\ell ; S\}$ be a proper system, and this ends the argument.

## V. Schmidt's Theorem. End of the Proof.

Let $L_j = \alpha_{j1}X_1 + \ldots + \alpha_{j\ell}X_\ell$ be linear forms of determinent 1 and let E be the corresponding automorphism of $\mathbb{R}^\ell$. For $1 \leq p \leq \ell$ the exterior power $\Lambda^p E$ defines an automorphism of

$$\Lambda^p \mathbb{R}^\ell \simeq \mathbb{R}^{\binom{\ell}{p}} .$$

Expressing $\Lambda^p \mathbb{R}^\ell$ by means of a standard basis of $\mathbb{R}^\ell$ one obtains a set of $\binom{\ell}{p}$ linear forms $L_\sigma^{(p)}$ indexed by ordered p-subsets $\sigma$ of $\{1, \ldots, \ell\}$ ; explicitly

$$L_\sigma^{(p)} = \sum_\tau \alpha_{\sigma\tau} x_\tau$$

where

$$\alpha_{\sigma\tau} = \det(\alpha_{ij})_{i \in \sigma,\, j \in \tau} .$$

Let $A_1, \ldots, A_\ell$ be positive numbers with

$$A_1 \ldots A_\ell = 1 ,$$

let also

$$A_\sigma = \prod_{i \in \sigma} A_i$$

and consider the convex set $K^{(p)}$ :

$$|L_{\sigma}^{(p)}(x)| \leq A_{\sigma} , \qquad \mathrm{Card}(\sigma) = p .$$

This is called the p-compound of the set $K^{(1)}$ :

$$|L_j(x)| \leq A_j , \qquad j = 1,\dots,\ell .$$

Let $\nu_1,\dots,\nu_{\binom{n}{p}}$ be the successive minima of $K^{(p)}$ and $\lambda_1,\dots,\lambda_\ell$ those of $K^{(1)}$. Put also

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

MAHLER'S THEOREM.- There is an ordering $\sigma_j$ of the $\sigma$ such that

$$\nu_j \ll \lambda_{\sigma_j} \ll \nu_j , \qquad \text{all } j .$$

For Mahler's proof, see Mahler [2].

Now Schmidt's idea is to apply the previous theorem to get a non-trivial lower bound for $\nu_{\binom{\ell}{p}-1}$ and then use Mahler's theorem to deduce a non-trivial lower bound for the first minimum $\lambda_1$ .

One needs a lemma.

Lemma 1.- Let $L_j$ , $\lambda_i$ be as in the theorem of the previous section. Then if $A_1 \dots A_\ell = 1$ and

$$\lambda_1 A_i > Q^{-\delta/2\ell} , \qquad i \in S$$

we have

$$\lambda_{\ell-1} > \lambda_\ell Q^{-\delta}$$

provided $\quad Q \geq \max(A_1,\dots,A_\ell ; Q_1)$ .

(Note that the condition $A_i \geq 1$ for $i \in S$ is not needed.)

The proof goes as follows. Put

$$\rho_o = (\lambda_1 \cdots \lambda_{\ell-2}\lambda_{\ell-1}^2)^{1/\ell} ,$$

$$\rho_i = \rho_o/\lambda_i \quad , \quad i = 1,\dots,\ell-1 \quad , \quad \rho_\ell = \rho_{\ell-1} .$$

By a general result of Davenport there is a permutation $\{p_j\}$ of $\{1,\dots,\ell\}$ such that the successive minima $\lambda_j'$ of

$$|L_i(x)| \le \rho_{p_i}^{-1} A_i = A_i'$$

satisfy

$$\rho_j\lambda_j \ll \lambda_j' \ll \rho_j\lambda_j ;$$

note that $\rho_j\lambda_j = \rho_o$ for $j = 1,\dots,\ell-1$ , and

$$\rho_1 \cdots \rho_\ell = 1 .$$

If $A_i' \le 1$ for some $i \in S$ then since

$$A_i' = A_i\rho_{p_i}^{-1} \ge A_i\rho_1^{-1} = \lambda_1 A_i \rho_o^{-1} \ge Q^{-\delta/2\ell}\rho_o^{-1}$$

we have

$$\rho_o \ge Q^{-\delta/2\ell}$$

therefore

$$\lambda_\ell\lambda_{\ell-1}^2 \cdots \lambda_1 \gg \lambda_\ell Q^{-\delta/2} .$$

By Minkowski's theorem, $\lambda_1 \cdots \lambda_\ell \ll 1$ and we deduce

$$\lambda_{\ell-1} \gg \lambda_\ell Q^{-\delta/2} .$$

Now suppose $A_i' > 1$ for every $i \in S$ . Then we may apply the theorem of the next to last minimum and find

$$\lambda_{\ell-1}' \ge Q^{-\delta C}$$

provided

$$Q^C \ge \max(A_1',\dots,A_\ell' ; Q_2) .$$

By Davenport's lemma one has $\lambda_{\ell-1}' \ll \rho_o$ therefore $\rho_o \gg Q^{-\delta C}$ and as before we get

$$\lambda_{\ell-1} \gg \lambda_\ell Q^{-\ell C\delta}$$

hence the result (taking a smaller $\delta$ if necessary). It remains to show that if $Q \geq \max(A_1, \ldots, A_\ell ; Q_1)$ then for some $C$ we have $Q^C \geq \max(A'_1, \ldots, A'_\ell ; Q_2)$. This is easy :

$$\max A'_i \leq \rho_{\ell-1}^{-1} \max A_i = \lambda_{\ell-1} \rho_o^{-1} \max A_i$$
$$\leq \lambda_{\ell-1} \lambda_1^{-1} \max A_i \ll \lambda_i^{-\ell} \max A_i$$

(since $\lambda_1^{\ell-1} \lambda_{\ell-1} \leq \lambda_1 \ldots \lambda_\ell \ll 1$ )

$$\ll (\lambda_1 \max A_i)^{-\ell} (\max A_i)^{\ell+1}$$
$$\ll Q^{\delta/2 + \ell + 1} ,$$

whence the result with $C = 2\ell$.

The proof of Schmidt's theorem now ends as follows. Firstly one proves

Lemma 2.- Let $1, \alpha_1, \ldots, \alpha_{\ell-1}$ be real algebraic linearly independent over $\mathbb{Q}$. Write

$$L_j(X) = X_j - \alpha_j X_\ell , \qquad j \leq \ell - 1 ,$$
$$L_\ell(X) = X_\ell$$

and for $1 \leq p \leq \ell - 1$ let $S^{(p)}$ be the set of ordered p-uples $\sigma \subset \{1, \ldots, \ell\}$ with $\ell \in \sigma$.

Then the forms $L_\sigma^{(p)}$ together with $S^{(p)}$ form a proper system.

Now let $A_1 \ldots A_\ell = 1$, $A_\ell > 1$, $0 < A_i < 1$, $i = 1, \ldots, \ell - 1$ and let $\lambda_1, \ldots, \lambda_\ell$ be the successive minima of $|L_j(x)| \leq A_j$. One now proves that

(5.1) $\qquad \lambda_1 \geq Q^{-\delta}$

for $\qquad Q \geq \max(A_\ell , Q_3)$

and some $\qquad Q_3 = Q_3(\alpha , \delta)$.

The theorem of the next to last minimum gives the result for $\lambda_{\ell-1}$ and so our statement is true if $\ell = 2$. Now suppose $\ell > 2$. We shall show that

$$(5.2) \qquad \lambda_{\ell-p} > \lambda_{\ell-p+1} Q^{-\delta}$$

for $p = 1, 2, \ldots, \ell - 1$, $Q \geq \max(A_\ell, Q_4)$ and the result will follow.

Let $\sigma = \{1, \ldots, p-1, \ell\}$. We shall prove that

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

In fact, let $B_i = A_i / A_\sigma^{1/p}$, $i \in \sigma$. Since $A_1 \ldots A_\ell = 1$ we have $A_\sigma \geq 1$ and

$A_\ell \geq B_\ell > 1$ , $B_i < 1$ for $i = 1, \ldots, p-1$ , $B_1 \ldots B_{p-1} B_\ell = 1$ .

By definition of $\lambda_1$ there is a non-zero integral point $x^o \in \mathbf{Z}^\ell$ with

$$|L_i(x^o)| \leq \lambda_1 A_i , \qquad i = 1, \ldots, \ell$$

and by Minkowski's theorem $\lambda_1 \leq 1$. Hence $\lambda_1 A_i < 1$, $i = 1, \ldots, \ell - 1$, and thus the last coordinate $x_\ell^o$ of $x^o$ is not $0$. Hence the vector $y^o = (x_1^o, \ldots, x_{p-1}^o, x_\ell^o)$ is not $0$ and regarding $L_i$, $i \in \{1, \ldots, p-1, \ell\}$ as forms in $p$ variables we get

$$|L_i(y^o)| \leq \lambda_1 A_i = \lambda_1 A_\sigma^{1/p} B_i .$$

Hence the first minimum $\mu_1$ of

$$|L_i(y)| \leq B_i , \qquad i \in \{1, \ldots, p-1, \ell\}$$

satisfies

$$\mu_1 \leq \lambda_1 A_\sigma^{1/p} .$$

Since $B_1 \ldots B_{p-1} B_\ell = 1$, $B_\ell > 1$, $B_i < 1$ for $i = 1, \ldots, p-1$, and since $p \leq \ell - 1$ we may use induction and apply (5.1). We get

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

provided $Q \geq \max(B_\ell, Q_5)$ ; since $B_\ell \leq A_\ell$, it suffices $Q \geq \max(A_\ell, Q_5)$.

Clearly the argument applies to every $\sigma \in S^{(p)}$, hence

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

for all $\sigma \in S^{(p)}$. By Mahler's theorem the first minimum $\nu_1$ of the p-compound $L_\sigma^{(p)}$ of the linear forms $L_j$ satisfies

$$\nu_1 \gg \lambda_1\lambda_2 \cdots \lambda_p \geq \lambda_1^p ,$$

therefore

$$\nu_1 A_\sigma \gg Q^{-p\delta} \qquad\qquad \text{for } \sigma \in S^{(p)} .$$

Hence taking a smaller $\delta$ if necessary, we may apply Lemma 1 and Lemma 2 and get

(5.3) $$\nu_{\binom{\ell}{p}-1} > \nu_{\binom{\ell}{p}} Q^{-\delta}$$

provided

$$Q \geq \max(A_\sigma , Q_6) \qquad\qquad \text{where Card } \sigma = p .$$

Since $A_\sigma \leq A_\ell$ , it suffices $Q \geq \max(A_\ell , Q_6)$ . By Mahler's theorem again, we have

$$\nu_{\binom{\ell}{p}} \gg \lambda_{\ell-p+1}\lambda_{\ell-p+2} \cdots \lambda_\ell$$

$$\nu_{\binom{\ell}{p}-1} \ll \lambda_{\ell-p}\,\lambda_{\ell-p+2} \cdots \lambda_\ell$$

and by (5.3) we deduce (5.2). Clearly (5.2) implies $\lambda_1 > \lambda_\ell Q^{-\ell\delta}$ and since $\lambda_1 \cdots \lambda_\ell \gg 1$ by Minkowski's theorem, we have also $\lambda_\ell \gg 1$ and (5.1) follows, by taking a smaller $\delta$ if necessary.

Schmidt's Theorem 1 is almost immediate from (5.1). In fact, by definition of first minimum, (5.1) implies that the inequalities

(5.4) $$|x_1 - \alpha_1 x_\ell| \leq Q^{-\delta}A_1 , \ldots, |x_{\ell-1} - \alpha_{\ell-1}x_\ell| \leq Q^{-\delta}A_{\ell-1} , |x_\ell| \leq Q^{-\delta}A_\ell$$

are insoluble if $A_1 < 1 , \ldots, A_{\ell-1} < 1$ , $A_\ell > 1$ and $A_1 \ldots A_\ell = 1$ , for

$$Q \geq \max(A_\ell , Q_3) ,$$

unless all the $x_i$ are $0$ . By Liouville's theorem, there is $C$ such that

$$|x_i - \alpha_i x_\ell| > |x_\ell|^{-C}$$

if $x_\ell$ is large enough ; now take

$$A_i = |x_i - \alpha_i x_\ell| Q^\delta ,$$

$$A_\ell = (A_1 \ldots A_{\ell-1})^{-1}$$

so that

$$A_\ell < |x_\ell|^{C\ell} Q^{\ell\delta} .$$

If $Q > \max(|x_\ell|^{C\ell} Q^{\ell\delta}, Q_3)$ and if

$$A_i = |x_i - \alpha_i x_\ell| Q^\delta < 1$$

we deduce that we must have (the inequalities (5.4) are insoluble)

$$|x_\ell| > Q^{-\delta} A_\ell ,$$

hence

$$|x_1 - \alpha_1 x_\ell| \ldots |x_{\ell-1} - \alpha_{\ell-1} x_\ell| \, |x_\ell| > Q^{-\ell\delta} .$$

Since the only restriction on $Q$ is

$$Q \gg |x_\ell|^C$$

for some $C$, we deduce that the inequalities

$$\begin{cases} \|q\alpha_1\| \ldots \|q\alpha_{\ell-1}\| \, q^{1+\ell\varepsilon} < 1 \\ \|q\alpha_i\| < q^{-\varepsilon} , \qquad\qquad i = 1, \ldots, \ell-1 \end{cases}$$

have only a finite number of solutions.

Clearly the conditions $\|q\alpha_i\| < q^{-\varepsilon}$ are not restrictive, because if say $\|q\alpha_{\ell-1}\| \geq q^{-\varepsilon}$ it is sufficient to show that

$$\|q\alpha_1\| \ldots \|q\alpha_{\ell-2}\| \, q^{1+(\ell-1)\varepsilon} < 1$$

has only a finite number of solutions, and Schmidt's theorem follows by an obvious inductive argument.

The proof of Schmidt's second theorem is essentially identical and therefore will be omitted.

## REFERENCES

[1] J.W.S. CASSELS - An Introduction to Diophantine Approximation, Cambridge Univ. Press, 1957.

[2] K. MAHLER - On Compound Convex Bodies I., Proc. London Math. Soc., (3) 5 (1955), 358-379.

[3] E. WIRSING - On Approximations of Algebraic Numbers by Algebraic Numbers of Bounded Degree, Proc. Symposia Pure Math., XX (1969), 213-247.

Schmidt's proof appears in three papers

W.M. SCHMIDT - Zür simultanen Approximation algebraischer Zahlen durch rationale, Acta Math., 114 (1965), 159-209.

\- On simultaneous approximations of two algebraic numbers by rationals, Acta Math., 119 (1967), 27-50.

\- Simultaneous approximation to algebraic numbers by rationals, Acta Math., 125 (1970), 189-201.