*REPRINT FROM THE*

PROCEEDINGS

OF THE

ROYAL SOCIETY OF EDINBURGH

Section A (Mathematics)

VOL. LXIV—PART III (No. 17)

# Invariant Matrices and the Geometry of Numbers

By K. Mahler, F.R.S.,
Department of Mathematics, Manchester University

PUBLISHED BY
OLIVER & BOYD
EDINBURGH: TWEEDDALE COURT
LONDON: 39A WELBECK STREET, W.1
1956

*Price 4s. 0d.*

# XVII.—Invariant Matrices and the Geometry of Numbers. By

**K. Mahler,** F.R.S., Department of Mathematics, Manchester University. *Communicated by* Professor A. C. AITKEN, F.R.S.

(MS. received December 13, 1955. Revised MS. received February 22, 1956. Read March 5, 1956)

## SYNOPSIS

With every matrix representation of the (real) full linear group can be associated a multi-linear mapping of one affine space, $R_n$, into another, $R_N$. This mapping is studied from the viewpoint of the geometry of numbers of convex bodies, and a general arithmetical property of such mappings is proved. The result generalizes my recent work on compound convex bodies.

IN a paper which has just appeared (Mahler 1955) I applied compound matrices to the study of convex bodies and their geometry of numbers. Professor A. C. Aitken, after reading this paper, suggested to me in a letter of 23rd August that a similar theory should also hold for other kinds of invariant matrices corresponding to the matrix representations of the full linear group.

I show now that this conjecture is correct and establish a general transfer principle for convex bodies. The method used generalizes that in the paper cited, but is perhaps even a little simpler. There is, fortunately, no need to use the complicated explicit formulæ for invariant matrices.

It is nearly unnecessary to express my indebtedness to Professor Aitken. This paper would scarcely have been written without his suggestion.

1. Let $R_n$ and $R_N$ denote the real affine spaces of all points

$$X=(x_1, \ldots, x_n) \quad \text{and} \quad \Xi=(\xi_1, \ldots, \xi_N),$$

respectively. We consider these points as vectors and use the ordinary notation for sums of vectors, or for the product of a vector by a scalar. If $S$ is any point set in $R_n$, then $tS$ means the set of all points $tX$ where $X \in S$; and correspondingly for point sets in $R_N$.

2. It is assumed that a mapping $M$ of $R_n \times \ldots \times R_n$ ($p$ factors) into $R_N$ is given which has the following properties:

$(M_1)$: *To every system of $p$* (equal or distinct) *points*

$$X^{(1)}=(x_1^{(1)}, \ldots, x_n^{(1)}), \ldots, X^{(p)}=(x_1^{(p)}, \ldots, x_n^{(p)})$$

*in $R_n$ there corresponds a unique point*

$$\Xi=[X^{(1)}, \ldots, X^{(p)}]=(\xi_1, \ldots, \xi_N)$$

*in $R_N$*, which is called the *associated point* of $X^{(1)}, \ldots, X^{(p)}$.

$(M_2)$: *The mapping is linear in each point,*

$$[X^{(1)}, \ldots, \lambda_1 X_1^{(\pi)}+\lambda_2 X_2^{(\pi)}, \ldots, X^{(p)}]=\lambda_1[X^{(1)}, \ldots, X_1^{(\pi)}, \ldots, X^{(p)}] + \lambda_2[X^{(1)}, \ldots, X_2^{(\pi)}, \ldots, X^{(p)}].$$

$(M_3)$: *The set $\Omega$ of all points $\Xi=[X^{(1)}, \ldots, X^{(p)}]$ in $R_N$ obtained by this mapping contains $N$ linearly independent points.*

$(M_4)$: *To every non-singular affine transformation $T$ of $R_n$ there corresponds a non-singular affine transformation $T^*$ of $R_N$ such that*

$$[TX^{(1)}, \ldots, TX^{(p)}]=T^*[X^{(1)}, \ldots, X^{(p)}]$$

*identically in $X^{(1)}, \ldots, X^{(p)}$.* We call $T^*$ the *associated transformation of $T$.*

$(M_5)$: *There is a constant $P$ such that, for all $T$, the determinants* det $T$ *and* det $T^*$ *satisfy the equation*

$$|\det T^*| = |\det T|^P.$$

$(M_6)$: *The associated point of $p$ points with rational coordinates has itself rational coordinates.*

Mappings $M$ of this kind exist; *e.g.* the mapping of systems of $p$ points in $R_n$ on their compound point in $R_N$ where $N=\binom{n}{p}$ has the required properties. Other examples corresponding to matrix representations of the full linear group will be mentioned at the end of this paper.

3. Denote by

$$U_1=(1, 0, \ldots, 0), \qquad U_2=(0, 1, \ldots, 0), \ldots, \qquad U_n=(0, 0, \ldots, 1)$$

the $n$ unit points in $R_n$. Every point $X=(x_1, \ldots, x_n)$ in $R_n$ has thus the form $X=x_1U_1+\ldots+x_nU_n$. It follows therefore from $(M_2)$ that every point $\Xi$ on $\Omega$ can be written as a linear combination with real coefficients of the special associated points

$$[U_{\nu_1}, U_{\nu_2}, \ldots, U_{\nu_p}],$$

where $\nu_1, \nu_2, \ldots, \nu_p$ separately run over all suffixes

$$1, 2, \ldots, n.$$

Hence, by $(M_3)$, *there exist* $N$ *sets of* $p$ *such suffixes*, the sets

$$\nu_{H1}, \nu_{H2}, \ldots, \nu_{Hp} \qquad (H=1, 2, \ldots, N)$$

say, *such that the* $N$ *associated points*

$$\Upsilon_H=[U_{\nu_{H1}}, U_{\nu_{H2}}, \ldots, U_{\nu_{Hp}}] \qquad (H=1, 2, \ldots, N)$$

*on* $\Omega$ *are linearly independent and therefore form a basis of* $R_N$.

4. Let $t_1>0, \ldots, t_n>0$, and let $T_t$ denote the affine transformation defined by

$$T_tU_1=t_1U_1, \ldots, T_tU_n=t_nU_n.$$

By $(M_2)$, the associated transformation $T_t^*$ satisfies

$$T_t^*\Upsilon_1=\tau_1\Upsilon_1, \ldots, T_t^*\Upsilon_N=\tau_N\Upsilon_N,$$

where, for shortness, we have put

$$\tau_H=t_{\nu_{H1}} \ldots t_{\nu_{Hp}} \qquad (H=1, 2, \ldots, N).$$

From the form of $T_t$ and $T_t^*$ evidently

$$\det T_t=t_1 \ldots t_n, \qquad \det T_t^*=\tau_1 \ldots \tau_N.$$

It follows therefore from $(M_5)$ that

$$\tau_1 \ldots \tau_N=(t_1 \ldots t_n)^P \qquad \textit{identically in } t_1, \ldots, t_n. \tag{1}$$

Hence, on comparing on both sides of this equation the exponents of $t_1, \ldots, t_n$, we find that

*Exactly* $P$ *of the suffixes*

$$\nu_{H\pi} \qquad (H=1, 2, \ldots, N;\ \pi=1, 2, \ldots, p) \tag{2}$$

*are equal to each of the values* $1, 2, \ldots, n$.

Thus, in particular, the exponent $P$ in $(M_5)$ is a positive integer and $P=pN/n$.

5. We need two further simple properties of the mapping $M$. Let again $\Xi=[X^{(1)}, \ldots, X^{(p)}]$. Then, by $(M_2)$, the coordinates $\xi_1, \ldots, \xi_N$ of $\Xi$ are homogeneous multilinear forms in the coordinates of $X^{(1)}, \ldots, X^{(p)}$. They are thus *continuous functions* of these coordinates.

It is further obvious that if one of the points $X^{(1)}, \ldots, X^{(p)}$ is replaced by its negative, then $\Xi$ likewise becomes $-\Xi$.

6. By "body" we always mean a *set with inner points*, and by "convex body" we mean a *closed bounded convex body symmetric in the coordinate origin.*

Let $K^{(1)}, \ldots, K^{(p)}$ be any $p$ (distinct or identical) convex bodies in $R_n$. The associated points

$$\Xi=[X^{(1)}, \ldots, X^{(p)}], \quad \text{where} \quad X^{(1)} \in K^{(1)}, \ldots, X^{(p)} \in K^{(p)},$$

form a certain point set

$$\Sigma = < K^{(1)}, \ldots, K^{(p)} >$$

on $\Omega$; denote by

$$\mathrm{K}=[K^{(1)}, \ldots, K^{(p)}]$$

the convex hull of $\Sigma$. We call $\mathrm{K}$ the *associated set of* $K^{(1)}, \ldots, K^{(p)}$.

THEOREM I.—*The associated set* $\mathrm{K}$ *is a convex body.*

*Proof.*—From the continuity of $M$, $\Sigma$ and so also $\mathrm{K}$ are bounded sets, and, being a convex hull, $\mathrm{K}$ is closed and convex. Next, $\mathrm{K}$ is symmetrical in the origin $O$. For $K^{(1)}$ contains with each point $X^{(1)}$ also the symmetric point $-X^{(1)}$, and hence $\Sigma$ and so also $\mathrm{K}$ contain with $\Xi$ the symmetric point $-\Xi$.

Finally, $O$ is an inner point of $\mathrm{K}$, hence $\mathrm{K}$ is a body. For each of $K^{(1)}, \ldots, \mathrm{K}^{(p)}$ contains a neighbourhood of $O$, and therefore a positive number $\delta$ exists such that the sphere $|X| \leqslant \delta$ lies in all $p$ bodies $K^{(1)}, \ldots, K^{(p)}$. The $2n$ points

$$\mp\delta U_1, \ldots, \mp\delta U_n$$

are then elements of these bodies, and it follows from $(M_2)$ that $\Sigma$ and hence also $\mathrm{K}$ contain the $2N$ associated points

$$\mp\delta^p \mathrm{Y}_1, \ldots, \mp\delta^p \mathrm{Y}_N.$$

Therefore, by convexity, the "octahedron"

$$\Xi=s_1\mathrm{Y}_1+ \ldots +s_N\mathrm{Y}_N, \quad \text{where} \quad |s_1|+ \ldots +|s_N| \leqslant \delta^p,$$

forms a subset of $\mathbf{K}$. But this octahedron contains a certain neighbourhood of the origin because $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ form a basis of $R_N$, whence the assertion.

The following two properties are immediate consequences of the definition of the associated body:

$$[s_1 K^{(1)}, \ldots, s_p K^{(p)}] = s_1 \ldots s_p [K^{(1)}, \ldots, K^{(p)}] \quad \textit{for positive } s_1, \ldots, s_p. \qquad (3)$$

$$[k^{(1)}, \ldots, k^{(p)}] \subseteq [K^{(1)}, \ldots, K^{(p)}] \quad \textit{if} \quad k^{(1)} \subseteq K^{(1)}, \ldots, k^{(p)} \subseteq K^{(p)}. \qquad (4)$$

7. From now on only the associated body

$$\mathbf{K} = \underbrace{[K, \ldots, K]}_{p \text{ times}} = [K^p]$$

of a single convex body $K$, $p$ times repeated, will be considered. We shall establish a relation between the volume $V(K)$ of $K$ in $R_n$ and the volume $V(\mathbf{K})$ of $\mathbf{K}$ in $R_N$, and we begin with a simple special case.

Let $E$ be any ellipsoid in $R_n$ with centre at the origin, and let $\mathbf{E} = [E^p]$ be the associated body in $R_N$. In general, $\mathbf{E}$ is a rather complicated convex body.

THEOREM 2.—*A positive constant $c_1$ depending only on the mapping $M$ exists such that*

$$V(\mathbf{E}) = c_1 V(E)^P.$$

*Proof.*—Denote by $G_n$: $|X| \leqslant 1$ the unit sphere in $R_n$, and by $\Gamma_N^{(p)} = [G_n^p]$ its associated body in $R_N$. There is a non-singular affine transformation $T$ of $R_n$ such that $E = TG_n$ and therefore

$$V(E) = |\det T|\ V(G_n).$$

The associated affine transformation $T^*$ of $R_N$ has then, by $(M_4)$, the property that $\mathbf{E} = T^*\Gamma_N^{(p)}$, so that

$$V(\mathbf{E}) = |\det T^*|\ V(\Gamma_N^{(p)}).$$

Further, by $(M_5)$,

$$|\det T^*| = |\det T|^P.$$

Therefore, finally,

$$\frac{V(\mathbf{E})}{V(E)^P} = \frac{V(\Gamma_N^{(p)})}{V(G_n)^P}, \quad = c_1 \quad \text{say},$$

as was to be proved.

8. Let now $K$ and $\mathbf{K} = [K^p]$ be a convex body in $R_n$, and its associated body in $R_N$, respectively. By the theorem of Fritz John (1948) there exists in $R_n$ an ellipsoid $E$ such that

$$n^{-\frac{1}{2}} E \subseteq K \subseteq E.$$

Let again $\mathbf{E} = [E^p]$ be the associated body of $E$. Then, by the rules (3) and (4), also

$$n^{-\frac{p}{2}}\mathbf{E} \subseteq \mathbf{K} \subseteq \mathbf{E}.$$

Hence

$$n^{-\frac{n}{2}}V(E) \leqslant V(K) \leqslant V(E)$$

and

$$n^{-\frac{Np}{2}}V(\mathbf{E}) \leqslant V(\mathbf{K}) \leqslant V(\mathbf{E}),$$

so that

$$n^{-\frac{Np}{2}}\frac{V(\mathbf{E})}{V(E)^P} \leqslant \frac{V(\mathbf{K})}{V(K)^P} \leqslant n^{\frac{nP}{2}}\frac{V(\mathbf{E})}{V(E)^P}.$$

Theorem 2 leads therefore to the following result:

THEOREM 3.—*Two positive constants $c_2$ and $c_3$ depending only on the mapping $M$ exist such that*

$$c_2 V(K)^P \leqslant V(\mathbf{K}) \leqslant c_3 V(K)^P.$$

9. We introduce now the distance functions $F(X)$ of $K$ and $\Phi(\Xi)$ of $\mathbf{K} = [K^p]$.

If $X \neq O$ is any point in $R_n$, then there is a unique positive number $F(X)$ such that

$$X \in sK \text{ if } s \geqslant F(X), \quad \text{but} \quad X \notin sK \text{ if } s < F(X).$$

Put $F(O) = 0$. Then $F(X)$, the *distance function of* $K$, has the following two properties:

$$F(tX) = |t| F(X) \quad \text{for all real } t;$$
$$F(X+Y) \leqslant F(X) + F(Y).$$

The distance function $\Phi(\Xi)$ of $\mathbf{K}$ is defined analogously.

THEOREM 4.—*If $X^{(1)}, \ldots, X^{(p)}$ are any $p$ points in $R_n$, and if $\Xi = [X^{(1)}, \ldots, X^{(p)}]$ is the associated point in $R_N$, then*

$$\Phi(\Xi) \leqslant F(X^{(1)}) \ldots F(X^{(p)}).$$

*Proof.*—The assertion is obvious if the points $X^{(1)}, \ldots, X^{(p)}$ are not all distinct from $O$; let this case be excluded, and let

$$Y^{(1)} = F(X^{(1)})^{-1}X^{(1)}, \ldots, Y^{(p)} = F(X^{(p)})^{-1}X^{(p)}.$$

Then

$$F(Y^{(1)}) = \ldots = F(Y^{(p)}) = 1,$$

so that $Y^{(1)}, \ldots, Y^{(p)}$ lie in $K$. Therefore the associated point

$$[Y^{(1)}, \ldots, Y^{(p)}] = \{F(X^{(1)}) \ldots F(X^{(p)})\}^{-1}\,\Xi$$

belongs to $\mathbf{K}$. But then $\Xi \in s\mathsf{K}$ for $s \geqslant F(X^{(1)}) \ldots F(X^{(p)})$, giving the assertion.

10. The results so far obtained will now be applied to the geometry of numbers.

Let $L_0$ be the lattice of all points in $R_n$ with integral coordinates, and let similarly $\Lambda_0$ be the lattice of such points in $R_N$. The successive minima $m_1, \ldots, m_n$ of $K$ in $L_0$ are defined as follows.

First, there is a point $X_1 \neq O$ in $L_0$ such that $F(X_1) = m_1$ is a minimum, called *the first minimum of $K$ in $L_0$*. Secondly, let $2 \leqslant k \leqslant n$, and assume that the points $X_1, \ldots, X_{k-1}$ in $L_0$ and the corresponding successive minima $F(X_h) = m_h$ $(h = 1, \ldots, k-1)$ have already been defined. Then there is a point $X_k$ in $L_0$ linearly independent of $X_1, \ldots, X_{k-1}$ such that $F(X_k) = m_k$ is as small as possible; $m_k$ is called *the kth minimum of $K$ in $L_0$*. Thus $X_1, \ldots, X_n$ are linearly independent, and

$$0 < m_1 \leqslant \ldots \leqslant m_n < \infty.$$

If $Y_1, \ldots, Y_n$ are any $n$ linearly independent points in $L_0$ arranged such that $F(Y_1) \leqslant \ldots \leqslant F(Y_n)$, then always

$$F(Y_1) \geqslant m_1, \ldots, F(Y_n) \geqslant m_n.$$

Further Minkowski, in his *Geometrie der Zahlen*, proved that

$$2^n n!^{-1} \leqslant m_1 \ldots m_n V(K) \leqslant 2^n. \tag{5}$$

Naturally, these results have their analogues for $\mathbf{K}$ and $\Lambda_0$. There exist $N$ linearly independent points $\Xi_1, \Xi_2, \ldots, \Xi_N$ in $\Lambda_0$ such that $\Phi(\Xi_K) = \mu_K$ $(K = 1, \ldots, N)$ are the successive minima of $\mathbf{K}$ in $\Lambda_0$. Again,

$$0 < \mu_1 \leqslant \ldots \leqslant \mu_N < \infty.$$

If $\mathrm{H}_1, \ldots, \mathrm{H}_N$ are $N$ linearly independent points of $\Lambda_0$ arranged such that $\Phi(\mathrm{H}_1) \leqslant \ldots \leqslant \Phi(\mathrm{H}_N)$, then

$$\Phi(\mathrm{H}_1) \geqslant \mu_1, \ldots, \Phi(\mathrm{H}_N) \geqslant \mu_N.$$

Finally, by Minkowski's theorem,

$$2^N N!^{-1} \leqslant \mu_1 \ldots \mu_N V(\mathbf{K}) \leqslant 2^N. \tag{6}$$

11. Our aim is to find inequalities between the minima $m_k$ and the minima $\mu_K$. One such inequality is obtained on dividing (6) by the $P$th power of (5), viz.

$$\frac{2^{N-nP}}{N!} \leqslant \frac{\mu_1 \ldots \mu_N V(\mathsf{K})}{\{m_1 \ldots m_n V(K)\}^P} \leqslant 2^{N-nP} n!^P.$$

Here $V(\mathrm{K})/V(K)^P$ lies, by Theorem 3, between the lower and upper bounds $c_2$ and $c_3$. Thus, on putting

$$c_4 = \frac{2^{N-nP}}{N!\,c_3}, \qquad c_5 = \frac{2^{N-nP}n!^P}{c_2},$$

we obtain the even simpler inequality

$$c_4(m_1 \ldots m_n)^P \leqslant \mu_1 \ldots \mu_N \leqslant c_5(m_1 \ldots m_n)^P \tag{7}$$

which involves only the successive minima.

As we shall see, (7) implies $N$ separate inequalities for the $\mu_K$. But, in order to obtain these, it is first necessary to derive upper bounds for these minima from Theorem 4.

12. By construction, the $n$ lattice points $X_1, \ldots, X_n$ are linearly independent, and therefore can be written as

$$X_1 = TU_1, \ldots, X_n = TU_n,$$

where $T$ is a certain non-singular affine transformation of $R_n$. Let $T^*$ as usual be the associated affine transformation in $R_N$. Further, let

$$\nu_{H\pi} \quad (H = 1, 2, \ldots, N;\ \pi = 1, 2, \ldots, p)$$

be the same sets of suffixes, and

$$\mathrm{Y}_H = [U_{\nu_{H1}}, \ldots, U_{\nu_{Hp}}] \qquad (H = 1, 2, \ldots, N)$$

the same base points of $R_N$, as in § 3. Finally, put

$$\mathrm{Z}_H = [X_{\nu_{H1}}, \ldots, X_{\nu_{Hp}}] \qquad (H = 1, 2, \ldots, N)$$

Then, by $(M_4)$,

$$\mathrm{Z}_1 = T^*\mathrm{Y}_1, \ldots, \mathrm{Z}_N = T^*\mathrm{Y}_N.$$

Since also $T^*$ is non-singular, it follows that the new points $\mathrm{Z}_1, \ldots, \mathrm{Z}_N$ are likewise linearly independent.

13. Being an element of $L_0$, each point $X_h$ is of the form

$$X_h = \sum_{k=1}^{n} x_{hk} U_k$$

with integral coefficients $x_{hk}$. The associated points $\mathrm{Z}_H$ may therefore be written as sums

$$\mathrm{Z}_H = \sum_{(\nu)} z_{H,(\nu)} [U_{\nu_1}, \ldots, U_{\nu_p}]$$

where $\nu_1, \ldots, \nu_p$ independently run over the suffixes $1, \ldots, n$, and where $z_{H,(\nu)}$ are certain integers. Applying now $(M_6)$ for the first time, we see that the $n^p$ points $[U_{\nu_1}, \ldots, U_{\nu_p}]$ all have rational coordinates. Hence there exists a positive integer $g$ such that all points

$$g[U_{\nu_1}, \ldots, U_{\nu_p}] \qquad (\nu_1, \ldots, \nu_p = 1, \ldots, n)$$

belong to the lattice $\Lambda_0$. This, however, implies that also

$$g\mathbf{Z}_1, \ldots, g\mathbf{Z}_N \in \Lambda_0.$$

14. Denote, from now on, by $M_1, M_2, \ldots, M_N$ the $N$ products

$$m_{\nu_{H1}} \ldots m_{\nu_{Hp}} \qquad (H = 1, 2, \ldots, N)$$

arranged in order of increasing size,

$$M_1 \leqslant M_2 \leqslant \ldots \leqslant M_N.$$

We call these numbers $M_H$ the *associated products of* $m_1, \ldots, m_n$.

Similarly, denote by $\mathrm{H}_1, \ldots, \mathrm{H}_N$ the $N$ lattice points $g\mathbf{Z}_1, \ldots, g\mathbf{Z}_N$ arranged in order of increasing distance function,

$$\Phi(\mathrm{H}_1) \leqslant \Phi(\mathrm{H}_2) \leqslant \ldots \leqslant \Phi(\mathrm{H}_N).$$

By Theorem 4,

$$\Phi(\mathbf{Z}_H) \leqslant F(X_{\nu_{H1}}) \ldots F(X_{\nu_{Hp}}) = m_{\nu_{H1}} \ldots m_{\nu_{Hp}},$$

and so evidently

$$\Phi(\mathrm{H}_H) \leqslant gM_H \qquad (H = 1, 2, \ldots, N). \qquad (8)$$

But, as was proved in the last two sections, $\mathrm{H}_1, \ldots, \mathrm{H}_N$ are $N$ linearly independent points of $\Lambda_0$. It follows then from the properties of the successive minima quoted in § 10 that

$$\Phi(\mathrm{H}_H) \geqslant \mu_H \qquad (H = 1, 2, \ldots, N). \qquad (9)$$

15. The inequalities (8) and (9) together imply that

$$\mu_H \leqslant gM_H \qquad (H = 1, 2, \ldots, N). \qquad (10)$$

Here, by (2), the identity

$$M_1 \ldots M_N = \prod_{H=1}^{N} (m_{\nu_{H1}} \ldots m_{\nu_{Hp}}) = (m_1 \ldots m_n)^P$$

holds. Hence

$$\mu_1 \ldots \mu_N \leqslant \mu_H . \frac{g^{N-1}}{M_H} M_1 \ldots M_N = \mu_H \frac{g^{N-1}}{M_H} (m_1 \ldots m_n)^P;$$

whence, finally, by (7),

$$\mu_H \geqslant c_4 g^{-(N-1)} M_H \qquad (H = 1, 2, \ldots, N). \qquad (11)$$

On combining (10) with (11), the following result has been obtained.

THEOREM 5.—*Two positive constants $c_6$ and $c_7$ depending only on the mapping $M$ exist which have the following property:*

*If $K$ and $\mathbf{K}=[K^p]$ are a convex body in $R_n$ and the associated body in $R_N$; if $m_1, \ldots, m_n$ are the successive minima of $K$ in $L_0$ and $\mu_1, \ldots, \mu_N$ are those of $\mathbf{K}$ in $\Lambda_0$; if, finally, $M_1, \ldots, M_N$ are the associated products of $m_1, \ldots, m_n$, then*

$$c_6 M_H \leqslant \mu_H \leqslant c_7 M_H \qquad (H=1, 2, \ldots, N).$$

16. Of the successive minima of $K$ and $\mathbf{K}$, the first minima $m_1$ and $\mu_1$ are the most important ones. It is therefore of interest to establish simple inequalities connecting these two numbers in which the other minima do not occur. One naturally cannot expect these inequalities to be quite as sharp as those given by Theorem 5.

17. A *lower* estimate for $\mu_1$ in terms of $m_1$ is easily found. For, from the definition of the associated products,

$$M_1 \geqslant m_1^p,$$

and so it follows, since $\mu_1 \geqslant c_6 M_1$, that

$$\mu_1 \geqslant c_6 m_1^p. \tag{12}$$

Here the exponent $p$ cannot be replaced by any smaller number. For the successive minima $m_1, m_2, \ldots, m_n$ may all be equal, *e.g.* if $L_0$ is a critical lattice of $K$, and then $M_1 = m_1^p$.

18. It is not quite so easy to determine an *upper* bound for $\mu_1$, and further properties of the mapping $M$ are needed for this purpose.

In § 4, the products

$$\tau_H = t_{\nu_{H1}} \ldots t_{\nu_{Hp}} \qquad (H=1, 2, \ldots, N)$$

were introduced. These may also be written in the form

$$\tau_H = t_1^{\alpha_{H1}} \ldots t_n^{\alpha_{Hn}} \qquad (H=1, 2, \ldots, N)$$

with exponents $\alpha_{Hh}$ that are non-negative integers such that

$$\alpha_{H1} + \ldots + \alpha_{Hn} = p \qquad (H=1, 2, \ldots, N).$$

Let $q$ denote the largest of all these exponents; we call $q$ the *type* of $M$. There is no loss of generality in assuming that

$$q = \max(\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n}).$$

Choose a permutation

$$\varpi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \kappa_1 & \kappa_2 & \ldots & \kappa_n \end{pmatrix}$$

of 1, 2, . . ., $n$ such that

$$a_{1\kappa_1} \geqslant a_{1\kappa_2} \geqslant \ldots \geqslant a_{1\kappa_n}.$$

Hence, on putting

$$q_1 = a_{1\kappa_1}, \quad q_2 = a_{1\kappa_2}, \quad \ldots, \quad q_n = a_{1\kappa_n},$$

we have

$$q_1 \geqslant q_2 \geqslant \ldots \geqslant q_n \geqslant 0, \qquad \sum_{h=1}^{n} q_h = p, \quad q_1 = q,$$

and exactly $q_1$ of the suffixes $\nu_{11}, \nu_{12}, \ldots, \nu_{1p}$ are equal to $\kappa_1$, $q_2$ are equal to $\kappa_2$, etc., and finally $q_n$ are equal to $\kappa_n$.

19. Let, as before, $X_1, X_2, \ldots, X_n$ be the lattice points at which the successive minima of $K$ in $L_0$ are attained. Denote by

$$\varpi^{-1} = \begin{pmatrix} \kappa_1 & \kappa_2 & \ldots & \kappa_n \\ 1 & 2 & \ldots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \ldots & n \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \end{pmatrix}$$

the permutation inverse to $\varpi$, and by $T$ the affine transformation of $R_n$ given by

$$TU_1 = X_{\lambda_1}, \quad TU_2 = X_{\lambda_2}, \quad \ldots, \quad TU_n = X_{\lambda_n}.$$

This transformation is non-singular because $X_1, X_2, \ldots, X_n$ are linearly independent, and so the same is true for the associated transformation $T^*$ in $R_N$.

Next write

$$\begin{aligned} \mathbf{X} &= g[TU_{\nu_{11}}, \quad TU_{\nu_{12}}, \quad \ldots, \quad TU_{\nu_{1p}}] \\ &= g[X_{\lambda_{\nu_{11}}}, \quad X_{\lambda_{\nu_{12}}}, \quad \ldots, \quad X_{\lambda_{\nu_{1p}}}], \end{aligned}$$

where $g$ is the positive integer defined in § 13. Hence, by the same proof as in §§ 12 and 13, $\mathbf{X}$ belongs to the lattice $\Lambda_0$. Moreover, $\mathbf{X} \neq O$ because $\mathbf{X} = gT^*\mathbf{Y}_1$ and $\mathbf{Y}_1 \neq O$, and $T^*$ is non-singular.

Therefore, from the definition of the first minimum,

$$\mu_1 \leqslant \Phi(\mathbf{X}),$$

where, by Theorem 4,

$$\Phi(\mathbf{X}) \leqslant g\, F(X_{\lambda_{\nu_{11}}}) F(X_{\lambda_{\nu_{12}}}) \ldots F(X_{\lambda_{\nu_{1p}}}).$$

20. This expression can be simplified. From the definition of $\varpi^{-1}$,

$$\lambda_{\kappa_1}=1, \quad \lambda_{\kappa_2}=2, \quad \ldots, \quad \lambda_{\kappa_n}=n.$$

Hence just $q_1$ of the points $X_{\lambda_{\nu_{11}}}, X_{\lambda_{\nu_{12}}}, \ldots, X_{\lambda_{\nu_{1p}}}$ are equal to $X_1$, $q_2$ are equal to $X_2$, etc., and finally $q_n$ are equal to $X_n$, and so

$$\Phi(\mathbf{X}) \leqslant g\, F(X_1)^{q_1} \ldots F(X_n)^{q_n} = g\, m_1^{q_1} m_2^{q_2} \ldots m_n^{q_n}.$$

We apply now the relations

$$q_1=q, \quad q \geqslant q_2 \geqslant \ldots \geqslant q_n, \quad q_1+q_2+\ldots+q_n=p, \quad m_1 \leqslant m_2 \leqslant \ldots \leqslant m_n,$$

and the obvious identity

$$m_2^{q_2} m_3^{q_3} \ldots m_n^{q_n} = m_2^{q_2-q_3}(m_2m_3)^{q_3-q_4} \ldots (m_2m_3 \ldots m_{n-1})^{q_{n-1}-q_n}(m_2m_3 \ldots m_n)^{q_n}.$$

Then evidently

$$m_2 \leqslant (m_2m_3 \ldots m_n)^{\frac{1}{n-1}}, \quad m_2m_3 \leqslant (m_2m_3 \ldots m_n)^{\frac{2}{n-1}}, \ldots, \quad m_2m_3 \ldots m_{n-1} \leqslant (m_2m_3 \ldots m_n)^{\frac{n-2}{n-1}},$$

and therefore

$$m_2^{q_2} m_3^{q_3} \ldots m_n^{q_n} \leqslant (m_2m_3 \ldots m_n)^{\delta},$$

where, for shortness,

$$\delta = \frac{1}{n-1}\{1 . (q_2-q_3)+2(q_3-q_4)+\ldots+(n-2)(q_{n-1}-q_n)+(n-1)q_n\}.$$

This sum simplifies to

$$\delta = \frac{1}{n-1}\{q_2+q_3+\ldots+q_n\} = \frac{p-q}{n-1},$$

whence

$$m_1^{q_1} m_2^{q_2} \ldots m_n^{q_n} \leqslant m_1^{q-\frac{p-q}{n-1}}(m_1m_2 \ldots m_n)^{\frac{p-q}{n-1}},$$

where, by Minkowski's theorem,

$$m_1m_2 \ldots m_n V(K) \leqslant 2^n.$$

Hence, finally,

$$\mu_1 \leqslant g m_1^{q_1} m_2^{q_2} \ldots m_n^{q_n} \leqslant g m_1^{\frac{nq-p}{n-1}}\{2^n V(K)^{-1}\}^{\frac{p-q}{n-1}}. \tag{13}$$

In this inequality, the exponent $\dfrac{nq-p}{n-1}$ cannot be replaced by a larger number. For we may choose $m_1 < m_2 = m_3 = \ldots = m_n$, and then

$$M_1 = m_1^q m_2^{p-q} = m_1^{q-\frac{p-q}{n-1}}(m_1m_2 \ldots m_n)^{\frac{p-q}{n-1}}.$$

On combining (12) and (13) the following result is obtained.

THEOREM 6.—*Two positive constants $c_8$ and $c_9$ depending only on the mapping $M$ exist, as follows:*

*Let $q$ be the type of $M$; let $K$ and $\mathrm{K}=[K^p]$ be a convex body in $R_n$ and its associated body in $R_N$; and let $m_1$ be the first minimum of $K$ in $L_0$ and $\mu_1$ that of $\mathrm{K}$ in $\Lambda_0$. Then*

$$m_1 \leqslant c_8 \mu_1^{\frac{1}{p}}, \qquad \mu_1 \leqslant c_9 V(K)^{-\frac{p-q}{n-1}} m_1^{\frac{nq-p}{n-1}}.$$

21. One special case of this theorem has interest in itself. Let $X=(x_1, \ldots, x_n)$ and $\Xi=(\xi, \ldots, \xi_N)$ be again the general points in $R_n$ and $R_N$, respectively. Denote by $G(X)$ and $\Psi(\Xi)$ the distance functions

$$G(X)=\max(|x_1|, \ldots, |x_n|), \qquad \Psi(\Xi)=\max(|\xi_1|, \ldots, |\xi_N|),$$

so that

$$K_0\colon G(X) \leqslant 1 \quad \text{and} \quad \mathrm{K}_0\colon \Psi(\Xi) \leqslant 1$$

are the generalized cubes of sides length 2 with centres at the origins of the two spaces. In general, $\mathrm{K}_0$ is distinct from the associated body $\mathrm{K}_0'=[K_0^p]$ of $K_0$. Since, however, both $\mathrm{K}_0$ and $\mathrm{K}_0'$ are bounded and contain $O$ as an inner point, there exist two positive constants $c_{10}$ and $c_{11}$ depending only on $M$ such that

$$c_{10}\mathrm{K}_0 \subseteq \mathrm{K}_0' \subseteq c_{11}\mathrm{K}_0.$$

The distance function of $\mathrm{K}_0'$, $\Psi'(\Xi)$, say, satisfies therefore the inequalities

$$c_{10}\Psi'(\Xi) \leqslant \Psi(\Xi) \leqslant c_{11}\Psi'(\Xi).$$

Denote by

$$T=(a_{hk}) \quad \text{and} \quad T^*=(\alpha_{HK})$$

a non-singular affine transformation of $R_n$ and its associated transformation in $R_N$, both given by their matrices, and assume, for simplicity, that both are unimodular,

$$\det T = \det T^* = 1.$$

Then the new functions

$$F(X)=G(TX), \qquad \Phi(\Xi)=\Psi(T^*\Xi), \qquad \Phi'(\Xi)=\Psi'(T^*\Xi)$$

evidently are the distance functions of the convex bodies

$$K=T^{-1}K_0, \qquad \mathrm{K}=T^{*-1}\mathrm{K}_0, \qquad \mathrm{K}'=T^{*-1}\mathrm{K}_0',$$

respectively. In particular, $\mathbf{K}'$, $=[K^p]$, is the associated body of $K$, as follows easily from the definition of the associated body and from $(M_4)$. It is further evident that

$$c_{10}\Phi'(\Xi) \leqslant \Phi(\Xi) \leqslant c_{11}\Phi'(\Xi). \tag{14}$$

22. Let again $m_1$ be the first minimum of $K$ in $L_0$, and let similarly $\mu_1$ and $\mu_1'$ be the first minima of $\mathbf{K}$ and $\mathbf{K}'$ in $\Lambda_0$, respectively. It is obvious from (14) that

$$c_{10}\,\mu_1' \leqslant \mu_1 \leqslant c_{11}\,\mu_1',$$

while, by Theorem 6,

$$m_1 \leqslant c_8\,\mu_1'^{\frac{1}{p}}, \qquad \mu_1' \leqslant c_9 V(K)^{-\frac{p-q}{n-1}} m_1^{\frac{nq-p}{n-1}}.$$

Furthermore, now $V(K)=V(K_0)=2^n$. We thus have found the

THEOREM 7.—*Two positive constants $c_{12}$ and $c_{13}$ depending only on the mapping $M$ exist, as follows:*

*Let $q$ be the type of $M$; let $T=(a_{hk})$ and $T^*=(\alpha_{HK})$ be a unimodular affine transformation in $R_n$ and the associated transformation in $R_N$, respectively; let*

$$F(X)=\max_{h=1,2,\ldots,n}\left(\left|\sum_{k=1}^{n} a_{hk}x_k\right|\right), \qquad \Phi(\Xi)=\max_{H=1,2,\ldots,N}\left(\left|\sum_{K=1}^{N} \alpha_{HK}\xi_K\right|\right)$$

*be the corresponding distance functions; and let $m_1$ and $\mu_1$ be the minima of $F(X)$ and of $\Phi(\Xi)$ for all sets of integral variables $x_1, \ldots, x_n$ and $\xi_1, \ldots, \xi_N$ that do not all vanish. Then*

$$m_1 \leqslant c_{12}\mu_1^{\frac{1}{p}}, \qquad \mu_1 \leqslant c_{13} m_1^{\frac{nq-p}{n-1}}.$$

23. We conclude this paper with some short remarks on the connections to representation theory.

The property $(M_4)$ connects with the mapping $M$ a certain homogeneous integral representation of degree $p$ by matrices of order $N$ of the full linear group in $n$ variables, $T \rightarrow T^*$. The last theorem makes thus a statement on the arithmetic of such representations.

We have nowhere assumed that this representation is irreducible. For instance, $M$ may be that mapping where $[X^{(1)}, \ldots, X^{(p)}]$ is the point in $n^p$-dimensional space which has as its coordinates the products $x_{\nu_1}^{(1)} x_{\nu_2}^{(2)} \ldots x_{\nu_p}^{(p)}$, where $\nu_1, \nu_2, \ldots, \nu_p$ run independently over the suffixes $1, 2, \ldots, n$. In this case evidently

$$N=n^p, \qquad P=pn^{p-1}, \qquad q=p,$$

and $R_N$ is simply the $p$th Kronecker power of $R_n$.

This representation is reducible. Among its irreducible factors are the representation by the $p$th compound matrices studied in my earlier paper, as well as the symmetric representation. For the latter, $[X^{(1)}, \ldots, X^{(p)}]$ denotes the point in $\binom{n+p-1}{p}$-dimensional space which has as its coordinates all the sums

$$\sum_{(i)} x_{\nu_1}^{(i_1)} x_{\nu_2}^{(i_2)} \ldots x_{\nu_p}^{(i_p)},$$

where $\nu_1, \nu_2, \ldots, \nu_p$ run over all distinct systems of $p$ integers 1, 2, . . ., $n$ with $\nu_1 \leqslant \nu_2 \leqslant \ldots \leqslant \nu_p$, while $i_1, i_2, \ldots, i_p$ run over the $p$! permutations of 1, 2, . . ., $p$. In this particular case,

$$N=\binom{n+p-1}{p}, \qquad P=\binom{n+p-1}{p-1}, \qquad q=p.$$

For general $M$, the corresponding representation can be split into a sum of irreducible representations, and to each of these there belongs an invariant subspace of $R_N$. The component of $[X^{(1)}, \ldots, X^{(p)}]$ in this subspace $R^*$, defined, say, by parallel projection, generates a mapping $M^*$ of $R_n$ on $R^*$ which is of the same kind as $M$ itself. Thus $M$ can likewise be split into components. Therefore those mappings $M$ which correspond to *irreducible* representations of the full linear group deserve particular interest.

24. It is known that all such irreducible representations can be obtained from the Young diagrams belonging to the various partitions of $p$. By way of example, let $p=3$. Now there are three partitions, viz. $p=3=2+1=1+1+1$, two of which correspond to the compound and the symmetrical representations already mentioned. The remaining partition $p=2+1$ gives a Young diagram of the form

$$\begin{matrix} 1 & 2 \\ 3 & \end{matrix}$$

In this case, the associated point $[X^{(1)}, X^{(2)}, X^{(3)}]$ is found to have coordinates of the form

$$\xi_{\nu_1\nu_2\nu_3} = x_{\nu_1}^{(1)} x_{\nu_2}^{(2)} x_{\nu_3}^{(3)} + x_{\nu_1}^{(2)} x_{\nu_2}^{(1)} x_{\nu_3}^{(3)} - x_{\nu_1}^{(3)} x_{\nu_2}^{(2)} x_{\nu_3}^{(1)} - x_{\nu_1}^{(3)} x_{\nu_2}^{(1)} x_{\nu_3}^{(2)},$$

where the suffixes $\nu_1, \nu_2, \nu_3$ assume again the values 1, 2, . . ., $n$. But not all systems of three such suffixes need be considered, because

$$\xi_{\nu_1\nu_2\nu_3} + \xi_{\nu_3\nu_2\nu_1} = 0,$$

$$\xi_{\nu_1\nu_2\nu_3} + \xi_{\nu_2\nu_3\nu_1} + \xi_{\nu_3\nu_1\nu_2} = 0,$$

so that, in particular, $\xi_{\nu_1\nu_2\nu_3}=0$ if $\nu_1=\nu_3$. There is no difficulty in selecting a full system of linearly independent coordinates $\xi_{\nu_1\nu_2\nu_3}$. The result is that there are for $n=3$ the following $N=8$ coordinates:

$$\xi_{112},\quad \xi_{113},\quad \xi_{122},\quad \xi_{123},\quad \xi_{132},\quad \xi_{133},\quad \xi_{223},\quad \xi_{233}.$$

Similarly, there are $N=20$ coordinates for $n=4$, etc. The mapping that corresponds to this Young diagram has the type $q=2$.

For general Young diagrams, one deduces easily from the rule defining the irreducible representation and the corresponding mapping that *the type q equals the number of columns of the diagram.*

---

## References to Literature

JOHN, F., 1948. "Extremum problems with inequalities as subsidiary conditions", *Courant Anniversary Volume*, 187–204.

MAHLER, K., 1955. "On compound convex bodies (I)", *Proc. Lond. Math. Soc.*, (3), **5**, 358–379.

(*Issued separately August* 3, 1956)