Lectures on Quadratic Forms

By

C.L. Siegel

Tata Institute of Fundamental Research, Bombay
1957
(Reissued 1967)

Lectures on Quadratic Fomrs

By

C.L. Siegel

Notes by

K. G. Ramanathan

Tata Institute of Fundamental Research, Bombay
1955 – 56
(Reissumed 1967)

Contents

Chapter 1

Vector groups and linear inequalities

1 Vector groups

1 Let K be the field of real numbers and V a vector space of dimension $\underline{n}$ over K. Let us denote element of V by small Greek letters and elements of K by small Latin letters. The identity element of V will be denoted by $\underline{0}$ and will be called the zero element of V. We shall also denote by $\underline{0}$ the zero element in K.

Let $\varepsilon_1, \ldots, \varepsilon_n$ be a base of V so that for any $\xi \in V$

$$\xi = \sum_i \xi_i \varepsilon_i x_i, \quad x_i \in K.$$

We call $x_1, \ldots, x_n$ the *coordinates* of ξ. Suppose $\varepsilon'_1, \ldots, \xi'_n$ is another basis of V, then

$$\varepsilon'_i = \sum_j \varepsilon_j a_{ji}, \quad i = 1, \ldots, n$$

where $a_{ji} \in K$ and the matrix $M = (a_{ji})$ is non-singular. If in terms of $\varepsilon'_1, \ldots, \varepsilon'_n$

$$\xi = \sum_i \varepsilon'_i y_i, \quad y_i \in K$$

then it is easy to see that

$$\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = M \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \tag{1}$$

Suppose $\alpha_1, \ldots, \alpha_m$ is any finite set of elements of V. We denote by
2 $L(\alpha_1, \ldots, \alpha_m)$ the linear subspace generated in V by $\alpha_1, \ldots, \alpha_m$. This
means that $L(\alpha_1, \ldots, \alpha_m)$ is the set of elements of the form

$$\alpha_1 x_1 + \cdots + \alpha_m x_m, \quad x_i \in X.$$

It is clear that $L(\alpha_1, \ldots, \alpha_m)$ has dimension $\leqslant \text{Min}(n, m)$.

Let R_n denote the Euclidean space of n dimensions, so that every point P in R_n has coordinates $x_1, \ldots, x_n$, $x_i \in K$. Let $\varepsilon_1, \ldots, \varepsilon_n$ be a basis of V and let $x_1, \ldots, x_n$ be the coordinates of ξ in V with regard to this basis. Make correspond to ξ, the point in R_n with coordinates $x_1, \ldots, x_n$. It is then easily seen that this correspondence is $(1, 1)$. For any $\xi \in V$ define the *absolute value* $|\xi|$ by

$$|\xi|^2 = \sum_{i=1}^{n} x_i^2$$

where $x_1, \ldots, x_n$ are coordinates of ξ. Then $|\ |$ satisfies the axioms of a distance function in a metric space. We introduce a topology in V by prescribing a fundamental system of neighbourhoods of ξ to be the set of $\{S_d\}$ where S_d is the set of η in V with

$$|\xi - \eta| < d \tag{2}$$

S_d is called a sphere of radius d and center ξ. The topology above makes V a locally compact abelian group. The closure $\overline{S}_d$ of S_d is a compact set. From (1), it follows that the topologies defined by different bases of V are equivalent.

3 A subgroup G of V is called a *vector group*. The closure $\overline{G}$ of G in V is again a vector group. We say that G is *discrete* if G has no limit points in V. Clearly therefore a discrete vector group is closed.

Suppose G is discrete, then there is a neighbourhood of zero which has no elements of G not equal to zero in it. For, if in every neighbourhood of zero there exists an element of G, then zero is a limit point of G in V. This contradicts discreteness of G. Since G is a group, it follows that all elements of G are isolated in V. As a consequence we see that every compact subset of V has only finitely many elements of G in it.

We now investigate the structure of discrete vector groups. We shall omit the completely trivial case when the vector group G consists only of the zero element.

Let $G \neq \{0\}$ be a discrete vector group. Let $\xi \neq 0$ be an element of G. Consider the intersection

$$G_1 = G \cap L(\xi).$$

Let $d > 0$ be a large real number and consider all the $y > 0$ for which ξy is in G_1 and $y \leqslant d$. If d is large, then this set is not empty. Because G is discrete, it follows that there are only finitely many y with this property. Let $q > 0$ be therefore the smallest real number such that $\xi_1 = \xi \cdot q \in G_1$. Let $\beta = \xi x$ be any element in G_1. Put $x = hq + k$ where h is an integer and $0 \leqslant k < q$. Then ξx and $\xi_1 h$ are in G_1 and so ξk is in **4** G_1. But from definition of q, it follows that $k = 0$ or

$$\beta = \xi_1 h, \quad h \text{ integer}.$$

This proves that

$$G_1 = \{\xi_1\},$$

the infinite cyclic group generated by ξ_1.

If in G there are no elements other than those in G_1, then $G = G_1$. Otherwise let us assume as induction hypothesis that in G we have found $m(\leqslant n)$ elements $\xi_1, \ldots, \xi_m$ which are linearly independent over K and such that $G \cap L(\xi_1, \ldots, \xi_m)$ consists precisely of elements of the form $\xi_1 g_1 + \cdots + \xi_m g_m$ where $g_1, \ldots, g_m$ are integers. This means that

$$G_m = G \cap L(\xi_1, \ldots, \xi_m) = \{\xi_1\} + \cdots + \{\xi_m\}$$

is the direct sum of m infinite cyclic groups. If in G there exist no other elements than in G_m then $G = G_m$. Otherwise let $\beta \in G, \beta \notin G_m$. Put

$$G_{m+1} = G \cap L(\xi_1, \ldots, \xi_m, \beta).$$

Consider the elements λ in $G_{m+1} \subset G$ of the form

$$\lambda = \xi_1 x_1 + \cdots + \xi_m x_m + \beta y, \quad x_i \in K.$$

where $y \neq 0$ and $y \leqslant d$ with d a large positive real number. This set C of elements λ is not empty since it contains β. Put now $x_i = g_i + k_i$ where g_i is an integer and $0 \leqslant k_i < 1$, $i = 1, \ldots, m$. Let $\mu = \xi_1 g_1 + \cdots + \xi_m g_m$, then $\mu \in G_m$ and so

$$\lambda - \mu = \xi_1 k_1 + \cdots + \xi_m k_m + \beta y$$

5 is an element of G_{m+1}. Thus for every $\lambda \in G_{m+1}$ there exists a $\lambda = \lambda - \mu \in G$ with the property

$$\lambda' = \xi_1 k_1 + \cdots + \xi_m k_m + \beta y$$

$0 \leqslant k_i < 1$, $y \leqslant d$. Thus all those λ's lie in a closed sphere of radius $(m + d^2)^{\frac{1}{2}}$. Since G is discrete, this point set has to be finite. Thus for the λ's in G the y can take only finitely many values.

Therefore let $q > 0$ be the smallest value of y for which $\xi_{m+1} = \xi_1 t_1 + \cdots + \xi_m t_m + \beta q$ is in G. Let

$$\lambda = \xi_1 x_1 + \cdots + \xi_m x_m + \beta y$$

be in G_{m+1}. Put $y = qh + k$ where h is an integer and $0 \leqslant k < q$. Then

$$\lambda - \xi_{m+1} h = \xi_1 (x_1 - t_1 h) + \cdots + \xi_m (x_m - t_m h) + \beta k$$

is in G_{m+1}. By definition of q, $k = 0$. But in that case by induction hypothesis $x_i - t_i h = h_i$ is an integer. Thus

$$\lambda = \xi_1 h_1 + \cdots + \xi_m h_m + \xi_{m+1} h$$

$h_1, \ldots, h$ are integers. This proves that

$$G_{m+1} = \{\xi_1\} + \cdots + \{\xi_{m+1}\}$$

is a direct sum of $m + 1$ infinite cyclic groups.

We can continue this process now but not indefinitely since $\xi_1, \ldots,$ $\xi_{m+1}, \ldots$ are linearly independent. Thus after $r \leqslant n$ steps, the process ends. We have hence the

6 **Theorem 1.** *Every discrete vector group $G \neq \{0\}$ in V is a direct sum of r infinite cyclic groups,* $0 < r \leqslant n$.

Conversely the direct sum of cyclic infinite groups is a discrete vector group. We have thus obtained the structure of all discrete vector groups.

We shall now study the structure of all closed vector groups.

Let G be a closed vector group. Let S_d be a sphere of radius d with the zero element of G as centre. Let $r(d)$ be the maximum number of elements of G which are linearly independent and which lie in S_d. Clearly $r(d)$ satisfies

$$0 \leqslant r(d) \leqslant n.$$

Also $r(d)$ is an increasing function of d and since it is integral valued it tends to a limit when $d \to 0$. So let

$$r = \lim_{d \to 0} r(d).$$

This means that there exists a $d_0 > 0$ such that for $d \leqslant d_0$

$$r = r(d).$$

We call $\underline{r}$ the *rank* of G.

Clearly $0 \leqslant r \leqslant n$. Suppose $r = 0$, then we maintain that G is discrete; for if not, there exists a sequence $\gamma_1, \ldots, \gamma_n, \ldots$ of elements of G with a limit point in V. Then the differences $\{\gamma_k - \gamma_1\}$, $k \neq 1$ will form a set of elements of G with zero as a limit point and so in every neighbourhood of zero there will be elements of G which will mean that $r > 0$.

Conversely if G is discrete there exists a sphere S_d, $d > 0$ which 7 does not contain any point of G not equal to zero and containing zero. This means $r = 0$. Hence

$$r = 0 \Leftrightarrow G \text{ is discrete.}$$

Let therefore $r > 0$ so that G is not discrete. Let d be a real number $0 < d < d_0$ so that $r(d) = r$. Let S_d be a sphere around the zero element of G and of radius d. Let $\alpha_1, \ldots, \alpha_r$ be elements of G in S_d which are linearly independent. Let $t > 0$ be any real number and let $d_1 > 0$ be chosen so that $d_1 < \text{Min}(d, \frac{t}{n})$. Then $r(d_1) = r$. If $\beta_1, \ldots, \beta_r$ be elements of G which are linearly independent and which are contained in the sphere S_{d_1} around the zero element of G, then $L(\beta_1, \ldots, \beta_r) \subset L(\alpha_1, \ldots, \alpha_r)$ since $S_{d_1} \subset S_d$. But since both have dimension r,

$$L(\beta_1, \ldots, \beta_r) = L(\alpha_1, \ldots, \alpha_r).$$

Since $\beta_1, \ldots, \beta_r$ are in S_{d_1} we have

$$|\beta_i| \leqslant d_1 \leqslant \frac{t}{n}, \quad i = 1, \ldots, r.$$

Let $\xi \in L(\alpha_1, \ldots, \alpha_r)$. Then by above

$$\xi = \beta_1 x_1 + \cdots + \beta_r x_r.$$

Put $x_i = g_i + k_i$ where g_i is an integer and $0 \leqslant k_i < 1$. Put $\beta = \beta_1 g_1 + \cdots + \beta_r g_r$. Since $\beta_1, \ldots, \beta_r \in G, \beta$ will also be in G. Now

$$|\xi - \beta| = |\beta_1 k_1 + \cdots + \beta_r k_r|$$
$$\leqslant |\beta_1 k_1| + \cdots + |\beta_r k_r| < \frac{t}{n} \cdot n = t.$$

8

Since t is arbitrary, it means that in every neighbourhood of ξ there are elements of G. Hence $\xi \in \overline{G}$. But G is closed and ξ is arbitrary in $L(\alpha_1, \ldots, \alpha_r)$. Thus

$$L(\alpha_1, \ldots, \alpha_r) \subset G.$$

We have now two possibilities; $r = n$ or $r < n$. If $r = n$ then $V = L(\alpha_1, \ldots, \alpha_r) \subset G \subset V$, which means $G = V$. So let $r < n$. Complete $\alpha_1, \ldots, \alpha_r$ into a basis $\alpha_1, \ldots, \alpha_n$ of V. In terms of this basis, any $\gamma \in G$ may be written

$$\gamma = \alpha_1 x_1 + \cdots + \alpha_n x_n, \quad x_i \in K.$$

But $\lambda = \alpha_1 x_1 + \cdots + \alpha_r x_r$ is an element of $L(\alpha_1, \ldots, \alpha_r)$ and so of G. Thus

$$\delta = \gamma - \lambda = \alpha_{r+1} x_{r+1} + \cdots + \alpha_n x_n$$

is in G. Also $\delta \in L(\alpha_{r+1}, \ldots, \alpha_n)$. It is to be noted that γ determines δ uniquely. The δ's that arise in this manner clearly form a vector group contained in $L(\alpha_{r+1}, \ldots, \alpha_n)$ and isomorphic to the factor group $G - L(\alpha_1, \ldots, \alpha_r)$. We contend that this subgroup of δ's is discrete. For, if not let $\delta_1, \ldots$ be a sequence of elements with a limit point in V. Then in every arbitrary neighbourhood of zero there are elements of the set $\{\delta_k - \delta_l\}$, $k \neq 1$. Since $\delta_k - \delta_l$ is an element of $L(\alpha_{r+1}, \ldots, \alpha_n)$, this means that the rank of G is $\geqslant r + 1$. This contradiction proves our contention. Using theorem 1, it follows that there exist s elements $\delta_1, \ldots, \delta_s$ is G such that every $\xi \in G$ can be written uniquely in the form 9

$$\xi = \alpha_1 x_1 + \cdots + \alpha_r x_r + \delta_1 g_1 + \cdots + \delta_s g_s \tag{$*$}$$

where $x_i \in K$ and g's are integers. The uniqueness of the above form implies that $\delta_1, \ldots, \delta_s$ are linearly independent. We have hence the

Theorem 2. *Let G be a closed vector group. There exist integers r and s, $0 \leq r \leq r + s \leq n$ and $r + s$ independent elements $\alpha_1, \ldots, \alpha_r, \delta_1, \ldots, \delta_s$ in G such that every element G can be uniquely expressed in the form* $(*)$.

It is easy to see that if G is a vector group such that G consists of all elements ξ of the form $(*)$ then G is closed. In particular if $r = 0$, we have discrete groups as a special case.

It can be seen that $L(\alpha_1, \ldots, \alpha_r)$ is the connected component of the zero element in G.

2 Lattices

Let G be a discrete vector group. There exist $r \leq n$ elements $\alpha_1, \ldots, \alpha_r$ of V such that

$$G = \{\alpha_1\} + \cdots + \{\alpha_r\}$$

is a direct sum of r infinite cyclic groups. If $\beta_1, \ldots, \beta_{r+1}$ are any $r + 1$ elements of G, then there is a non-trivial relation.

$$\beta_1 h_1 + \cdots + \beta_{r+1} h_{r+1} = 0$$

where $h_1, \dots, h_{r+1}$ are integers. For let $\beta_i = \sum_{j=1}^{r} \alpha_j a_{ji}$, $i = 1, \dots, r+1$. Then the matrix $A = (a_{ji})$ has r rows and $r+1$ columns and is an integral
10 matrix. There exist therefore rational numbers $h_1, \dots, h_{r+1}$ not all zero such that

$$A \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_{r+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

This means that there are rational numbers $h_1, \dots, h_{r+1}$ not all zero such that $\beta_1 h_1 + \cdots + \beta_{r+1} h_{r+1} = 0$. Multiplying by a common denominator we obtain the result stated.

Let us now make the

Definition. *A vector group G is said to be a lattice if G is discrete and contains a basis of V.*

This means that there exists a basis $\alpha_1, \dots, \alpha_n$ of V such that

$$G = \{\alpha_1\} + \cdots + \{\alpha_n\}. \tag{3}$$

The quotient group $V - G$ is clearly compact. Conversely suppose G is a discrete vector group such that $V - G$ is compact. If $\alpha_1, \dots, \alpha_r$ are independent elements of G generating G, complete $\alpha_1, \dots, \alpha_r$ to a basis $\alpha_1, \dots, \alpha_n$ of V. A set of representatives of Vmod G is then given by

$$\alpha = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

where $0 \leq x_i < 1$, $i = 1, \dots, r$. Since Vmod G is compact, it follows that $r = n$. Thus a lattice is a discrete vector group G with $V - G$ compact.

A set of elements $\alpha_1, \dots, \alpha_n$ of G generating G is said to be a *base*
11 *of the lattice G*. If $\beta_1, \dots, \beta_n$ is another base of G then

$$\begin{aligned} (\beta_1, \dots, \beta_n) &= (\alpha_1, \dots, \alpha_n) A \\ (\alpha_1, \dots, \alpha_n) &= (\beta_1, \dots, \beta_n) B \end{aligned} \tag{4}$$

where A and B are n rowed integral matrices. Because of (4), it follows that $AB = E$, E being the unit matrix of order n. Thus $|A| = \pm 1$, $|B| = \pm 1$.

We call a matrix A *unimodular* if A and A^{-1} are both integral. The unimodular matrices form a group Γ. (4) shows that a transformation of a base of G into another base is accomplished by means of a unimodular transformation.

Conversely if $\alpha_1, \ldots, \alpha_n$ is a base of G and A is a unimodular matrix, then $\beta_1, \ldots, \beta_n$ defined by

$$(\beta_1, \ldots, \beta_n) = (\alpha_1, \ldots, \alpha_n)A$$

form again a base of G as can be easily seen. Thus Γ is the group of automorphisms of a lattice.

Let G be a lattice and $\alpha_1, \ldots, \alpha_n$ a base of it. Let β be any element in G. Then β can be completed into a base of G if and only if

$$G \cap L(\beta) = \{\beta\}$$

as is evident from section 1. Let $\beta = \alpha_1 g_1 + \cdots + \alpha_n g_n$ where $g_1, \ldots, g_n$ are integers. If β can be completed into a base $\beta, \beta_2, \ldots, \beta_n$ of G then, by above, the transformation taking $\alpha_1, \ldots, \alpha_n$ to $\beta, \beta_2, \ldots, \beta_n$ is unimodular. This means that

$$(g_1, g_2, \ldots, g_n) = 1.$$

Conversely let $\hat{\beta} = \alpha_1 g_1 + \cdots + \alpha_n g_n$ with $(g_1, \ldots, g_n) = 1$. Let $\beta \in G$ **12**
and

$$G \cap L(\rho) = \{\beta_1\}$$

where $\beta_1 = \alpha_1 t_1 + \cdots + \alpha_n t_n$. Since $\beta \in L(\beta)$, it follows that $\beta \in \{\beta_1\}$ and $\beta = \beta_1 q$ for some integer q. Because of independence of $\alpha_1, \ldots, \alpha_n$, it follows that q divides $(g_1, \ldots, g_n)$. This means that $q = 1$, that is

$$G \cap L(\beta) = \{\beta\}$$

Therefore β can be completed to a base of G. Hence the

Theorem 3. *Let G be a lattice with a base $\alpha_1, \ldots, \alpha_n$. Let $\beta = \alpha_1 g_1 + \cdots + \alpha_n g_n$ be an element in G. Then β can be completed to a base of G if and only if $(g_1, \ldots, g_n) = 1$.*

From the relation between bases of G and unimodular matrices, we have

Corollary. *Let $g_1, \ldots, g_n$ be n integers. They can be made the first column of a unimodular matrix if and only if* $(g_1, \ldots, g_n) = 1$.

3 Characters

Let G be a vector group. A character χ of G is a real valued function on V with the properties

1) $\chi(\alpha)$ is an integer for $\alpha \in G$

2) χ is continuous on V

3) $\chi(\alpha + \beta) = \chi(\alpha) + \chi(\beta)$, $\alpha, \beta \in V$

It follows trivially therefore that

$$\chi(0) = 0.$$

13 Since χ is a continuous function, we have

$$\lim_n \chi(\lambda_n) = \chi(\lambda)$$

where $\lambda_1, \lambda_2, \ldots$ is a sequence of elements in V converging to λ.

If p is an integer then $\chi(\omega p) = p\chi(\omega)$. If r is a rational number, say $r = \frac{a}{b}$, a, b integers, then $b\chi(\omega r) = \chi(\omega a) = a\chi(\omega)$ so that

$$\chi(\omega r) = r\chi(\omega).$$

By continuity it follows that if r is real

$$\chi(\omega r) = r\chi(\omega).$$

Suppose χ_1 and χ_2 are two characters of G. Define $\chi = \chi_1 + \chi_2$ by

$$\chi(\omega) = \chi_1(\omega) + \chi_2(\omega).$$

It is then trivial to verify that χ is a character of G. It then follows that the characters of G form a group G^*, called the *character group* or the dual of G.

Let G be a vector group and $\overline{G}$ its closure. Then

Lemma. *G and $\overline{G}$ have the same character group.*

Proof. A character of $\overline{G}$ is already a character of G. □

Conversely let χ be a character of G. Then χ satisfies properties 2) and 3). We have only to verify the property 1). Let $\omega \in \overline{G}$. Then there is a sequence of elements $\omega_1, \omega_2, \ldots$ in G with ω as the limit. Since χ is continuous

$$\lim_n \chi(\omega_n) = \chi(\omega).$$

But $\chi(\omega_n)$ are all integers. Thus $\chi(\omega)$ is integral. Thus χ is a character of $\overline{G}$. 14

The interest in lemma is due to the fact that in order to study the structure of G^*, it is enough to consider G^* as the dual of the closed vector group $\overline{G}$ whose structure we had investigated earlier.

Let $\overline{G}$ be the closure of the vector group G and G^* its character group. By theorem 2 there exists a base $\omega_1, \ldots, \omega_n$ of V such that

$$\xi = \omega_1 x_1 + \cdots + \omega_n x_n, \quad x_i \in K$$

belongs to $\overline{G}$ if and only if x_i is integral for $r < i \leq r + s$ and $x_i = 0$ for $i > r + s$, r and s being integers determined by theorem 2. If $\chi \in G^*$ then for $\xi \in V$

$$\chi(\xi) = x_1 \chi(\omega_1) + \cdots + x_n \chi(\omega_n).$$

If however $\xi \in \overline{G}$ then $\chi(\xi)$ is integral. Therefore

$$\chi(\omega_i) = \begin{cases} 0 & i \leq r \\ \text{integer} & r < i \leq r + s \\ \text{arbitrary real} & i > r + s. \end{cases}$$

Thus for $\xi \in \overline{G}$

$$\chi(\xi) = \sum_{i=r+1}^{r+s} \chi(\omega_i) \cdot x_i$$

If $\xi \notin \overline{G}$, then because of definition of $\omega_1, \ldots, \omega_n$ it follows that either at least one of $x_{r+1}, \ldots, x_{r+s}$ is not an integer or at least one of

$x_{r+s+1}, \ldots, x_n$ is not zero. Suppose that $\xi = \sum\limits_i \omega_1 x_i$, $x_{r+1} \neq 0 (\text{mod } 1)$. Define the linear function χ on V by 15

$$\chi(\omega_i) = \begin{cases} 1 & \text{if } i = r+1 \\ 0 & \text{if } i \neq r+1. \end{cases}$$

Then χ is a character of G and

$$\chi(\xi) = \chi(\omega_{r+1}) x_{r+1} = x_{r+1} \not\equiv 0 (\text{mod } 1)$$

The same thing is true if $x_{r+i} \not\equiv 0 (\text{mod } 1)$, $1 \leq i \leq s$. Suppose now that $\xi = \sum\limits_i \omega_i x_i$ and one of $x_{r+s+1}, \ldots, x_n$ say $x_n \neq 0$. Define χ linear on V by

$$\chi(\omega_i) = \begin{cases} 0 & \text{if } i \neq n \\ \dfrac{1}{2x_n} & \text{if } i = n. \end{cases}$$

Then χ is a character of G and $\chi(\xi) = \dfrac{1}{2} \not\equiv 0 (\text{mod } 1)$. Hence if $\xi \notin \overline{G}$ there is a character of G which is not integral for ξ. We have thus proved.

Theorem 4. *Let $\xi \in V$. Then $\xi \in \overline{G}$ if and only if for every character χ of G, $\chi(\xi)$ is integral.*

Let us fix a basis $\omega_1, \ldots, \omega_n$ of V so that $\omega_1, \ldots, \omega_{r+s}$ is a basis of $\overline{G}$. If $\chi \in G^*$ then $\chi(\omega_i) = c_i$ where

$$c_i = \begin{cases} 0 & i \leq r \\ \text{integer} & r < i \leq r+s \\ \text{real} & i > r+a \end{cases}$$

If $(c_1, \ldots, c_n)$ is any set of n real numbers satisfying the above conditions, then the linear function χ defined on V by

$$\chi(\xi) = \sum_{i=1}^{n} c_i x_i$$

16 where $\xi = \sum_i \omega_i x_i$, is a character of G. If R_n denotes the space of real n-tuples $(x_1, \ldots, x_n)$ then the mapping

$$\chi \to (c_1, \ldots, c_n)$$

is seen to be an isomorphism of G^* into R_n. Thus G^* is a closed vector group of rank $n - r - s$.

It can be proved easily that G^{**} the character group of G^* is isomorphic to $\overline{G}$.

4 Diophantine approximations

We shall study an application of the considerations in §3 to a problem in linear inequalities.

Let

$$L_i(h) = \sum_{j=1}^{m} a_{ij} h_j, \quad (i = 1, \ldots, n)$$

be n linear forms in m variables $h_1, \ldots, h_m$ with real coefficient a_{ij}. Let $b_1, \ldots, b_n$ be n arbitrarily given real numbers. We consider the problem of ascertaining necessary and sufficient conditions on the a_{ij}'s so that given $a > 0$ there exist integers $h_1, \ldots, h_m$ such that

$$|L_i(h) - b_i| < \alpha, \quad (i = 1, \ldots, n).$$

In order to study this problem, let us introduce the vector space V of all a rowed real columns

$$\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad a_i \in K.$$

V has then dimension n over K. Let $\alpha_1, \ldots, \alpha_n$ be elements of V defined by 17

$$\alpha_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, \quad i = 1, \ldots, m$$

and let G be the vector group consisting of all sums $\sum\limits_{i=1}^{m} \alpha_i g_i$ where g_i's are integers. Let γ be the vector

$$\gamma = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Then our problem on linear forms is seen to be equivalent to that of obtaining necessary and sufficient conditions that there be elements in G as close to γ as one wishes; in other words the condition that γ be in $\overline{G}$. Theorem 4 now gives the answer, namely that

$$\chi(\gamma) \equiv 0 (\text{mod } 1)$$

for every character χ of G.

Let us choose a basis $\varepsilon_1, \ldots, \varepsilon_n$ of V where

$$\varepsilon_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad i = 1, \ldots, n$$

with zero everywhere except at the i th place. Now in terms of this basis

$$\alpha_k = \varepsilon_1 a_{1k} + \cdots + \varepsilon_n a_{nk}, \quad k = 1, \ldots, m$$

18 Therefore if χ is a character of G

$$\chi(\alpha_k) = \sum_{i=1}^{n} a_{ik} c_i$$

where $\chi(\varepsilon_i) = c_i$, $i = 1, \ldots, n$. Also $\chi(\varepsilon_k) \equiv 0 (\text{mod } 1)$. Furthermore if $c_1, \ldots, c_n$ be any real numbers satisfying

$$\sum_{i=1}^{n} c_i a_{ik} \equiv 0 (\text{mod } 1), \quad \text{k} = 1, \ldots, \text{m},$$

then the linear function χ defined on V by $\chi(\varepsilon_i) = c_i$ is a character of G. By theorem 4 therefore

$$\sum_{i=1}^{n} c_i b_i \equiv 0 (\text{mod } 1)$$

We have therefore the theorem due to *Kronecker*.

Theorem 5. *A necessary and sufficient condition that for every $t > 0$, there exist integers $h_1, \ldots, h_m$ satisfying*

$$|L_i(h) - b_i| < t, \quad i = 1, \ldots, n,$$

is that for every set $c_1, \ldots, c_n$ of real numbers satisfying

$$\sum_{i=1}^{n} c_i a_{ik} \equiv 0 (\text{mod } 1), \quad \text{k} = 1, \ldots, \text{m},$$

we should have

$$\sum_{i=1}^{n} a_i b_i \equiv 0 (\text{mod } 1).$$

We now consider the special case $m > n$. Let $m = n + q$, $q \geq 1$. Let the linear forms be

$$\sum_{j=1}^{q} a_{ij} h_j + g_i, \quad i = 1, \ldots, n$$

in the m variables $h_1, \ldots, h_q, g_1, \ldots, g_n$. Then the vectors $\alpha_1, \ldots, \alpha_m$ 19
above are such that

$$\alpha_{q+i} = \varepsilon_i, \quad i = 1, \ldots, n.$$

This means that if χ is a character of G, $c_i = \chi(\varepsilon_i)$ is an integer. Thus

Corollary 1. *The necessary and sufficient condition that for every $t > 0$, there exist integers $h_1, \ldots, h_q, g_1, \ldots, g_n$ satisfying*

$$\left| \sum_{j=1}^{q} a_{ij} h_j + g_i - b_i \right| < t, \quad i = 1, \ldots, n$$

is that for every set $c_1, \ldots, c_n$ of integers satisfying

$$\sum_i c_i a_{ij} \equiv 0(\text{mod } 1), \quad \text{j} = 1, \ldots, \text{q}$$

we have

$$\sum_i c_i b_i \equiv 0(\text{mod } 1).$$

We now consider another special case $q = 1$. The linear forms are of the type

$$a_i h + g_i - b_i \quad i = 1, \ldots, n$$

$a_1, \ldots, a_n$, $b_1, \ldots, b_n$ being real numbers. Suppose now we insist that the condition on $b_1, \ldots, b_n$ be true *whatever* $b_1, \ldots, b_n$ are. This will mean that from above Corollary $c_1 = c_2 = \ldots = c_n = 0$ or, in other words, that $a_1, \ldots, a_n$ have to satisfy the condition that

$$\sum_i c_i a_i \equiv 0(\text{mod } 1), \quad \text{c}_\text{i} \ \text{ integral}$$

20 if and only if $c_i = 0$, $i = 1, \ldots, n$. This is equivalent to saying that the real numbers 1, $a_1, \ldots, a_1$ are linearly independent over the field of rational numbers.

Let us denote by R_n the Euclidean space of n dimensions and by F_n the unit cube consisting of points $(x_1, \ldots, x_n)$ with

$$0 \leq x_i < 1 \quad i = 1, \ldots, n.$$

For any real number x, let $((x))$ denote the fractional part of x, i.o. $((x)) = x - [x]$. Then

Corollary 2. *If* $1, a_1, \ldots, a_n$ *are real numbers linearly independent over the field of rational numbers, then the points* $(x_1, \ldots, x_n)$ *where*

$$x_i = ((ha_i)) \quad i = 1, \ldots, n$$

are dense in the unit cube, if h runs through all integers.

We consider now the homogeneous problem namely of obtaining integral solutions of the inequalities

$$|L_i(h)| < t, \quad i = 1, \ldots, n$$

$t > 0$ being arbitrary. Here we have to insist that $h_1, \ldots, h_m$ should not all be zero.

We study only the case $m > n$. As before introduce the vector space V of n-tuples. Let $\alpha_1, \ldots, \alpha_m$ and G have the same meaning as before. If the group G is not discrete, it will mean that the inequalities will have solutions for any t, however small. If however G is discrete then since **21** $m > n$ the elements $\alpha_1, \ldots, \alpha_m$ have to be linearly integrally dependent. Hence we have integers $h_1, \ldots, h_m$ not all zero such that

$$\alpha_1 h_1 + \cdots + \alpha_m h_m = 0.$$

We have hence the

Theorem 6. *If $m > n$, the linear inequalities*

$$|L_i(h)| < t, \quad i = 1, \ldots, n$$

have for every $t > 0$, a non-trivial integral solution.

Bibliography

[1] O. Perron : *Irrationalzahlen* Chelsea, 1948.

[2] C. L. Siegel : *Lectures on Geometry of Numbers*, New York University, New York, 1946.

Chapter 2

Reduction of positive quadratic forms

1 Quadratic forms

22 Let V be a vector space of dimension n over the field K of real numbers. Define an inner product $\xi\eta$ between vectors ξ, η of V by

i) $\xi\eta \in K$

ii) $\xi\eta = \eta\xi$

iii) $\xi(\eta + \zeta) = \xi\eta + \xi\zeta$

iv) $\xi(\eta a) = (\xi\eta)a,\ a \in K$.

Obviously if $\varepsilon_1, \ldots, \varepsilon_n$ is a base of V and ξ, η have the expression $\xi = \sum\limits_i \varepsilon_i a_i$, $\eta = \sum\limits_i \varepsilon_i b_i$ then

$$\xi\eta = \sum_{i,j=1}^{n} a_i b_j (\varepsilon_i \cdot \varepsilon_j).$$

If we denote by S the n-rowed real matrix $S = (s_{ij})$, $s_{ij} = \varepsilon_i \varepsilon_j$ then S is symmetric and

$$\xi\eta = \underline{a}' S \underline{b} \tag{1}$$

where $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ and $\underline{a}'$ denotes the transpose of the column vector $\underline{a}$. (1) is a bilinear form in the $2n$ quantities $a_1, \ldots, a_n, b_1, \ldots, b_n$. In particular

$$\xi^2 = \underline{a}' S\, \underline{a}$$

is a quadratic form in $a_1, \ldots, a_n$.

23 Suppose that $\varepsilon'_1, \ldots, \varepsilon'_n$ is another base of V. Then

$$\varepsilon'_i = \sum_j \varepsilon_j a_{ij} \quad i = 1, \ldots, n.$$

and the matrix $A = (a_{ij})$ is non-singular. If $S_1 = (\varepsilon'_i \cdot \varepsilon'_j)$ then one sees easily that

$$S_1 = S[A] = A' S A.$$

Thus if S with regard to one base is non-singular, then the S corresponding to any other base is also non-singular.

Conversely let S by and real n-rowed symmetric matrix and $\varepsilon_1, \ldots, \varepsilon_n$ a base of V over K. Put

$$\varepsilon_i \cdot \varepsilon_j = s_{ij} \quad (j, i = 1, \ldots, n)$$

and extend it by linearity to any two vectors of V. Then we have in V an inner product defined.

If $\xi = \sum\limits_i \varepsilon_i x_i$ is a generic vector of V over K,

$$\xi^2 = \underline{x}' S\, \underline{x} = S[x] = \sum_{i,j} x_i x_j s_{ij}.$$

The expression on the right is a quadratic form in the n variables $x_1, \ldots, x_n$ and we call S its matrix. The quadratic form is *degenerate* or *non-degenerate* according as its matrix S is or is not singular.

Let $\underline{x}' S\, \underline{x} = \sum\limits_{k,l=1}^{n} s_{kl} x_k x_l$ be a quadratic form in the n variables $x_1, \ldots, x_n$ and let $s_1 = s_{11} \neq 0$. We may write

$$\underline{x}' S\, \underline{x} = s_1 x_1^2 + 2 s_{12} x_1 x_2 + \cdots + 2 s_{1n} x_1 x_n + Q(x_2, \ldots, x_n)$$

24 so that $Q(x_2, \ldots, x_n)$ is a quadratic form in the $n-1$ variables $x_2, \ldots, x_n$. We now write, since $s_1 \neq 0$,

$$\underline{x}' S \underline{x} = s_1 \left(x_1 + \frac{s_{12}}{s_1} x_2 + \cdots + \frac{s_{1n}}{s_1} x_n \right)^2 - \frac{s_{12}^2}{s_1} x_2^2 - \ldots$$
$$- \frac{{s_{1n}}^2}{s_1} x_n^2 + Q(x_2, \ldots, x_n).$$

We have thus finally

$$\underline{x}' S \underline{x} = s_1 y_1^2 + R(x_2, \ldots, x_n)$$

where $y_1 = x_1 + \frac{s_{12}}{s_1} x_2 + \cdots + \frac{s_{1n}}{s_1} x_n$ and $R(x_2, \ldots, x_n)$ is a quadratic form in the $n-1$ variables $x_2, \ldots, x_n$. If we make a change of variables

$$\left. \begin{aligned} y_1 &= x_1 \frac{s_{12}}{s_1} x_2 + \cdots + \frac{s_{1n}}{s_1} x_n \\ y_1 &= x_i \qquad\qquad\qquad i > 1 \end{aligned} \right\} \tag{2}$$

then we may write

$$\underline{x}' S \underline{x} = \begin{pmatrix} s_1 & 0 \\ 0 & S_1 \end{pmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

where S_1 is the matrix of the quadratic form $R(x_2, \ldots, x_n)$. Using matrix notation we have

$$S = \begin{pmatrix} s_1 & \underline{q}' \\ \underline{q} & S_2 \end{pmatrix} = \begin{pmatrix} s_1 & \underline{0} \\ \underline{0} & S_1 \end{pmatrix} \begin{bmatrix} 1 & s_1^{-1} q' \\ \underline{0} & E \end{bmatrix} \tag{3}$$

where E is the unit matrix of order $n-1$, $\underline{q}$ is a column of $n-1$ rows and 25

$$S_1 = S_2 - s_1^{-1} \underline{qq'};$$

which, incidentally gives an expression for the matrix of R.

More generally suppose $S = \left(\begin{smallmatrix} S_1 & Q \\ Q' & S_2 \end{smallmatrix}\right)$ where S_1 is a k-rowed matrix and is non-singular. Put $\underline{x} = \left(\begin{smallmatrix} \underline{y} \\ \underline{z} \end{smallmatrix}\right)$ where $\underline{y}$ is a column of k rows and $\underline{z}$ has $n-k$ rows. Then

$$S[\underline{x}] = S_1[y] + \underline{y}' Q \underline{z} + z' Q' \underline{y} + S_2[\underline{z}],$$

which can be written in the form

$$S[x] = S_1[\underline{y} + S_1^{-1}Q\underline{z}] + W[z] \tag{4}$$

where $W = S_2 - Q'S^{-1}Q$. In matrix notation we have

$$S = \begin{pmatrix} S_1 & 0 \\ 0 & W \end{pmatrix} \begin{bmatrix} E & S_1^{-1}Q \\ 0 & E \end{bmatrix} \tag{5}$$

the orders of the two unit matrices being evident. In particular, we have

$$|S| = |S_1|\ |W|.$$

Let S be a real, non-singular, n-rowed, symmetric matrix. It is well known that there exists an orthogonal matrix V such that

$$S[V] = V'SV = D$$

where $D = [d_1, \ldots, d_n]$ is a real diagonal matrix. The elements $d_1, \ldots,$ 26 d_n of D are called the eigen-values of S. Let $\mathscr{L}$ denote the unit sphere

$$\mathscr{L} : \underline{x}'\ \underline{x} = 1$$

so that a generic point $\underline{x}$ on $\mathscr{L}$ is an n-tuple $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ of real numbers. Let m and N denote the smallest and largest of the eigen values of S. Then for any $\underline{x}$ on $\mathscr{L}$.

$$m \leq S[\underline{x}] \leq M$$

For, if we put $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = y - V^{-1}\underline{x}$, then $\underline{y}'\underline{y} = 1$ and

$$S[\underline{x}] = D[V^{-1}\underline{x}] = D[\underline{y}] = d_1 y_1^2 + \cdots + d_n y_n^2.$$

But then

$$S[\underline{x}] = (d_1 - M)y_1^2 + \cdots + (d_n - M)y_n^2 + M \leq M.$$

The other inequality is obtained by changing S to $-S$.

More generally we have, for any arbitrary real vector $\underline{x}$

$$m\underline{x}'\ \underline{x} \leq S[x] \leq M\underline{x}'\underline{x}. \tag{6}$$

If $\underline{x} = \underline{0}$, the statement is obvious. Let $\underline{x} \neq \underline{0}$. Then $t^2 = \underline{x}'\underline{x} \neq 0$. Put $\underline{y} = t^{-1}\underline{x}$. Then $\underline{y}'y = 1$ and so $m \leq S[\underline{y}] \leq M$. Multiplying throughout by t^2 we get the result in (6).

We now define a quadratic form $\underline{x}'S\,\underline{x}$ to be *positive definite* (or simply positive) if $S[\underline{x}] > 0$ for all vectors $\underline{x} \neq \underline{0}$. It is *positive semi-definite* if $S[\underline{x}] \geq 0$ for real $\underline{x} \neq 0$. We shall denote these by $S > 0$ and $S \geq 0$ respectively. If $S > 0$, then obviously $|S| \neq 0$. For, if $|S| = 0$, then there exists $\underline{x} \neq 0$ such that $S\,\underline{x} = \underline{0}$. Then 27

$$0 = \underline{x}'S\,\underline{x} > 0$$

which is absurd.

If $S > 0$ and $|A| \neq 0$ and A is a real matrix, then $T = S[A]$ is again positive. For, if $\underline{x} = \underline{0}$, the $A\underline{x} \neq y \neq 0$ and so

$$T[\underline{x}] = S[A\underline{x}] = S[\underline{y}] > 0.$$

We now prove two lemmas for later use.

Lemma 1. *A matrix S is positive definite if and only if $|S_r| > 0$ for $r = 1, \ldots, n$, where S_r is the matrix formed by the first r rows and columns of S.*

Proof. We shall use induction on $\underline{n}$. If $n = 1$, the lemma is trivial. Let therefore lemma be proved for matrices of order $n-1$ instead of n. Let

$$S = \begin{pmatrix} S_{n-1} & \underline{q} \\ \underline{q}' & a \end{pmatrix}$$

If $S > 0$ then $S_{n-1} > 0$ and so $|S_{n-1}| \neq 0$. We can therefore write

$$S = \begin{pmatrix} S_{n-1} & \underline{0}' \\ \underline{0} & l \end{pmatrix} \begin{bmatrix} E & S_{n-1}^{-1}\underline{q} \\ 0 & l \end{bmatrix} \tag{7}$$

so that $|S| = |S_{n-1}|l$. Induction hypothesis shows that $|S_{n-1}| > 0$ and $l > 0$ so that $|S| > 0$ and $|S_r| > 0$ for all r. □

The converse also follows since by hypothesis $|S| > 0$ and $|S_{n-1}| > 0$. So $1 > 0$. But by induction hypothesis $S_{n-1} > 0$. 28

Lemma 2. *If $S > 0$ and $S = (s_{kl})$, then*

$$|S| \leq s_1 \dots s_n$$

where $s_{kk} = s_k$, $k = 1, \dots, n$.

Proof. We again use induction on $\underline{n}$. From the equation (7) we have

$$|S| = |S_{n-1}| \cdot l.$$

But $l = s_n - \underline{q}' S_{n-1}^{-1} \underline{q} > 0$ since $S_{n-1}^{-1} > 0$ and $s_n > 0$. If we assume lemma proved for $n - 1$ instead of n we get

$$|S| \leq s_1 \dots s_{n-1} l \leq s_1 \dots s_n.$$

More generally we can prove that if $S > 0$ and $S = \begin{pmatrix} S_1 & S_{12} \\ S_{12}' & S_2 \end{pmatrix}$ then

$$|S| \leq |S_1| \cdot |S_2| \tag{8}$$

It is easy to see that equality holds in (8) if and only if $S_{12} = 0$.

Let $S > 0$, then $s_1, \dots, s_n$ are all positive. We can write as in (3)

$$S = \begin{pmatrix} s_1 & \underline{0} \\ \underline{0} & W \end{pmatrix} \begin{bmatrix} 1 & s_1^{-1} \underline{q}' \\ \underline{0} & E \end{bmatrix}$$

29 But since now $W > 0$, its first diagonal element is different from zero and we can write W also in the form (3). In this way we get

$$S = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \begin{bmatrix} 1, & d_{12}, \dots, d_{1n} \\ 0, & 1, d_{23}, \dots, d_{2n} \\ \vdots & \dots\dots \\ 0 & \dots\dots 1 \end{bmatrix} = D[V] \tag{9}$$

where $D = [d_1, \dots, d_n]$ is a diagonal matrix and $V = (d_{kl})$ is a triangle matrix with $d_{kk} = 1$, $k = 1, \dots, n$, and $d_{kl} = 0$, $k > 1$. We can therefore write

$$S[\underline{x}] = \sum_{k=1}^{n} d_k (x_k + d_{k\,k+1} x_{k+1} + \dots + d_{kn} x_n)^2$$

The expression $S = D[V]$ is *unique*. For if $S = D_1[V_1]$ where D_1 is a diagonal matrix and V_1 is triangular, then

$$D[W] = D_1$$

where $W = VV_1^{-1}$ is also a triangular matrix. In this case, it readily follows that $W = E$ and $D = D_1$.

In general we have the fact that if

$$S = \begin{pmatrix} S & 0 \\ 0 & S_2 \end{pmatrix} \begin{bmatrix} E & T \\ 0 & E \end{bmatrix} \tag{10}$$

where S_1 has order k then S_1, S_2 and T are unique.

We call the decomposition (9) of S the *Jacobi transformation* of S. □

2 Minima of definite forms

30 Let S and T be two real, non-singular n-rowed symmetric matrices. They are said to be *equivalent* (denoted $S \sim T$) if there exists a unimodular matrix U such that

$$S[U] = T.$$

Since the unimodular matrices form a group, the above relation is an equivalence relation. We can therefore put the n-rowed real symmetric matrices into *classes* of equivalent matrices. Evidently, two matrices in a class have the same determinant.

If $S = S'$ is real and t is a real number, we say that S *represents* t *integrally*, if there is an integral vector $\underline{x}$ such that

$$S[\underline{x}] = t.$$

In case $t = 0$, we insist that $\underline{x} \neq \underline{0}$. The representation is said to be *primitive*, if $\underline{x}$ is a primitive vector. Obviously if $S \sim T$ then S and T both represent the same set of real numbers.

If $S > 0$, then all the eigen values of S are positive. Let $m > 0$ be the smallest eigen value of S. Let $t > 0$ be a large real number. Then if

$S[\underline{x}] < t$ then $m\underline{x}'\underline{x} < t$ and so the elements of $\underline{x}$ are bounded. Therefore there exist only finitely many integral vectors $\underline{x}$ satisfying

$$S[\underline{x}] < t.$$

This means that if $\underline{x}$ runs through all non-zero integral vectors, $S[\underline{x}]$ has
31 a minimum. We denote this minimum by $\mu(S)$. There is therefore an
integral $\underline{x}$ such that

$$S[\underline{x}] = \mu(S)., \quad \underline{x} \neq \underline{0}.$$

Moreover $\underline{x}$ is a primitive vector. For if $\underline{x}$ is not primitive, then $\underline{x} = q\underline{y}$ where $q > 1$ is an integer, and $\underline{y}$ is a primitive vector. Then

$$\mu(S) = S[\underline{x}] = q^2 S[\underline{y}] > S[\underline{y}]$$

which is impossible. Furthermore if $S \sim T$ then $\mu(S) = \mu(T)$. For, let $S = T[U]$ where U is unimodular. If $\underline{x}$ is a primitive vector such that $\mu(S) = S[\underline{x}]$, then

$$\mu(S) = S[\underline{x}] = T[U\underline{x}] \geq \mu(T).$$

Also if $\mu(T) = T[\underline{y}]$, then

$$\mu(T) = T[\underline{y}] = S[U^{-1}\underline{y}] \geq \mu(S).$$

This proves the contention.

If $S > 0$ and t is a real number, then $\mu(tS) = t\mu(S)$. But $|tS| = t^n|S|$ so that it seems reasonable to compare $\mu(S)$ with $|S|^{1/n}$.

We not prove the following important theorem due to *Hermite*.

Theorem 1. *If $\mu(S)$ is the minimum of the positive matrix S of n rows, there exist a constant c_n depending only on n, such that*

$$\mu(S) \leq c_n |S|^{1/n}$$

Proof. We use induction on n. □

32 If $n = 1$, then S is a positive real number s. If $x \neq 0$, and integral, then $sx^2 > s$ unless $x = \pm 1$ so that

$$c_1 = 1.$$

Let us assume theorem proved for $n - 1$ instead of n. Let $\underline{x}$ be the primitive integral vector such that $\mu(S) = S[\underline{x}]$. Complete $\underline{x}$ into a unimodular matrix U. Then $T = S[U]$ has first diagonal element equal to $\mu(S)$. Also $\mu(S) = \mu(T)$ by our remarks above. Furthermore $|S| = |T|$. Therefore in order to prove the theorem we may assume that the first diagonal element s_1 of S is equal to $\mu(S)$.

$$\text{Let} \quad S = \begin{pmatrix} s_1 & \underline{q}' \\ \underline{q} & S_1 \end{pmatrix}. \text{ Then}$$

$$S = \begin{pmatrix} s_1 & \underline{0}' \\ \underline{0} & W \end{pmatrix} \begin{bmatrix} 1 & s_1^{-1}\underline{q}' \\ \underline{0} & E \end{bmatrix}$$

where $W = S_1 - \underline{q} s_1^{-1} \underline{q}!$ Also $|S| = s_1 |W|$.

Let $\underline{x} = \binom{x_1}{\underline{y}}$ be a vector and let $\underline{y}$ have $n - 1$ rows, so that

$$S[\underline{x}] = s_1(x_1 + s_1^{-1}\underline{q}'\underline{y})^2 + W[y]. \tag{11}$$

Since $W > 0$, we can choose an integral $\underline{y}$ so that $W[\underline{y}]$ is minimum. x_1 can now be chosen integral in such a manner that

$$-\frac{1}{2} \leq x_1 + s_1^{-1}\underline{q}'\underline{y} \leq \frac{1}{2} \tag{12}$$

using (11) and (12) and induction hypothesis we get 33

$$\mu(S) \leq S[\underline{x}] \leq \frac{\mu(S)}{4} + c_{n-1}\left|W\right|^{1/(n-1)}$$

Substituting for $|W|$ we get

$$\mu(S) \leq \left(\frac{4}{3} c_{n-1}\right)^{\frac{n-1}{n}} |S|^{\frac{1}{n}}$$

which proves the theorem.

Using $c_1 = 1$ and computing successively from the recurrence formula $c_n = \left(\frac{4}{3}c_{n-1}\right)^{\frac{n-1}{n}}$ we see that

$$c_n = (4/3)^{\frac{n-1}{2}} \tag{13}$$

is a possible value of c_n. This estimate is due to *Hermite*.

The best possible value for c_n is unknown except in a few cases. We shall show that $c_2 = \sqrt{\frac{4}{3}}$ and that it is the best possible for $n = 2$. From Hermite's estimate (13), we see that for a positive binary matrix S,

$$\mu(S) \leq \left(\frac{4}{3}\right)^{\frac{1}{2}} |S|^{\frac{1}{2}}.$$

Consider now the positive quadratic from $x^2 + xy + y^2$ whose matrix

$$S = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$

For integral x, y not both zero, $x^2 + xy + y^2 \geq 1$ so that $\mu(S) = 1$. Also $|S| = \frac{3}{4}$. We have

$$1 = \left(\frac{4}{3}\right)^{\frac{1}{2}} |S|^{\frac{1}{2}}$$

34 which proves that $\sqrt{\frac{4}{3}}$ is the best possible value of c_2.

We shall now obtain a 'finer' estimate for c_n due to *Minkowski*. This estimate is better than Hermite's for *large* values of n. To this end we make the following consideration.

Let R_n denote the Euclidean space of n dimensions regarded as a vector space of ordered n-tuples $(x_1, \ldots, x_n)$. A point set $\mathscr{L}$ in R_n is said to be *convex* if whenever A and B are two points of it, $\frac{A+B}{2}$, the mid point of the line joining A and B, is also a point of $\mathscr{L}$. It is said to be *symmetric* about the origin if whenever x belongs to it, $-x$ also belongs to it. Obviously if $\mathscr{L}$ is both convex and symmetric, it contains the origin.

If $\mathscr{L}$ is a point set in R_n and h is any point in R_n we denote by $\mathscr{L}_h$ the set of points x such that $x \in \mathscr{L}_h$ if and only if $x - h$ is a point of $\mathscr{L}$. With this notation $\mathscr{L} = \mathscr{L}_0$.

If $\mathscr{L}$ is an open, bounded symmetric convex set, the $\mathscr{L}$ has a measure $v(\mathscr{L})$ in the Jordon sense and for $h \in R_n$

$$v(\mathscr{L}) = v(\mathscr{L}_h).$$

We call a point $P = (x_1, \ldots, x_n)$ in R_n a *lattice point* if $x_1, \ldots, x_n$ are all integers. The lattice points form a lattice in R_n considered as a vector group. We shall denote points of this lattice by the letters $g, g', \ldots$.

The following lemma, due to *Minkowski*, shows the relationship between convex sets and lattices.

Lemma 3. *If $\mathscr{L}$ is an open, bounded, symmetric and convex set of volume $> 2^n$, then $\mathscr{L}$ contains a lattice point other than the origin.* 35

Proof. We shall assume that $\mathscr{L}$ has no lattice point in it other than the origin and then prove that $v(\mathscr{L}) \leq 2^n$. □

So let $\mathscr{L}$ have no lattice point in it other than the origin. Define the point set $\mathscr{M}$ by $x \in \mathscr{M}$ if and only if $2x \in \mathscr{L}$. Then $\mathscr{M}$ is an open, symmetric, bounded and convex set. Also

$$v(\mathscr{L}) = 2^n v(\mathscr{M}). \tag{14}$$

Consider now the translates $\mathscr{M}_g$ of $\mathscr{M}$ by the lattice points. If $g \neq g'$ then $\mathscr{M}_g$ and $\mathscr{M}_{g'}$ are disjoint sets. For, if $x \in \mathscr{M}_g \cap \mathscr{M}_{g'}$ then $x - g$ and $x - g'$ are points of $\mathscr{M}$. Since $\mathscr{M}$ is symmetric and convex.

$$\frac{g - g'}{2} = \frac{(x - g') + (g - x)}{2}$$

is a point of $\mathscr{M}$. By definition of $\mathscr{M}$, $g - g'$ is a point of $\mathscr{L}$. But $g \neq g'$. Thus $\mathscr{L}$ has a lattice point other than the origin. This contradicts our assumption. Thus the $\mathscr{M}_g$ for all g are distinct.

Let ε denote the unit cube, that is the set of points $x = (x_1, \ldots, x_n)$ with $0 \leq x_i < 1$, $i = 1, \ldots, n$. By the property of $\mathscr{M}_g$'s above

$$\sum_g v(\mathscr{M}_g \cap \varepsilon) = v(\varepsilon \cap \sum_g \mathscr{M}_g) \leq v(\varepsilon) = 1 \tag{15}$$

But $v(\mathscr{M}_g \cap \varepsilon) = v(\mathscr{M} \cap \varepsilon_{-g})$ so that by (15)

$$1 \geq \sum_g v(\varepsilon \cap \mathscr{M}_g) = \sum_g v(\varepsilon_{-g} \cap \mathscr{M}) = v(\sum_g \varepsilon_{-g} \cap \mathscr{M}).$$

36 But the ε_{-g} cover R_n completely without gaps or overlapping when g runs over all lattice points. Hence

$$v(\mathscr{M}) \leq 1.$$

Using (14) our theorem follows.

We can now prove the following theorem due to *Minkowski*.

Theorem 2. *If $S > 0$ and $\mu(S)$ is its minimum, then*

$$\mu(S) \leq \frac{4}{\pi} \cdot \left\{\Gamma\left(\frac{n}{2} + 1\right)\right\}^{2/n} |S|^{1/n}$$

Proof. In R_n let us consider the point set $\mathscr{L}$ defined by the set of $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ with

$$S[\underline{x}] < \rho$$

It is trivially seen to be open and symmetric. Also since $S > 0$, $\mathscr{L}$ is bounded. To see that it is convex, write $S = A'A$ and put $A\underline{x}_1 = \underline{y}_1$, $A\underline{x}_2 = \underline{y}_2$. Then a simple calculation proves that

$$2\left(\frac{y_1 + y_2}{2}\right)' \left(\frac{\underline{y}_1 + \underline{y}_2}{2}\right) \leq \underline{y}_1'\underline{y}_1 + \underline{y}_2'\underline{y}_2.$$

This shows that $\mathscr{L}$ is a convex set. The volume of $\mathscr{L}$ is

$$v(\mathscr{L}) = \frac{\rho^{n/2}\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} |S|^{-1/2}$$

37 If we put $\rho = \mu(S)$, then $\mathscr{L}$ contains no lattice point other than the origin. Minkowski's lemma then proves theorem 2. □

Denote the constants in Hermite's and Minkowski's theorems by c_n and c'_n respectively. If we use stirling's formula for the Γ-function in the form

$$\log\Gamma(x) \sim x\log x.$$

We get $\log \text{’¸}_n = \log\dfrac{4}{\pi} + \dfrac{2}{n}\log\Gamma\left(\dfrac{n}{2}+1\right) \sim \log n$ whereas $\log c_n = \dfrac{n-1}{2}\log 4/3 \sim \lambda n$ where λ is an absolute constant. This shows that for large *n*, Minkowski's estimate is better than Hermite's.

3 Half reduced positive forms

We now consider the space R_h, $h = \dfrac{n(n+1)}{2}$ of real symmetric n-rowed matrices and impose on it the topology of the h-dimensional real Euclidean space. Let $\mathscr{P}$ denote the subspace of positive matrices. If $S \in \mathscr{P}$ then all the principal minors of S have positive determinant. This shows that $\mathscr{P}$ is the intersection of a finite number of open subsets of R_h and hence is open.

Let S be a matrix in the frontier of $\mathscr{P}$ in R_h. Let $S_1, S_2, \ldots$ be a sequence of matrices in $\mathscr{P}$ converging to S. Let $\underline{x} \neq \underline{0}$ be *any* real column vector. Then $S_k[\underline{x}] > 0$ and hence by continuity $S[\underline{x}] \geq 0$. From the arbitrariness of $\underline{x}$, it follows that $S \geq 0$. On the other hand let S be any positive semi-definite matrix in R_h. Let E denote the unit matrix of order n. Then for $\varepsilon > 0$, $S + \varepsilon E$ is a positive matrix, which shows that in every neighbourhood of S there are points of $\mathscr{P}$. This proves that the 38 frontier of $\mathscr{P}$ in R_h consists precisely of positive semi-definite matrices.

Let Γ denote the group of unimodular matrices. We represent Γ in R_h as a group of transformations $S \rightarrow S[U]$, $S \in R_h$. Also U and $-U$ load to the same representation in R_h. It is easy to see that the only elements in Γ which keep every element of R_h fixed are $\pm E$. Thus if we identify in Γ, the matrices U and $-U$ then $S \rightarrow S[U]$ gives a faithful representation of Γ_0 in R_h, $\Gamma_0 = \Gamma/\pm E$. If U runs over all elements of Γ and $S \in R_h$, $S[U]$ runs through all matrices in the class of S. We shall now find in each class of positive matrices, a matrix having certain 'nice' properties.

Let $T \in \mathscr{P}$ and let $\underline{u}$ run over the first columns of all the matrices in Γ. There $\underline{u}$ are precisely all the primitive vectors. Consider the values $T[\underline{u}]$ as $\underline{u}$ runs over these first columns. Then $T[\underline{u}]$ has a minimum, which is none other than $\mu(T)$. Let this be attained for $\underline{u} = \underline{u}_1$. It is obvious that $\underline{u}_1$ is not unique for, $-\underline{u}_1$, also satisfies this condition. In any case, since $T > 0$, there are only finitely many $\underline{u}$'s with the property $T[\underline{u}] = T[\underline{u}_1]$. Let $\underline{u}_1$ be fixed and let $\underline{u}$ run over the second columns of all unimodular matrices whose first column is $\underline{u}_1$. The $\underline{u}$'s now are *not* all the primitive vectors (for instance $\underline{u} \neq \underline{u}_1$). $T[\underline{u}]$ again has a minimum say for $\underline{u} = \underline{u}_2$ and by our remark above 39

$$T[\underline{u}_1] \leq T[\underline{u}_2]$$

Also there are only finitely many $\underline{u}$ with $T[\underline{u}] = T[\underline{u}_2]$. Consider now all unimodular matrices whose first two columns are $\underline{u}_1$, $\underline{u}_2$ and determine a $\underline{u}_3$ such that $T[\underline{u}_3]$ is minimum. Continuing in this way one finally obtains a unimodular matrix

$$U = (\underline{u}_1, \ldots, \underline{u}_n)$$

and a positive matrix $S = T[U]$.

$S \sim T$ and by our construction, it is obvious, that S is not unique in the class of T. We shall study the matrices S and U more closely.

Suppose we have constructed the columns $\underline{u}_1, \ldots, \underline{u}_{k-1}$. In order to construct the k-th column we consider all unimodular matrices V whose first $k-1$ columns are $\underline{u}_1, \ldots, \underline{u}_{k-1}$ in that order. Using the matrix U above which has this property,

$$U^{-1}V = \begin{pmatrix} E_{k-1} & A \\ 0 & B \end{pmatrix} \tag{16}$$

where E_{k-1} is the unit matrix of order $k-1$ and A and B are integral matrices. Since U and V are unimodular, B is unimodular. If $\underline{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ denotes the first column of the matrix $\binom{A}{B}$ then, since B is unimodular

$$(w_k, w_{k+1}, \ldots, w_n) = 1. \tag{17}$$

40 The k-th column $\underline{y}_k$ of V is $U\underline{w}$. Conversely let $\underline{w}$ be any integral column satisfying (17). Then $w_k, \ldots, w_n$ can be made the first column of a unimodular matrix B of order $n-k+1$. Choosing *any* integral matrix A of $k-1$ rows and $n-k+1$ columns, whose first column is $w_1, \ldots, w_{k-1}$, we get a matrix V whose first $k-1$ columns are $\underline{u}_1, \ldots, \underline{u}_{k-1}$ (by means of the equation (16)). Thus the k-th column of all the unimodular matrices with first $k-1$ columns equal to $\underline{u}_1, \ldots, \underline{u}_{k-1}$ is of the form $U\underline{w}$, where $\underline{w}$ is an arbitrary integral vector with $(w_k, \ldots, w_n) = 1$.

Consider the matrix $S = T[U]$. By the choice of $\underline{u}_k$, we have if $\underline{w}$ satisfies (17), then

$$S[\underline{w}] = T[U\underline{w}] \geq T[\underline{u}_k] = s_k$$

where $S = (s_{kl})$. We have thus proved that in each class of T there exists a matrix S satisfying

$$\left.\begin{aligned} &\text{I)}\ \ s_1 > 0 \\ &\text{II)}\ \ S[\underline{w}] \geq s_k, \quad k = 1, \ldots, n \end{aligned}\right\}$$

for every integral column $\underline{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ with $(w_k, \ldots, w_n) = 1$.

Matrices which satisfy (I) and (II) shall be called *half reduced* and the subset of $\mathscr{P}$ of matrices S, half reduced, shall be denoted $\mathscr{R}_0$.

In the sequel we shall denote by $\underline{e}_1, \ldots, \underline{e}_n$, the n columns in order of the unit matrix of order n and by an *admissible* k-vector $\underline{w}$ we shall understand an integral vector $\underline{w}$ of n rows, satisfying (17). $\underline{e}_k$ is clearly 41 an admissible k-vector.

Since $\underline{e}_{k+1}$ is an admissible $k+1$-vector, we have

$$s_{k+1} = S[\underline{e}_{k+1}] \geq s_k$$

which shows that

$$s_1 \leq s_2 \leq \ldots \leq s_n. \tag{18}$$

Let $\underline{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ be an integral vector with $x_k = 1$, $x_l = 1$, $x_i = 0$ for $i \neq k$, $i \neq l$ and $k < l$. Then $\underline{u}$ is an admissible l-vector and so

$$s_k + 2s_{kl} + s_l = S[\underline{u}] \geq s_l.$$

This means that $-2s_{kl} \leq s_k$. Changing the sign of x_k we get $2s_{kl} \leq s_k$. Hence

$$-s_k \leq 2s_{kl} \leq s_k, \quad 1 \leq k < l \leq n \tag{19}$$

Remark. Suppose S is a real symmetric matrix satisfying (II). Let S_1 be the matrix obtained from S by deleting the h_1-th, h_2-th,...,h_l-th rows and columns from S. Then S_1 also has properties similar to S since we have only to consider such admissible vectors $\underline{w}$ for which the $h_1, \ldots, h_l$-th elements are zero.

We now prove the

Theorem 3. *Let S be a real, symmetric n-rowed matrix with the property* (II). *Then $S \geq 0$. If, in addition, it satisfies* (I), *then $S > 0$.* 42

Proof. Suppose $s_1 = 0$. Then by (19) we have

$$0 = -s_1 \leq 2s_{1l} \leq s_1 = 0$$

which shows that S has the form

$$S = \begin{pmatrix} 0 & \underline{0}' \\ \underline{0} & S_1 \end{pmatrix}$$

If $s_2 = 0$, we again have a similar decomposition for S_1, since S_1, by our remark above, also satisfies II. Thus either $S = 0$ or else there is a first diagonal element s_k, such that $s_k \neq 0$. Then

$$S = \begin{pmatrix} 0 & \underline{0} \\ \underline{0} & S_k \end{pmatrix}$$

S_k having s_k for its first diagonal element. We shall now show that $S_k > 0$. Observe that S_k satisfies both I) and II) and therefore for proving the theorem it is enough to show that if S satisfies I and II, then $S > 0$. □

If $n = 1$, the theorem is trivially true. Let therefore theorem proved for $n - 1$ instead of n. Put

$$S = \begin{pmatrix} S_1 & \underline{q} \\ \underline{q}' & s_n \end{pmatrix}$$

43 where $\underline{q}$ is a column of $n-1$ rows. S_1 satisfies I and II and so by induction hypothesis $S_1 > 0$. Also since $s_n \geq s_1$, therefore $s_n > 0$.

Let $\underline{x} = \begin{pmatrix} \underline{y} \\ z \end{pmatrix}$ be a column of n rows, $\underline{y}$ having $n-1$ rows and let z be a real number. Then

$$S[\underline{x}] = S_1[\underline{y} + S_1^{-1}\underline{q}z] + (s_n - \underline{q}'S_1^{-1}\underline{q})z^2.$$

We assert that $s_n - \underline{q}'S_1^{-1}\underline{q} > 0$. For let $s_n \leq \underline{q}'S_1^{-1}\underline{q}$. Then for $\varepsilon > 0$ and every $\underline{x} \neq \underline{0}$

$$S[\underline{x}] \leq S_1[\underline{y} + S_1^{-1}\underline{q}z] + \varepsilon z^2 \tag{20}$$

Consider the quadratic form on the right side of the inequality above. It is of order n, positive and has a determinant $|S_1|\varepsilon$. Therefore we may find a column vector $\underline{x} = \begin{pmatrix} \underline{y} \\ z \end{pmatrix}$, integral, such that the value of the right side is a minimum and so by Hermite's theorem

$$S_1[\underline{y} + S_1^{-1}\underline{q}z] + \varepsilon z^2 \leq c_n |S_1|^{1/n} \varepsilon^{1/n}.$$

Using (20) and observing that s_1 is the minimum of $S[\underline{x}]$ we get, for this $\underline{x}$,

$$0 < s_1 \leq S[\underline{x}] \leq c_n |S_1|^{1/n} \varepsilon^{1/n} \tag{21}$$

Since ε can be chosen arbitrarily small we get a contradiction from (21). Thus $s_n - \underline{q}'S_1^{-1}\underline{q} > 0$. This means the $S > 0$.

We have thus shown that all matrices satisfying (I) and (II) are in $\mathscr{P}$.

We prove now the following important theorem due to *Minkowski*. 44

Theorem 4. *If S is a positive half-reduced matrix, then*

$$1 \leq \frac{s_1 \ldots s_n}{|S|} \leq b_n$$

where b_n is a constant depending only on n.

Proof. The left hand side inequality has already been proved in lemma 2 even for all matrices in $\mathscr{P}$. In order to prove the right hand side inequality we use induction. □

Consider now the ratios

$$\frac{s_n}{s_{n-1}}, \frac{s_{n-1}}{s_{n-2}}, \ldots, \frac{s_2}{s_1}.$$

Since S is half-reduced, all these ratios are ≥ 1. Let $\gamma = \frac{n(n-1)}{4}$. For the above ratios, therefore, one of two possibilities can happen. Either there exists a k, $2 \leq k \leq n$ such that

$$\left.\begin{array}{c} \frac{s_n}{s_{n-1}} < \gamma, \frac{s_{n-1}}{s_{n-2}} < \gamma, \ldots, \frac{s_{k+1}}{s_k} < \gamma \\ \frac{s_k}{s_{k-1}} \geq \gamma \end{array}\right\} \tag{22}$$

or that

$$\frac{s_n}{s_{n-1}}, \ldots, \frac{s_2}{s_1} < \gamma \tag{23}$$

Note that in the case $n = 2$, the second possibility cannot occur since then $\gamma = \frac{1}{2}$ and $\frac{s_2}{s_1} \geq 1$.

45 Consider (23) first. We have

$$\frac{s_1 \ldots s_n}{s_1^n} < \gamma \frac{n(n-1)}{2}$$

and since

$$\frac{s_1 \ldots s_n}{|S|} = \frac{s_1 \ldots s_n}{s_1^n} \cdot \frac{s_1^n}{|S|}$$

we get, using Hermite's inequality

$$\frac{s_1 \ldots s_n}{|S|} < c_n^n \cdot \gamma \frac{n(n-1)}{2}$$

which proves theorem.

Suppose now that (22) is true and so $k \geq 2$. Write

$$S = \begin{pmatrix} S_{k-1} & Q \\ Q^1 & R \end{pmatrix}$$

where S_{k-1} has $k-1$ rows. Let $\underline{x} = \binom{\underline{y}}{\underline{z}}$ where $\underline{y}$ is a column with $k-1$ rows. We have, by completion of squares

$$S[\underline{x}] = S_{k-1}[\underline{y} + S_{k-1}^{-1}Q\underline{z}] + (R - Q'S_{k-1}^{-1}Q)[\underline{z}] \tag{24}$$

Also $|R - Q'S_{k-1}^{-1}Q| = |S|/|S_{k-1}|$. Choose $\underline{z}$ to be an integral primitive vector such that $(R - Q'S_{k-1}^{-1}Q)[\underline{z}]$ is minimum. By Hermite's theorem therefore

$$(R - Q'S_{k-1}^{-1}Q)[\underline{z}] \leq c_{n-k+1}(|S|/|S_{k-1}|)^{1/n-k+1} \tag{25}$$

Put $\underline{y} + S_{k-1}^{-1}Q\underline{z} = \underline{w}$ so that $\underline{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_{k-1} \end{pmatrix}$. Choose now $\underline{y}$ to be an integral vector such that

$$-\frac{1}{2} \leq w_i \leq \frac{1}{2}, \quad i = 1, \ldots, k-1. \tag{26}$$

By the choice of $\underline{z}$, it follows that $\underline{x} = \binom{\underline{y}}{\underline{z}}$ is an admissible k-vector. **46** Hence

$$s_k \leq S[\underline{x}]. \tag{27}$$

Also since S_{k-1} is half-reduced, we get

$$S_{k-1}[\underline{w}] = \sum_{p,q=1}^{k-1} s_{pq}w_p w_q \leq \frac{k(k-1)}{8} s_{k-1}.$$

Using (22) we get

$$S_{k-1}[\underline{w}] \leq \frac{s_k}{2} \tag{28}$$

From (24), (25), (27) and (28) we get

$$s_k \leq 2c_{n-k+1}(|S|/|S_{k-1}|)^{1/(n-k+1)} \tag{29}$$

Since

$$\frac{s_1 \ldots s_n}{|S|} = \frac{s_1 \ldots s_{k-1}}{|S_{k-1}|} \frac{|S_{k-1}|}{|S|} \cdot s_k^{n-k+1} \frac{s_k \ldots s_n}{s_k^{n-k+1}}$$

we get by induction hypothesis on S_{k-1}, that

$$\frac{s_1 \ldots s_n}{|S|} \leq b_{k-1} \cdot (2c_{n-k+1})^{n-k+1} \cdot \gamma \frac{(n-k)(n-k+1)}{2}$$

which proves the theorem completely.

The best possible value of b_n is again unknown except in a few simple cases. We shall prove that

$$b_2 = 4/3 \tag{30}$$

and it is the best possible value.

Let $ax^2 + 2bxy + cy^2$ be a half-reduced positive form. Then $2b \leq a \leq$
47 c. The determinant of the form is $d = ac - b^2$. Thus

$$ac = ac - b^2 + b^2 \leq d + \frac{a^2}{4} \leq d + \frac{ac}{4}$$

which gives

$$ac \leq \frac{4}{3} d \tag{31}$$

Consider the binary quadratic form $x^2 + xy + y^2$. It is half-reduced because if x and y are two integers not both zero, then $x^2 + xy + y^2 \geq 1$. The determinant of the form is $3/4$. Product of diagonal elements is unity. Hence

$$1 = \frac{4}{3} d$$

and this shows that $4/3$ is the best possible value.

4 Two auxiliary regions

Let $\mathscr{R}_0$ denote the space of half-reduced matrices. Define the point set $\mathscr{R}_t^*$ for $t > b_n \geq 1$ as the set of S satisfying

$$\left.\begin{array}{ll} 0 < s_k < ts_{k+1} & k = 1, \ldots, n-1 \\ -t < \dfrac{s_{kl}}{s_k} < t & 1 \leq k < 1 \leq n \\ \dfrac{s_1 \ldots s_n}{|S|} < t & \end{array}\right\} \tag{32}$$

Because of (18), (19) and theorem 4, it follows that

$$\mathscr{R}_0 \subset \mathscr{R}^*_t. \tag{33}$$

But what is more important is that

$$\lim_{t\to\infty} \mathscr{R}^*_t = \mathscr{P} \tag{34}$$

This is easy to see. For, if $S \in \mathscr{P}$, let t be chosen larger than the maximum of the finite number of ratios $\dfrac{s_k}{s_{k+1}}$, $k = 1, \ldots, n-1$; $\pm\dfrac{s_{kl}}{s_k}$, $1 \leq k < 1 \leq n$, $\dfrac{s_1 \ldots s_n}{|S|}$ and b_n. Then $S \in \mathscr{R}^*_t$ for this value of t. 48

Let $S \in \mathscr{R}^*_t$ and consider the Jacobi transformation of S; namely

$$S = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \begin{bmatrix} 1, t_{12}, \ldots t_{1n} \\ \cdots\cdots \\ 0 \ldots\ldots 1 \end{bmatrix} = D[T] \tag{35}$$

Then

$$s_{kl} = d_k t_{kl} + \sum_{h=1}^{k-1} d_h t_{hk} t_{hl}, \quad 1 \leq k \leq l \leq n.$$

In particular, putting $k = 1$, and using the fact that $d_1, \ldots, d_n$ are all positive, we get

$$\frac{s_k}{d_k} \geq 1. \tag{36}$$

Also since $|S| = d_1 \ldots d_n$, we have $\prod\limits_{k=1}^{n} \dfrac{s_k}{d_k} = \dfrac{s_1 \ldots s_n}{|S|} < t$. Since $t > 1$, we have

$$\frac{s_k}{d_k} < t \quad (k = 1, \ldots, n).$$

Using (32) we get

$$\frac{d_k}{d_{k+1}} = \frac{d_k}{s_k} \cdot \frac{s_k}{s_{k+1}} \cdot \frac{s_{k+1}}{d_{k+1}} < t^2. \tag{37}$$

Now $s_{1l} = d_1 \cdot t_{1l}$ so that

$$|t_{1l}| = \frac{|s_{1l}|}{d_1} = \frac{|s_{1l}|}{s_1} \cdot \frac{s_1}{d_1} < t^2$$

Let us assume that we have proved that 49

$$abs\, t_{gl} < u_0, \quad 1 \leq g \leq k-1, \quad g < l \leq n \tag{38}$$

for a constant u_0 depending on t and n. Then

$$abs\, t_{kl} \leq \frac{abs\, s_{kl}}{d_k} + \sum_{h=1}^{k-1} \frac{d_h}{d_k} abs\, t_{hk} \cdot abs\, t_{hl} < u_1,$$

because of (37) and (38), u_1 depending only on t and n. It therefore follows that if u is the maximum of u_0, u_1, t^2, then for the elements of D and T in (35) we have

$$\left.\begin{array}{ll} 0 < d_k < ud_{k+1}, & k = 1, \ldots, n-1 \\ abs\, t_{kl} < u, & k < l. \end{array}\right\} \tag{39}$$

We now define $\mathscr{R}_u^{**}$ to be the set of points $S \in \mathscr{P}$ such that if $S = D[T]$ where $D = [d_1, \ldots, d_n]$ is a diagonal matrix and $T = (t_{kl})$ is a triangle matrix then D and T satisfy (39) for some u. Since the Jacobi transformation is unique, this point set is well defined.

From what we have seen above, it follows that given $\mathscr{R}_t^*$, there exists a $u = u(t, n)$ such that

$$\mathscr{R}_t^* \subset \mathscr{R}_u^{**}$$

Conversely one sees easily that given $\mathscr{R}_u^{**}$ there exists a $t = t(u, n)$ such that

$$\mathscr{R}_u^{**} \subset \mathscr{R}_t^*.$$

In virtue of (34), it follows that

$$\lim_{u \to \infty} \mathscr{R}_u^{**} = \mathscr{P}. \tag{40}$$

50

We now prove two lemmas useful later.

Let $S \in \mathscr{P}$ and let t be a real number such that $S \in \mathscr{R}_t^*$. Let S_0 denote the matrix

$$S_0 = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix} \tag{41}$$

We prove

Lemma 4. *There exists a constant $c = c(t, n)$ such that whatever be the vector $\underline{x}$,*

$$\frac{1}{c} S_0[\underline{x}] \leq S[\underline{x}] \leq c S_0[\underline{x}].$$

Proof. Let P^{-1} denote the diagonal matrix $P^{-1} = [\sqrt{s_1}, \ldots, \sqrt{s_n}]$. Put $W = S[P]$. In order to prove the lemma, it is enough to show that if $\underline{x}'\underline{x} = 1$ then

$$\frac{1}{c} \leq W[\underline{w}] \leq c.$$

□

Let $W = (w_{kl})$. Then $w_{kl} = s_{kl}/\sqrt{s_k s_l}$. Because $S \in \mathscr{R}_t^*$ we have

$$abs\, w_{kl} = abs \frac{s_{kl}}{s_k} \sqrt{\frac{s_k}{s_1}} < t \cdot c_1, \quad k \leq l \tag{42}$$

where c_1 depends only on t and n. W being symmetric, it follows that the elements of W are in absolute value less than a constant $c_2 = c_2(t, n)$.

Consider now the characteristic polynomial $f(\lambda) = |\lambda E - W|$. By (42) all the coefficients of the polynomial $f(\lambda)$ are bounded in absolute value by a constant $c_3 = c_3(t, n)$. Also since $W > 0$, the eigen values 51 of W are bounded by $c_4 = c_4(t, n)$. Let $\lambda_1, \ldots, \lambda_n$ be these eigen values. Then

$$\lambda_1 \ldots \lambda_n = |W| = \frac{|S|}{s_1 \ldots s_n} > t^{-1}$$

which means that there exists a constant $c_5 = c_5(t, n)$ such that

$$\lambda_i > c_5(t, n), \quad i = 1, \ldots, n.$$

(6) then gives the result of lemma 4.

Next we prove

Lemma 5. *If $S \in \mathscr{R}_t^*$ and $S = \begin{pmatrix} S_1 & S_{12} \\ S_{12}' & S_2 \end{pmatrix}$, then $S_1^{-1} S_{12}$ has all its elements bounded in absolute value by a constant depending only on t and n.*

Proof. By the Jacobi transformation we have $S = D[T]$. Since $\mathscr{R}_t^* \subset \mathscr{R}_u^{**}$ for $u = u(t, n)$, the elements of T are $\leq u$ in absolute value. Write

$$T = \begin{pmatrix} T_1 & T_{12} \\ 0 & T_2 \end{pmatrix}, \quad D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$$

where T_1 and D_1 have the same number of rows and columns as S_1. We have $S_1 = D_1[T_1]$ and $S_{12} = T_1' D_1 T_{12}$ so that

$$S_1^{-1} S_{12} = T_1^{-1} T_{12}.$$

Since T_1 is a triangle matrix, so is T_1^{-1} and its elements are $\leq u_1$ in absolute value, $u_1 = u_1(t, n)$. The elements of T_{12} are already $\leq u$. Our lemma is proved. □

52 We are now ready to prove the following important

Theorem 5. *Let S and T be two matrices in $\mathscr{R}_t^*$. Let G be an integral matrix such that* 1) $S[G] = T$ *and* 2) *abs* $|G| < t$. *Then the elements of G are less, in absolute value, then a constant c depending only on t and n.*

Proof. The constants $c_1, c_2, \ldots$ occurring in the following proof depend only on t and n. Also 'bounded' shall mean bounded in absolute value by such constants. □

Let $G = (g_{kl})$ and let $\underline{g}_1, \ldots, \underline{g}_n$ denote the n columns of G. We then have

$$S[\underline{g}_l] = t_l \quad l = 1, \ldots, n.$$

Introducing the positive diagonal matrix of lemma 4, we obtain

$$S_0[\underline{g_l}] \leq c_1 S[\underline{g_l}] = c_1 t_l.$$

But $S_0[\underline{g_l}] = \sum_k s_k g_{kl}^2$ so that

$$s_k g_{kl}^2 \leq c_1 t_1 \quad k, l = 1, \ldots, n \tag{43}$$

Consider now the matrix G. Since $|G| \neq 0$, there exists in its expansion a non-zero term. That means there is a permutation $l_1, \ldots, l_n$ of 1, 2, 3, $\ldots, n$ such that

$$g_{1l_1} g_{2l_2} \ldots g_{nl_n} \neq 0.$$

From (43) therefore we get

$$s_k \leq s_k g^2_{kl_k} \leq c_1 t_{l_k} \quad k = 1, \ldots, n.$$

Consider now the integers $k, k+1, \ldots$ and $l_k, l_{k+1}, \ldots, l_n$. All of the 53
latter cannot be $> k$. So there is an $i \geq k$ such that $l_i \leq k$. Hence

$$s_i \leq c_1 t_{l_i}.$$

So, since S and T are in $\mathscr{R}^*_0$,

$$s_k \leq c_2 t_k, k = 1, \ldots, n. \tag{44}$$

On the other hand

$$\prod_{k=1}^{n} \frac{t_k}{s_k} = \frac{t_1 \ldots t_n}{|T|} \cdot \frac{|S|}{s_1 \ldots s_n} |G|^2$$

and all the factors on the right are bounded. Therefore

$$\prod_{k=1}^{n} \frac{t_k}{s_k} < c_3.$$

Using (44), it follows that

$$t_k \leq c_4 s_k, \quad (k = 1, 2, \ldots, n). \tag{45}$$

Combining (43) and (45) we have the inequality

$$s_k g^2_{kl} < c_5 s_l \quad k, l = 1, \ldots, n. \tag{46}$$

Let p now be defined to be the largest integer such that

$$s_l \geq c_5 s_l, k \geq p, l \leq p - 1. \tag{47}$$

If $p = 1$, this condition does not exist. From the definition of p, it follows that for every integer g with $p + 1 \leq g \leq n$, there exists a $k_g \geq g$ and an $l_g < g$ such that

$$s_{k_g} < c_5 s_{l_g} \tag{48}$$

This holds for $p = 1$, but if $p = n$, it does not exist. 54

Let c_6 be a constant such that

$$s_k < c_6 s_l \quad k \leq l. \tag{49}$$

This exists since $S \in \mathscr{R}_t^*$. Using (48) and (49) and putting $c_7 = c_5 c_6^2$ we have

$$s_g < c_7 s_{g-1} \quad g \geq p + 1 \tag{50}$$

(49) and (50) give the important inequality

$$\frac{1}{c_8} < \frac{s_k}{s_l} < c_8 \quad k \geq p, l \leq p \tag{51}$$

Using (46) and (47), we have if $k \geq p$ and $l \leq p - 1$

$$s_k g_{kl}^2 < c_5 s_1 \leq s_k.$$

Since $s_k \neq 0$, we have $g_{kl}^2 < 1$. But g_{kl} are integers. Hence

$$g_{kl} = 0 \quad k \geq p, \quad l \leq p - 1 \tag{52}$$

Let us split G up into 4 parts by

$$G = \begin{pmatrix} G_1 & G_{12} \\ G_{21} & G_2 \end{pmatrix}$$

where G_1 is a square matrix of order $p - 1$. (52) then shows that

$$G_{21} = 0. \tag{53}$$

Let now $k \geq p$, $1 \geq p$. Then from (51) we have

$$g_{kl}^2 < c_5 \frac{s_1}{s_k} < c_5 c_8 \tag{54}$$

55 which means that the elements of G_2 are bounded.

Note that if $p = 1$, our theorem is already proved by (54). So we may assume $p > 1$.

In order to prove the theorem we use induction. If $n = 1$, the theorem is trivially true. Assume theorem therefore proved for $n - 1$ instead of n. Split S and T in the form

$$S = \begin{pmatrix} S_1 & S_{12} \\ S'_{12} & S_2 \end{pmatrix} \quad T = \begin{pmatrix} T_1 & T_{12} \\ T'_{12} & T_2 \end{pmatrix}$$

where S_1 and T_1 are $p - 1$ rowed square matrices. Because $S[G] = T$, we get

$$\begin{aligned} &S_1[G_1] = T_1 \\ &G'_1 S_1 G_{12} + G'_1 S_{12} G_2 = T_{12} \end{aligned} \tag{55}$$

By considerations above $G_{21} = 0$ therefore $|G| = |G_1| \cdot |G_2|$. Since G is integral it follows that abs $|G_1| < t$. Also S_1 and T_1 are $p - 1$ rowed square matrices which are in $\mathscr{R}^*_{t,p-1}$, where $\mathscr{R}^*_{t,p-1}$ is the same as $\mathscr{R}^*_t$ with $p - 1$ instead of n. E_y induction hypothesis and (55) we see that G_1 is bounded.

Using the fact that $G'_1 S_1 = T_1 G_1^{-1}$ we get

$$G_{12} = G_1 T_1^{-1} T_{12} - S_1^{-1} S_{12} G_2.$$

Using lemma 5, it follows that the elements of G_{12} are bounded.

Our theorem is completely proved.

In particular,

Corollary. *If S and T are in $\mathscr{R}^*_t$ and $S[U] = T$ for a unimodular U, then U belongs to a finite set of unimodular matrices determined completely by t and n.* **56**

5 Space of reduced matrices

We have seen that given any matrix $T > 0$, there exists in the class of T, a half-reduced matrix S. Consider now the 2^n unimodular matrices of the form

$$A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

where $a_i = \pm 1$. If S is half-reduced, then $S[A]$ also is half-reduced. For, if $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is an admissible k-vector, then $A\underline{x} = \begin{pmatrix} \pm x_1 \\ \vdots \\ \pm x_n \end{pmatrix}$ is also an admissible k-vector. Also, the diagonal elements of S and $S[A]$ are the same. We shall choose A properly so that $S[A]$ satisfies some further conditions.

Since $S[A] = S[-A]$, there is no loss in generality if we assume $a_1 = 1$. Denote by $\underline{\alpha}_1, \ldots, \underline{\alpha}_n$ the n columns of the matrix A. Consider now $\underline{\alpha}'_1 S \underline{\alpha}_2$. This equals $a_2 s_{12}$. If $s_{12} \neq 0$ choose a_2 so that

$$a_2 s_{12} \geq 0.$$

If $s_{12} = 0$, a_2 may be chosen arbitrarily. Having chosen $a_1, \ldots, a_k$ consider $\underline{\alpha}'_k S \underline{\alpha}_{k+1} = a_k a_{k+1} s_{kk+1}$. Since a_k has been chosen, we choose $a_{k+1} = \pm 1$ by the condition

$$a_k a_{k+1} s_{kk+1} \geq 0,$$

57 provided $s_{kk+1} \neq 0$. If $s_{kk+1} = 0$, a_{k+1} may be arbitrarily chosen. We have thus shown that in each class of equivalent matrices, there is a matrix S satisfying

α) $s_1 > 0$

β) $s_{kk+1} \geq 0$, $k = 1, \ldots, n-1$.

γ) $S[\underline{x}] - s_k \geq 0$, $k = 1, \ldots, n$ for every admissible k-vector.

We shall call a matrix satisfying the above conditions a *reduced matrix*, reduced in the sense of *Minkowski*. Let $\mathscr{R}$ denote the set of reduced matrices, then

$$\mathscr{R} \subset \mathscr{R}_0. \tag{56}$$

Since the elements of $S \in \mathscr{P}$ are coordinates of the point S, the conditions β) and γ) above show that $\mathscr{R}$ is defined by the intersection of an infinity of closed half spaces of $\mathscr{P}$. We shall denote the linear functions in β) and γ) by $L_r, r = 1, 2, 3, \ldots$. It is to be noted that we exclude the case when an L_r is identically zero. This happens when in

γ), $\underline{x}$ is the admissible k-vector equal to $\pm \underline{e_k}$. We may therefore say that $\mathscr{R}$ is defined by

$$\overline{\alpha}) \ s_1 > 0, \quad \overline{\beta}) \ L_r \geq 0 \ r = 1, 2, 3, \ldots \tag{57}$$

We shall see presently that the infinite system of linear inequalities can be replaced by a finite number of them.

In order to study some properties of the reduced space $\mathscr{R}$, we first make some definitions. 58

Definition. *i) S is said to be an* inner point *of $\mathscr{R}$ if $s_1 > 0$ and $L_r(S) > 0$ for all r.*

ii) It is said to be a boundary point *of $\mathscr{R}$ if $s_1 > 0$ $L_r(S) \geq 0$ for all r and $L_r(S) = 0$* at least *for one r.*

iii) It is said to be an outer point *of $\mathscr{R}$ if $s_1 > 0$ and $L_r(S) < 0$ at least for one r.*

We first show that $\mathscr{R}$ *has* inner points.

Consider the quadratic form

$$S\,|\underline{x}| = x_1^2 + \cdots + x_n^2 + (p_1x_1 + \cdots + p_nx_n)^2$$

where $p_1, \ldots, p_n$ are n real numbers satisfying

$$0 < p_1 < p_2 \ldots < p_n < 1.$$

The matrix $S = (s_{kl})$ is then given by

$$\begin{aligned} s_k &= 1 + p_k^2, & k &= 1, \ldots, n \\ s_{kl} &= p_kp_l, & k &\neq l. \end{aligned}$$

We assert that S is an inner point of $\mathscr{R}$. In the first place

$$s_1 > 0, s_{kk+1} = p_kp_{k+1} > 0; k = 1, \ldots, n-1.$$

Next let $\underline{x}$ be an admissible k-vector not equal to $\pm \underline{e_k}$. Then at least one of $x_k, \ldots, x_n$ has to be different from zero. If at least two of then, say x_j, x_1 are different from zero, so that $k \leq 1 < j \leq n$, then

$$S\,[\underline{x}] \geq x_1^2 + x_j^2 + \cdots \geq 2 > 1 + p_k^2 = s_k.$$

The worst case is when all of $x_k, \ldots, x_n$ except one are zero. If $x_k = \pm 1$ 59 and $x_1 = 0$ for $l > k$, then $x_i \neq 0$ for some $i < k$ since $\underline{x} \neq \pm \underline{e_k}$ and (then) $S[\underline{x}] \geq 2$. Let $x_i = 0$ for all i except $i = l > k$ so that $\underline{x} = \pm \underline{e_l}$. Then

$$S[\underline{x}] = 1 + p_l^2 > 1 + p_k^2 = s_k.$$

This proves that S is an inner point.

We now prove

Theorem 6. *The set of inner points of $\mathscr{R}$ is an open set in $\mathscr{P}$.*

Proof. Let S be an inner point of $\mathscr{R}$. Then $s_1 > 0$ and $L_r > 0$ for all r. The inequalities $s_{kk+1} > 0$ being finitely many can be satisfied also at all points of a sufficiently small neighbourhood of S. We therefore consider the other infinitely many inequalities. Let S^* be a point close to S so that the elements of $S^* - S$ are near zero. Let $\underline{x}$ be an admissible k-vector $\neq \pm \underline{e_k}$. Consider $S^*[\underline{x}] - s_k^*$ where $S^* = (s_{kl}^*)$. Let $\varepsilon > 0$ be a real number. We can choose S^* close to S so that $(S^* - S)[\underline{x}] \geq -\varepsilon \underline{x}' \underline{x}$. Now

$$\begin{aligned} S^*[\underline{x}] - s_k^* &= (S^* - S)[\underline{x}] + S\,\underline{x} - s_k^* \\ &\geq -\varepsilon \underline{x}' \underline{x} + S[\underline{x}] - s_k^*. \end{aligned}$$

If $\lambda > 0$ is the smallest eigen value of S, then we may assume ε small enough so that

$$S^*[\underline{x}] - s_k^* \geq \frac{\lambda}{2} \underline{x}' \underline{x} - s_k^*.$$

60 There are only finitely many integral vectors $\underline{x}$ with $\frac{\lambda}{2} \underline{x}' \underline{x} \leq s_k^*$. We may therefore choose S^* close enough to S such that

$$S^*[\underline{x}] - s_k^* \geq \frac{\lambda}{2} \underline{x}' \underline{x} - s_k^* > 0$$

for all admissible k-vectors $\underline{x}$. Doing this for $k = 1, \ldots, n$ we see that there is a small sphere containing S and which consists entirely of points S^*. These, by our construction, are inner points. This proves our contention. □

Consider now the outer points of $\mathscr{R}$. Let S be one such. Then at least for one r, $L_r(S) < 0$. Since the L_r are linear functions of the coordinates and hence continuous, we may choose a neighbourhood of S consisting of points for all of which $L_r < 0$. This means that the set of outer points of $\mathscr{R}$ is open. Note that here it is enough to deal with one inequality alone unlike the previous one where one had to deal with all the L_r's.

Let now S be a boundary point of $\mathscr{R}$. Let S^* be an inner point. Consider the points T_λ defined by

$$T_\lambda = \lambda S^* + (1 - \lambda)S.$$

These are points on the line joining S and S^* and every neighbourhood of S contains points T_λ with $\lambda > 0$ and points T_λ with $\lambda < 0$.

Consider the points T_λ with $0 < \lambda \leq 1$. These are the points between S and S^*. Let L_r be one of the linear polynomials defining $\mathscr{R}$. Now $L_r(S) \geq 0$, and $L_r(S^*) > 0$, for all r. Thus

$$L_r(T_\lambda) = \lambda L_r(S^*) + (1 - \lambda)L_r(S) > 0.$$

Hence T_λ is an inner point. 61

Let now T be a point with $\lambda < 0$. Since S is a boundary point, there is an r such that $L_r(S) = 0$. For this r

$$L_r(T_\lambda) = \lambda L_r(S^*) < 0$$

which proves that T_λ is an outer point.

Since linear functions are continuous, the limit of a sequence of points of $\mathscr{R}$ is again a point of $\mathscr{R}$. This proves

Theorem 7. *$\mathscr{R}$ is a closed set in $\mathscr{P}$ and the boundary points of $\mathscr{R}$ constitute the frontier of $\mathscr{R}$ in the topology of $\mathscr{P}$.*

We now prove the following

Theorem 8. *Let S and S^* be two points of $\mathscr{R}$ such that $S[U] = S^*$ for a unimodular $U \neq \pm E$. Then S and S^* are boundary points of $\mathscr{R}$ and U belongs to a finite set of unimodular matrices determined completely by the integer n.*

Proof. The second part of the theorem follows readily from the Corollary to Theorem 5. To prove the first part, we consider two cases: (1) U is a diagonal matrix, and (2) U is not a diagonal matrix. □

Let U be a diagonal matrix, $U = (a_1, \ldots, a_n)$, with $a_i = \pm 1$. We may assume, since $S[U] = S[-U]$ that $a_1 = 1$. Let a_{k+1} be the first element $= -1$. Then, with usual notation,

$$s^*_{kk+1} = -s_{kk+1}.$$

62 But S and S^* being points of $\mathscr{R}$ we have

$$0 \leq s^*_{kk+1} = -s_{kk+1} \leq 0$$

which means that $s_{kk+1} = 0 = s^*_{kk+1}$. Hence S and S^* are both boundary points of $\mathscr{R}$.

Suppose U is not a diagonal matrix and denote its columns by $\underline{u_1}, \ldots, \underline{u_n}$. Let $\underline{u_k}$ be the first column different from the corresponding column of a diagonal matrix. Hence $u_i = \pm e_i$, $i = 1, \ldots, k-1$. (Note that k may very well be equal to 1). Then

$$U = \begin{pmatrix} D & * \\ 0 & V \end{pmatrix}$$

where D is a diagonal matrix which is unimodular. V is a unimodular matrix. Furthermore

$$U^{-1} = \begin{pmatrix} D^{-1} & * \\ 0 & V^{-1} \end{pmatrix}$$

is unimodular. Let $\underline{w_k}$ be the k-th column of U^{-1}. Then $\underline{w_k} \neq \pm \underline{e_k}$. Now

$$s^*_k = S[\underline{u_k}] \geq s_k$$

and

$$s_k = S^*[\underline{w_k}] \geq s^*_k$$

which proves that $S[\underline{u_k}] - s_k = 0 = S^*[\underline{w_k}] - s^*_k$ and therefore S and S^* are boundary points of $\mathscr{R}$.

Suppose now that S is a boundary point of $\mathscr{R}$. By Theorem 7, therefore, there exists a sequence of outer points $S_1, S_2, \ldots$ converging to S. 63

If the suffix k is sufficiently large, then all the S_k's lie in a neighbourhood of S. Therefore they are all contained in an $\mathscr{R}^*_t$ for some t. For each k let U_k be a unimodular matrix such that $S_k[U_k]$ is in $\mathscr{R}$. Since $\mathscr{R} \subset \mathscr{R}^*_t$, we have for all sufficiently large k, S_k and $S_k[U_k]$ are both in $\mathscr{R}^*_t$. It follows therefore by Theorem 5, that U_k's belong to a finite set of matrices. There exists therefore a subsequence $S_{k_1}, S_{k_2}, \ldots$ converging to S such that one unimodular matrix U, among these finitely many, carries S_{k_i} into $\mathscr{R}$. Also $\underset{n\to\infty}{\text{Lim}} S_{k_n} = S$ and therefore $\lim S_{k_n}[U] = S[U]$ is a point of $\mathscr{R}$. Since S is a point of $\mathscr{R}$, it follows from the above theorem that $S[U]$ is also a boundary point of $\mathscr{R}$. Furthermore $U \neq \pm E$ since S_k are all outer points and $S_k[U] \in \mathscr{R}$. Hence

Theorem 9. *If S is a boundary point of $\mathscr{R}$, there exists a unimodular matrix $U \neq \pm E$ and belonging to the finite set determined by Theorem 8, such that $S[U]$ is again a boundary point of $\mathscr{R}$.*

By Theorem 8, there exist finitely many unimodular matrices say $U_1, \ldots, U_g$ which occur in the transformation of boundary points into boundary points. If $\underline{u_k}$ is the k-th column of one of these matrices, then $\underline{u_k}$ is an admissible k-vector. Suppose it is $\neq \pm e_k$. Then for all $S \in \mathscr{R}$, $S[\underline{u_k}] - s_k \geq 0$. Let us denote by $L_1, L_2, \ldots, L_h$ all the linear forms, not identically zero, which result from all the u_k's $k = 1, \ldots, n$ occurring in the set $U_1, \ldots, U_g$. Let $L_1, \ldots, L_h$ also include the linear forms s_{kk+1}, $k = 1, \ldots, n-1$; then from above we see that for a boundary point S of $\mathscr{R}$, there is an $r \leq h$ such that $L_r(S) = 0$ (not identically). Also for all points of $\mathscr{R}$ 64

$$s_1 > 0, L_1(S) \geq 0, \ldots, L_h(S) \geq 0. \tag{59}$$

But what is more important, we have

Theorem 10. *A point S of $\mathscr{P}$ belongs to $\mathscr{R}$ if and only if $s_1 > 0$ and $L_r(S) \geq 0$ for $r = 1, \ldots, h$.*

Proof. The interest in the theorem is in the *sufficiency* of the conditions (59). □

Let S be a point of $\mathscr{P}$ satisfying (59). Suppose S is not in $\mathscr{R}$. Since it is in $\mathscr{P}$, it is an outer point of $\mathscr{R}$. Therefore $L_r(S) < 0$ for some $r > h$.

Let S^* be an inner point of $\mathscr{R}$. Consider the points T_λ,

$$T = \lambda S + (1-\lambda)S^*$$

for $0 < \lambda < 1$, in the open segment joining S and S^*. Since the set of inner points of $\mathscr{R}$ is open and S is assumed to be an outer point, there exists a λ_0 such that T_{λ_0} is on the frontier of $\mathscr{R}$ and $0 < \lambda_0 < 1$. By our remarks above, there exists for T_{λ_0} an $s \leq h$ such that $L_s(T_{\lambda_0}) = 0$. This means that

$$0 = L_s(T_{\lambda_0}) = \lambda_0 L_s(S) + (1-\lambda_0)L_s(S^*).$$

But $(1-\lambda_0)L_s(S^*) > 0$ so that $L_s(T_{\lambda_0}) > 0$. This is a contradiction. Therefore $S \in \mathscr{R}$.

65 We have therefore proved that $\mathscr{R}$ is bounded by a finite number of planes all passing through the origin. $\mathscr{R}$ is thus a pyramid.

Let now $\overline{\mathscr{R}}$ denote the closure of $\mathscr{R}$ in the space R_h. At every point S or $\overline{\mathscr{R}}$ one has, because of continuity of linear functions,

$$s_1 \geq 0, \quad L_r(S) \geq 0, \quad r = 1, 2, 3, \ldots$$

If $S \in \overline{\mathscr{R}}$ but not in $\mathscr{R}$, then $s_1 = 0$. In virtue of the other inequalities, we see that

$$S = \begin{pmatrix} 0 & 0 \\ 0 & S_1 \end{pmatrix}.$$

S_1 again has similar properties. Thus either $S = 0$ or

$$S = \begin{pmatrix} 0 & 0 \\ 0 & S_k \end{pmatrix}$$

where S_k is non-singular and is a reduced matrix of order r, $0 < r < n$. We thus see that the points of $\overline{R}$ which are not in $\mathscr{R}$ are the semi-positive reduced matrices.

Consider now the space $\mathscr{P}$ and the group Γ. If $U \in \Gamma$, the mapping $S \to S[U]$ is topological and takes $\mathscr{P}$ onto itself. For $U \in \Gamma$ denote by $\mathscr{R}_U$ the set of matrices $S[U]$ with $S \in \mathscr{R}$. Because U and $-U$ lead to the same mapping, we have $\mathscr{R}_U = \mathscr{R}_{-U}$. Since in every class of matrices there is a reduced matrix we see that

1) $\sum\limits_{U \in \Gamma} \mathscr{R}_{\pm U} = \mathscr{P}$

66 where in the summation we identify U and $-U$. Thus the $\mathscr{R}_{\pm U}$'s cover $\mathscr{P}$ without gaps.

Let U and V be in Γ and $U \neq \pm V$. Consider the intersection of $\mathscr{R}_U$ and $\mathscr{R}_V$. Let $S \in \mathscr{R}_U \cap \mathscr{R}_V$. Then $T_1 = S[U^{-1}]$ and $T_2 = S[V^{-1}]$ are both points of $\mathscr{R}$. Moreover $T_1 = T_2[VU^{-1}]$ and $VU^{-1} \neq \pm E$ so that T_1 is a boundary point of $\mathscr{R}$. Since the mapping $S \to S[U]$ is topological S is a boundary point of $\mathscr{R}_U$ and also of $\mathscr{R}_V$. Hence

2) If $UV^{-1} \neq \pm E$ *and U and V are unimodular, then $\mathscr{R}_U$ and $\mathscr{R}_V$ can have at most boundary points in common.*

In particular, if $U \neq \pm E$, $\mathscr{R}$ and $\mathscr{R}_U$ can have only boundary points in common. If $S \in \mathscr{R} \cap \mathscr{R}_U$ then S and $S[U^{-1}]$ are in $\mathscr{R}$ and by Theorem 9, U belongs to a finite set of matrices depending only on n. If we call $\mathscr{R}_U$ a *neighbour* of $\mathscr{R}$ if $\mathscr{R} \cap \mathscr{R}_U$ is not empty, then we have proved

3) *$\mathscr{R}$ has only finitely many neighbours.*

Let K now be a compact subset of $\mathscr{P}$. It is therefore bounded in $\mathscr{P}$ and hence there exists a $t > 0$ such that $K \subset \mathscr{R}_t^*$. Suppose $\mathscr{R}_U$, for a unimodular U, intersects K. Let $S \in \mathscr{R}_U \cap K$. There is then a $T \in \mathscr{R}$ such that $T[U] = S$. For large t, $\mathscr{R} \subset \mathscr{R}_t^*$. Then T and S are both in $\mathscr{R}_t^*$ and $S = T[U]$. Therefore U belongs to a finite set of matrices. Hence there exist a finite number of unimodular matrices, say $U_1, \ldots, U_p$ such that

$$K \subset \sum_{i=1}^{p} \mathscr{R}_{U_i}$$

Hence 67

4) *Every compact subset of $\mathscr{P}$ is covered by a finite number of images $\mathscr{R}_U$ of $\mathscr{R}$.*

We have thus obtained the fundamental results of Minkowski's reduction theory.

We now give a simple application.

Suppose S is a positive, reduced, *integral* matrix. Then since $s_1 s_2 \ldots s_n \leq b_n |S|$, $s_1, \ldots, s_n$ are positive and b_n depends only on n, it follows that for a given $|S|$, there exist only finitely many integer values for $s_1, \ldots, s_n$. Also

$$-s_k \leq 2s_{kl} \leq s_k, \quad k < l$$

so that s_{kl} being integers, there are finitely many values of s_{kl} satisfying the above inequalities. We have therefore the

Theorem 11. *There exist only finitely many positive, integral, reduced matrices with a given determinant and number of rows.*

Since all matrices in a class have the same determinant, and in each class there is at least one reduced matrix, we get the

Theorem 12. *There exist only a finite number of classes of positive integral matrices with given determinant and number of rows.*

It has to be noticed, that in virtue of property 3) above, one has, in general, only one reduced matrix in a class.

6 Binary forms

68

We now study the particular case $n = 2$.

Let $S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ be a positive binary matrix and $\underline{x} = \binom{x}{y}$ a vector. The quadratic form $S[\underline{x}] = ax^2 + 2bxy + cy^2$ is positive definite. By the results of the previous section, we see that, if S is reduced then

$$a > 0, \quad 0 \leq 2b \leq a \leq c. \tag{60}$$

We shall now prove that *any* matrix S satisfying (60) is reduced.

Let $\underline{x} = \binom{x}{y}$ be an admissible one-vector. If $y = 0$, then $x = \pm 1$. If $y \neq 0$, then x and y are coprime integers. Consider the value $ax^2 + 2bxy + cy^2$ for admissible one-vectors. We assert that $ax^2 + 2bxy + cy^2 \geq a$. In the first case $S[\underline{x}] = a$. In the second case, because of (60)

$$ax^2 + 2bxy + cy^2 \geq a(x^2 - xy + y^2).$$

But x and y are not both zero. Thus $x^2 - xy + y^2 \geq 1$ which means that $S[\underline{x}] \geq a$.

Let now $\underline{x} = \binom{x}{y}$ be an admissible two-vector. Then $y = \pm 1$. If $x = 0$, then $S[\underline{x}] = c$. Let $x \neq 0$, then

$$S[\underline{x}] = ax^2 \pm 2bx + c = c + x(ax \pm 2b).$$

Because of (60), it follows that $x(ax \pm 2b) \geq 0$. Thus S satisfies conditions I) and II) of half reduction. Also $b \geq 0$. This proves that $S > 0$ and reduced.

(60) thus gives the necessary and sufficient conditions for a binary quadratic form to be reduced. 69

In the theory of binary quadratic forms, one discusses some-times equivalence not under all unimodular matrices, but only with respect to those unimodular matrices whose determinant is unity. We say that two binary matrices S and T are *properly equivalent* if there is a unimodular matrix U such that

$$S = T[U], \quad |U| = 1. \tag{61}$$

The properly equivalent matrices constitute a *proper class*. Note that the properly unimodular matrices form a group. Two matrices S and T which are equivalent in the sense of the previous sections, but which do not satisfy (61) are said to be *improperly equivalent.* Note that improper equivalence is *not* an equivalence relation.

In order to obtain the reduction theory for proper equivalence we proceed thus: If $S_1 = \left(\begin{smallmatrix} a_1 & b_1 \\ b_1 & c_1 \end{smallmatrix}\right)$ is positive, then there is a unimodular matrix U such that $S = S_1[U] = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ satisfies (60). If $|U| = 1$ we call S a *properly reduced* matrix. If $|U| = -1$, then consider W

$$W = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{62}$$

Then $V = UW$ has the property $|V| = 1$. Now $S[W] = \left(\begin{smallmatrix} a & -b \\ -b & c \end{smallmatrix}\right)$ and we call this properly reduced. In any case we see that S is properly reduced means 70

$$a > 0, \quad 0 \leq |2b| \leq a \leq c. \tag{63}$$

If we denote by $\mathscr{R}$ the reduced domain, that is the set of reduced matrices in the old sense and $\mathscr{R}^*$ the properly reduced domain, one sees immediately that

$$\mathscr{R}^* = \mathscr{R} + \mathscr{R}_W$$

where W has the meaning in (62).

We shall now give two applications.

Let $S = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ be a positive *integral* matrix. Because of conditions (63) and the additional condition (31), it follows that for given $|S|$, there exist only finitely many properly reduced integral matrices. Consider now the case $|S| = 1$. Then because of (31),

$$ac \leq \frac{4}{3} \tag{64}$$

and hence the only integers a, b, c, satisfying (63) and (64) are $a = c = 1, b = 0$. This proves

i) *Every binary integral positive quadratic form of determinant unity is properly equivalent to $x^2 + y^2$.*

Let now p be a prime number > 2. Let p be representable by the quadratic form $x^2 + y^2$. We assert that then $p \equiv 1 (\text{mod } 4)$. For, if x and y are integers such that

$$x^2 + y^2 = p$$

71 then x and y cannot be congruent to each other mod 2. So let x be odd and y even. Then $p = x^2 + y^2 \equiv 1 (\text{mod } 4)$.

We will now prove that conversely if $p \equiv 1 (\text{mod } 4)$, the form $x^2 + y^2$ represents p (integrally). For, let ρ be a primitive root mod p. There is then an integer k, $1 \leq k < p - 1$ such that

$$\rho^k \equiv -1 (\text{mod p}).$$

This means that $\rho^{2k} \equiv 1 (\text{mod p})$ and by definition of primitive root, we get $k = p - 1/2$. But $p \equiv 1 (\text{mod } 4)$ so that k is an even integer. Therefore

$$-1 \equiv (\rho^{k/2})^2 (\text{mod p}).$$

There is thus an integer b, $1 \leq b \leq p-1$ such that $b^2 \equiv -1 (\text{mod p})$. Put $b^2 = -1 + \lambda p$, $\lambda \geq 1$ an integer.

Consider the binary form $px^2 + 2bxy + \lambda y^2$. Its determinant is $p\lambda - b^2 = 1$. By the result obtained in i), this form is equivalent to $x^2 + y^2$. But $px^2 + 2bxy + \lambda y^2$ represents p, $(x = 1, y = 0)$. Therefore $x^2 + y^2$ represents p. Thus

ii) *If p is a prime > 2, then $x^2 + y^2 = p$ has a solution if and only if $p \equiv 1 (\text{mod } 4)$.*

Results i) and ii) are due originally to *Lagrange*.

Let $S[\underline{x}] = ax^2 + 2bxy + cy^2$ be a real, positive, binary quadratic form. We can write

$$S[\underline{x}] = a(x - \tau y)(x - \overline{\tau} y) \tag{65}$$

where τ is a root, necessarily complex, of the polynomial $az^2 + 2bz + c$ **72**
and $\overline{\tau}$ is its conjugate. Let $\tau = \xi + i\tau$ have positive imaginary part.

Let $V = \left(\begin{smallmatrix} \lambda & \mu \\ \nu & \rho \end{smallmatrix}\right)$ be a real matrix of unit determinant and consider the mapping

$$S \to S[V].$$

Then $S[V\underline{x}]$ is given by

$$S[V\underline{x}] = a'(x - \tau' y)(x - \overline{\tau}' y) \tag{66}$$

where $a' = a(\lambda - \nu\tau)\,(\lambda - \nu\overline{\tau})$ is necessarily real and positive, and

$$\tau' = V^{-1}(\tau) = \frac{\rho\tau - \mu}{-\nu\tau + \lambda}. \tag{67}$$

It is easy to see that τ' also has positive imaginary part. Let us also observe that $\tau = \dfrac{-b + i\sqrt{|S|}}{a}$.

Consider now the relationship between S and τ. If S is given, then (65) determines a τ with positive imaginary part. Now given τ, (65) itself shows that S is determined only upto a real factor. This real factor can be determined by insisting that the associated quadratic forms have a given determinant. In particular, if $|S| = 1$ then the τ is uniquely determined by S and conversely. If $\tau = \xi + i\eta$, $\eta > 0$, then the S is given by

$$S = \begin{pmatrix} \eta^{-1} & 0 \\ 0 & \eta \end{pmatrix} \begin{bmatrix} 1 & -\xi \\ 0 & 1 \end{bmatrix} \tag{68}$$

73 Let $\mathscr{P}$ denote the space of positive binary forms of unit determinant and $\mathscr{G}$ the upper half complex τ-plane. By what we have seen above the mapping $S \to \tau$ in (65) is $(1, 1)$ both ways. Let Γ denote the group of proper unimodular matrices. It acts on $\mathscr{G}$ as a group of mappings

$$\tau \to U(\tau) = \frac{\lambda\tau + \mu}{\nu\tau + \rho}, \quad U = \begin{pmatrix} \lambda & \mu \\ \nu & \rho \end{pmatrix} \tag{69}$$

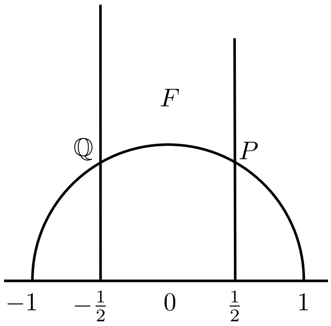
of $\mathscr{G}$ onto itself. If we define two points τ_1, τ_2 in $\mathscr{G}$ as *equivalent* if there is a $U \in \Gamma$ such that $\tau_1 = U(\tau_2)$, then the classical problem of constructing a fundamental region in $\mathscr{G}$ for Γ, is seen to be the same as selecting from each class of equivalent points one point so that the resulting point set has 'nice' properties.

By means of the $(1, 1)$ correspondence, we have established in (68) between $\mathscr{P}$ and $\mathscr{G}$, we have $S_1 = S_2[U]$ if and only if the corresponding points τ_1, τ_2 respectively satisfy

$$\tau_1 = U^{-1}(\tau_2).$$

We define the fundamental region F in $\mathscr{G}$ to be the set of points τ such that the matrices corresponding to them are properly reduced; in other words, they satisfy (63). For the S in (68), $S[\underline{x}] = \dfrac{1}{\eta}(x^2 - 2\xi xy + (\xi^2 + \eta^2)y^2)$. Therefore F consists of points $\tau = \xi + i\eta$ for which

$$\begin{aligned} |2\xi| &\leq 1 \\ \xi^2 + \eta^2 &\geq 1 \end{aligned} \tag{70}$$



This is the familiar modular region in the upper half τ-plane. That it is a fundamental region follows from the properties of the space of reduced matrices in $\mathscr{P}$. The points P and Q are the complex numbers $\dfrac{\pm 1 + i\sqrt{3}}{2}$, and so for any point in F, $\eta \geq \dfrac{\sqrt{3}}{2}$. This means that for a positive reduced binary form $ax^2 + 2bxy + cy^2$ of determinant $\underline{d}$ 74

$$\frac{a}{\sqrt{d}} \leq \frac{2}{\sqrt{3}},$$

which we had already seen in Theorem 1.

7 Reduction of lattices

Let V be the Euclidean space of n dimensions formed by n-rowed real columns

$$\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Let $\alpha_1, \ldots, \alpha_n$ be a basis of V so that

$$\alpha_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, \quad i = 1, \ldots, n.$$

Denote by A the matrix (a_{kl}). Obviously $|A| \neq 0$.

Let L be a lattice in V and let $\alpha_1, \ldots, \alpha_n$ be a basis of this lattice. L then consists of elements $\alpha_1 g_1 + \cdots + \alpha_n g_n$ where $g_1, \ldots, g_n$ are integers. We shall call A the *matrix of the lattice.*

75 Conversely if A is any non-singular n-rowed matrix, then the columns of A, as elements of V are linearly independent and therefore determine a lattice.

Let L be the lattice above and let $\beta_1, \ldots, \beta_n$ be any other base of L and B its matrix, then

$$B = AU$$

where U is a unimodular matrix. Also if U runs through all unimodular matrices, then AU runs through all bases of L. We now wish to single out among these bases one which has some distinguished properties.

Let us introduce in V, the inner product $\alpha \cdot \beta$ of two vectors α and β by

$$\alpha \cdot \beta = a_1 b_1 + \cdots + a_n b_n$$

where $\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \beta = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. The square of the length of the vector α is given by

$$\alpha^2 = a_1^2 + \cdots + a_n^2.$$

Let A be the matrix of a base $\alpha_1, \ldots, \alpha_n$ of L. Consider the positive matrix $S = A'A$. If S is given A is determined only upto an orthogonal matrix P on its left. For, if $A'A = A_1' A_1$ then $AA_1^{-1} = P$ is orthogonal. But multiplication on the left by an orthogonal matrix implies a rotation in V about the origin.

We shall call a base B of L *reduced* if $S_1 = B'B$ is a reduced matrix. Obviously in this case

$$\begin{gathered} 0 < \beta_1^2 \leq \ldots \leq \beta_n^2 \\ \beta_k \beta_{k+1} \geq 0, \quad k = 1, \ldots, n-1. \end{gathered}$$

76 From the way reduced matrices are determined we see that a reduced base $\beta_1, \ldots, \beta_n$ of L may be defined to be a base such that for every set of integers $x_1, \ldots, x_n$ such that $(x_k, \ldots, x_n) = 1$ the vector

$$\beta = \beta_1 x_1 + \cdots + \beta_n x_n$$

satisfies

$$\beta^2 \geq \beta_k^2 \quad (k = 1, \ldots, n.)$$

Also

$$\beta_k \cdot \beta_{k+1} \geq 0 (k = 1, \ldots, n + 1).$$

If follows therefore that

$$\beta_1^2 \ldots \beta_n^2 \leq c_n |A'A| = c_n |A|^2$$

c_n being a constant depending only on n. Also abs $|A|$ is the volume of the parallelopiped formed by the vectors $\beta_1, \ldots, \beta_n$.

consider the case $n = 2$.

We have, because of (30)

$$\beta_1^2 \cdot \beta_2^2 \leq \frac{4}{3} |A|^2 \tag{72}$$

Let now Θ denote the acute angle between the vectors β_1 and β_2. Since the area of the parallelogram formed by β_1 and β_2 on the one hand equals abs $|A|$ and on the other $\sqrt{\beta_1^2 \cdot \beta_2^2} \cdot \sin\theta$, we see that 77

$$\sin^2 \Theta \geq \frac{3}{4} \tag{73}$$

Since $0 \leq \Theta \leq \dfrac{\pi}{2}$, it follows from (73) that

$$\frac{\pi}{3} \leq \theta \leq \frac{\pi}{2}.$$

Hence for a two dimensional lattice we may choose a basis in such a manner that the angle (acute) between the basis vectors is between 60° and 90°.

Bibliography

[1] C.F. Gauss *Disquisitiones Arithmeticae* Ges. Werke 1, (1801).

[2] C. Hermite *Oeuvres* Vol.1, Paris (1905) P. 94-273.

[3] H. Minkowski *Geometrie der Zahlen*, Leipzig (1896).

[4] H. Minkowski Discontinuitätsbereich für arithmetische Aquivalenz. *Ges. Werke*, Bd.2, (1911), P.53 - 100.

[5] C.L. Siegel Einheiten quadratischer Formen *Abh. Math. Sem. Hansischen Univ.* 13, (1940), P. 209-239.

Chapter 3

Indefinite quadratic forms

1 Discontinuous groups

78

In the previous chapter we had met with the situation in which a group of transformations acts on a topological space and we constructed, by a certain method, a subset of this space which has some distinguished properties relative to the group. We shall now study the following general situation.

Let Γ be an abstract group and T a Hausdorff topological space on which Γ has a representation

$$t \to t\gamma, \quad t \in T, \quad \gamma \in \Gamma \tag{1}$$

carrying T into itself. We say that this representation of Γ is *discontinuous* if for every point $t \in T$, the set of points $\{t\gamma\}$ for $\gamma \in \Gamma$ has no limit point in T. The problem now is to determine, for a given Γ, all the spaces T on which Γ has a discontinuous representation. For an arbitrarily given group, this problem can be very difficult. We shall, therefore, impose certain restrictions on Γ and T. Let us assume that there is a group Ω, of transformations of T onto itself, which is *transitive* on T. This means that if t_1 and t_2 are any two elements of T, there exists $\omega \in \Omega$ such that

$$t_1 = t_2\omega. \tag{2}$$

Let us further assume that Γ is a subgroup of Ω. Let t_0 be a point in T
79 and consider the subgroup Λ of Ω consisting of $\lambda \in \Omega$ such that

$$t_0 = t_0\lambda. \tag{3}$$

If t is any point of T, we have because of transitivity,

$$t = t_0\rho$$

for some $\rho \in \Omega$. Because of (3), we get

$$t = t_0\Lambda\rho.$$

Conversely if $\rho' \in \Omega$ is such that $t = t_0\rho'$, then $t_0\rho' = t_0\rho$ or $\rho' \in \Lambda_\rho$. Thus every point $t \in T$ determines a coset Λ_ρ of $\Lambda\backslash\Omega$ that is, the space of right cosets of Ω modulo Λ. Conversely if Λ_ρ is any coset, then $t = t_0\rho$ is a point determined by Λ_ρ. Hence the mapping

$$t \to \Lambda_\rho \tag{4}$$

of T on $\Lambda\backslash\Omega$ is $(1, 1)$ both ways. In order to make this correspondence topological, let us study the following situation.

Let Ω be locally compact topological group and T a Hausdorff topological space on which Ω has a representation

$$t \to t\omega \tag{5}$$

as a transitive group of mappings. Let us assume that this representation is open and continuous. We recall that (5) is said to be *open* if for every open set P in Ω and every $t \in T$ the set $\{t\omega\}$, $\omega \in P$ is an open set in T. Then it follows that the subgroup Λ of Ω leaving $t_0 \in T$ fixed is not only a closed subgroup but that the mapping (4) of T on $\Lambda\backslash\Omega$ is a homeomorphism.

80 Let Γ be a subgroup of Ω which has on T a discontinuous representation. Then Γ has trivially a representation in $\Lambda\backslash\Omega$. By the remarks above, the representation

$$\Lambda\omega \to \Lambda\omega\rho, \quad \rho \in \Gamma \tag{6}$$

is discontinuous in $\Lambda\backslash\Omega$.

On the other hand, let Λ be *any* closed subgroup of Ω. Then the representation

$$\omega' \to \Lambda\omega\omega'$$

of Ω on $\Lambda\backslash\Omega$ is open and continuous. It is clearly transitive. In order, therefore, to find all spaces on which Γ has a discontinuous representation, it is enough to consider the spaces of right cosets of Ω with regard to closed subgroups Λ of Ω.

Suppose Λ is a closed subgroup of Ω and Γ has a discontinuous representation on $\Lambda\backslash\Omega$. Let K be a closed subgroup of Ω contained in Λ. Then Γ has a discontinuous representation on $K\backslash\Omega$. For, if $K\omega$ is a coset such that the set of cosets $\{K\omega\rho\}$, $\rho \in \Gamma$ has a limit point in $K\backslash\Omega$, then the set $\{\Lambda\omega\rho\}$, $\rho \in \Gamma$ also has a limit point in $\Lambda\backslash\Omega$ and so (6) would not be discontinuous. In particular, if we take for K the subgroup consisting only of the identity element $\underline{e}$, then Γ is discontinuous in Ω is clearly equivalent to Γ is a *discrete* subgroup of Ω.

Thus if there exists some subgroup Λ of Ω such that Γ is discontinuous in $\Lambda\backslash\Omega$, then necessarily Γ has to be discrete. It can be proved that if Ω has a countable basis of open sets, then Γ is enumerable.

Suppose now that Ω is a locally compact group with a countable basis of open sets. Let Γ be a discrete subgroup of Ω. If Λ is any compact, hence closed, subgroup of Ω then it follows that the representation (6) of Γ in $\Lambda\backslash\Omega$ is discontinuous. This can be seen by assuming that for a certain ω, the set $\Lambda\omega\rho_n$, $\rho_n \in \Gamma$ has limit point and this will lead to a contradiction because of the discreteness of Γ. **81**

In general the fact (6) is discontinuous in $\Lambda\backslash\Omega$ does not entail that Λ is compact. Let us, therefore, consider the following situation.

Let Ω be a locally compact group possessing a countable basis of open sets. Then there exists in Ω a right invariant Haar measure $d\omega$ which is determined uniquely upto a positive multiplicative factor. Let Γ be a discrete subgroup of Ω. There exists then in Ω a subset F possessing the following properties: 1) $\bigcup\limits_{a\in\Gamma} Fa = \Omega$, 2) the sets $\{Fa\}$ for $a \in \Gamma$ are mutually disjoint and 3) F is measurable in terms of the Haar measure $d\omega$. F is then said to be a *fundamental* set relative to Γ. Note that if F is a fundamental set then so if Fa for any $a \in \Omega$ so that a fundamental set is not unique. 1) and 2) assert that F intersects each coset of $\Gamma\backslash\Omega$ in exactly one point so that F has to be formed in Ω by choosing one element from each coset $\Gamma\backslash\Omega$. The interesting point is that, under the conditions on Ω, this can be done in such a way that the resulting set F is measurable. Let us now assume that

$$\int\limits_F d\omega < \infty. \tag{7}$$

82 It can then be shown that the value of the integral in (7) is independent of the choice of F. We now state, without proof, the important

Theorem 1. *Let Ω be a locally compact topological group with a countable basis of open sets. Let Γ be a discrete subgroup of Ω and F a fundamental set in Ω relative to Γ. Let F have finite Haar measure in Ω. If Λ is any closed subgroup of Ω, then Γ has a discontinuous representation in $\Lambda\backslash\Omega$ if and only if Λ is compact.*

The interest in the theorem lies in the *necessity* part of it.

Let us assume that Ω is, as will be in the applications, a Lie group. Let Γ be a discrete subgroup of Ω. For any closed subgroup Λ of Ω, the dimensions of Λ, $\Lambda\backslash\Omega$ and Ω are connected by

$$\dim\Lambda + \dim\Lambda\backslash\Omega = \dim\Omega.$$

If F is a fundamental set in Ω with regard to Γ and is of finite measure, in terms of the invariant measure in Ω, then by Theorem 1, Γ will be discontinuous in $\Lambda\backslash\Omega$ if and only if Λ is compact. In order, therefore, to obtain a space $T = \Lambda\backslash\Omega$ of smallest dimension in which Γ has a discontinuous representation, one has to consider a Λ which is compact and maximal with this property.

83 Let us consider the following example.

Let Ω be the group of n-rowed real matrices $A \cdot \Omega$ is a Lie group. Let us determine first all compact subgroups of Ω. Let K be a compact subgroup of Ω. If $C \in K$, then $|C| = \pm 1$. For, the mapping $C \to |C|$ of K into the multiplicative group of real numbers is clearly topological and since K is compact, the set of images $|C|$ is a compact and so bounded subgroup of the multiplicative group of real numbers. Thus $|C| = \pm 1$. In order to study K therefore, it is enough to study the group Ω_0 of real matrices A with $|A| = \pm 1$. Let $\{dA\}$ denote the volume measure in Ω_0 so that

$$\{dAB\} = \{dA\}$$

for $B \in \Omega_0$. Let $\mathscr{M}$ be an open bounded subset of Ω_0. Consider the set

$$\mathscr{G} = \bigcup_{C \in K} \mathscr{M} C.$$

Since the sets $\mathscr{M}C$ are open, $\mathscr{G}$ is open. Since K is compact, it follows that $\mathscr{G}$ is bounded. Consider the integral

$$P = \int_{\mathscr{G}} A'A\{dA\}.$$

Since $A'A > 0$, it follows that P is positive. Also if C is in K,

$$\begin{aligned} P[C] &= \int_{\mathscr{G}} C'A'AC\{dA\} \\ &= \int_{\mathscr{G}} A'A\{dAC^{-1}\} = P \end{aligned}$$

This proves that there exists a $P > 0$ such that $P[C] = P$ for $C \in K$. **84** Since $P > 0$, there exists $B \in \Omega$ such that $P = B'B$. Hence if $Q = BCB^{-1}$ then $Q'Q = E$ or Q is orthogonal. Hence BKB^{-1} is a subgroup of the orthogonal group. We have hence proved

Theorem 2. *All maximal compact subgroups of Ω are conjugates of the real orthogonal group.*

Let $\mathscr{P}_0$ denote the space of all positive real n-rowed matrices of determinant 1. Ω_0 has in $\mathscr{P}_0$ a representation $P \to P[A]$, $P \in \mathscr{P}_0$, $A \in \Omega_0$ and this representation is both open and continuous. Also Ω_0 is transitive on $\mathscr{P}_0$. The set of elements $A \in \Omega_0$ which fix the matrix E_n is precisely the orthogonal group Λ. By our considerations above,

$\Lambda\backslash\Omega_0$ is homeomorphic to $\mathscr{P}_0$. So every discrete subgroup Γ of Ω_0 has discontinuous representation in $\mathscr{P}_0$. We shall consider the subgroup Γ consisting of unimodular matrices. That this is discrete is clear. In the previous chapter we constructed for Γ in $\Lambda\backslash\Omega$ a fundamental domain $\mathscr{R}$. We shall now construct a fundamental set for Γ in Ω_0.

In Ω_0, Γ is represented as a group of translations $A \to AU$. Let us define the point set F_1 in Ω_0 to consist of matrices A such that $A'A = P$ is reduced in the sense of Minkowski and so is in $\mathscr{R}$. Clearly if $A \in F_1$ then BA is also an element of F_1 for arbitrary orthogonal B. Because of the properties of $\mathscr{R}$, the point set F_1 satisfies

$$F_1\Gamma = \Omega_0.$$

85 Since $P[\pm E] = P$, we shall take the subset F_0 of F_1 consisting of A with $a_{11} \geq 0$ where $A = (a_{kl})$. It is easy to see from the properties of $\mathscr{R}$, that F_0 and $F_0\gamma$ for $\gamma \in \Gamma$ have non-empty intersection only for finitely many γ. By removing from F_0 a suitably chosen set of points, one obtains a fundamental set in Ω_0 for Γ. Minkowski proved that the volume of F_0 is finite.

For more details we refer to the papers [6], [7] and [8].

2 The $\mathfrak{H}$ - space of a symmetric matrix

We now consider another important application of the previous considerations.

Let S be a non-singular n-rowed symmetric matrix of signature p, q where $p + q = n$ and $0 \leq p \leq n$. This means that there exists a real matrix L such that

$$S[L] = S_0 = \begin{pmatrix} E_p & 0 \\ 0 & -E_q \end{pmatrix} \tag{8}$$

Let Ω denote the group of real matrices C such that

$$S[C] = S. \tag{9}$$

Ω is called the *orthogonal group* of S. Ω is a Lie group. We shall now determine all compact subgroups of Ω. Let K be a compact subgroup of Ω. Then there exists a positive matrix P such that

$$P[V] = P, \quad V \in K \tag{10}$$

Since $P > 0$ and S is symmetric, there exists a matrix L such that 86

$$S[L] = S_0, P[L] = [d_1, \ldots, d_n] = D, \tag{11}$$

D being a diagonal matrix with positive diagonal elements. Let $B = L^{-1}VL$. Then since $V \in K$

$$S_0[B] = S_0, \quad D[B] = D.$$

Put $T = S_0 D = DS_0$. Then from above we have $TB = BT$. Therefore $T^2 B = T \cdot TB = TB \cdot T = BT^2$. But $T^2 = D^2$. Therefore

$$BD^2 = D^2 B. \tag{12}$$

Let $B = (b_{kl})$, then (12) gives

$$b_{kl} d_l^2 = b_{kl} d_k^2, \quad l \leq k, \quad l \leq n \tag{13}$$

so that either $b_{kl} = 0$ or $d_l^2 = d_k^2$. In any case since the $d_k > 0$ for all k, we get

$$BD = DB.$$

This means that $D = B'DB = B'BD$ and as $D > 0$, we see that B ir orthogonal.

If Λ is the orthogonal group, then K is a subgroup of $\Omega \cap L\Lambda L^{-1}$. This shows that all maximal compact subgroup of Ω are conjugates of each other and conjugate to $L\Lambda L^{-1} \cap \Omega$ Call this subgroup $\Lambda_0 \cdot \Lambda_0$ is a maximal compact subgroup of Ω.

Put now $P = (LL')^{-1}$. Then for $V \in \Lambda_0$.

$$P[V] = P.$$

Also P and S are connected by the relation 87

$$PS^{-1}P = S. \tag{14}$$

Denote by $\mathfrak{H}$ the space of symmetric matrices $P > 0$ satisfying (14) for a fixed S. For any $H \in \mathfrak{H}$ there exists a matrix M such that

$$H[M] = D, \quad S[M] = S_0 \tag{15}$$

where D is a diagonal matrix. Because of (14) we see that $DS_0D = S_0$ or since $D > 0$, $D = E$, the unit matrix. Hence

$$H = (MM')^{-1}. \tag{16}$$

But from (11), $S[L] = S_0$ which proves that $ML^{-1} \in \Omega$ or $M = CL$ for $C \in \Omega$. From (16) therefore

$$H = P[C'^{-1}].$$

Conversely for any $C \in \Omega$, $P[C] = H$ also satisfies (14). Thus the totality of positive solutions H of (14) is given by

$$H = P[C]$$

where C runs through all matrices in Ω and P is a fixed solution of (14).

This proves that the representation $H \to H[C]$ of Ω in $\mathfrak{H}$ is transitive.

Consider now the space of right cosets of Ω module Λ_0. If for a H in $\mathfrak{H}$, $H = P[C] = P[C_1]$, then by definition of Λ_0, $CC_1^{-1} \in \Lambda_0$ so that H determines a unique right coset $\Lambda_0 C_1$ of $\Lambda_0\backslash\Omega$. Also every right coset determines uniquely an element $H = P[C_1]$ in $\mathfrak{H}$. By the considerations in the previous section $\Lambda_0\backslash\Omega$ and $\mathfrak{H}$ are homeomorphic. 88 Since Λ_0 is a maximal compact subgroup, every discrete subgroup of Ω has a discontinuous representation in $\mathfrak{H}$.

We call $\mathfrak{H}$ the *representation space* of the orthogonal group Ω of S.

We remark that if S is definite, that is $p = 0$ or n, Ω is compact and so the $\mathfrak{H}$ space consists only of one point namely S if $S > 0$ and $-S$ if $-S > 0$.

We shall now obtain a parametrical representation for the space $\mathfrak{H}$ which is defined by

$$\boxed{H > 0, \quad HS^{-1}H = S.} \tag{17}$$

Let H be any solution. Put

$$K = \frac{1}{2}(H + S), \quad -L = \frac{1}{2}(H - S). \tag{18}$$

Using the matrix M in (15) we have

$$\left.\begin{aligned} K[M] &= \frac{1}{2}(S_0 + E) = \begin{pmatrix} E_p & 0 \\ 0 & 0 \end{pmatrix} \\ -L[M] &= \frac{1}{2}(E - S_0) = \begin{pmatrix} 0 & 0 \\ 0 & E_q \end{pmatrix} \end{aligned}\right\} \tag{19}$$

which shows at once that $K \geq 0$ and has rank p and $-L \geq 0$ and has rank q. Furthermore because of (17) and (18), we get

$$\left.\begin{aligned} &KS^{-1}K = K, \quad LS^{-1}L = L \\ &KS^{-1}L = 0 = LS^{-1}K. \end{aligned}\right\} \tag{20}$$

Suppose now that K is any matrix satisfying

$$KS^{-1}K = K \tag{21}$$

with $K \geq 0$ and K having rank p. Define then two matrices H and L by **89**

$$H = 2K - S, \quad -L = \frac{H - S}{2}.$$

Then $K + L = S$ so that by the law of inertia, L has rank $\geq q$. Also because of (21), H satisfies the equation $HS^{-1}H = S$. So $|H| \neq 0$ K and L satisfy the equation (20). From the equation

$$S^{-1}[K, L] = \begin{pmatrix} K & 0 \\ 0 & L \end{pmatrix}$$

and from the signature of S we have rank $L = q$ and $-L \geq 0$. Since $H = K - L$ and $|H| \neq 0$, it follows that $H > 0$ or that H is a solution of (17). We have thus reduced the solution of the inhomogeneous problem (17) to that of the homogeneous problem (21). Therefore let $K \geq 0$ be an n-rowed matrix of rank p. There exists a non-singular matrix F such that

$$K = F'\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} F$$

where $D > 0$ and has p rows. If G denotes the matrix formed by the first p-rows of F then

$$K = G'DG$$

and G has rank p. Thus the most general form of a semi-positive matrix K of n-rows and columns and of rank p is

$$K = QT^{-1}Q', \quad Q = Q^{(n,p)}$$

90 where $T = T^{(p)} > 0$, and Q has rank p. Let K satisfy (21). Then

$$Q(T^{-1}Q'S^{-1}QT^{-1} - T^{-1})Q' = 0. \tag{22}$$

But since Q has rank p, there is a submatrix of Q of p rows which is non-singular. Using this, it follows from (22) that

$$T = S^{-1}[Q] > 0.$$

The most general solution of (21) therefore is given by

$$K = T^{-1}[Q'], \quad T = S^{-1}[Q] > 0.$$

We thus obtain the homogeneous parametric representation of $\mathfrak{H}$ by

$$H = 2K - S, \quad K = T^{-1}[Q'], \quad T = S^{-1}[Q] > 0, \quad Q = Q^{(n,p)} \tag{23}$$

It is obvious that Q determines H uniquely whereas if W is a p-rowed non-singular matrix, then Q and QW determine the same H. In order to obtain the inhomogeneous parametrical representation, we consider the special case $S = S_0$ given by (8). Let us denote the corresponding H by H_0. Write

$$Q = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}, \quad Q_1 = Q_1^{(p,p)} \tag{24}$$

Then

$$T_0 = Q_1'Q_1 - Q_2'Q_2 > 0.$$

This means that $|Q_1| \neq 0$. For, if $|Q_1| = 0$, there is a column $\underline{x}$ of p rows such that $\underline{x} \neq \underline{0}$ and $Q_1\underline{x} = \underline{0}$. Then

$$0 < T_0[\underline{x}] = -\underline{x}'Q_2'Q_2x \leq 0$$

which is absurd. We can therefore put

$$Q = \begin{pmatrix} E \\ -X' \end{pmatrix} Q_1 \tag{25}$$

where $X = X^{(p,q)}$ and E is the unit matrix of order $p \cdot T_0 > 0$ then means (since $|Q| \neq 0$) that 91

$$E - XX' > 0. \tag{26}$$

Thus $K_0 = T_0^{-1}[Q']$ is given by

$$K_0 \begin{pmatrix} (E - XX')^{-1} & -(E - XX')^{-1}X \\ -X'(E - XX')^{-1} & X'(E - XX')^{-1}X \end{pmatrix} \tag{27}$$

where $E = E_p$. In order to compute $H_0 = 2K_0 - S_0$ we put

$$W = \begin{pmatrix} E_p & -X \\ -X' & E \end{pmatrix}, \quad F = \begin{pmatrix} E_p & X \\ 0 & E \end{pmatrix}$$

W and F being n-rowed matrices. We then have

$$W[F] = \begin{pmatrix} E & 0 \\ 0 & E - X'X \end{pmatrix}, \quad W[F'] = \begin{pmatrix} E - XX' & 0 \\ 0 & E_q \end{pmatrix} \tag{28}$$

Since $T_0 > 0$, (26) shows that $W > 0$ and therefore $E - X'X > 0$. We can therefore write

$$H_0 = \begin{pmatrix} \dfrac{E + XX'}{E - XX'} & \dfrac{-2X}{E - XX'} \\ \dfrac{-2X'}{E - XX'} & \dfrac{E + X'X}{E - X'X} \end{pmatrix} \tag{29}$$

$\mathfrak{H}_0$ is thus the space of matrices H_0 with X satisfying the condition (26). This shows that $\mathfrak{H}_0$ has the topological dimension pq.

In order to obtain the inhomogeneous parametrical representation for $\mathfrak{H}$ from that of $\mathfrak{H}_0$ we observe that if the matrix L is such that $S[L] = S_0$, then $H_0 = H[L]$. 92

3 Geometry of the $\mathfrak{H}$-space

Consider the space $\mathscr{P}$ of all positive n-rowed matrices P. Let $P = (p_{kl})$ and dP denote the matrix of the differentials dp_{kl}. If A is any non-singular matrix, then $d(A'PA) = A'dPA$ so that

$$ds^2 = \sigma(P^{-1}dPP^{-1}dP) \tag{30}$$

where σ denotes the trace, is invariant under the transformations $P \to P[A]$ of $\mathscr{P}$ into itself. ds^2 is a quadratic differential form in the $\dfrac{n(n+1)}{2}$ independent differentials dp_{kl}. In order to see that this is a positive definite form, observe that since the group of non-singular matrices acts transitively on $\mathscr{P}$, it is enough to verify the positivity of (30) at some particular point, say $P = E$. At $P = E$ we see that the quadratic form (30) equals

$$\sigma((dP)^2) = \sum_{k,l}(dp_{kl})^2$$

which is positive. This shows that $\mathscr{P}$ is a Riemannian space with the metric (30).

It can be shown that joining any two points P_1, P_2 of $\mathscr{P}$ there exists one geodesic only. Since P_1 and P_2 can be transformed simultaneously into the unit matrix E and a positive diagonal matrix $D = [d_1, \ldots, d_n]$, it is enough to show this for points E and D. One can see that if for $0 \leq \lambda \leq 1$

$$D^\lambda = \begin{pmatrix} d_1^\lambda & & \\ & \ddots & \\ & & d_n^\lambda \end{pmatrix}$$

93 be defined symbolically then the geodesic line joining E and D consists precisely of these points D^λ.

Consider now the $\mathfrak{H}$ space. It is a subspace of $\mathscr{P}$. The quadratic differential form

$$ds^2 = \sigma(H^{-1}dHH^{-1}dH)$$

defines in $\mathfrak{H}$ a line element invariant under the mappings $H \to H[C]$, $C \in \Omega$. It is practical to take

$$ds^2 = \frac{1}{s}\sigma(H^{-1}dHH^{-1}dH)$$

as the line element. Since this is again positive definite, it follows that $\mathfrak{H}$ is a Riemannian space of pq real dimensions.

One can also express the line element in terms of the parameter X. We obtain

$$ds^2 = \sigma\{(E - XX')^{-1}dX(E - X'X)^{-1}dX'\}.$$

(28) shows that $|E - XX'| = |E - X'X|$ and so we obtain the invariant volume element under this metric as

$$dv = |E - XX'|^{-n/2}\{dX\}$$

where $\{dX\} = \prod_{a=1}^{p}\prod_{b=1}^{q} dx_{ab}$, $X = (x_{ab})$, is the Euclidean volume element in the pq dimensional X space.

As before one can construct geodesics joining two points H_1 and H_2 in $\mathfrak{H}$. As $\mathfrak{H}$ is a subspace of $\mathscr{P}$ and since the metric on $\mathfrak{H}$ is the induced metric from $\mathscr{P}$ one can construct a geodesic joining the two points H_1 and H_2 considered as points of $\mathscr{P}$. It is interesting to note that all points of the geodesic lie in $\mathfrak{H}$ showing that $\mathfrak{H}$ is a geodesic submanifold of $\mathscr{P}$. **94** See [7].

The spaces $\mathscr{P}$ and $\mathfrak{H}$ come under a remarkable class of spaces studied by E. Cartan. They are the *'symmetric spaces'*. According to E. Cartan a topological space T is said to be symmetric about a point P, if there exists an analytic automorphism σ of T onto itself which has P as the only fixed point and whose square is the identity automorphism. T is said to be a symmetric space, if it is symmetric with regard to every point of T. If the space T is homogeneous, that is if on T there acts a transitive group Ω of analytic automorphisms, then symmetry with regard to one point implies that T is a symmetric space.

Let us now consider the space $\mathscr{P}$ and let $P_0 \in \mathscr{P}$. Let P be any matrix in $\mathscr{P}$. Define

$$P^* = P_0P^{-1}P_0.$$

Then $P^* \in \mathscr{P}$ and $P \to P^*$ is clearly an analytic homeomorphism of $\mathscr{P}$ into itself. Also

$$P^{**} = P_0P^{*-1}P_0 = P_0(P_0^{-1}PP_0^{-1})P_0 = P.$$

By the remark in §2, the only solution $P > 0$ of $P = P_0P^{-1}P_0$ is P_0 itself. Thus $\mathscr{P}$ is symmetric about P_0. P_0 being arbitrary, it shows that $\mathscr{P}$ is a symmetric space.

Consider now the space $\mathscr{M}$ and let H_0 be a point in it. For any H in $\mathfrak{H}$ define

$$H^* = H_0H^{-1}H_0.$$

95 Then H^* is also in $\mathfrak{H}$ because

$$H^*S^{-1}H^* = H_0H^{-1}H_0S^{-1}H_0H^{-1}H_0 = S$$

since H and H_0 are in $\mathfrak{H}$. Thus $H \to H^*$ is an analytic automorphism of $\mathfrak{H}$ onto itself and the previous considerations show that $\mathfrak{H}$ is a symmetric space.

The 'symmetrie' $H \to H^*$ is isometric. For,

$$dH^* = d(H_0H^{-1}H_0) = -H_0H^{-1}dH \cdot H^{-1}H_0,$$

as can be seen from differentiating the equation $HH^{-1} = E$. Therefore $dH^* \cdot H^{*-1} = -H_0H^{-1}dH \cdot H^{-1}H_0 \cdot H_0^{-1}HH_0^{-1} = -H_0H^{-1}dHH_0^{-1}$. Hence

$$\sigma(dH^*H^{*-1}dH^*H^{*-1}) = \sigma(dH \cdot H^{-1} \cdot dHH^{-1})$$

which proves our contention.

4 Reduction of indefinite quadratic forms

Let S be a real, non-singular, symmetric matrix of n-rows. If it has signature p, q with $pq > 0$, then there is associated with it a space $\mathfrak{H}$ of positive matrices H satisfying

$$HS^{-1}H = S.$$

$\mathfrak{H}$ has the dimension $pq > 0$.

We now say that S is *reduced* if the $\mathfrak{H}$ space of S has a non-empty intersection with the Minkowski reduced space $\mathscr{R}$. Note that this has a meaning since $\mathfrak{H}$ is a subspace of $\mathscr{P}$. This means that there is *at least* one H in the $\mathfrak{H}$ space of S which is reduced in the sense of Minkowski. 96 Obviously if S is definite, then $\mathfrak{H}$ reduces to the single point $\pm S$ and this definition coincides with that of Minkowski's for definite forms.

We recall that the class of S is defined to be the set of all matrices $S[U]$ where U runs through all unimodular matrices. Our definition shows that for any $S = S'$, there exists an element in the class of S which is reduced. For, let H be any element in the $\mathfrak{H}$ space of S. By the Minkowski theory, there exists a unimodular U such that $H[U] \in \mathscr{R}$. Put $H[U] = H_1$ and $S[U] = S_1$. Then

$$H_1 S_1^{-1} H_1 = S_1$$

which shows that the $\mathfrak{H}$ space of S_1 intersects $\mathscr{R}$ in a non-empty set. Also $S[U]$ is in the class of S.

In general, there will be an infinity of reduced matrices in the class of S. We, however, have the following

Theorem 3. *There exist only finitely many integral symmetric reduced matrices with given determinant.*

Proof. If S is definite, the theorem is already proved in the last chapter. So let S be indefinite. Let S be reduced in the above sense. Let H be in the $\mathfrak{H}$ space of S which is reduced in the sense of Minkowski. By the Jacobi transformation

$$H = D[V]$$

where

$$D = [d_1, \ldots, d_n], \quad V = (v_{kl})$$

where $v_{kk} = 1$, $v_{kl} = 0$ if $k > l$. From Minkowski's reduction theory, it follows that there is a constant c depending only on n such that 97

$$\begin{aligned} 0 < d_k < c d_{k+1}, k &= 1, \ldots, n-1 \\ -c < v_{kl} < c, \ k &\leq 1 \leq n. \end{aligned}$$

Introduce the matrix W given by

$$W \equiv (w_{kl}) = \begin{pmatrix} 0 & . & . & . & . & 0 & 1 \\ 0 & . & . & . & . & . & 0 \\ . & . & . & . & . & . & 0 \\ 1 & . & . & . & . & . & 0 \end{pmatrix} \tag{31}$$

so that $w_{kl} = 0$ if $k + l \neq n + 1$ and equal to 1 otherwise. It then follows that

$$W^2 = E.$$

Put $D_1 = D^{-1}[W] = [d'_1, \dots, d'_n]$. Then

$$d_{n-k+1} d'_k = 1, \quad k = 1, \dots, n,$$

so that

$$0 < d'_k < c\, d'_{k+1}, \quad k = 1, \dots, n-1. \tag{32}$$

□

Let $V_1 = WV'^{-1}_W$. Then because of the choice of W, $V_1 = (v'_{kl})$ is again a triangle matrix. Because the elements of V satisfy the above conditions and W is a constant matrix, we see that there is a constant c_1, depending only on n such that

$$-c_1 < v'_{kl} < c_1, \quad 1 \leq k \leq l \leq n. \tag{33}$$

If we put $c_0 = \max(c, c_1)$ then the matrix $H_1 = D_1[V_1]$ satisfies the condition

$$H_1 \in \mathscr{R}^{**}_{c_0} \tag{34}$$

98 with the notation of the previous chapter. But then

$$H_1 = D_1[V_1] = D^{-1}[WV_1] = D^{-1}[W^2 V'^{-1}_W] = H^{-1}[W].$$

Since $HS^{-1}H = S$, we see that if $WS = S_1$, then

$$H = SH^{-1}S = S'_1 H_1 S_1 = H_1[S_1].$$

H and H_1 are both in $\mathscr{R}^{**}_{c_0}$ and so by theorem 5 of the previous chapter, there are only finitely many S_1 integral and with determinant equal to $|S|$ satisfying the above condition. This proves the theorem.

Since in each class of matrices there is at least one reduced matrix we have

Corollary 1. *There exist only finitely many classes of integral matrices with given determinant.*

If S is rational, then for a certain integer a, aS in integral. Hence from Theorem 3 we get

Corollary 2. *In each class of rational matrices, there exist only finitely many reduced matrices.*

Let S be a rational non-singular symmetric matrix and let S be the matrix of an indefinite quadratic form. Let Ω be the orthogonal group of S, that is the group of real matrices C with $S[C] = S$. A unimodular matrix U satisfying the condition

$$S[U] = S$$

is said to be a *unit* of S. The units of S clearly form a group $\Gamma(S)$ called the *unit group* of S.

Let us consider the $\mathfrak{H}$ space of positive matrices H with $HS^{-1}H = S$. 99 The orthogonal group Ω of S has, in $\mathfrak{H}$, a representation as a transitive group of mappings

$$H \to H[C] \quad C \in \Omega.$$

Since $\Gamma(S)$ is a subgroup of Ω, $\Gamma(S)$ has a representation in the $\mathfrak{H}$ space. Since the unimodular group Γ is discontinuous in $\mathscr{P}$, the representation of $\Gamma(S)$ in $\mathfrak{H}$ is discontinuous. Clearly U and $-U$ lead to the same representation. Therefore if we identify $-U$ and U in $\Gamma(S)$, the $H \to H[U]$, $U \in \Gamma(S)$ gives a faithful and discontinuous representation of $\Gamma(S)$ in $\mathfrak{H}$. Thus $\Gamma(S)$ is a discrete subgroup of Ω. $\mathfrak{H}$ will be the space of smallest dimension in which $\Gamma(S)$ is discontinuous. We shall construct for $\Gamma(S)$ in $\mathfrak{H}$ a fundamental domain F.

In the class of the rational non-singular symmetric indefinite S there exist finitely many reduced matrices, say $S_1, \ldots, S_l$. Let $U_1, \ldots, U_l$ be unimodular matrices so that $S_i = S[U_i]$, $i = 1, \ldots, l$.

Let H be any matrix in the $\mathfrak{H}$ space of S. There exists then a unimodular matrix U so that $H[U] \in \mathscr{R}$. By definition of reduction for indefinite matrices $S[U]$ is reduced. Thus $S[U]$ has to be one of the finitely many $S_1, \ldots, S_l$, say S_k. Then

$$S[U] = S_k = S[U_k],$$

or $S[UU_k^{-1}] = S$ which means that $UU_k^{-1} \in \Gamma(S)$.

Let us denote by $\mathscr{R}_{U_k^{-1}}$ the set of matrices $P[U_k^{-1}]$ for $P \in \mathscr{R}$. Then 100 for the H above $H[U] \in \mathscr{R}$ or $H[UU_k^{-1}] \in \mathscr{R}_{U_k^{-1}}$. Since UU_k^{-1} is a unit of S and $H \in \mathfrak{H}$, it follows that if $V = UU_k^{-1}$ then

$$H[V] \in \mathfrak{H} \cap \mathscr{R}_{U_k^{-1}}.$$

If we therefore take F as the set

$$F = \bigcup_{k=1}^{l} (\mathfrak{H} \cap R_{U_k^{-1}}) \tag{35}$$

then for every point H in $\mathfrak{H}$ there is a unit V such that $H[V] \in F$. If for any unit A of S we put F_A as the image of F under A we have proved that

$$\mathfrak{H} \subset \sum_A F_A.$$

Note that $F_A = F_{-A}$. We shall sketch a proof that F is indeed a fundamental region for $\Gamma(S)$ in $\mathfrak{H}$.

Let U and V be two units of S such that $UV^{-1} \neq \pm E$ and F_U and F_V have a non-empty intersection. Then F and $F_{UV^{-1}}$ have a non-empty intersection. We call $F_{UV^{-1}}$ a *neighbour* of F. Let therefore F_A be a neighbour of F so that $A \neq \pm E$ is a unit of S and F_A intersects F in a non-empty set. Because of the definition (35) of F we see that $\bigcup_{k=1}^{l} (\mathfrak{H} \cap \mathscr{R}_{U_k^{-1}A})$ and $\bigcup_{k=1}^{l} (\mathfrak{H} \cap \mathscr{R}_{U_k^{-1}})$ have a non-empty intersection. (Note that $\mathfrak{H} = \mathfrak{H}[A]$). This means that for two integers i, j, $1 \leq i, j \leq l$, $\mathscr{R}_{U_i^{-1}A}$ and $\mathscr{R}_{U_j^{-1}}$ have a non-empty intersection or

$$\mathscr{R} \cap \mathscr{R} U_i^{-1} A U_j$$

101 is not empty. Now $U_1^{-1} A U_j \neq \pm E$. For, this means that $AU_j = \pm U_i$ which is the same thing as $i = j$ or $A = \pm E$. Then $U_i^{-1} A U_j$ (by Minkowski's theory) belongs to a finite set of matrices depending only on n. Hence F has a finite number of neighbours.

In order to study the points of intersection of F and F_A where F_A is a neighbour of F, we remark that since $\mathscr{R}$ and $\mathscr{R}_V$, for a unimodular

V, have only boundary points in common, it is enough to show that the boundary points of F relative to $\mathfrak{H}$ are the intersection of the boundaries of $\mathscr{R}_{U_i}$ with $\mathfrak{H}$. This would be achieved if we prove that $\mathfrak{H}$ does not lie on a boundary plane of $\mathscr{R}_{U_i}$ for any $\underline{i}$. For a proof of this non-trivial fact we refer to [8].

In the $\mathfrak{H}$ space we have seen that there is a volume element dv invariant under the mappings $H \to H[C]$, $C \in \Omega$. In the next chapter we shall prove that

$$\int\limits_F dv \tag{36}$$

if finite except in the case of binary zero form. This will show incidentally that $\Gamma(S)$ is an infinite group if S is not the matrix of a binary zero form. For, one can show, rather easily, that $\int\limits_{\mathfrak{H}} dv$ is infinite.

5 Binary forms

We shall now study the case of binary quadratic forms systematically.

Let $ax^2 + 2bxy + cy^2$ be a real binary quadratic from whose matrix 102

$$S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

is non-singular and let $a \neq 0$. Write

$$S[\underline{x}] = ax^2 + 2bxy + cy^2 = a(x - \tau_1 y)(x - \tau_2 y)$$

where $\tau_1 + \tau_2 = -\dfrac{2b}{a}$, $\tau_1\tau_2 = c/a$. Thus τ_1 and τ_2 are roots of the equation

$$a\lambda^2 + 2b\lambda + c = 0. \tag{37}$$

If $|S| = ac - b^2 > 0$, then τ_1 and τ_2 are both complex. If $|S| < 0$, then τ_1 and τ_2 are real.

Let

$$L = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

be the matrix of the linear transformation $x \to \alpha x + \beta y$, $y \to \gamma x + \delta y$. Then

$$S[L\underline{x}] = a'(x - \tau_1' y)(x - \tau_2' y)$$

where $a' = a(\alpha - \tau_1\gamma)(\alpha - \tau_2\gamma)$. We shall assume that $a' \neq 0$. We have then

$$\tau_i' = \frac{\delta\tau_i - \beta}{-\gamma\tau_i + \alpha} i = 1, 2. \tag{38}$$

This shows that the transformation $S \to S[L]$ results in transformation (38) in the roots of the equation (37), the matrix of the transformation for the roots being $|L|L^{-1}$.

103 We shall now consider the case $ae - b^2 < 0$ so that $S[\underline{x}]$ is an indefinite binary form. Our object is to study the $\mathfrak{H}$ space of S. This is the set of two-rowed symmetric matrices H satisfying

$$H > 0, \quad HS^{-1}H = S.$$

If C is a non-singular matrix such that $S[C^{-1}] = S_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, then by our considerations before, H can be obtained from the H_0 satisfying

$$H_0 S_0^{-1} H_0 = S_0$$

by taking $H = H_0[C]$. This space $\mathfrak{H}_0$ of H_0 has the parametrical representation

$$H_0 = \begin{pmatrix} \dfrac{1+y^2}{1-x^2} & \dfrac{2x}{1-x^2} \\ \dfrac{2x}{1-x^2} & \dfrac{1+x^2}{1-x^2} \end{pmatrix}, \quad |x| < 1.$$

We put

$$H_0 = \begin{pmatrix} p & q \\ q & p \end{pmatrix};$$

then $|H_0| = p^2 - q^2 = 1$.

In the second chapter we had associated uniquely with every positive matrix of determinant 1 a point z in the upper half of the complex z-plane. If z is the representative point for H_0, then $z = \xi + i\eta$, $\eta > 0$ and

$$z = \frac{-q}{p} + \frac{i}{p}$$

104 which, in terms of the parameter x is

$$z = \frac{1}{i}\frac{x-i}{x+i} \tag{39}$$

(39) shows that z lies on the semi-circle in the upper half plane with unit radium around the origin as centre. Since the linear transformation (39) takes the points $x = -1, 0, 1$ into the points $z = 1, i, -1$ it follows that as x runs through all x in the interval $(-1, 1)$, z traces the above semi-circle of unit radius. We may take this semi-circle as the $\mathfrak{H}_0$ space.

From our general results we see that the line element is

$$ds = \frac{dx}{1-x^2} \tag{40}$$

so that the distance between two points z_1, z_2, with values of the parameters x_1, x_2 respectively, is

$$\delta(z_1, z_2) = \int\limits_{x_1}^{x_2} ds = \frac{1}{2}\log\left(\frac{1+x_2}{1-x_2} : \frac{1+x_1}{1-x_1}\right) \tag{41}$$

From (39), z determines x uniquely, namely

$$x = \frac{1}{i}\frac{z-i}{z+i} \tag{42}$$

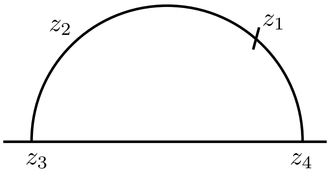
so that

$$dx = \frac{2dz}{(z+i)^2}$$

and hence the distance $\delta(z_1, z_2)$ is also

$$\delta(z_1, z_2) = \frac{1}{4}\log\left(\frac{z_2+1}{z_2-1} : \frac{z_1+1}{z_1-1}\right) \tag{43}$$

In Poincare's model of non-Euclidean geometry, See [11], the distance $\delta_0(z_1, z_2)$ between two points z_1, z_2 in the upper half plane is given by 105



$$\delta_0(z_1, z_2) = \log[D(z_1, z_2, z_3, z_4)]$$

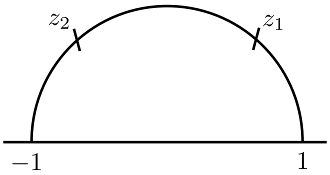
where z_3 and z_4 are the points in which the unique circle through z_1 and z_2 orthogonal to the real line, intersects the real line: z_1, z_2, z_3, z_4 being in cyclical order and $D(z_1, z_2, z_3, z_4)$ is the cross ratio

$$D(z_1, z_2, z_3, z_4) = \frac{z_1 - z_3}{z_1 - z_4} : \frac{z_2 - z_3}{z_2 - z_4}.$$

With this notation we see that (43) can be written as

$$\delta(z_1, z_2) = \frac{1}{2} \log[D(z_1, z_2, -1, 1)] \tag{45}$$

Let us now go back to the $\mathfrak{H}$ space. This consists of points $H = H_0[C]$. We shall assume $|C| = 1$. For if not, we interchange the columns of C which merely means taking $-S$ instead of S. By (38) it follows that



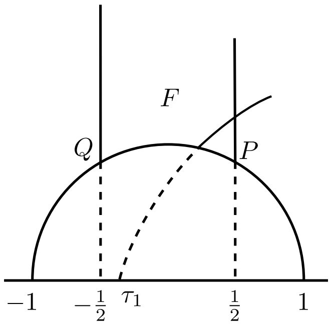
the $\mathfrak{H}$ space is precisely the semi-circle on τ_1, τ_2 as diameter, where τ_1, τ_2 are the roots of (37). The equation of this semi-circle is, as can be seen

$$a(\xi^2 + \eta^2) + 2b\xi + c = 0 \tag{46}$$

with centre on the real line at the point $-\dfrac{b}{a}$ and radius $\sqrt{\dfrac{\|S\|}{a}}$.

106 In the previous chapter we had seen that the modular region F de-

fined by



$$z\bar{z} \geq 1$$
$$-\frac{1}{2} \leq \xi \leq \frac{1}{2}$$

is a fundamental region for the proper unimodular group. Analogous to our definition of reduction of indefinite form in the last section, we define a binary form $S[\underline{x}]$ *reduced* if its $\mathfrak{H}$ space intersects F in a non-empty set. Since transformations in the upper half plane are by means of matrices of determinant unity, this can be called *proper reduction*. Since the $\mathfrak{H}$ space is given by (46), the fact that S is reduced means that at least one of the vertices P, Q lies within the circle (46). Since P, Q are the points $\dfrac{\pm 1 + i\sqrt{3}}{2}$ we see that if S is reduced, then

$$a\left(\xi^2 + \eta^2\right) + 2b\xi + c \leq 0$$

where $\xi = \pm\frac{1}{2}$ and $\eta = \dfrac{\sqrt{3}}{2}$. This gives

$$a \pm b + c \leq 0. \tag{47}$$

We may assume $a > 0$. For, if $a < 0$ (we have already assumed it is not equal to zero) there exists a properly unimodular matrix U such that the first diagonal element of $S[U]$ is positive. In this case, we have since $a + c \leq |b|$ we get

$$\frac{1}{4}(4a - 2|b|)^2 + 3b^2 = (2|b| - a)^2 + 3a^2 = 4d + 4(a^2 - a|b| + ac) \leq 4d$$

107 where $d = \|S\|$. This gives at once

$$a^2 \leq \frac{4}{3}d, \quad b^2 \leq \frac{4}{3}d, \quad 3ac \leq d$$

which at once shows that the number of integral properly reduced forms of given determinant is finite.

It is to be *noted* that the reduction conditions above are not the same as those of Gauss.

Let us now construct a fundamental region for the unit group in this $\mathfrak{H}$ space. Before doing this, we shall first study the structure of the unit group $\Gamma = \Gamma(S)$ of S.

Let $S[\underline{x}] = ax^2 + bxy + cy^2$ be a form representing integers for integral values of x and y. Then

$$S = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

is a semi-integral matrix. Let

$$b^2 - 4ac = d.$$

$d > 0$ and not a square. Then $S[\underline{x}]$ is not a zero form. Also let $(a, b, c) = 1$. Since $S[\underline{x}]$ is not a zero form, neither $\underline{a}$ nor c is zero. We shall consider the proper group of units U of S, namely the group $\Gamma_0 = \Gamma_0(S)$ of unimodular matrices U such that

$$S[U] = S, \quad |U| = 1. \tag{48}$$

Clearly Γ_0 is a subgroup of Γ of index 2. Let $U = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ be an element of Γ_0. Let τ_1, τ_2 be the roots of

$$a\lambda^2 + b\lambda + c = 0. \tag{49}$$

108 Then it can be seen that the mapping $S \to S[U]$ keeps τ_1 and τ_2 fixed. Thus

$$\tau_i = \frac{\delta\tau_i - \beta}{-\gamma\tau_i + \alpha}, i = 1, 2.$$

which means that τ_1 and τ_2 are again roots of the polynomial

$$\gamma\lambda^2 + (\delta - \alpha)\lambda - \beta = 0 \tag{50}$$

(49) is irreducible in the rational number field since otherwise τ_1 and τ_2 will both be rational and so $\sqrt{\|S\|}$ is rational, which is a contradiction to our assumption that $S[\underline{x}]$ is not a zero form. Therefore (49) and (50) give

$$\frac{\gamma}{a} = \frac{\delta - \alpha}{b} = \frac{-\beta}{c} = q \tag{51}$$

where q is an integer. If $U \neq \pm E$ then $\gamma \neq 0, \beta \neq 0$. Let us put $\delta + \alpha = p$. Then

$$\delta = \frac{p + bq}{2}, \quad \alpha = \frac{p - bq}{2} \tag{52}$$

Since $\alpha\delta - \beta\gamma = 1$, we get the relation $\dfrac{p^2 - q^2b^2}{4} + q^2ac = 1$ or

$$\boxed{p^2 - dq^2 = 4} \tag{53}$$

which is the well-known Pell's equation. Thus every unit of S in Γ_0 gives rise to a solution of Pell's equation. Conversely every solution of (53) gives rise to a unit U of S, as can be seen from (51) and (52). This 109
unit is

$$U = \begin{pmatrix} \dfrac{p - bq}{2} & -cq \\ aq & \dfrac{p + bq}{2} \end{pmatrix}$$

Consider the representation $H \to H[U]$, $U \in \Gamma_0$ in the $\mathfrak{H}$ space. Let the representative point of H_0 be z_0 on the semi-circle which defines the $\mathfrak{H}$ space. Denote by w_0 the point $H_0[U]$. Let z be any variable point on this semi-circle and w the image by the transformation $H \to H[U]$. If $U = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \beta \end{smallmatrix}\right)$ then

$$w_0 = \frac{\delta z_0 - \beta}{-\gamma z_0 + \alpha}, \quad W = \frac{\delta z - \beta}{-\gamma z + \alpha}.$$

Because of the fact that U fixes both τ_1 and τ_2 and cross-ratio is unaltered by linear transformation, we see that

$$\frac{z-\tau_1}{z-\tau_2} : \frac{z_0-\tau_1}{z_0-\tau_2} = \frac{w-\tau_1}{w-\tau_2} : \frac{w_0-\tau_1}{w_0-\tau_2}$$

or that

$$\frac{w-\tau_2}{w-\tau_1} = \mu\frac{z-\tau_2}{z-\tau_1} \tag{54}$$

where $\mu = D(w_0, z_0, \tau_1, \tau_2)$. But equation (54) shows that $\mu = D(w, z, \tau_1, \tau_2)$, which means that μ is a constant independent of the point z. This shows that in the non-Euclidean geometry of ours, the mapping $H \to H[U]$ corresponds to a translation. Also (See [11]) the quantity μ
110 has the property that if λ_1 and λ_2 are the eigen values of the matrix of the transformation

$$w = \frac{\delta z - \beta}{-\gamma z + \alpha} \tag{55}$$

then, by proper ordering of λ_1 and λ_2 we have

$$\mu = \frac{\lambda_1}{\lambda_2}$$

and therefore the non-Euclidean distance $\delta_0(z, w)$ is given by

$$\delta_0(z, w) = \log\frac{\lambda_1}{\lambda_2}$$

where $\lambda_1 \geq \lambda_2$. Now λ_1 and λ_2 are characteristic roots of the mapping (55) and hence they satisfy

$$\lambda^2 - (\alpha + \delta)\lambda + 1 = 0$$

which shows that

$$\lambda_1, \lambda_2 = \frac{\alpha+\delta}{2} \pm \sqrt{\left(\frac{\alpha+\delta}{2}\right)^2 - 1.}$$

Substituting from (52), $\alpha + \delta = p$, we get

$$\lambda_1, \lambda_2 = \frac{p \pm q\sqrt{d}}{2} \tag{56}$$

where p and q are the unique solutions of Pell's equation corresponding to the unit U.

Let R be the field of rational numbers and $R(\sqrt{d})$ the real quadratic field. The element ε of $R(\sqrt{d})$ defined by

$$\varepsilon = \frac{p + q\sqrt{d}}{2}$$

where p and q are a solution of Pell's equation, is a unit of norm 1. This 111 is seen from the fact that ε is a root of

$$\lambda^2 - p\lambda + 1 = 0.$$

(If d is square-free the converse is also true).

In (56), the quantities λ_1, λ_2 are ε and ε^{-1} in some order. If U is changed to U^{-1} or to $-U$, then ε gets changed to ε^{-1} or $-\varepsilon$ respectively. We therefore choose among the four quantities ε, $-\varepsilon$, ε^{-1}, $-\varepsilon^{-1}$, one (and there is only one in general), call it ε^* which is such that

$$\varepsilon^* \geq 1$$

and put

$$\delta_0(z, w) = \log \varepsilon^{*2}.$$

(Note that $\lambda_1/\lambda_2 = \varepsilon^{*2}$ with $\lambda_1 \geq \lambda_2$). This will then mean that the translation in the $\mathfrak{H}$ space is by the amount

$$\delta(z, w) = \log \varepsilon^*.$$

Since the representation of Γ_0 in $\mathfrak{H}$ is discontinuous it follows that the translations form a discrete subgroup in the group of non-Euclidean motions on $\mathfrak{H}$. There is thus a U_0 and a corresponding ε_0^* such that $\log \varepsilon_0^*$ is the smallest. Hence any U will be of the form

$$U = \pm U_0^n \quad (n = 0, \pm 1, \ldots)$$

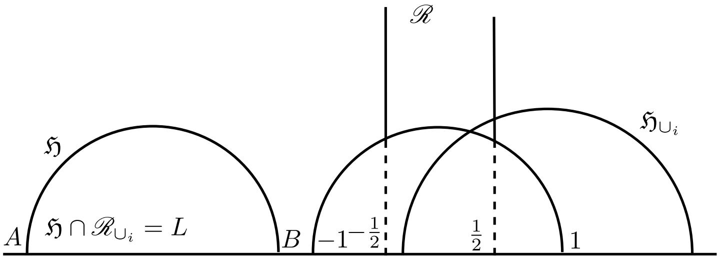
Similarly any ε which arises from a U in Γ_0 is of the form

$$\varepsilon = \pm \varepsilon_0^{*n} \quad (n = 0, \pm 1, \ldots)$$

112 If instead of S being semi-integral, it was rational symmetric we could, by multiplying it by an integer, make it integral satisfying the condition $(a, b, c) = 1$. But this multiplication does not affect the unit group. Hence the

Theorem 4. *The group of proper units of an indefinite, binary, non-zero, rational quadratic form is an infinite cyclic group whose elements stand in a*$(1, 1)$ *correspondence with the solutions of Pell's equation.*

Let $S = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ be a rational symmetric, non-singular indefinite matrix and let $S[\underline{x}]$ be not a zero form. Let the



semi-circle ALB denote the $\mathfrak{H}$ space of the matrix S so that A and B are points on the real axis with coordinates τ_2 and τ_1, $\tau_2 \neq \tau_1$. Furthermore since $S[\underline{x}]$ is not a zero form, the quantities τ_1 and τ_2 are irrational. Let $\mathscr{R}$ denote the fundamental region of the proper unimodular group in the upper half z-plane. Let $U_1, \ldots, U_g$ be the finitely many reducing 113 properly unimodular matrices. If $\mathfrak{H}_U$ for $U = U_1, \ldots, U_g$ denotes the image of $\mathfrak{H}$ under the transform $H \to H[U]$ (this means for the points z on $\mathfrak{H}$ the transformation $z \to \dfrac{\varepsilon_z - \beta}{-\gamma z + \alpha}$, $U = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$), then

$$G = \sum_{i=1}^{g} (\mathfrak{H}_{U_i} \cap \mathscr{R})_{U_i^{-1}}$$

is a point set on $\mathfrak{H}$ and is a fundamental region for Γ_0 in $\mathfrak{H}$. It is important to notice that $\mathfrak{H}_{U_i}$ does not lie completely on the boundary of $\mathscr{R}$. For then $\mathfrak{H}_{U_i}$ would have to be the unit circle or one of the lines $\xi = \pm 1/2$.

In the first case this would mean that

$$\frac{\delta z - \beta}{-\gamma z + \alpha} = \pm 1$$

where $z = \tau_1$ or τ_2. This is the same as saying τ_1 and τ_2 are rational, which they are not. Then same happens if the second case is true.

This shows that none of the arcs $(\mathfrak{H}_{U_i} \cap \mathscr{R})_{U_i}$ has an end point at τ_1 or τ_2. Hence G is compact. Its volume therefore in the measure induced by the invariant metric is finite.

Bibliography

[1] E. Cartan : Sur une classe remarquable d' espaces de Riemann, *Bull. Math. Soc. France*, Tome 55 (1927) P. 114 - 134.

[2] Fricke-Klein : *Vorlesungen über die theorie der Modulfunktionen*, Bd.1, Leipsig (1890), P.243-269. **114**

[3] C. F. Gauss : *Disquisitiones Arithmeticae.*

[4] C. Hermite : *Oeuvres*, Tome 1, Paris (1905), P.164-293.

[5] L. Pontrjagin : *Topological groups*, Princeton (1940).

[6] C. L. Siegel : Discontinuous groups, *Annals of Math.*, 44 (1943), P. 674-689.

[7] C. L. Siegel : Some remarks on discontinuous groups, *Annals of Math.*, 46(1945), P.708-718.

[8] C. L. Siegel : Einheiten quadratischer Formen, *Abh. aus. dem. Math. Semi. Hansischen Univ.*, 13(1940), P.209-239.

[9] C. L. Siegel : Indefinite quadratische Formen und Funktionentheorie I, *Math. Annalen*, 124 (1951), P.17-54.

[10] C. L. Siegel : The average measure of quadratic forms with given determinant and signature, *Annals of Math.*, 45(1944), P.667-685.

[11] C. L. Siegel : *Ausgewählte Fragen der Funktionentheorie II* Göttingen, 1953.

[12] A. Weil : *L'intégration dans les groupes topologiques et ses applications*, Paris (1940).

Chapter 4

Analytic theory of Indefinite quadratic forms

1 The theta series

115 Let

$$f(x_1, \ldots, x_n) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + b_1x_1 + \cdots + b_nx_n + c = 0 \qquad (1)$$

be a Diophantine equation with integral coefficients. Let S denote the matrix of the homogeneous part. We consider integral linear homogeneous transformations

$$x_i \to \sum_{j=1}^{n} q_{ij}x_j, \quad i = 1, \ldots, n$$

where the matrix $Q = (q_{ij})$ is unimodular. Let the resulting function be $Q(x_1, \ldots, x_n)$. Then $f = 0$ has an integral solution if and only if $Q = 0$ has an integral solution.

Suppose the matrix S has rank r, $0 < r < n$ so that $|S| = 0$. Let $\underline{p}$ be a primitive vector such that $S\underline{p} = \underline{0}$. Let U be the unimodular matrix with $\underline{p}$ as its first column. Then

$$S[U] = \begin{pmatrix} 0 & \underline{0}' \\ \underline{0} & S_1 \end{pmatrix}.$$

We may repeat the process with S_1 instead of S. Finally therefore we arrive at a unimodular matrix V so that

$$S[V] = \begin{pmatrix} 0 & 0 \\ 0 & S_r \end{pmatrix}$$

$|S_r| \neq 0$. Put now

$$(b_1, b_2, \ldots, b_n)V = (c_1, \ldots, c_n).$$

116 If $c_1, \ldots, c_{n-r}$ are zero it means that by a unimodular transformation we can bring f into a quadratic form in r-variables. Suppose now that $c_1, \ldots, c_{n-r}$ are not all zero. Since they are integers, there exists a unimodular matrix V_1 of $n-r$ rows such that

$$(c_1, \ldots, c_{n-r})V_1 = (0, 0, \ldots, d).$$

Put now

$$V_2 = \begin{pmatrix} V_1 & 0 \\ 0 & E_r \end{pmatrix}.$$

Then $S[VV_2] = S[V]$ and $f(x_1, \ldots, x_n)$ becomes transformed into

$$\begin{aligned} \varphi(x_{n-r}, \ldots, x_n) = d_{11}x_{n-r+1}^2 + \cdots d_{rr}x_n^2 \\ + dx_{n-r} + d_1x_{n-r+1} + \cdots + d_rx_n + d'. \end{aligned}$$

This is the form into which f can, in general, be transformed by unimodular transformation.

We shall hereafter assume $|S| \neq 0$ and that the quadratic form is integral valued, that is, that for $x_1, \ldots, x_n$ integral, $f(x_1, \ldots, x_n)$ is an integer. This means

$$\alpha) \ S \text{ is semi-integral} \tag{2}$$

that is that its diagonal elements are integers and twice the non-diagonal elements are integers. (1) can now be written in the form

$$S[\underline{x}] + \underline{b}'\underline{x} + c = S[\underline{x} + \frac{1}{2}S^{-1}\underline{b}] + c - \frac{1}{4}S^{-1}[\underline{b}].$$

If we put $t = c - \frac{1}{4}S^{-1}[\underline{b}]$ and $2S\ \underline{a} = \underline{b}$, then (1) takes the simple form

$$S[\underline{x} + \underline{a}] - t = 0. \tag{3}$$

117 Obviously

$$\beta)\ 2S\,\underline{a} \text{ is integral.} \tag{4}$$

We shall therefore consider the diophantine problem in which the left side is $S\,[\underline{x}+\underline{a}]$. Clearly and rational number t which can be represented by $S\,[\underline{x}+\underline{a}]$ satisfies

$$t \equiv S\,[\underline{a}] (\text{mod } 1). \tag{5}$$

Consider the diophantine equation $S\,[\underline{x}+\underline{a}] = t$ under the conditions (2) and (4). If $S > 0$, the number of integral solutions, denoted $A(S, \underline{a}, t)$, is finite. We now form the *generating series*

$$\sum_t A(S, \underline{a}, t) e^{2\pi i t z} \tag{6}$$

where $z = \xi + i\eta$, $\eta > 0$. It follows that

$$\sum_t A(S, \underline{a}, t) e^{2\pi i t z} = \sum_{\underline{x} \equiv -\infty}^{\infty} e^{2\pi i S\,[x+a] z} \tag{7}$$

and since $S > 0$, the series on the right of (7) converges. The right side of (7) is a so-called *theta series* studied extensively by *Jacobi*. In particular if $S = E_4$, the unit matrix of order 4 and $\underline{a} = \underline{0}$, we have

$$\sum_{t=0}^{\infty} A_4(t) e^{2\pi i t z} = \left(\sum_{x=-\infty}^{\infty} e^{2\pi i x^2 z} \right)^4$$

where $A_4(t)$ is the number of integral solutions of

$$t = x_1^2 + x_2^2 + x_3^2 + x_4^2. \tag{8}$$

It was conjectured by *Fermat* and proved by *Lagrange* that for every 118 $t \geq 1$, (8) has a solution. Jacobi proved, by using the theory of Elliptic theta series, that

$$A_4(t) = 8 \sum_{\substack{d/t \\ 4 \nmid d}} d.$$

Since for every t, unity divider t, we find that $A_4(t) \geq 1$, for all $t \geq 1$. This, besides proving Lagrange's theorem, is a quantitative improvement of it.

If S is indefinite, $A(S, \underline{a}, t)$, if different from zero, is infinite. This means that the right side of (7) diverges. It is therefore desirable to define an analogue of $A(S, \underline{a}, t)$ for indefinite S. To this end we shall introduce the theta series.

Let $z = \xi + i\eta$, $\eta > 0$ be a complex parameter. Let S be n-rowed symmetric and H any matrix in the representation space $\mathfrak{H}$ of the orthogonal group of S. Then $H > 0$ and $HS^{-1}H = S$. Put

$$R = \xi S + i\eta H. \tag{9}$$

Then $R = zK + \bar{z}L$ where $K = \frac{1}{2}(S + H)$, $L = \frac{1}{2}(S - H)$. R is now a complex symmetric matrix whose imaginary part ηH is positive definite. Let $\underline{a}$ be a rational column satisfying (4). Put $\underline{y} = \underline{x} + \underline{a}$, $\underline{x}$ being an integral column. Define

$$f_{\underline{a}}(z, H) = \sum_{\underline{y} \equiv a (\text{mod } 1)} e^{2\pi i R[\underline{y}]Z} \tag{10}$$

where the summation runs over all rational columns $\underline{y} \equiv \underline{a} (\text{mod } 1)$. Since $H > 0$, (10) is absolutely convergent for every H in $\mathfrak{H}$.

119 For our purposes, it seems practical to consider a more general function $f_{\underline{a}}(z, H, \underline{w})$ defined by

$$f_{\underline{a}}(z, H, \underline{w}) = \sum_{\underline{y} \equiv \underline{a} (\text{mod } 1)} e^{2\pi i R[\underline{y}] + 2\pi i \underline{y}'\overline{\omega}} \tag{11}$$

where $\underline{w}$ is a complex column of n rows of elements $w_1, \ldots, w_n$. It is clear that $f_{\underline{a}}(z, H, \underline{w})$ are still convergent. (11) is the general theta series.

It is to be noticed that if $S > 0$, (10) coincides with (7).

Consider now all the rational vectors $\underline{a}$ which satisfy (4), namely that $2S\underline{a}$ is integral. It is clear that there are only finitely many such $\underline{a}$ incongruent (mod 1). The number l of such residue classes is clearly at most equal to

$$d = abs2|S|. \tag{12}$$

Let $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_l$ be the complete set of these l residue classes incongruent (mod 1). For each class $\underline{a}_i$ form the function $f_{\underline{a}_i}(z, H, \underline{w})$. We denote by $f(x, H, \underline{w})$ the functional vector

$$f(z, H, \underline{w}) = \begin{pmatrix} f_{\underline{a}_1}(z, H, \underline{w}) \\ \vdots \\ f_{\underline{a}_l}(z, H, \underline{w} \end{pmatrix}. \tag{13}$$

2 Proof of a lemma

Let $P > 0$ be an n-rowed real matrix and $\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ a column of n real numbers. The function

$$f(u_1, \ldots, u_n) = \sum_{\underline{x}} e^{-\pi P[\underline{x}+\underline{u}]}$$

where $\underline{x}$ runs through all integral n-rowed columns, is a continuous function of the n-variables $u_1, \ldots, u_n$ and has in each variable the period 1. It has the fourier series 120

$$\sum_{\underline{l}} c_{\underline{l}} e^{2\pi i \underline{l}'\underline{u}}$$

$\underline{l}$ running through all integral n-rowed vectors and

$$c_{\underline{l}} = \int_{\sigma} f(u_1, \ldots, u_n) e^{-2\pi i \underline{l}'\underline{u}} du_1 \ldots du_n \tag{14}$$

where σ is the unit cube in the n dimensional space R_n of $u_1, \ldots, u_n$.

Since $P > 0$, we can write $P = M'M$ for a non-singular M. Put

$$M\underline{u} = \underline{v}, \quad M'^{-1}\underline{l} = \underline{k}. \tag{15}$$

Then

$$c_{\underline{l}} : |P|^{-\frac{1}{2}} \int_{R_n} e^{-\pi \underline{v}'\underline{v} - 2\pi i \underline{k}'\underline{v}} d\underline{v}$$

where we take the positive value of the square root. If $\underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $\underline{k} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$, then $\underline{v}'\underline{v} = v_1^2 + \cdots + v_n^2$ and $\underline{k}'\underline{v} = k_1 v_1 + \cdots + k_n v_n$ so that

$$c_{\underline{l}} = |P|^{-\frac{1}{2}} e^{-\pi P^{-1}[\underline{l}]} \left(\int\limits_{-\infty}^{\infty} e^{-\pi t^2} dt \right)^n \tag{16}$$

That the value of the integral on the right of (16) is unity is seen as follows: In the first place from the uniform convergence of the series $f(u_1, \ldots, u_n)$, it follows that

$$\sum_{\underline{x}} e^{-\pi P[\underline{x}+\underline{u}]} = |P|^{-\frac{1}{2}} \lambda^n \sum_{\underline{l}} e^{-\pi P^{-1}[\underline{l}] + 2\pi i \underline{l}'\underline{u}} \tag{17}$$

121 where

$$\lambda = \int\limits_{-\infty}^{\infty} e^{-\pi t^2} dt.$$

Secondly, λ is independent of P and so putting $n = 1$, $P = 1$ and $u = 0$, we see from (17) that $\lambda = 1$.

Suppose now that in (17), $u_1, \ldots, u_n$ are complex variables. Since $f(u_1, \ldots, u_n)$ is absolutely convergent which moreover is uniformly convergent in every compact subset of the n-complex space of $u_1, \ldots, u_n$, it follows that $f(u_1, \ldots, u_n)$ is an analytic function of the n-complex variables $u_1, \ldots, u_n$. The same is true of the right side of (17) also. Since (17) holds for all real $u_1, \ldots, u_n$, it holds, by analytic continuation, for complex $u_1, \ldots, u_n$ also.

Suppose now that P is a complex symmetric matrix $P = X + iY$ whose real part X is positive definite. Then P^{-1} also has this property. For, since X and Y are real symmetric and $X > 0$, there exists a non-singular real C such that

$$X = C'C, \quad Y = C'DC$$

where $D = [d_1, \ldots, d_n]$ is a real diagonal matrix with diagonal elements $d_1, \ldots, d_n$. Now $P = X + iY = C'(E + iD)C$ so that

$$P^{-1}[C'] = \left[\frac{1 - id_1}{1 + d_1^2}, \ldots, \frac{1 - id_n}{1 + d_n^2}\right] \tag{18}$$

which shows, since C is real, that the real part of P^{-1} is positive definite symmetric. Incidentally we have shown that P is non-singular.

If we now take $u_1, \ldots, u_n$ to be fixed complex numbers, then $f(P) = \sum\limits_{\underline{x}} e^{-\pi[\underline{x}+\underline{u}]}$ is an analytic function of the $\dfrac{n(n+1)}{2}$ complex variables constituting the matrix P. Since (17) is true for P real, by analytic continuation, it is true also if P is complex symmetric with positive real part. For $|P|^{-\frac{1}{2}}$ one takes that branch of the algebraic function which is positive for real P. We thus have the 122

Lemma 1. *Let P be a complex n-rowed symmetric matrix with real part positive. Let $\underline{u}$ be any complex column. Then*

$$\sum_{\underline{x}} e^{-\pi P[\underline{x}+\underline{u}]} = |P|^{-\frac{1}{2}} \sum_{\underline{x}} e^{-\pi P^{-1}[\underline{x}]+2\pi i \underline{x}'\underline{u}}$$

where $\underline{x}$ runs through all integral columns and $|P|^{-\frac{1}{2}}$ is that branch of the algebraic function which is positive real P.

3 Transformation formulae

We now wish to study the transformation theory of the theta series defined in (11), under the modular substitutions

$$z \to z_M = \frac{\alpha z + \beta}{\gamma z + \delta}, \quad M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad |M| = 1 \tag{19}$$

where M is an integral matrix. Since H will be fixed throughout this section we shall write $f_{\underline{a}}(z, \underline{w})$ instead of $f_{\underline{a}}(z, H, \underline{w})$. Following *Hermite*, we first consider the case $\gamma \neq 0$, and write

$$\frac{\alpha z + \beta}{\gamma z + \delta} = \frac{\alpha}{\gamma} + \gamma^{-2} z_1, \quad -z_1^{-1} = z_2, \quad z_2 = z + \frac{\delta}{\gamma} \tag{20}$$

Clearly 123

$$f_{\underline{a}}(z_M, \underline{w}) = f_{\underline{a}}\left(\frac{\alpha}{\gamma} + \gamma^{-2} z_1, \underline{w}\right). \tag{21}$$

Denote by R_M the matrix R in (9) with z replaced by z_M, then

$$R_M = \frac{\alpha}{\gamma} S + \gamma^{-2} R_1$$

where

$$R_1 = z_1 \frac{S+H}{2} + \bar{z}_1 \frac{S-H}{2} \tag{22}$$

By definition of y, $\underline{y} - \underline{a}$ is integral. We may therefore write $\underline{v} - \underline{a} = \underline{x}\gamma + \underline{g}$ where y belongs to the finite set of residue classes of integral vectors (mod γ). When $\underline{x}$ runs through all integral vectors and $\underline{g}$ through a complete system of (mod γ) incongruent integral vectors, then $\underline{y}$ runs through all rational vectors $\equiv \underline{a}$(mod 1). We have therefore

$$f_{\underline{a}}(z_M, \underline{w}) = \sum_{\underline{g}((\mathrm{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \sum_{\underline{x}} e^{2\pi i(\gamma^{-2} R_1[\underline{x}\gamma+\underline{g}+\underline{a}]+\underline{w}'(\underline{w}\gamma+\underline{g}+\underline{a}))}$$

R_1 being non-singular, we can complete squares in the exponent in the inner sum and obtain,

$$\left.\begin{aligned} f_{\underline{a}}(z_M, \underline{w}) = e^{-\frac{\pi i}{2} R_1^{-1}[\underline{w}\gamma]} \sum_{\underline{g}((\mathrm{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \\ \sum_{\underline{x}} e^{2\pi i R_1[\underline{x}+(\underline{g}+\underline{a})\gamma^{-1}+R_1^{-1}\underline{w}\gamma/2]} \end{aligned}\right\} \tag{23}$$

In order to be able to apply lemma 1 to the inner sum in (23) we first compute R_1^{-1}. Since S and H are symmetric and $H > 0$, there exists a real non-singular C such that

$$S = C' \begin{pmatrix} E_p & 0 \\ 0 & -E_q \end{pmatrix} C, H = C'C.$$

124 Then R_1 is given by

$$R_1 = C' \begin{pmatrix} z_1 E_p & 0 \\ 0 & -\bar{z}_1 E_q \end{pmatrix} C \tag{24}$$

where z_1 is given by (20). It readily follows that

$$-R_1^{-1} = -z_1^{-1}\frac{S^{-1}+H^{-1}}{2} - \overline{z}_1^{-1}\frac{S^{-1}-H^{-1}}{2}. \tag{25}$$

Using (20), we finally have

$$-R_1^{-1} = \left(\frac{\delta}{\gamma}S + R\right)[S^{-1}] \tag{26}$$

Applying lemma 1 to the inner sum we now find

$$\left.\begin{aligned} & f_{\underline{a}}(z_M, \underline{w}) = |-2iR_1|^{-\frac{1}{2}} e^{-\frac{\pi i}{2}R_1^{-1}[w\gamma]} \sum_{\underline{g}(\mathrm{mod}\ \gamma)} e^{2\pi i\frac{\alpha}{\gamma}S[\underline{g}+\underline{a}]} \\ & \sum_{l} e^{\frac{\pi i}{2}\left(\frac{\delta}{\gamma}S+R\right)[S^{-1}\underline{l}]+2\pi i\underline{l}'((\underline{g}+\underline{a})\gamma^{-1}+R_1^{-1}\underline{w}\frac{\gamma}{2})} \end{aligned}\right\} \tag{27}$$

where the square root has to be taken according to the prescription in lemma 1. It follows then, using (24), that

$$|-2iR_1|^{-\frac{1}{2}} = \epsilon d^{-\frac{1}{2}}\left(z+\frac{\delta}{\gamma}\right)^{\frac{p}{2}}\left(\overline{z}+\frac{\delta}{\gamma}\right)^{\frac{q}{2}} \tag{28}$$

where

$$\epsilon = e^{\frac{\pi i}{4}(q-p)} \tag{29}$$

is an eighth root of unity and d is given by (12).

In order to simplify the inner sum in (27), we prove the following

Lemma 2. *If $\underline{a}$ and $\underline{b}$ are two rational columns such that $2S\,\underline{a}$ and $2S\,\underline{b}$* 125
are integral, α, γ, δ integers such that $\alpha\delta \equiv 1((\mathrm{mod})\ \gamma)$ and x is any integral column, then

$$\sum_{\underline{g}((\mathrm{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\left\{\alpha S[\underline{g}+\underline{a}]-2(\underline{x}+\underline{b})'S(\underline{g}+\underline{a})+\delta S[\underline{x}+\underline{b}]\right\}}$$

is independent of $\underline{x}$.

Proof. We have only to consider the exponent in each term (mod γ). In the first place we have

$$\begin{gathered}\alpha S[\underline{g}+\underline{a}] = \alpha S[\underline{g}+\underline{a}-\delta\underline{x}] + 2\alpha\delta\underline{x}'S(\underline{g}+\underline{a}) - \alpha\delta^2 S[\underline{x}]\\ \delta S[\underline{x}+\underline{b}] = \delta S[\underline{b}] + \delta S[\underline{x}] - 2\underline{b}'S(\underline{g}+\underline{a}-\delta\underline{x}) + 2\underline{b}'S(\underline{g}+\underline{a})\\ 2(\underline{x}+\underline{b})'S(\underline{g}+\underline{a}) = 2\underline{b}'S(\underline{g}+\underline{a}) + 2\underline{x}'S(\underline{g}+\underline{a}).\end{gathered}$$

Using the fact that $\alpha\delta \equiv 1((\text{mod})\ \gamma)$ we see that the exponent in each term is congruent (mod γ) to

$$\alpha S[\underline{g}+\underline{a}-\delta\underline{x}] - 2\underline{b}'S(\underline{g}+\underline{a}-\delta\underline{x}) + \delta S[\underline{b}].$$

Since now δ and $\underline{x}$ are fixed and $\underline{g}$ runs over a complete system of residue classes (mod γ), it follows that $\underline{g}-\delta\underline{x}$ also runs through a complete system (mod γ). This proves the lemma.

In the inner sum in (27), $\underline{l}$ runs through all integral vectors, so that we may write

$$\frac{1}{2}S^{-1}\underline{l} = -(\underline{x}+\underline{b})$$

where $\underline{x}$ is an integral column and $\underline{b}$ is one of the finitely many representatives $\underline{a}_1,\ldots,\underline{a}_n$ of the residue classes (mod 1) given in (13). Also when $\underline{x}$ runs through all integral vectors and $\underline{b}$ through these residue class representatives, $-(\underline{x}+\underline{b})$ runs through all rational columns of the type $\frac{1}{2}S^{-1}\underline{l}$. We have thus

$$\left.\begin{aligned}&f_{\underline{a}}(z_M,\underline{w}) = |-2iR_1|^{-\frac{1}{2}} e^{-\frac{\pi i}{2}R_1^{-1}[\underline{W}\gamma]}\\ &\sum_{\underline{b}}\sum_{\underline{x}}\sum_{\underline{g}((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\left\{\alpha S[\underline{g}+\underline{a}]-2(\underline{x}+\underline{b})'S(\underline{g}+\underline{a})+\delta S[\underline{x}+\underline{b}]\right\}}\\ &e^{2\pi iR[\underline{x}+\underline{b}]+2\pi i(\underline{x}+b)'SR_1^{-1}\underline{w}\gamma}\end{aligned}\right\}$$

126 Let us use the abbreviation

$$\lambda_{\underline{ab}}(M) = \sum_{g((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\left\{\alpha S[\underline{g}+\underline{a}]-2\underline{b}'S(\underline{g}+a)+\delta S[\underline{b}]\right\}} \tag{30}$$

Then we have

$$f_{\underline{a}}(z_M, \underline{w}) = |-2iR_1|^{-\frac{1}{2}} e^{-\frac{\pi i}{2} R_1^{-1}[\underline{w}\gamma]}$$
$$\sum_{b} \sum_{\underline{x}} \lambda_{\underline{a},\underline{b}}(M) e^{2\pi i R[\underline{x}+b]+2\pi i(\underline{x}+\underline{b})' S R_1^{-1} \underline{w}\gamma} \tag{31}$$

Let us now define the vector $\underline{w}_{M,z}$ by

$$-S R_1^{-1} \underline{w}_{M,z} = \underline{w}\gamma^{-1}. \tag{32}$$

Using (22) for R_1 we get

$$\underline{w}_{M,z} = \left(\frac{1}{\gamma z + \delta} K + \frac{1}{\gamma \overline{z} + \delta} L\right) S^{-1} \underline{w}. \tag{33}$$

With this definition of $\underline{w}_{M,z}$ we see that

$$R_1^{-1}[\underline{w}_{M,z}\gamma] = R_1[S^{-1}\underline{w}].$$

Use now the abbreviation

$$\rho(M, z, \underline{w}) = e^{-\frac{\pi i}{2} R_1[S^{-1}\underline{w}]}; \tag{34}$$

then substituting $\underline{w}_{M,z}$ for $\underline{w}$ in (27), we get the formula

$$\left.\begin{aligned} f_{\underline{a}}(z_M, \underline{w}_{M,z}) = \epsilon d^{-\frac{1}{2}} \left(z + \frac{\delta}{\gamma}\right)^{p/2} \left(\overline{z} + \frac{\delta}{\gamma}\right)^{q/2} \\ \sum_{\underline{b}} \lambda_{\underline{a},\underline{b}}(M) f_{\underline{b}}(z, \underline{w}) \rho(M, z, \underline{w}) \end{aligned}\right\} \tag{35}$$

□

Till now we considered the case $\gamma \neq 0$. Let now $\gamma = 0$. Then 127

$$M = \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix}$$

and $\alpha\delta = 1$. Also $z_M = \dfrac{\alpha z + \beta}{\delta} = z + \alpha\beta$. The definition of $\underline{w}_{M,z}$ given in (33) is valid even if $\gamma = 0$. Thus

$$e^{-2\pi i\alpha\beta S[\underline{a}]} f_{\underline{a}}(z_M, \underline{w}_{M,z}) = \sum_{\underline{x}} e^{2\pi i R[\alpha\underline{x}+\underline{a}\alpha]+2\pi i(\underline{x}+\underline{a})'\underline{w}_{M,z}}.$$

Since $\alpha = \pm 1$, $\alpha\underline{x}$ also runs through all integral vectors and so

$$f_{\underline{a}}(z_M, \underline{w}_{M,z}) = e^{2\pi i\alpha\beta S[\underline{a}]} f_{\underline{a}\alpha}(z, \underline{w}) \tag{36}$$

$\underline{a}\alpha$ being again some one of the $\underline{a}_1, \ldots, \underline{a}_l$ determined by $\underline{a}$ and α.

For any two rational columns $\underline{a}$, $\underline{b}$ with $2S\,\underline{a}$ and $2S\,\underline{b}$ integral, let us define

$$e_{\underline{ab}} = \begin{cases} 1 & \text{if } \underline{a} \equiv \underline{b} (\text{mod } 1) \\ 0 & \text{otherwise.} \end{cases}$$

Define now the l-rowed matrix $G(M, z)$ by

$$G(M, z) = \begin{cases} \epsilon d^{-\frac{1}{2}} \left(z + \frac{\delta}{\gamma}\right)^{\frac{p}{2}} \left(\bar{z} + \frac{\delta}{\gamma}\right)^{\frac{q}{2}} (\lambda_{\underline{ab}}(M)), \text{ if } \gamma \neq 0 \\ \left(e_{a\ a_\alpha} e^{2\pi i\alpha\beta S[\underline{a}]}\right), \text{ if } \gamma = 0 \end{cases} \tag{37}$$

Also put

$$\rho(M, z, \underline{w}) = \begin{cases} e^{-\frac{\pi i}{2} R_1[S^{-1}\underline{w}]} & \text{if } \gamma \neq 0 \\ 1 & \text{if } \gamma = 0 \end{cases} \tag{38}$$

128 We then have the following fundamental formula for the vector $\underline{f}(z, M, \underline{w})$ defined in (13):

Theorem 1. *If $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ is any modular matrix and $z_M = \dfrac{\alpha z + \beta}{\gamma z + \delta}$, then*

$$\boxed{\underline{f}(z_M, \underline{w}_{M,z}) = G(M, z)\underline{f}(z, \underline{w})\rho(M, z, \underline{w})}$$

where $\underline{w}_{M,z}$ is defined by (33) *and $G(M, z)$ and $\rho(M, z, \underline{w})$ by* (37) *and* (38) *respectively.*

We shall now obtain a composition formula for the l-rowed matrices $G(M, z)$.

Let M and M_1 be two modular matrices

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad M_1 = \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix}, \quad MM_1 = \begin{pmatrix} \alpha_2 & \beta_2 \\ \gamma_2 & \delta_2 \end{pmatrix}$$

By the definition of z_M, it follows that

$$(z_{M_1})_M = z_{MM_1} \tag{39}$$

From definition (33), it follows that

$$\left.\begin{aligned} (\underline{w}_{M_1,z})_{M,z_{M_1}} = &\left(\frac{1}{\gamma z_{M_1} + \delta} K + \frac{1}{\gamma \bar{z}_{M_1} + \delta} L\right) S^{-1}. \\ &\left(\frac{1}{\gamma_1 z + \delta_1} K + \frac{1}{\gamma_1 \bar{z} + \delta_1} L\right) S^{-1} \overline{w}. \end{aligned}\right\}$$

Using the properties of the matrices K and L we get

$$(\underline{w}_{M_1,z})_{M,z_{M_1}} = \left(\frac{1}{\gamma_2 z + \delta_2} K + \frac{1}{\gamma_2 \bar{z} + \delta_2} L\right) S^{-1} \underline{w}$$

which gives the formula

$$\underline{w}_{MM_1,z} = (\underline{w}_{M_1,z})_{M,z_{M_1}}. \tag{40}$$

Using the definition of R_1 and of $\underline{w}_{M,z}$ we get 129

$$R_1[S^{-1}\underline{w}] = -\underline{w}' S^{-1} \underline{w}_{M,z} \gamma \tag{41}$$

Let us now assume, for a moment, that γ, γ_1, γ_2 are all different from zero. Using definition (38) let us write

$$\rho(M_1, z, \underline{w}) \cdot \rho(M, z_{M_1}, \underline{w}_{M_1,z}) = e^{-\frac{\pi i}{2}\varphi}.$$

Then using (41), it follows that

$$-\rho = \underline{w}' S^{-1} \left\{ \gamma_1 + \left(\frac{1}{\gamma_1 z + \delta_1} K + \frac{1}{\gamma_1 \bar{z} + \delta_1} L\right) \gamma S^{-1} \right.$$

$$\left(\frac{1}{\gamma z_{M_1}+\delta}K+\frac{1}{\gamma\bar{z}_{M_1}+\delta}L\right)S^{-1}\biggr\}\underline{w}_{M_1},z$$

Using again the properties of K and L we obtain

$$-\varphi=\underline{w}'S^{-1}\gamma_2\left(\frac{\gamma_1 z+\delta_1}{\gamma_2 z+\delta_2}K+\frac{\gamma_1\bar{z}+\delta_1}{\gamma_2\bar{z}+\delta_2}L\right)S^{-1}\underline{w}_{M_1,z}$$

which is seen to be equal to

$$\underline{w}'S^{-1}\gamma_2\underline{w}_{MM_1,z}.$$

By (41) therefore we get the formula

$$\rho(M_1,z,\underline{w})\cdot\rho(M,z_{M_1},\underline{w}_{M_1,z})=\rho(MM_1,z,\underline{w})\tag{42}$$

We can now release the condition on γ, γ_1, γ_2. If some or all of them are zero, then using definition (38), we can uphold (42). Thus (42) is true for any two modular matrices M, M_1.

If we now use theorem 1 we have

$$\underline{f}(z_{MM_1},\underline{w}_{MM_1,z})=G(MM_1,z)\underline{f}(z,\underline{w})\rho(MM_1,z,\underline{w}).\tag{43}$$

Using (39) and (40) we have

$$\underline{f}(z_{MM_1},\underline{w}_{MM_1,z})=G(M,z_{M_1})\underline{f}(z_{M_1},\underline{w}_{M_1,z})\rho(M,z_{M_1},\underline{w}_{M_1,z})$$

130 which again gives the formula

$$\begin{aligned}&\underline{f}(z_{MM_1},\underline{w}_{MM_1,z})=\\&=G(M,z_{M_1})G(M_1,z)\underline{f}(z,\underline{w})\rho(M,z_{M_1},\underline{w}_{M_1,z})\rho(M_1z,\underline{w})\end{aligned}\tag{44}$$

Using (42), (43) and (44) and observing that $\rho(MM_1,z,\underline{w})\neq 0$, we get the matrix equation

$$(G(MM_1,z)-G(M,z_{M_1})G(M_1,z))\underline{f}(z,\underline{w})=\underline{0}.\tag{45}$$

We remark that the 1-rowed matrix on the left hand side of equation (45) is independent of $\underline{w}$. Let us now prove

Lemma 3. *The l functions $f_{\underline{a}_1}(z,\underline{w}),\ldots,f_{\underline{a}_1}(z,\underline{w})$ are linearly independent over the field of complex numbers.*

Proof. By definition $f_{\underline{a}}(z,\underline{w}) = \sum_{\underline{x}} e^{2\pi i R[\underline{x}+\underline{a}]+2\pi i \underline{w}'(\underline{x}+\underline{a})}$ so that it is a fourier series in the n variables $w_1,\ldots,w_n$. We may write

$$f_{\underline{a}}(z,\underline{w}) = \sum_{\underline{r}} c_r e^{2\pi i \underline{w}'\underline{r}}$$

where $\underline{r}$ runs through all rational vectors $\equiv \underline{a}(\text{mod } 1)$. If $\alpha_1,\ldots,\alpha_r$ be complex numbers such that

$$\sum_{i=1}^{l} \alpha_i f_{\underline{a}_i}(z,\underline{w}) = 0$$

then by uniqueness theorem of fourier series, every fourier coefficient must vanish. But since $\underline{a}_1,\ldots,\underline{a}_l$ are all distinct (mod 1), it follows that the exponents in the l series $f_{\underline{a}}(z,\underline{w})$ are all distinct. Hence $\alpha_i = 0, i = 1,\ldots,n$, and our lemma is proved. □

Using (45) and lemma 3, it follows that the l-rowed matrix on the left of (45) is identically zero. Hence the 131

Theorem 2. *For any two modular matrices M, M_1 we have the composition formula*

$$\boxed{G(MM_1,z) = G(M,z_{M_1})G(M_1,z)}$$

Let $M = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ be a modular matrix so that

$$M^{-1} = \begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}$$

Let us assume that $\gamma \neq 0$. Let as before $\underline{a}$, $\underline{b}$ be two rational columns chosen from the set $\underline{a}_1,\ldots,\underline{a}_l$. We shall prove

$$\overline{\lambda_{\underline{a}\,\underline{b}}(M)} = \lambda_{\underline{b}\,\underline{a}}(M^{-1}) \tag{46}$$

where $\lambda_{\underline{a}\,\underline{b}}(M)$ is the sum defined in (30).

In order to prove (46), put $t = \lambda_{\underline{a}\,\underline{b}}(M)$ and $t' = \lambda_{\underline{b}\,\underline{a}}(M^{-1})$. Because of lemma 2, we have

$$t = \sum_{\underline{y}((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\left\{\alpha S\,[\underline{y}+\underline{a}]-2(\underline{y}+\underline{b})'S\,(\underline{x}+\underline{a})+\delta S\,[\underline{x}+\underline{b}]\right\}}$$

Taking the sum over all integral $\underline{x}(\text{mod } \gamma)$ we have

$$\text{t.abs}\gamma^n = \sum_{\underline{x}((\text{mod})\ \gamma)} \sum_{\underline{y}((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\{\alpha S\,[\underline{y}+\underline{a}]-2(\underline{y}+\underline{b})'S\,(\underline{x}+\underline{a})+\delta S\,[\underline{x}+\underline{b}]\}}$$

Interchanging the two summations we have

$$t\ abs\gamma^n = \sum_{\underline{y}((\text{mod})\ \gamma)} \sum_{\underline{x}((\text{mod})\ \gamma)} e^{-\frac{2\pi i}{-\gamma}\{\delta S\,[\underline{x}+\underline{b}]-2(\underline{x}+\underline{a})'S\,(\underline{y}+\underline{b})+\alpha S\,[\underline{y}+\underline{a}]\}}$$

132 But by lemma 2 again we see that the inner sum is independent of $\underline{y}$ and equal to $\overline{t'}$, the complex conjugate of t'. Thus

$$t\ abs\gamma^n = \overline{t'} abs\gamma^n$$

and since $\gamma \neq 0$, it follows that $\bar{t} = t'$ and (46) is proved.

In the composition formula of theorem 2, let us put $M_1 = M^{-1}$. Then $G(E, z) = E$ is the unit matrix of order 1. From the definition of $G(M, z)$ we have

$$G(M^{-1}, z) = \epsilon d^{-\frac{1}{2}} \left(z - \frac{\alpha}{\gamma}\right)^{\frac{p}{2}} \left(\bar{z} - \frac{\alpha}{\gamma}\right)^{\frac{q}{2}} (\lambda_{\underline{a}\,\underline{b}}(M^{-1}))$$
$$G(M, z_{M^{-1}}) = \epsilon d^{-\frac{1}{2}} (\gamma^{-1}(-\gamma z + \alpha))^{\frac{p}{2}} (\gamma^{-1}(-\gamma \bar{z} + \alpha))^{\frac{q}{2}} (\lambda_{\underline{ab}}(M))$$

Let us put

$$\Lambda(M) = \epsilon d^{-\frac{1}{2}} abs\gamma^{-\frac{n}{2}} (\lambda_{\underline{ab}}(M)). \tag{47}$$

Then we get from the previous equations

$$\Lambda(M) \cdot \overline{\Lambda(M)}' = E$$

which shows that $\Lambda(M)$ is unitary.

In case $\gamma = 0$, from (37), $G(M,z)$ is clearly unitary. We therefore put

$$\Lambda(M) = G(M,z), \quad \text{if } \gamma = 0. \tag{48}$$

Let us now put $\underline{w} = \underline{0}$ in theorem 1. Then $\rho(M,z,\underline{w}) = 1$ so that if we write as in § 1, $\underline{f}(z,H)$ instead of $\underline{f}(z,H,\underline{w})$, when $\underline{w} = \underline{0}$, we get

$$\underline{f}(z_M,H) = G(M,z)\underline{f}(z,H). \tag{49}$$

Using the definitions (47) and (48), we get

Theorem 3. *If $M = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ is a modular matrix, then*

$$(\gamma_z + \delta)^{-\frac{p}{2}}(\gamma\bar{z} + \delta)^{-\frac{q}{2}}\underline{f}(z_M,H) = \Lambda(M)\underline{f}(z,H)$$

where $\Lambda(M)$ is a certain unitary matrix and the radical scalar factors on the left side are taken with their principal parts. 133

We remark that we introduced the vector $\underline{w}$ only to prove the composition formula in theorem 2. Hereafter we will have only $\underline{f}(z,H)$, the column consisting of $f_{\underline{a}_i}(z,H)$ defined in (10).

From the composition formula we get

$$\Lambda(MM_1) = \Lambda(M) \cdot \Lambda(M_1) \tag{50}$$

which shows that the mapping $M \to \Lambda(M)$ is a unitary representation of the modular group.

4 Convergence of an integral

Let S be the matrix of a quadratic form and let S be non-singular and semi-integral. Let S have signature p, q so that $p+q = n$. Let us assume that $pq > 0$. With S we had associated the $\mathfrak{H}$ space of matrices H with $H > 0$, $HS^{-1}H = S$. Let Γ be the group of units of S. Let $\underline{a}$ denote a rational column vector with $2S\underline{a}$ integral. Denote by $\Gamma_{\underline{a}}$ the group of units of S satisfying

$$U\underline{a} \equiv \underline{a} (\text{mod } 1) \tag{51}$$

$\Gamma_{\underline{a}}$ is obviously a subgroup of Γ of finite index. Let $U_1, \ldots, U_s$ denote a complete system of representatives of left cosets of $\Gamma \bmod \Gamma_{\underline{a}}$ so that

$$\Gamma = \sum_{i=1}^{s} U_i \Gamma_{\underline{a}}, \quad s = (\Gamma : \Gamma_{\underline{a}}).$$

134 Denote by F a fundamental region of Γ in $\mathfrak{H}$ and by F_k the image by U_k of F in $\mathfrak{H}$. Put

$$F_{\underline{a}} = \bigcup_k F_k;$$

then it is easy to verify that $F_{\underline{a}}$ is a fundamental domain for $\Gamma_{\underline{a}}$ in $\mathfrak{H}$.

For every H in $\mathfrak{H}$ we had defined the theta series

$$f_{\underline{a}}(z, H) = \sum_{\underline{y} \equiv \underline{a} (\text{mod } 1)} e^{2\pi i R[\underline{y}]}$$

so that regarded as a function of H, $f_{\underline{a}}(z, H)$ is a function on the manifold $\mathfrak{H}$. If $U \in \Gamma_{\underline{a}}$ then

$$f_{\underline{a}}(z, H[U]) = \sum_{\underline{y} \equiv \underline{a} (\text{mod } 1)} e^{2\pi i (\xi S + i\eta H[U])[\underline{y}]}.$$

Writing $S[U]$ instead of S and observing that $U\underline{y} \equiv U\underline{a} \equiv \underline{a} (\text{mod } 1)$ and that $U\underline{y}$ runs through all rational columns $\equiv \underline{a} (\text{mod } 1)$ if $\underline{y}$ does, we have

$$f_{\underline{a}}(z, H[U]) = f_{\underline{a}}(z, H) \tag{52}$$

so that we may regard $f_{\underline{a}}(z, H)$ as a function on $F_{\underline{a}}$. Let dv be the invariant volume measure in the $\mathfrak{H}$ space. We shall now prove that

$$\int_{F_{\underline{a}}} f_{\underline{a}}(z, H) dv$$

converges, in particular, if $n > 4$ and that

$$\int_{F_{\underline{a}}} f_{\underline{a}}(z, H) dv = \sum_{\underline{y} \equiv \underline{a} (\text{mod } 1)} \int_{F_{\underline{a}}} e^{2\pi i R[\underline{y}]} dv \tag{53}$$

For proving this it is enough to show that the series of absolute values of the terms of $f_{\underline{a}}(z, H)$ converges uniformly in every compact subset of $F_{\underline{a}}$ and that the integral over $F_{\underline{a}}$ of this series of absolute values converges. 135

Because of the property (52) and the invariance of the volume measure it is enough to consider the integral over F instead of $F_{\underline{a}}$. By our method of construction

$$F = \bigcup_k (\mathfrak{H} \cap \mathscr{R}_k)$$

where $\mathscr{R}_k$ is obtained from the Minkowski fundamental domain $\mathscr{R}$. It is therefore enough to consider the integral

$$\int\limits_{\mathfrak{H}\cap\mathscr{R}} f_{\underline{a}}(z, H) dv$$

and prove (53) for $\mathfrak{H} \cap \mathscr{R}$ instead of $F_{\underline{a}}$.

The general term of the integrand is $e^{2\pi i R[\underline{y}]}$ and its absolute value is

$$e^{-2\pi\eta H[\underline{x}+\underline{a}]}$$

where $\eta > 0$, $H > 0$ is reduced in the sense of Minkowski, $\underline{x}$ an integral column and $\underline{a}$ a rational column with $2S\underline{a}$ integral. If $H = (h_{kl})$ then H being reduced, there exists a constant c_1, such that

$$H[\underline{y}] > c_1(h_1 y_1^2 + \cdots + h_n y_n^2)$$

$\underline{y}$ being a real column with n elements $y_1, \ldots, y_n$. Therefore

$$\prod_{i=1}^{n} \sum_{y_i} e^{-2\pi c_1 \eta h_i y_i^2}$$

is a majorant for the sum of the absolute values of the terms of $f_{\underline{a}}(z, H)$. Since, for a constant $c_2 > 0$,

$$\sum_{t=-\infty}^{\infty} e^{-c_2 h t^2} < c_3(1 + h^{-\frac{1}{2}})$$

where $h > 0$ is a positive real number and c_3 is a constant depending on c_2, it follows that it is enough to prove, for our purpose, the convergence of 136

$$\int\limits_{\mathfrak{H}\cap\mathscr{R}} \prod_{k=1}^{n} (1 + h_k^{-\frac{1}{2}}) dv \tag{54}$$

If D is any compact subset of $\mathfrak{H} \cap \mathscr{R}$, then because H is reduced, $\prod\limits_{k=1}^{n} (1 + h_k^{-\frac{1}{2}})$ is uniformly bounded in D. This will prove (53) as soon as (54) is proved.

The proof depends on an application of Minkowski's reduction theory.

Consider any element $H = (h_{kl})$ in $\mathfrak{H} \cap \mathscr{R}$ and consider the products $h_k h_{n-k}$, $k = 1, 2, \ldots, n-1$. There exists an integer r

$$0 \leq r \leq \frac{n}{2} \tag{55}$$

such that

$$h_k h_{n-k} \geq \frac{1}{4} \qquad r < k < n - r \tag{56}$$

$$h_r h_{n-r} < \frac{1}{4} \tag{57}$$

If $r = 0$, (57) is empty and if $r = \frac{n}{2}$ (which implies that n is even) (56) is empty. Let us denote by M_r the subset of $\mathfrak{H} \cap \mathscr{R}$ consisting of those H which have the same integer r associated with them. Clearly $\mathfrak{H} \cap \mathscr{R} = \bigcup\limits_{r} M_r$. It is enough therefore to prove that for every r,

$$\int\limits_{M_r} \prod_{k=1}^{n} (1 + h_k^{-\frac{1}{2}}) dv$$

converges.

137 We first obtain a parametrical representation for the matrices H in M_r.

Let $K = \dfrac{H+S}{2} = (u_{kl})$ and $-L = \dfrac{H-S}{2} = (v_{kl})$; then K and $-L$ are non-negative matrices so that

$$\pm u_{kl} \leq \sqrt{u_k u_1}, \quad \pm v_{kl} \leq \sqrt{v_k v_1}$$

where for a real number g, $\pm g$ denotes its absolute value. Since $K+L=S$ we get

$$\pm s_{kl} \leq \sqrt{u_k u_l} + \sqrt{v_k v_l}.$$

But since $u_k + v_k = h_k$ we obtain, by using Schwarz's inequality,

$$\pm s_{kl} \leq \sqrt{h_k h_1}.$$

H being reduced we get for $k \leq r$, $l \leq n-r$, using (57),

$$\pm s_{kl} \leq \sqrt{h_k h_1} \leq \sqrt{h_r h_{n-r}} < \frac{1}{2} \tag{58}$$

Since S is a semi-integral matrix, it follows that

$$s_{kl} = 0, \quad k \leq r, \quad l \leq n-r.$$

We have therefore a decomposition of S into the form

$$S = \begin{pmatrix} 0 & 0 & P \\ 0 & F & Q \\ P' & Q' & G \end{pmatrix} \tag{59}$$

where P is an r-rowed non-singular matrix. It has to be noted that if $r=0$, then

$$S = F \tag{60}$$

and if $r = \dfrac{n}{2}$ (n is then even),

$$S = \begin{pmatrix} 0 & P \\ P' & G \end{pmatrix} \tag{61}$$

138

We now put $S^* = S[C]$ where

$$S^* = \begin{pmatrix} 0 & 0 & P \\ 0 & F & 0 \\ P' & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} E & -P'^{-1}Q' & -\frac{1}{2}P'^{-1}G \\ 0 & E & 0 \\ 0 & 0 & E \end{pmatrix} \tag{62}$$

We split up H also in the same fashion, by the Jacobi transformation, $H = H_0[C_0]$ where

$$H_0 = \begin{pmatrix} H_1 & 0 & 0 \\ 0 & H_2 & 0 \\ 0 & 0 & H_3 \end{pmatrix}, \quad C_0 = \begin{pmatrix} E & L_1 & L_2 \\ 0 & E & L_3 \\ 0 & 0 & E \end{pmatrix} \tag{63}$$

where H_1 and H_3 are r-rowed symmetric matrices. Put

$$C_0 C = L = \begin{pmatrix} E & Q_1 & Q_2 \\ 0 & E & Q_3 \\ 0 & 0 & E \end{pmatrix} \tag{64}$$

If we put $H^* = H[C]$, then since $S^{-1}[H] = S$, it follows that $S^{*-1}[H^*] = S^*$. Using the matrix L we have

$$(LS^{*-1}L')[H_0] = S^*[L^{-1}].$$

Substituting for the various matrices above, we get

$$F^{-1}[H_2] = F \tag{65}$$

$$H_3 P^{-1} H_1 = P' \tag{66}$$

$$Q_3 = -F^{-1} Q_1' P \tag{67}$$

$$Q_2 = (A - \frac{1}{2} F^{-1}[Q_1'])P \tag{68}$$

139 where A is a skew symmetric matrix of r-rows. It is obvious that if H_1, H_2, H_3, Q_1, Q_2 and Q_3 satisfy the above conditions, then the corresponding H is in $\mathfrak{H}$. We therefore choose the parameters for the M_r space in the following manner: We have H_1 is arbitrary, r-rowed and positive. From this H_3 is uniquely fixed. Q_1 is an arbitrary matrix of r-rows and $n - 2r$ columns. Q_3 is then determined uniquely. Choose A to be arbitrary skew symmetric. Then (68) determines Q_2. H_2 is now a positive matrix satisfying (65). Thus the parameters are H_1, Q_1, A and the parameters required to parametrize the space of positive H_2 satisfying (65). A simple calculation shows that the number of parameters is

$$\frac{r(r+1)}{2} + r(n-2r) + \frac{r(r-1)}{2} + (p-r)(q-r). \tag{69}$$

We now compute the volume element in terms of these parameters. The metric in the $\mathscr{H}$ space is

$$ds^2 = \frac{1}{8}\sigma(H^{-1}dHH^{-1}dH).$$

We substitute for H in terms of these new parameters. We denote differentiation by $(\cdot)$ dot. Since C is a constant matrix we get

$$ds^2 = \frac{1}{8}\sigma(H^{*-1}dH^{*}H^{*-1}dH^{*}) \tag{70}$$

As $H^{*} = H_0[L]$ we get

$$H^{*-1}\dot{H}^{*} = H_0^{-1}[L'^{-1}](\dot{H}_0[L] + \dot{L}'H_0L + LH_0\dot{L}).$$

This gives the expression

$$\left.\begin{aligned}\sigma(H^{*-1}\dot{H}^{*})^2 = \Big\{&\sigma(H_0^{-1}\dot{H}_0H_0^{-1}\dot{H}_0) + 4\sigma(H_0^{-1}\dot{H}_0\dot{L}L^{-1}\Big\}\\ &+ 2\sigma(\dot{L}L^{-1}\dot{L}L^{-1}) + 2\sigma(H_0[\dot{L}L^{-1}]H_0^{-1})\end{aligned}\right\} \tag{71}$$

We shall now simplify the expression on the right of (71). Since $L = C_0C$ and C is a constant matrix, we get **140**

$$\dot{L}L^{-1} = \begin{pmatrix} 0 & \dot{Q}_1 & \dot{Q}_2 \quad -\dot{Q}_1Q_3 \\ 0 & 0 & \dot{Q}_3 \\ 0 & 0 & 0 \end{pmatrix}$$

which shows that

$$\sigma(\dot{L}L^{-1}\dot{L}L^{-1}) = 0, \quad \sigma(H_0^{-1}\dot{H}_0\dot{L}L^{-1}) = 0. \tag{72}$$

Using the expression for H_0 in (63) we get

$$\sigma(H_0^{-1}\dot{H}_0H_0^{-1}\dot{H}_0) = \sum_{i=1}^{3}\sigma(H_i^{-1}\dot{H}_iH_i^{-1}\dot{H}_i).$$

Differentiating (66) with regard to the variables H_1 and H_3 we get

$$\dot{H}_3P^{-1}H_1 + H_3P^{-1}\dot{H}_1 = 0$$

which shows that $H_1^{-1}\dot{H}_1 = -P'^{-1}\dot{H}_3H_3^{-1}P'$ and therefore

$$\sigma(H_1^{-1}\dot{H}_1H_1^{-1}\dot{H}_1) = \sigma(H_3^{-1}\dot{H}_3H_3^{-1}\dot{H}_3) \tag{73}$$

Using the expressions for $\dot{L}L^{-1}$ and H_0, we obtain

$$\begin{aligned}\sigma(H_0[\dot{L}L^{-1}]H_0^{-1}) &= \sigma(H_1[\dot{Q}_1]H_2^{-1}) + \sigma(H_1[\dot{Q}_2 - \dot{Q}_1Q_3]H_3^{-1})\\ &\quad + \sigma(H_2[\dot{Q}_3]H_3^{-1}).\end{aligned}$$

Differentiating (67) and (68) with regard to the variables Q_1, Q_2, Q_3, we get

$$\begin{aligned}\dot{Q}_3 &= -F^{-1}\dot{Q}_1'P\\ \dot{Q}_2 &= (\dot{A} - \frac{1}{2}\dot{Q}_1F^{-1}Q_1' - \frac{1}{2}Q_1F^{-1}\dot{Q}_1')P\end{aligned} \tag{74}$$

We now introduce the matrix B defined by

$$B = \frac{1}{2}(\dot{Q}_1F^{-1}Q_1' - Q_1F^{-1}\dot{Q}_1').$$

141 We can then write

$$\dot{Q}_2 - \dot{Q}_1Q_3 = (\dot{A} + B)P.$$

We have finally, except for a positive constant, the metric in the space M_r

$$\begin{aligned}ds^2 = \sigma\left\{(H_2^{-1}\dot{H}_2)^2\right\} + 2\sigma\left\{(H_1^{-1}\dot{H}_1)^2\right\} + 4\sigma(H_1[\dot{Q}_1]H_2^{-1})\\ + 2\sigma(H_1[\dot{A} + B]H_1).\end{aligned} \tag{75}$$

The determinant of the quadratic differential form

$$2\sigma\left\{(H_1^{-1}\dot{H}_1)^2\right\} + 4\sigma(H_1[\dot{Q}_1]H_2^{-1}) + 2\sigma(H_1[\dot{A} + B]H_1)$$

is given by

$$2^{r(n-r-1)}|H_1|^{n-2r-2}|H_2|^{-r} \tag{76}$$

If dv_2 denotes the invariant volume measure in the space of H_2, satisfying (65), then we have to prove

$$\int\limits_{M_r}\prod_{k=1}^{n}(1+h_k^{-\frac{1}{2}})|H_1|^{\frac{n-2r-2}{2}}|H_2|^{-\frac{r}{2}}\{dH_1\}\{dQ_1\}\{dA\}dv_2 \tag{77}$$

is convergent.

The constants $c_3, \ldots$ appearing in the sequel all depend only on n and S. Moreover 'bounded' shall mean bounded in absolute value by such constants.

Since $|S| \neq 0$, at least one term in the expansion of $|S|$ does not vanish. This means there is a permutation

$$\begin{pmatrix} 1, 2, \ldots, n \\ l_1, l_2, \ldots, l_n \end{pmatrix}$$

such that $s_{kl_k} \neq 0$, $k = 1, \ldots, n$. 142

Since S is semi-integral, $\pm s_{kl_k} \geq \frac{1}{2}$ which shows that

$$h_k h_{l_k} \geq s^2_{kl_k} \geq \frac{1}{4} \tag{78}$$

Consider now the integers $1, 2, \ldots, a$ and the corresponding integers $l_1, \ldots, l_a$, $a \leq n$. At least one of the latter, say $l_t \leq n-a+1$. Therefore

$$t \leq a, \quad l_t \leq n-a+1.$$

Since H is reduced, $h_t \leq h_a$, $h_{l_t} \leq h_{n-a+1}$. Using (78) we get

$$h_a h_{n-a+1} \geq \frac{1}{4}, \quad a = 1, \ldots, n. \tag{79}$$

Let us consider the identity

$$\prod_{k=1}^{n}(h_k h_{n-k+1}) = \prod_{k=1}^{r}(h_k h_{n-k+1})^2 \cdot \frac{\prod\limits_{k=r+1}^{n}(h_k h_{n-k+1})}{\prod\limits_{k=1}^{r}(h_k h_{n-k+1})}$$

Since $r \leq n - r$, it follows using (56)

$$\frac{\prod\limits_{k=r+1}^{n} (h_k h_{n-k+1})}{\prod\limits_{k=1}^{r} (h_k h_{n-k+1})} \geq c_3 h_{n-r}^2$$

Therefore we obtain

$$\prod_{k=1}^{n} (h_k h_{n-k+1}) \geq c_3 h_{n-r}^2 \prod_{k=1}^{r} (h_k h_{n-k+1})^2$$

Using (79) and the fact that H is reduced, we get the inequality

$$h_{n-r} \leq c_4. \tag{80}$$

143 Since H is reduced, $H_0 \in \mathscr{R}_c^*$ for a $c > 0$ depending only on n and S. It follows from (80), therefore, that the elements of H_1 and H_2 are bounded. Also from (79) and (80), it follows that

$$h_k \geq c_5 \quad r < k \leq n \tag{81}$$

which shows that the elements of H_2^{-1} are bounded.

From equations (63) and (64) we get

$$Q_1 = L_1 - P'^{-1} Q', \quad Q_2 = L_2 - \frac{1}{2} P'^{-1} G.$$

Since P, Q and G are constant matrices and L_1, L_2 have bounded elements (since H is reduced), it follows that the elements of Q_1 and Q_2 are bounded.

From the definition of A in (68), it follows that its elements are also bounded. From (80) and the fact that H is reduced we get

$$h_1 \leq h_2 \leq h_3 \ldots \leq h_r \leq c_4 \tag{82}$$

and therefore

$$\prod_{k=1}^{r} (1 + h_k^{-\frac{1}{2}}) \leq c_6 (h_1 \ldots h_r)^{-\frac{1}{2}} \tag{83}$$

We therefore finally see that it is enough to prove

$$\int |H_1|^{\frac{n-2r-2}{2}}(h_1\dots h_r)^{-\frac{1}{2}}\{dH_1\}$$

converges, H_1 being reduced and satisfying (82). $dH_1 = \prod\limits_{1\leq i\leq j\leq r} dh_{ij}$. Since $-h_i \leq 2h_{ij} \leq h_i$, $i < j$, it follows that the variation of h_{ij} is h_i. Therefore it is enough to prove that the integral

$$\int (h_1\dots h_r)^{-\frac{1}{2}}(h_1\dots h_r)^{\frac{n}{2}-r-1}h_1^{r-1}\dots h_{r-1}dh_1\dots dh_r,$$

extended over the set $0 < h_1 \leq h_2 \leq h_3\dots \leq h_r \leq c_4$ converges. We 144
make a change of variables

$$\left.\begin{aligned} h_1 &= s_1\dots s_r\\ h_2 &= s_2\dots s_r\\ h_r &= s_r \end{aligned}\right\} \tag{84}$$

The integral then becomes transformed into

$$\int s_1^{\lambda_1}\dots s_r^{\lambda_\gamma}\cdot\frac{ds_1\dots ds_r}{s_1\dots s_r} \tag{85}$$

where since $s_k = \dfrac{h_k}{h_{k+1}}$, $0 < s_k < \text{Min}(1, c_4)$ and $\lambda_k = k\left(\dfrac{n-k}{2} - 1\right)$, $k = 1,\dots,r$.

Now $\lambda_k \geq \dfrac{n-r}{2} - 1$, $k = 1, 2, \dots, r$ so that if $n - r - 2 > 0$, the integral obviously converges. If $n > 4$, since $r \leq \dfrac{n}{2}$, this condition is satisfied and the integral converges.

Now the maximum value of r is $\leq \text{Min}(p, q)$. Let $n = 4$ and $S[\underline{x}]$ be not a quaternionic form, i.e., it is not the norm of a general element of a quaternion algebra over the field of rational numbers. In that case the maximum value of r is 0 or 1 so that $n - r - 2 > 0$ and the integral converges. If $n = 3$ and $S[\underline{x}]$ is not a zero form, then $r = 0$ and $n-r-2 > 0$.

If $r = 0$, then all elements of H in M_0 are bounded and the integral over M_0 converges. This shows that if $n = 2$ and $S[\underline{x}]$ is not a zero form, the integral again converges.

In particular, we have

145 **Theorem 4.** *If $n > 4$ the integral*

$$\int\limits_{F_{\underline{a}}} f_{\underline{a}}(z, H)dv$$

converges and

$$\int\limits_{F_{\underline{a}}} f_{\underline{a}}(z, H)dv = \sum_{\underline{y}} \int\limits_{F_{\underline{a}}} e^{2\pi i R[\underline{y}]} dv.$$

Let us now consider the integral $\int\limits_{F_{\underline{a}}} dv$. In order to prove it is finite, it is enough to prove $\int\limits_{M_r} dv$ is finite for every r. Thus we have to prove

$$\int\limits_{0 < h_i \leq h_2 \leq \ldots \leq h_r \leq c_4} h_1^{\frac{n-4}{2}} \ldots h_r^{\frac{n-2r-2}{2}} dh_1 \ldots dh_r$$

is finite. By the same change of variables we see that instead of λ_k, one has $\mu_k = k\left(\frac{n-k-1}{2}\right)$ so that since $\mu_k \geq \frac{n-k-1}{2}$, the integral converges if $n - r - 1 > 0$. Since $r \leq \frac{n}{2}$, the integral converges if $n > 2$. If $n = 2$ and $r = 0$, then again the integral converges. If $n = 2$ and $r = 1$, $S[\underline{x}]$ is a binary zero form and we had see in the previous chapter that $\int\limits_{F} dv$ diverges. We have thus proved

Theorem 5. *If $S[\underline{x}]$ is not a binary zero form*

$$\int\limits_{F_{\underline{a}}} dv$$

converges.

5 A theorem in integral calculus

146

For out later purposes we shall prove a theorem on multiple integrals.

Let R_m denote the Euclidean space of m dimensions with $x_1, \ldots, x_m$ forming a coordinate system. Let

$$y_k = f_k(x_1, \ldots, x_m), k = 1, \ldots, n,$$

be n differentiable functions with $n \leq m$. Let $a_1, \ldots, a_n$ be n real numbers and let F be the 'surface' determined by the n equations

$$y_k = a_k \quad k = 1, \ldots, n.$$

Let us moreover assume that the functional matrix

$$\left(\frac{\partial f_i}{\partial x_j}\right) \qquad \begin{cases} i = 1, \ldots, n \\ j = 1, \ldots, m \end{cases}$$

has the maximum rank n at every point of F. Introduce $m - n$ differentiable functions $y_{n+1}, \ldots, y_m$ of $x_1, \ldots, x_m$ so that the Jacobian

$$J = \left|\left(\frac{\partial y_i}{\partial x_j}\right)\right| \qquad i, j = 1, \ldots, m$$

is different from zero at every point of F. The $y_{n+1}, \ldots, y_m$ are the local coordinates of the 'surface' F. Let Δ denote the absolute value of J and put

$$d\omega = \Delta^{-1} dy_{n+1} \ldots dy_m. \tag{86}$$

The properties of Jacobians show that $d\omega$ is independent of the choice of $y_{n+1}, \ldots, y_m$. We shall denote $d\omega$ symbolically by

$$d\omega = \frac{\{dx\}}{\{dy\}} \tag{87}$$

and take $d\omega$ as the measure of volume on 'surface' F. 147

In case $m = n$, because of the conditions on the Jacobian, the point set F is zero dimensional and we define $d\omega$ to be the measure which assigns to each point the measure $\frac{1}{\Delta}$.

As an example put $m = 2$, and consider in R_2 the point set F defined by

$$y_1 = \sqrt{x_1^2 + x_2^2}, \quad y_1 = 1.$$

Then $\dfrac{\partial y_1}{\partial x_1}, \dfrac{\partial y_1}{\partial x_2}$ cannot both vanish at any point of F. Choose now y_2 as

$$y_2 = \tan^{-1} \frac{x_2}{x_1}$$

The Jacobian is

$$= \left|\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)}\right| = \frac{1}{y_1}.$$

and the volume element on the circle $F : x_1^2 + x_2^2 = 1$ is

$$d\omega = y_1 dy_2.$$

Let $X = X^{(r,s)}$ be a real matrix of r rows and s columns with elements x_{kl} constituting a coordinate system in R_{rs}. We denote by

$$\{dX\} = \prod_{k=1}^{r} \prod_{l=1}^{s} dx_{kl}$$

the Euclidean volume element in R_{rs}. If however $X = X'$ is r-rowed symmetric, then

$$\{dX\} = \prod_{1 \leq k \leq l \leq r} dx_{kl}$$

148 Let V be a k-rowed real non-singular symmetric matrix with signature $\alpha, k - \alpha$. Let F be a rectangular matrix with k-rows and β columns so that the matrix T defined by

$$V[F] = T$$

is non-singular and has signature $\alpha, \beta - \alpha$. Obviously $\alpha \leq \beta \leq k$. Let W be a fixed matrix of $\beta + \lambda$ rows and of signature $\alpha, \beta + \lambda - \alpha$. Then $\beta + \lambda \leq k$. Let D be the ‘surface consisting of real matrices X of k rows and λ columns satisfying

$$V[F, X] = W.$$

If we write

$$W = \begin{pmatrix} T & Q \\ Q' & R \end{pmatrix}$$

then D is the surface defined by the equations

$$\left.\begin{aligned} F'VX &= Q \\ X'VX &= R \end{aligned}\right\} \tag{88}$$

In conformity with our previous notation, let the volume element on the ‘surface’ D be denoted by

$$\frac{\{dX\}}{\{dQ\}\{dR\}}.$$

We have then the following

Theorem 6. $\displaystyle\int\limits_D \frac{\{dX\}}{\{dQ\}\{dR\}} = \frac{\rho_{k-\beta}}{\rho_{k-\beta} - \lambda} \|V\|^{\frac{-\lambda}{2}} \|T\|^{\frac{\beta-k+1}{2}} \|W\|^{\frac{k-\beta-\lambda-1}{2}}$

where $\rho_h = \displaystyle\prod_{i=1}^{h} \frac{\pi^{i/2}}{\Gamma(i/2)}$ *and* $\rho_0 = 1$. *Also if* $\beta = 0$, $\|T\|$ *has to be taken* 149
equal to 1.

Proof. First let $\beta > 0$. Denote by I the integral

$$I = \|V\|^{\frac{\lambda}{2}} \|T\|^{-\frac{\beta-k+1}{2}} \|W\|^{-\frac{k-\beta-\lambda-1}{2}} \int\limits_D \frac{\{dX\}}{\{dQ\}\{dR\}}$$

Let C be a k-rowed non-singular matrix. Consider the transformation,

$$X \to CX, \quad F \to CF, \quad V \to V[C^{-1}].$$

This leaves W unaltered. Also

$$\{d(CX)\} = \|C\|^{\lambda}\{dX\}$$

which shows that I is unaltered. We shall choose C in such a manner that the resulting integral can be easily evaluated. □

Since F has rank β, there exists a matrix C_0 such that

$$C_0F = \begin{pmatrix} E_\beta \\ 0 \end{pmatrix}$$

E_β being the unit matrix of order β. Since $V[F] = T$ is non-singular, we have

$$V[C_0^{-1}] = \begin{pmatrix} T & L \\ L' & N \end{pmatrix} = \begin{pmatrix} T & 0 \\ 0 & M \end{pmatrix} \begin{bmatrix} E & T^{-1}L \\ 0 & E \end{bmatrix}$$

As V has signature α, $k - \alpha$, it follows tht $-M > 0$. Put $-M = P'P$, where P is non-singular. Then

$$V[C_0^{-1}] = \begin{pmatrix} T & 0 \\ 0 & -E_{k-\beta} \end{pmatrix} \begin{bmatrix} E & T^{-1}L \\ 0 & P \end{bmatrix}$$

We now choose C so that

$$C = \begin{pmatrix} E_\beta & T^{-1}L \\ 0 & P \end{pmatrix} C_0$$

150 A simple computation of determinants now shows that I reduces to

$$\|T\|^{-\frac{\beta-k-\lambda+1}{2}} \|W\|^{-\frac{k-\beta-\lambda-1}{2}} \int\limits_D \frac{\{dX\}}{\{dQ\}\{dR\}}$$

where D is now the domain defined by $X = \binom{X_1}{X_2}$, $X_1 = X_1^{(\beta,\lambda)}$ satisfying

$$\left.\begin{aligned} &\begin{pmatrix} T & 0 \\ 0 & -E_{k-\beta} \end{pmatrix} \begin{bmatrix} E & X_1 \\ 0 & X_2 \end{bmatrix} = \begin{pmatrix} T & Q \\ Q' & R \end{pmatrix} = W \\ &Q = TX_1, \quad R = X_1'TX - X_2'X_2. \end{aligned}\right\} \tag{89}$$

Completing squares, we get

$$\begin{pmatrix} T & 0 \\ 0 & -E \end{pmatrix} \begin{bmatrix} E & X_1 - T^{-1}Q \\ 0 & X_2 \end{bmatrix} = \begin{pmatrix} T & 0 \\ 0 & R_1 \end{pmatrix} \tag{90}$$

where

$$R_1 = -X_2'X_2 \tag{91}$$

and $-R_1 > 0$.

$$\text{Now } \{dX\} = \{dX_1\}\{dX_2\} \text{ and from (89)}$$
$$\{dX_1\} = \|T\|^{-\lambda}\{dQ\}$$

Also (90) shows that

$$\|W\| = \|T\| \; \|R_1\|.$$

Therefore I reduces to

$$\|R_1\|^{-\frac{(k-\beta-\gamma-1)}{2}} \int\limits_D \frac{\{dX_2\}}{\{dR_1\}}$$

where D is the domain defined by X_2 satisfying (91).

Let G be a non-singular matrix such that 151

$$\left.\begin{aligned} X_2G &= Y \\ R_1[G] &= -S \end{aligned}\right\} \tag{92}$$

Then $\{dY\} = \|G\|^{k-\beta}\{dX_2\}$ and $\{dS\} = \|G\|^{\lambda+1}\{dR_1\}$ so that if we choose G such that $R_1[G] = -E_\lambda$, then in order to prove the theorem it is enough to prove

$$\int\limits_D \frac{\{dY\}}{\{dS\}} = \frac{\rho_k}{\rho_{k-\lambda}} \tag{93}$$

where we have written k instead of $k - \beta$ and D is the domain of Y with

$$Y'Y = S, \quad S = E_\lambda. \tag{94}$$

Note that (93) is a special case of the theorem we want to prove, namely with $V = E_k$, $W = E_\lambda, \beta = 0$.

In order to prove (93) we shall use induction on λ. Assume theorem 6 to have been proved for $\lambda - 1 \geq 1$. Let be an integer $0 < \beta < \lambda$. Put

$$Y = \left(Y_1^{(k,\beta)}, Y_2^{(k,\lambda-\beta)}\right)$$

and

$$S = \begin{pmatrix} T & Q \\ Q' & R \end{pmatrix}$$

where $T = Y_1'Y_1$, $Q = Y_1'Y_2$, $R = Y_2'Y_2$. Then $\{dY_1\}\{dY_2\} = \{dY\}$ and $\{dS\} = \{dT\}\{dQ\}\{dR\}$. Assume now Y_1 fixed. Varying Y_2 we get

$$\int\limits_D \frac{\{dY\}}{\{dS\}} = \int \frac{\{dY_1\}}{\{dT\}} \frac{\{dY_2\}}{\{dQ\}\{dR\}}$$

152 Induction hypothesis works and hence

$$\int\limits_D \frac{\{dY\}}{\{dS\}} = \frac{\rho_{k-\beta}}{\rho_{k-\lambda}} \int \frac{\{dY_1\}}{\{dT\}}$$

with $T = Y_1'Y_1$. Again induction hypothesis works since $0 < \beta < \lambda$ and we have

$$\int \frac{\{dY_1\}}{\{dT\}} = \frac{\rho_k}{\rho_{k-\beta}}$$

(93) is thus proved. In order to uphold induction, we have to prove (93) in case $\lambda = 1$ that is

$$\int\limits_D \frac{\{dX\}}{\{dt\}} = \frac{\rho_k}{\rho_{k-1}} \tag{95}$$

where D is the space

$$x_1^2 + \cdots + x_k^2 = t, \quad t = 1.$$

We now use induction on k. For $k = 1$, the proposition is trivial; so let $k > 1$ and (95) proved for $k-1$ instead of k. Introducing $x_1, \ldots, x_{k-1}$ as a coordinate system on D we get

$$\int \frac{\{d\underline{x}\}}{\{dt\}} = \int \frac{\{(dx_1 \ldots dx_{k-1})\}}{x_k}$$

since $2x_k\, dx_k = dt$ and we consider only positive values of x_k. Now

$$x_k = (1-u)^{\frac{1}{2}}$$

where $u = x_1^2 + \cdots + x_{k-1}^2$. We therefore have

$$\int \frac{\{dx_1 \ldots dx_k\}}{dt} = \int (1-u)^{-\frac{1}{2}} \frac{dx_1 \ldots dx_{k-1}}{du}$$

153 By induction hypothesis,

$$\frac{dx_1\dots dx_{k-1}}{du}=\frac{\pi^{\frac{k-1}{2}}}{\Gamma(\frac{k-1}{2})}u^{\frac{k-1}{2}-1}$$

Therefore we get

$$\int\frac{\{dx\}}{dt}=\int\limits_0^1(1-u)^{-\frac{1}{2}}u^{\frac{k-1}{2}-1}du\frac{\pi^{\frac{k-1}{2}}}{\Gamma(\frac{k-1}{2})}$$

Evaluating the beta integral, we get the result.

The case $\beta=0$ is also contained in the above discussion

Theorem 6 is now completely demonstrated.

6 Measure of unit group and measure of representation

Let S be the matrix of a non-degenerate real quadratic form with signature p, q, $(p+q=n)$. Let Ω denote the orthogonal group of S, hence the group of real matrices Y with

$$S[Y]=S,$$

Ω is then a locally compact group and there exists on Ω a left invariant Haar measure determined uniquely upto a positive multiplicative constant. Instead of Ω we shall consider the surface $\Omega(W)$ consisting of all solutions Y of the matrix equation

$$S[Y]=W,$$

where W is a fixed matrix, non-singular and of signature p, q. Clearly if Y_1 and Y_2 lie in $\Omega(W)$, $Y_1Y_2^{-1}\in\Omega$ so that $\Omega(W)$ consists of all CY where Y is a fixed solution of $S[Y]=W$ and C runs through all elements in Ω. According to the previous section, we can introduce on $\Omega(W)$, a volume 154 measure

$$\frac{\{dY\}}{\{dW\}}\tag{96}$$

The surface $\Omega(W)$ has the property that the orthogonal group of S acts as a transitive group of left translations

$$Y \to CY \tag{97}$$

$C \in \Omega$, on it. Also the measure (96) defined above is invariant under these left translations. Since Ω and $\Omega(W)$ are homeomorphic and $\Omega(W)$ is locally compact, (96) is the Haar measure in $\Omega(W)$ invariant under (97).

It is practical to consider on $\Omega(W)$ the measure

$$\|S\|^{-\frac{1}{2}}\|W\|^{\frac{1}{2}}\frac{\{dY\}}{\{dW\}} \tag{98}$$

instead of (96), for the following reason. (96) already has the invariance property under the transformations (97). Consider now the mapping

$$Y \to YP, \quad W \to W[P] \tag{99}$$

where P is an n-rowed non-singular matrix. Since $\{d(YP)\} = \|P\|^n\{dY\}$ and $\{dW[P]\} = \|P\|^{n+1}\{dW\}$, it follows that (98) remains unaltered by the transformations (99). Thus (98) is independent of W which means, we can choose for W a matrix suitable to us. In particular, if $W = S$, (98) gives the Haar measure on Ω required for our purposes.

Let now S be a rational matrix and $\mathfrak{H}$ the representation space of the unit group of S. The unit group $\Gamma(S)$ of S is a discrete subgroup of Ω **155** and is represented in $\mathfrak{H}$ by the mapping $H \to H[U]$, $H \in \mathfrak{H}$, $U \in \Gamma(S)$. We constructed in $\mathfrak{H}$ for $\Gamma(S)$ a fundamental domain F. By theorem 5 it follows that, if $S[x]$ is not a binary zero form,

$$V = \int_F dv < \infty \tag{100}$$

where dv is the invariant volume element in $\mathfrak{H}$.

$\Gamma(S)$ being a discrete subgroup of Ω, there exists a fundamental set F_0 for $\Gamma(S)$ in Ω. By means of the translation $Y \to CY$, $C \in \Omega$, we construct a fundamental set $\widetilde{F}$ for $\Gamma(S)$ in $\Omega(W)$. Let $\mu(S)$ denote

$$\mu(S) = \|S\|^{-\frac{1}{2}}\|W\|^{\frac{1}{2}}\int_{\widetilde{F}}\frac{\{dY\}}{\{dW\}} \tag{101}$$

It is to be noted that the value of $\mu(S)$ is independent of the way $\widetilde{F}$ is constructed. Since (98) is independent of W, $\mu(S)$ is actually the Haar measure of the fundamental set F_0 for $\Gamma(S)$ in Ω. We call $\mu(S)$, *the measure of the unit group* $\Gamma(S)$.

It is to be noticed that the mappings $H \to H[{}^{\pm}U]$ are identical in $\mathfrak{H}$, whereas for $U \in \Gamma(S)$, the representations $Y \to UY$ and $Y \to -UY$ are distinct in $\Omega(W)$. We now prove the important

Theorem 7. *If* $\rho_h = \prod\limits_{k=1}^{h} \dfrac{\pi^{k.2}}{\Gamma(k/2)}$, *then* $\mu(S)$ *and* V *are connected by the relation*

$$\boxed{2\mu(S) = \rho_p\rho_q \|S\|^{-(\frac{n+1}{2})} V}$$

provided S *is not the matrix of a binary zero form.* 156

Proof. In order to prove this we consider the homogeneous as well as the inhomogeneous parametrical representation of the $\mathfrak{H}$ space. In the homogeneous parametrization H in $\mathfrak{H}$ is given by

$$H = 2K - S, \quad K = T^{-1}[Z'S], \quad T = S[Z] > 0 \tag{102}$$

where $Z = Z^{(n,p)}$ is a real matrix. Z determines H uniquely, but H determines Z only upto a non-singular p-rowed matrix factor on the right. Let us put as before

$$S = S_0[C^{-1}]$$

where $S_0 = \left(\begin{smallmatrix} E_p & 0 \\ 0 & -E_q \end{smallmatrix}\right)$. Let

$$Z = C\begin{pmatrix} E_p \\ X \end{pmatrix} L$$

with $X = X^{(q,p)}$, and $|L| \neq 0$. The inhomogeneous parametrical representation is given by

$$T_0 = E - X'X > 0 \tag{103}$$

with X real and

$$T = T_0[L] \tag{104}$$

Let W be a symmetric n-rowed matrix of signature (p, q) and having the form

$$W = \begin{pmatrix} T & Q \\ Q' & R \end{pmatrix}, \quad T = T^{(p)} > 0 \tag{105}$$

$\Omega(W)$ will now be the space of solutions Y,

$$Y = (Y_1 Y_2), \quad Y_1 = Y_1^{(n,p)}, \quad Y_2 = Y_2^{(n,q)}$$

satisfying $W = S[Y]$ so that

$$T = S[Y_1] > 0, \quad Q = Y_1' S Y_2, \quad R = S[Y_2] \tag{106}$$

157 Every Y_1 satisfying (106) determines a H in $\mathfrak{H}$ uniquely by (102). Since with Y_1, UY_1 also is a solution where $U \in \Gamma(S)$, we construct a fundamental set $\widetilde{F}$ in $\Omega(W)$ to be the set consisting of those Y_1 for which the corresponding H determined by (102) lie in F, the fundamental domain for $\Gamma(S)$ in $\mathfrak{H}$. It is easy to verify that $\widetilde{F}$ is actually a fundamental set. □

Now

$$\{dY\} = \{dY_1\}\{dY_2\}; \quad \{dW\} = \{dT\}\{dQ\}\{dR\}.$$

Let Y_1 be fixed so that the corresponding H which it determines in in $\mathfrak{H}$ is in F. Let now Y_2 satisfy (106). We then have

$$Z\mu(S) = \|S\|^{-\frac{1}{2}} \|W\|^{\frac{1}{2}} \int\limits_{\tilde{F}(Y_1)} \frac{\{dY_1\}}{\{dT\}} \int\limits_{D} \frac{\{dY_2\}}{\{dQ\}\{dR\}}$$

where D is the domain determined by Y_2 satisfying (106) with Y_1 fixed and $\widetilde{F}(Y_1)$ is the set of Y_1 which determine H in F. For the inner integral we apply theorem 6 and so

$$2\mu(S) = \rho_q \|S\|^{-\frac{q+1}{2}} \|T\|^{\frac{1-q}{2}} \int\limits_{\widetilde{F}(Y_1)} \frac{\{dY_1\}}{\{dT\}}$$

Now X and L determine Y_1 uniquely, X satisfying (103) and L satisfying (104). Thus

$$\{dY_1\} = \|C\|^p \|L\|^q \{dX\}\{dL\}.$$

Expressing L in terms of T_0 (104), we get

$$2\mu(S) = \rho_q \|S\|^{-\frac{n+1}{2}} \|T\|^{\frac{1}{2}} \int\limits_{F(Y_1)} |T_0|^{-\frac{q}{2}} \frac{\{dX\}\{dL\}}{\{dT\}}$$

But since $T = T_0[L]$ we get 158

$$\int \frac{\{dL\}}{\{dT\}} = \rho_p \|T_0\|^{-p/2} \|T\|^{-\frac{1}{2}}$$

We therefore finally have the formula

$$2\mu(S) = \rho_p \rho_q \|S\|^{-\frac{n+1}{2}} \int |T_0|^{-\frac{n}{2}} \{dX\}$$

From the form of T_0 we see that the integral has the value V and our theorem is proved.

Let now S be the matrix of a non-degenerate, rational quadratic form of signature p, q so that $p+q = n$. Let t be a rational number represented by S so that

$$S[\underline{y}] = t \tag{107}$$

for an integral column $\underline{y}$. It is obvious that with $\underline{y}$, $U\underline{y}$ is also a solution of (107) where U is a unit of S. We shall associate with a given solution $\underline{y}$ of (107) a real number $\mu(\underline{y}, S)$ called the *measure of the representation* $\underline{y}$, which will allow us to generalize, later, to indefinite forms the notion of "number" of representations.

Let W be the real symmetric matrix of signature p, q given by

$$W = \begin{pmatrix} t & \underline{q}' \\ \underline{q} & R \end{pmatrix}$$

We consider all the real solutions $Y_0 = Y_0^{(n,n-1)}$ satisfying

$$S[Y] = W \tag{108}$$

where $Y = (\underline{y}\ Y_0)$. Let $\Omega(\underline{q}, R)$ be the surface determined by Y_0. Thus Y_0 satisfies

$$\underline{q}' = \underline{y}' S Y_0, \quad R = S[Y_0] \tag{109}$$

W being a fixed matrix. Clearly $\Omega(\underline{q}, R)$ is a locally compact topological space. 159

Let $\Omega(\underline{y})$ be the subgroup of the orthogonal group of S consisting of those matrices V in Ω with

$$V\underline{y} = \underline{y}. \tag{110}$$

Then $\Omega(\underline{y})$ is a locally compact topological group. Since with Y, VY for $V \in \Omega(\underline{y})$ is also a solution of (108), it follows that the mapping

$$Y_0 \to VY_0 \tag{111}$$

gives a representation of $\Omega(\underline{y})$ in $\Omega(\underline{q}, R)$. Clearly this representation is faithful. Also since W is fixed, the representation (111) of $\Omega(\underline{y})$ on $\Omega(\underline{q}, R)$ is transitive on $\Omega(\underline{q}, R)$. We introduce the volume element

$$||S||^{-\frac{1}{2}}||W||^{\frac{1}{2}} \frac{\{dY_0\}}{\{d\underline{q}\}\{dR\}} \tag{112}$$

which is clearly invariant under the mappings (111). Thus (112) gives the left invariant Haar measure in the locally compact space $\Omega(\underline{q}, R)$.

The volume element (112) introduced above has another property. Let P be a real matrix of the form

$$P = \begin{pmatrix} 1 & \underline{p}' \\ \underline{0} & P_0 \end{pmatrix}$$

where $|P_0| \neq 0$ so that P is non-singular. Consider the transformation

$$\left.\begin{aligned} Y &\to \underline{y}\underline{p}' + Y_0P_0 \\ W &\to W[P] \end{aligned}\right\} \tag{113}$$

160 Then

$$\begin{aligned} \{d(\underline{y}\underline{p}' + Y_0P_0)\} &= ||P_0||^n\{dY_0\} \\ \{dW[P_0]\} &= ||P_0||^{n+1}\{dW\} \end{aligned}$$

which shows that the transformations (113) leave (112) unaltered. Thus (112) is independent of W and we may therefore choose W a particular way suitable to us.

Put $Y = (\underline{y} Y_1 Y_2)$ where $Y_1 = Y_2^{(n,p)}$, $Y_2 = Y_2^{(n,q-1)}$ and write

$$W = \begin{pmatrix} W_1 & Q \\ Q' & R_1 \end{pmatrix}, \quad W_1 = W_1^{(p+1)}$$

where

$$W_1 = S[\underline{y} Y_1] = \begin{pmatrix} t & \underline{v}' \\ \underline{v} & T \end{pmatrix} \tag{114}$$

We now choose W so that

$$|W_1| \neq 0, \quad T > 0. \tag{115}$$

Since T has p rows and columns and S has signature p, q, it follows that W_1 has signature p, 1.

The subgroup $\Gamma(\underline{y})$ of units U of S with $U\underline{y} = \underline{y}$ is a discrete subgroup of $\Omega(\underline{y})$ and so the representation (111) with $\overline{V} \in \Gamma(\underline{y})$ is discontinuous in $\Omega(\underline{q}, R)$. Let $F(\underline{y})$ be a fundamental region in $\Omega(\underline{q}, R)$, for this discrete subgroup $\Gamma(\underline{y})$. We define the measure $\mu(\underline{y}, S)$ of the representation $\underline{y}$ by

$$\mu(\underline{y}, S) = \|S\|^{-\frac{1}{2}} \|W\|^{\frac{1}{2}} \int\limits_{\widetilde{F}(\underline{y})} \frac{\{dY_0\}}{\{d\underline{q}\}\{dR\}} \tag{116}$$

161 We shall first show how to construct the fundamental region $\widetilde{F}(\underline{y})$. Let Y be a solution of the equations (115), (114). According to (102), this determines uniquely a H in the $\mathfrak{H}$ space. If $U \in \Gamma(\underline{y})$, then UY_1 determines the point $H[U^{-1}]$ in $\mathfrak{H}$. Let $F(\underline{y})$ be the fundamental region in $\mathfrak{H}$ for the discrete subgroup $\Gamma(\underline{y})$ of $\Gamma(S)$, the unit group of S. This $F(\underline{y})$ can be constructed as follows: Let $\Gamma(S)$ be written as a union of left cosets modulo $\Gamma(\underline{y})$,

$$\Gamma(S) = \sum_i U_i \Gamma(\underline{y}).$$

Let F be the fundamental region for $\Gamma(S)$ in $\mathfrak{H}$. Let $F(\underline{y}) = \bigcup_i F(U_i)$. Then $F(\underline{y})$ is the required region. Since $Y_0 = (Y_1, Y_2)$ we define $\widetilde{F}(\underline{y})$ to be the set of Y_0 for which the Y_1 determines a point in $F(\underline{y})$. It can be easily verified that $\widetilde{F}(\underline{y})$ determined in this manner is a fundamental region for $\Gamma(\underline{y})$ in $\Omega(\underline{q}, R)$.

Because of (109) and (114) we may write,

$$\underline{q} = \begin{pmatrix} \underline{v} \\ \underline{v}_1 \end{pmatrix}, \quad R = \begin{pmatrix} T & T_1 \\ T_1' & R_1 \end{pmatrix} \tag{117}$$

Then

$$\{d\underline{q}\} = \{d\underline{v}\}\{d\underline{v}_1\}$$
$$\{dR\} = \{dT\}\{dT_1\}\{dR_1\}$$

162 Since $\{dY_0\} = \{dY_1\}\{dY_2\}$ we fix Y_1 so that the H that it determines in $\mathfrak{H}$ is in $F(\underline{y})$ and integrate over the space of Y_2 which clearly is determined by

$$S[\underline{y}Y_1, Y_2] = \begin{pmatrix} W_1 & Q \\ Q' & R_1 \end{pmatrix}$$

Since

$$\frac{\{dY_2\}}{\{dQ\}\{dR_1\}} = \frac{\{dY_2\}}{\{d\underline{v}_1\}\{dT_1\}\{dR_1\}}$$

We have, on using theorem 6,

$$\mu(\underline{v}, S) = \rho_{q-1}\|S\|^{-q/2}\|W_1\|^{1-\frac{q}{2}} \int \frac{\{dY_1\}}{\{d\underline{v}\}\{dT\}} \tag{118}$$

where the domain of integration is over those Y_1 which determine points H in $F(\underline{y})$. Since $T > 0$, (114) now gives

$$W = \begin{pmatrix} t - w & \underline{0}' \\ \underline{0} & T \end{pmatrix} \begin{bmatrix} 1 & \underline{0}' \\ T^{-1}\underline{v} & E \end{bmatrix}$$

where $w = T^{-1}[\underline{v}]$. Since $T > 0$ and W_1 has signature $(p, 1)$, it follows that

$$\left.\begin{aligned} w - t > 0 \\ w \geq 0 \end{aligned}\right\} \tag{119}$$

Substituting $|W_1| = (t-w)|T|$, we get from (118)

$$\mu(\underline{y}, S) = \rho_{q-1}\|S\|^{-q/2}\|T\|^{1-\frac{q}{2}}(w-t)^{1-\frac{q}{2}}\int\frac{\{dY_1\}}{\{d\underline{v}\}\{dT\}} \qquad (120)$$

We now remark that w depends only on the H which Y_1 determines in $\mathfrak{H}$. For,

$$w = T^{-1}[\underline{v}] = T^{-1}[Y_1'S\underline{y}] = T^{-1}[Y_1'S][\underline{y}]$$

But from (102), $T^{-1}[Y_1'S] = \dfrac{H+S}{2}$ so that 163

$$2W = (H+S)[\underline{y}] = H[\underline{y}] + t$$

or that

$$w = \frac{H[\underline{y}]+t}{2} \qquad (121)$$

Let now $g(w)$ be an integrable function of w to be chosen later. Multiply both sides of (120) by $g(w)\ (w-t)^{\frac{q}{2}-1}$ and integrate over the $\underline{v}$ space satisfying

$$T^{-1}[\underline{v}] = w > t.$$

We then get, by applying theorem 6 and using

$$\{d\underline{v}\} = \frac{\{d\underline{v}\}}{dw}\cdot dw$$

the result

$$\rho_1\mu(\underline{y}, S)\int\limits_{w>\mathrm{Max}(0,t)} g(w)w^{\frac{p}{2}-1}(\omega-t)^{\frac{q}{2}-1}dw$$
$$= \rho_{p-1}\rho_{q-1}\|S\|^{-q/2}\|T\|^{-\frac{1}{2}-\frac{q}{2}}\int g(w)\frac{\{dY_1\}}{\{dT\}} \qquad (122)$$

The function $g(w)$ has to be so chosen that the integrals are convergent. We will see later that this can be done. The domain of integration for the integral on the right of (122) is over that set of Y_1 which determine H in $F(\underline{y})$. Since every H in $F(\underline{y})$ determines a Y_1, we see that we have to apply the analysis in the proof of theorem 7 to obtain

Theorem 8. *Let $\mu(\underline{y}, S)$ be the measure of the representation $\underline{y}$ of $S[\underline{y}] = t$. Then*

$$\mu(\underline{y}, S)\int\limits_{\substack{w>0\\w>t}} g(w)w^{\frac{p}{2}-1}(w-t)^{\frac{q}{2}-1}dw = \rho_{p-1}\rho_{q-1}\|S\|^{-\frac{n}{2}}\int\limits_{F(\underline{y})} g\left(\frac{H[\underline{y}]+t}{2}\right)dv$$

164 *where $g(w)$ is an integrable function making the integrals converge and dv is the invariant volume element in the $\mathfrak{H}$ space.*

7 Integration of the theta series

We shall hereafter assume that $n > 4$.

Let us denote by $V_{\underline{a}}$ the volume of the fundamental region $V_{\underline{a}}$ for $\Gamma_{\underline{a}}$ in the $\mathfrak{H}$ space so that

$$V_{\underline{a}} = \int\limits_{F_{\underline{a}}} dv.$$

$F_{\underline{a}}$ is finite by theorem 5. We put

$$\varphi_{\underline{a}}(z) = V_{\underline{a}}^{-1}\int\limits_{F_{\underline{a}}} f_{\underline{a}}(z, H)dv \tag{123}$$

Then by theorem 4,

$$\varphi_{\underline{a}}(z) = V_{\underline{a}}^{-1}\sum_{\underline{y}\equiv\underline{a}(\text{mod } 1)}\int\limits_{F_{\underline{a}}} e^{2\pi i R[\underline{y}]}dv$$

If $\underline{a} \equiv \underline{0}(\text{mod } 1)$, then $\underline{y} \equiv \underline{0}$ is a possible value of $\underline{y}$ and then we have the term

$$V_{\underline{a}}^{-1}\int\limits_{F_{\underline{a}}} dv \tag{124}$$

By definition of $V_{\underline{a}}$, the value of (124) is unity. Let us therefore put

$$\gamma_{\underline{a}} = \begin{cases} 1 & \text{if } \underline{a} \equiv \underline{0}(\text{mod } 1) \\ 0 & \text{otherwise.} \end{cases}$$

165 Then we have

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + V_{\underline{a}}^{-1} \sum_{\substack{\underline{y}\equiv\underline{a}(\text{mod } 1)\\ \underline{y}\neq\underline{0}}} \int_{F_{\underline{a}}} e^{2\pi i R[\underline{y}]} dv.$$

Let us call two rational vectors $\underline{y}_1$ and $\underline{y}_2$ *associated*, if there exists a matrix U in $\Gamma_{\underline{a}}$ such that $\underline{y}_1 = U\underline{y}_2$. Otherwise they are said to be non-associated. For any $\underline{y}$ consider the subgroup $\Gamma_{\underline{a}}(\underline{y})$ of $\Gamma_{\underline{a}}$ with $U\underline{y} = \underline{y}$. We can write

$$\Gamma_{\underline{a}} = \sum_k U_k \Gamma_{\underline{a}}(\underline{y}) \tag{125}$$

as a union of left cosets. $U_k\underline{y}$ then run through all vectors associated with $\underline{y}$. Because of uniform convergence, we can write

$$\varphi_a(z) = \gamma_a + {\sum_{\underline{y}}}' \sum_k V_{\underline{a}}^{-1} \int_{F_{\underline{a}}} e^{2\pi i R[U_k\underline{y}]} dv$$

where the accent indicates that we should sum over all non-associate vectors $\underline{y}$ with $\underline{y} \neq \underline{0}$ and $y \equiv \underline{a}(\text{mod } 1)$. Since the volume element dv has the invariance property we may write

$$\varphi_{\underline{a}}(z) = \gamma_a + {\sum_{\underline{y}}}' \sum_k V_{\underline{a}}^{-1} \int_{F_{\underline{a}}[U_k]} e^{2\pi i R[\underline{y}]} dv$$

where $F_{\underline{a}}[U_k]$ is the image of the fundamental region $F_{\underline{a}}$ by the transformation $H \to H[U_k]$. Because of (125) a fundamental region $F(\underline{y})$ for $\Gamma_{\underline{a}}(\underline{y})$ in $\mathfrak{H}$ is given by

$$F(\underline{y}) = \sum_k F_{\underline{a}}[U_k].$$

Consider the group $\Gamma_{\underline{a}}(y)$. Now $-E$ is not an element of $\Gamma_{\underline{a}}(\underline{y})$ since that means $-\underline{y} = \underline{y}$ or $\underline{y} = \underline{0}$. But $-E$ may be in $\Gamma_{\underline{a}}$. This means that 166 $-\underline{a} \equiv \underline{a}(\text{mod } 1)$ or $2\underline{a} \equiv \underline{0}(\text{mod } 1)$. In this the U_k's in (125) may be so chosen that with U_k, $-U_k$ is also a representative of a coset. Since

$H \to H[U]$ and $H \to H[-U]$ define the same mapping in the $\mathfrak{H}$ space, it shows that if $2\underline{a} \equiv \underline{0} (\mathrm{mod}\ 1)$, the $F_{\underline{a}}[U_k]$ give a double covering of the fundamental region $F(\underline{y})$. So let us define

$$j_{\underline{a}} = \begin{cases} 2 & \text{if } 2\underline{a} \equiv \underline{0} (\mathrm{mod}\ 1) \\ 1 & \text{if } 2\underline{a} \not\equiv \underline{0} (\mathrm{mod}\ 1). \end{cases}$$

Then we can write

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + j_{\underline{a}} \sum_{\underline{y}} V_{\underline{a}}^{-1} \int\limits_{F(\underline{y})} e^{2\pi i R[\underline{y}]} dv$$

Let us now put in theorem 8

$$g(w) = e^{2\pi i t\bar{z} - 4\pi\eta w}$$

and use the abbreviation

$$h_t(z) = e^{2\pi i t\bar{z}} \int\limits_{w > \max(0,t)} w^{\frac{p}{2}-1} (w-t)^{\frac{q}{2}-1} e^{-4\pi\eta w} dw, \tag{126}$$

then, since (126) converges for $p > 0$, $q > 0$, $\eta > 0$, we get

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{j_{\underline{a}}}{V_{\underline{a}}} \sum_{\underline{y}}{}' \frac{\mu(\underline{y}, S)}{\rho_{p-1}\rho_{q-1}} \|S\|^{\frac{n}{2}} h_t(z) \tag{127}$$

It can be shown that for each rational number $t \neq 0$, the number of non-associate representations

$$S[\underline{y}] = t, \quad \underline{y} \equiv \underline{a} (\mathrm{mod}\ 1)$$

is finite. If $t = 0$, one has to consider only non-associate primitive representations. If therefore we put

$$M(S, \underline{a}, t) = \sum_{\underline{y}} \mu(\underline{y}, S) \tag{128}$$

167 where the summation runs on the right through the finitely many non-

associate representations of $S[\underline{y}] = t$, we can write (127) in the form

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{j_{\underline{a}}}{V_{\underline{a}}} \frac{\|S\|^{n/2}}{\rho_{p-1}\rho_{q-1}} \sum_{t \equiv S[\underline{a}] (\text{mod } 1)} M(S, \underline{a}, t) h_t(z) \tag{129}$$

Just as we defined $\mu(S)$ in theorem 7 for the unit group Γ, we can define $\mu_{\underline{a}}(S)$ for the subgroup $\Gamma_{\underline{a}}$ of Γ also. $\Gamma_{\underline{a}}$ is a subgroup of finite index $(\Gamma : \Gamma_{\underline{a}})$ in Γ. Let

$$\Gamma = \sum_U U\Gamma_{\underline{a}}$$

be a decomposition of Γ into left cosets mod $\Gamma_{\underline{a}}$. If F is a fundamental region for Γ, then

$$F_{\underline{a}} = \sum_U F[U]$$

is a fundamental region for $\Gamma_{\underline{a}}$. Since U and $-U$ give rise to the same mapping in $\mathfrak{H}$ space, we have to consider whether $-E$ belongs to $\Gamma_{\underline{a}}$; i.e., $2\underline{a} \equiv \underline{0} (\text{mod } 1)$, which means that U and $-U$ are in the same coset and so

$$V_{\underline{a}} = (\Gamma : \Gamma_{\underline{a}})V.$$

If however $2\underline{a} \not\equiv \underline{0} (\text{mod } 1)$, then U and $-U$ belong to different cosets and so $\sum\limits_U F[U]$ gives a double covering of $F_{\underline{a}}$. Thus

$$V_{\underline{a}} = \frac{1}{2}(\Gamma : \Gamma_{\underline{a}})V.$$

Using the definition of $j_{\underline{a}}$ we get

$$\frac{\mu(S)}{V} = \frac{\mu_{\underline{a}}(S)}{V_{\underline{a}}} \cdot \frac{j_{\underline{a}}}{2}$$

If we denote $\mu(S, \underline{a}, t)$ the quantity **168**

$$\mu(S, \underline{a}, t) = \frac{M(S, \underline{a}, t)}{\mu_{\underline{a}}(S)} \tag{130}$$

We get, on using theorem 7, the final formula

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{n/2}||S||^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_{t \equiv S[\underline{a}](\text{mod } 1)} \mu(S, \underline{a}, t) h_t(z) \tag{131}$$

We call $M(S, \underline{a}, t)$ the *measure of representation* of t by $S[\underline{x} + \underline{a}]$. (131) is the analogue, for indefinite forms, of the generating function (6).

Let us now consider the functional vector

$$\varphi(z) = \begin{pmatrix} \varphi_{\underline{a}_1}(z) \\ \vdots \\ \varphi_{\underline{a}_l}(z) \end{pmatrix} \tag{132}$$

$\underline{a}_1, \ldots, \underline{a}_l$ having the same meaning as before. Let $d = abs 2S$ and $\overline{\Gamma}$ the subgroup of units U in Γ satisfying

$$U \equiv E (\text{mod d}).$$

Since for every $\underline{a}_i$, $2S\underline{a}_i$ is an integral vector, it follows that

$$\Gamma_{\underline{a}_i} \supset \overline{\Gamma}, \quad (i = 1, 2, \ldots, 1).$$

Also $\Gamma/\overline{\Gamma}$ is a finite group. If F_0 is a fundamental region for $\overline{\Gamma}$ in $\mathfrak{H}$ and $\overline{V}$ its volume then because of invariance of volume element we have

$$\varphi_{\underline{a}}(z) = \overline{V}^{-1} \int_{F_0} f_{\underline{a}}(z, H) dv.$$

Let now $\underline{\mu}(S, t)$ and $\underline{\gamma}$ denote the vectors

$$\underline{\mu}(S, t) = \begin{pmatrix} \mu(S, \underline{a}_1, t) \\ \vdots \\ \mu(S, \underline{a}_1, t) \end{pmatrix}, \quad \underline{\gamma} = \begin{pmatrix} \gamma_{\underline{a}_l} \\ \vdots \\ \gamma_{\underline{a}_l} \end{pmatrix}$$

169 where $\mu(S, \underline{a}, t)$ is defined by (130) and $\gamma_{\underline{a}} = 0$ or 1 according as $\underline{a} \equiv \underline{0}(\text{mod } 1)$ or not. Then from (49) and (131) we have the

Theorem 9. *Let $n > 4$ and $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ be a modular matrix. Then*

$$\varphi(z) = \underline{\gamma} + \frac{\pi^{n/2} \| S \|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_{t} \underline{\mu}(S, t) h_t(z)$$

satisfies

$$\underline{\varphi}(z_M) = G(M, z)\underline{\varphi}(z).$$

The function $h_t(z)$ introduced in (126) can be expressed in terms of the confluent hypergeometric function $h(\alpha, \beta, \eta)$ defined by

$$h(\alpha, \beta, \eta) = \int_0^\infty w^{\alpha - 1}(w + 1)^{\beta - 1} e^{-w\eta})dw$$

where α and β are complex numbers with positive real parts and η is a positive real parameter. $h(\alpha, \beta, \eta)$ is a solution of the second order differential equation

$$\eta \frac{d^2 h}{d\eta^2} + (\alpha + \beta + \eta)\frac{dh}{d\eta} - \alpha h = 0.$$

From the definition of $h_t(z)$ we have

$$h_0(z) = \int_0^\infty W^{n/2 - 2} e^{-4\pi w \eta} dw$$

which reduces to the Γ-integral. We have hence

$$h_0(z) = (4\pi\eta)^{1 - n/2} \Gamma\left(\frac{n}{2} - 1\right). \tag{134}$$

Let now $t < 0$. Changing, in (126), the variable w to $-tw$ we get easily

$$h_t(z) = e^{2\pi i t \bar{z}} (-t)^{n/2 - 1} h(p/2, q/2, -4\pi t\eta) \tag{135}$$

170

In case $t > 0$, we make a change of variable $w \to wt + t$. One then obtains

$$h_t(z) = e^{2\pi i t \bar{z}} t^{n/2 - 1} h(q/2, p/2, 4\pi t\eta). \tag{136}$$

If we put $h_t(z) = u(\xi, \eta) = u$ as a function of the two real variables ξ and η, then u satisfies the partial differential equation

$$\Delta u = 0 \tag{137}$$

where

$$\Delta = \eta\left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right) + \frac{n}{2}\frac{\partial}{\partial\eta} + i\frac{(q-p)}{2}\frac{\partial}{\partial\xi} \tag{138}$$

The interesting fact to be noticed is that the differential operator Δ is *independent* of t. Since $\underline{\varphi}(z)$ in theorem 9 is a linear function in $h_t(z)$ we see that

$$\boxed{\Delta\underline{\varphi}(z) = \underline{0}} \tag{139}$$

It is to be noted also that $\underline{f}(z, H)$ is *not* a solution of the differential equation (137).

8 Eisenstein series

Let $M = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ be a modular matrix. By (30), for any two vectors $\underline{a}$, $\underline{b}$ among $\underline{a}_1, \ldots, \underline{a}_l$ we have

$$\lambda_{\underline{ab}}(M) = \sum_{\underline{g}((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}(\alpha S[\underline{g}+\underline{a}]-2\underline{b}'S(\underline{g}+\underline{a})+\delta S[\underline{b}])}$$

Let us consider the case $\underline{b} = \underline{0}$ which is a possible value of $a_1, \ldots, a_l$. Then

$$\lambda_{\underline{a},\underline{0}}(M) = \sum_{\underline{g}((\text{mod})\ \gamma)} e^{\frac{2\pi i}{\gamma}\alpha S[\underline{g}+\underline{a}]}$$

171 which is an ordinary Gaussian sum. It is to be noted that $\lambda_{\underline{a},\underline{0}}(M)$ depends only on the first column of the matrix M and is independent of β and δ.

Let G denote the group of proper unimodular matrices and G_0 the subgroup consisting of all modular matrices with $\gamma = 0$. Let

$$G = \sum_M MG_0 \tag{141}$$

be a decomposition of G as a sum of left cosets modulo G_0. If M and M_1 belong to the same left coset, then

$$M_1^{-1}M = \begin{pmatrix} \pm 1 & * \\ 0 & \pm 1 \end{pmatrix}$$

so that the values of $\lambda_{\underline{a},\underline{0}}(M)$ and $\lambda_{\underline{a},\underline{0}}(M_1)$ are equal. Also since G_0 contains the matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, we may choose the representatives in (141) so that $\gamma \geq 0$.

Let $\underline{g}(M, z)$ denote the first column of the matrix $G(M, z)$ defined in (37). Let M have $\gamma > 0$. Then because of theorem 2 we have

$$\left.\begin{aligned} \underline{g}(M, z_{M^{-1}}) &= \epsilon^{-1} d^{-\frac{1}{2}} \gamma^{-n/2} (\gamma z - \alpha)^{-p/2} (\gamma \bar{z} - \alpha)^{-q/2} \begin{pmatrix} \lambda_{\underline{a}_1} & \underline{0} \\ \vdots & \\ \lambda_{\underline{a}_l} & \underline{0} \end{pmatrix} \\ \underline{g}(E, z) &= \gamma \end{aligned}\right\} \tag{142}$$

We now form the series

$$\underline{\psi}(z) = \sum_M \underline{g}(M, z_{M^{-1}})$$

the sum taken over all representatives in (141). $\underline{\psi}(z)$ is a vector of functions $\psi_{\underline{a}_1}(z), \ldots, \psi_{\underline{a}_l}(z)$ where 172

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \epsilon^{-1} d^{-\frac{1}{2}} \sum_{\substack{(\alpha,\gamma)=1 \\ \gamma>0}} \gamma^{-n} \left(z - \frac{\alpha}{\gamma}\right)^{-p/2} \left(\bar{z} - \frac{\alpha}{\gamma}\right)^{-q/2} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \tag{143}$$

In order to prove the absolute convergence of the above series for $\psi_{\underline{a}}(z)$, observe that by (47) and (48), $\lambda(M)$ is unitary and so it is enough to prove the convergence of

$$\sum_{(\alpha,\gamma)=1} |\gamma_z - \alpha|^{-\frac{n}{2}}$$

It is well-known that this converges for $n > 4$. The convergence is even uniform in every compact subdomain of the z-plane.

From theorem 2 we have

$$\underline{g}(MM_1, z_{M_1^{-1}}) = G(M,z)\underline{g}(M_1, z_{M_1^{-1}})$$

If M is fixed and M_1 runs through a complete system of representatives in (141), then MM_1 also runs through the representatives in (141). This gives

$$\underline{\psi}(z_M) = G(M,z)\underline{\psi}(z). \tag{144}$$

Thus $\underline{\psi}(z)$ also satisfies the same transformation formula as $\underline{\varphi}(z)$. We shall now obtain a fourier expansion for the function $\psi_{\underline{a}}(Z)$. To this end we first prove

Lemma 4. *Let $a > 1$, $b > 1$* be two real numbers and

$$E(z) = \sum_{k=-\infty}^{\infty} (z-k)^{-a}(\overline{z}-k)^{-b}$$

Then

$$E(z) = \frac{i^{b-a}(2\pi)^{a+b}}{\Gamma(a)\Gamma(b)} \sum_{-\infty}^{\infty} e^{2\pi i l \overline{z}} \int\limits_{\max(0,1)}^{\infty} u^{a-1}(u-1)^{b-1}e^{-4\pi\eta u}du.$$

173 *where $z = \xi + i\eta$, $\eta > 0$.*

Proof. Since $a + b > 2$, it follows that $E(z)$ is absolutely convergent and is also uniformly convergent in every bounded domain of the z-plane. It is clearly periodic of period 1 in ξ. Hence

$$E(z) = \sum_{l=-\infty}^{\infty} e^{2\pi i l\xi}\left(\int\limits_0^1 \sum_{-\infty}^{\infty}(z-k)^{-a}(\overline{z}-k)^{-b}e^{-2\pi i l\xi}d\xi\right).$$

This shows that the fourier coefficient equals

$$\int\limits_{-\infty}^{\infty} z^{-a}\overline{z}^{-b}e^{-2\pi i l\xi}d\xi.$$

By means of the substitution $\xi \to -i\xi$ we get for this fourier coefficient the integral

$$i^{b-a-1}\int\limits_{-i\infty}^{i\infty}(\eta-\xi)^{-a}(\eta+\xi)^{-b}e^{-2\pi l\xi}d\xi.$$

We now write ξ instead of $\xi+\eta$ obtaining thus the integral

$$i^{b-a-1}e^{2\pi i\eta}\int\limits_{\eta-i\infty}^{\eta+i\infty}\xi^{-b}(2\eta-\xi)^{-a}e^{-2\pi l\xi}d\xi.$$

In order to evaluate the integral above we use the Γ-integral and obtain

$$\frac{1}{\Gamma(a)}\int\limits_{\eta-i\infty}^{\eta+i\infty}e^{-2\pi l\xi}\xi^{-b}\left(\int\limits_{0}^{\infty}u^{a-1}e^{-(2\eta-\xi)u}du\right)d\xi \tag{145}$$

We can change the order of integration and hence the above integral equals

$$\frac{1}{\Gamma(a)}\int\limits_{0}^{\infty}u^{a-1}e^{-2\eta u}\left(\int\limits_{\eta-i\infty}^{\eta+i\infty}\xi^{-b}e^{\xi(u-2\pi l)}d\xi\right)du$$

We now use the well-known Weierstrass' formula in Γ-functions, 174 namely

$$\frac{1}{2\pi i}\int\limits_{c-i\infty}^{c+i\infty}x^{-b}e^{\lambda x}dx=\begin{cases}\frac{\lambda^{b-1}}{\Gamma(b)} & \text{if } \lambda>0\\ 0 & \text{if } \lambda\leq 0\end{cases}$$

where $c>0$, $b>0$. □

From this formula, it follows that the integral in (145) equals

$$\frac{2\pi i}{\Gamma(a)\Gamma(b)}\int\limits_{u>\max(0,2\pi,l)}u^{a-1}(u-2\pi l)^{b-1}e^{-2\eta u}du.$$

We once again make a change of variable u to $2\pi u$. We then obtain the fourier coefficient as given in the lemma.

Actually the lemma can be seen to be true for $a > 0$, $b > 0$ and $a + b \succ 1$. In particular, if we put $a = p/2$ and $b = q/2$ and use the definition of $h_t(z)$ in (126) and ϵ in (29), we obtain the formula

$$\sum_{k=-\infty}^{\infty}(z-k)^{-p/2}(\overline{z}-k)^{-q/2} = \frac{(2\pi)^{n/2}\varepsilon}{\Gamma(p/2)\Gamma(q/2)}\sum_{1=-\infty}^{\infty}h_l(z) \qquad (146)$$

Let us now consider the expression for $\psi_{\underline{a}}(z)$, namely

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \epsilon^{-1}d^{-\frac{1}{2}}\sum_{\substack{(\alpha,\gamma)=1\\ \gamma>0}}\gamma^{-m}\left(z-\frac{\alpha}{\gamma}\right)^{-p/2}\left(\overline{z}-\frac{\alpha}{\gamma}\right)^{-q/2}\cdot\sum_{\underline{g}((\text{mod})\ \gamma)}e^{2\pi i\frac{\alpha}{\gamma}S[\underline{g}+\underline{a}]}$$

Put $D = 2d$. We shall prove that $\psi_{\underline{a}}(z)$ has the period D in ξ.

$$\psi_{\underline{a}}(z+D) = \gamma_{\underline{a}} + \epsilon^{-1}d^{-\frac{1}{2}}\sum_{(\alpha,\gamma)=1}\gamma^{-n}\left(z+D-\frac{\alpha}{\gamma}\right)^{-p/2}\left(\overline{z}+D-\frac{\alpha}{\gamma}\right)^{-q/2}$$
$$\gamma > 0\sum_{\underline{g}((\text{mod})\ \gamma)}e^{2\pi i\frac{\alpha}{\gamma}S[\underline{g}+\underline{a}]}$$

175 But since $2S\,\underline{a}$ is integral, it follows that $DS[\underline{a}]$ is an integer. Hence

$$DS[\underline{g}+\underline{a}] \equiv 0(\text{mod } 1)$$

We may therefore write

$$\psi_{\underline{a}}(z+D) = \gamma_{\underline{a}} + \epsilon^{-1}d^{-\frac{1}{2}}\sum_{(\alpha,\gamma)=1}\gamma^{-n}\left(z+D-\frac{\alpha}{\gamma}\right)^{-p/2}\left(\overline{z}+D-\frac{\alpha}{\gamma}\right)^{-q/2}$$
$$\gamma > 0\cdot\sum_{\underline{g}((\text{mod})\ \gamma)}e^{2\pi i}\left(\frac{\alpha}{\gamma}-D\right)S[\underline{g}+\underline{a}]$$

Since $\frac{\alpha}{\gamma} + D$ runs through all rational fractions when $\frac{\alpha}{\gamma}$ does so, we see that

$$\psi_{\underline{a}}(z+D) = \psi_{\underline{a}}(z).$$

Because of absolute convergence we can write the series for $\psi_{\underline{a}}(z)$ in the following way: We put all rational numbers $\dfrac{\alpha}{\gamma}$ into residue classes modulo D. If $0 \le \dfrac{\alpha}{\gamma} < D$ is a fixed rational number with $(\alpha, \gamma) = 1$, then all rational numbers in the class of $\dfrac{\alpha}{\gamma}$ are obtained in the form $\dfrac{\alpha}{\gamma} \pm kD$ where k runs through integers. Thus

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \epsilon^{-1} d^{-\frac{1}{2}} \sum_{0\le\frac{\alpha}{\gamma}<D} \gamma^{-n} D^{-\frac{n}{2}} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \sum_{k=-\infty}^{\infty} (\zeta - k)^{-\frac{p}{2}} (\bar{\zeta} - k)^{-\frac{q}{2}} \tag{147}$$

where $\zeta = \left(z - \dfrac{\alpha}{\gamma}\right) D^{-1}$. Using (146) we get

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{d^{-\frac{1}{2}} (2\pi)^{n/2}}{\Gamma(p/2)\Gamma(q/2)} \sum_{0\le\frac{\alpha}{\gamma}<D} \gamma^{-n} D^{-n/2} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \sum_{l=-\infty}^{\infty} h_l \left(D^{-1}\left(z - \frac{\alpha}{\gamma}\right)\right) \tag{148}$$

176 Using (126) we have

$$h_l \left(D^{-1}\left(z - \frac{\alpha}{\gamma}\right)\right) = D^{n/2-1} e^{-2\pi i \frac{\alpha}{\gamma}\frac{l}{D}} h_{l/D}(z).$$

We may therefore write (148) in the form

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{n/2} \|S\|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_{l=-\infty}^{\infty} h_{1/D}(z) \sum_{0\le\frac{\alpha}{\gamma}<D} D^{-1}\gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \left(S[\underline{g}+\underline{a}] - \frac{l}{D}\right)\frac{\alpha}{\gamma}}$$

We now contend that the inner sum is zero if

$$S[\underline{a}] - \frac{l}{D} \not\equiv 0 (\text{mod } 1)$$

For, from (147), it is obvious that instead of the summation over $0 \leq \dfrac{\alpha}{\gamma} < D$, we could equally well have the summation range as $1 \leq \dfrac{\alpha}{\gamma} < D + 1$. This means that the expression

$$\sum_{0 \leq \frac{\alpha}{\gamma} < D} D^{-1} \gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma}\left(S[\underline{g}+\underline{a}] - \frac{l}{D}\right)} \tag{149}$$

is unaltered by changing $\dfrac{\alpha}{\gamma}$ into $\dfrac{\alpha}{\gamma} + 1$. But this change multiplies (149) by

$$e^{2\pi i\left(S[\underline{a}] - \frac{l}{D}\right)}$$

This proves our contention. We can therefore write

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{n/2} \|S\|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_{t \equiv S[\underline{a}](\text{mod } 1)} h_t(z)$$
$$\sum_{0 \leq \frac{\alpha}{\gamma} < D} D^{-1} \gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma}(S[\underline{g}+\underline{a}] - t)}$$

177 We can now write all numbers $0 \leq \dfrac{\alpha}{\gamma} < D$ in the form $\dfrac{\alpha}{\gamma} + r$ where $0 \leq \dfrac{\alpha}{\gamma} < 1$ and $r = 0, 1, 2, \ldots, D - 1$. Because of the property of the expression (149) we get finally the

Theorem 10. *The function $\psi_{\underline{a}}(z)$ has the expansion*

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{n/2} \|S\|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_{t \equiv S[\underline{a}](\text{mod } 1)} h_t(z)$$
$$\sum_{0 \leq \frac{\alpha}{\gamma} < 1} \gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma}(S[\underline{g}+\underline{a}] - t)}$$

The expression

$$\beta_t = \sum_{0 \leq \frac{\alpha}{\gamma} < 1} \gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma}(S[\underline{g}+\underline{a}] - t)} \tag{150}$$

is a so-called 'singular series'. Series of this type were studied by Hardy and littlewood in their researches on Waring's problem. We shall now give some properties of this singular series.

Let $q > 0$ be an integer. Put

$$f_q = \sum_{\gamma|q} \left(\sum_{0\leq\frac{\alpha}{\gamma}<1} \gamma^{-n} \sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma}(S[\underline{g}+\underline{a}]-t)} \right)$$

Since $q = \gamma s$ where s is an integer, we may take the inner summation over a complete residue system modulo q. Then each of the terms will be repeated s^n times. This gives

$$f_q = q^{-n} \sum_{\lambda=0}^{q-1} \sum_{\underline{g}(\text{mod q})} e^{\frac{2\pi i\lambda}{q}(S[\underline{g}+a]-t)}$$

Interchanging the two summations above we have

$$f_q = q^{-n} \sum_{\underline{g}(\text{mod q})} \left(\sum_{\lambda=0}^{q-1} e^{\frac{2\pi i\lambda}{q}(S[\underline{g}+\underline{a}]-t)} \right) \tag{151}$$

Because of the well-known formula 178

$$\sum_{\mu=0}^{q-1} e^{\frac{2\pi i\lambda\mu}{q}} = \begin{cases} 0 & \text{if } q \nmid \lambda \\ q & \text{if } q|\lambda \end{cases}$$

We see that the inner sum in (151) vanishes if the congruence

$$S[\underline{x} + \underline{a}] \equiv t(\text{mod q}) \tag{152}$$

has no solution. If it has a solution $\underline{g}$, then the inner sum has the value q. Thus

$$f_q = \frac{A_0(S, \underline{a}, t)}{q^{n-1}} \tag{153}$$

where $A_q(S, \underline{a}, t)$ is the number of incongruent solutions mod q of the congruence (152). It will then follow from the definition of β_t that if

$q \to \infty$ through a sequence of integers q_1, q_2, $q_3, \ldots$ such that every natural number divides all but a finite number of these q's,

$$\beta_t = \lim_{q\to\infty} f_q = \lim_{q\to\infty} \frac{A_q(S, \underline{a}, t)}{q^{n-1}} \tag{154}$$

From the definition of $A_q(S, \underline{a}, t)$ and the Chinese-remainder theorem, it follows that

$$A_q(S, \underline{a}, t) \cdot A_{q'}(S, \underline{a}, t) = A_{qq'}(S, \underline{a}, t)$$

for two coprime integers q, q'. This shows that $f_q = \dfrac{A(S, \underline{a}, t)}{q^{n-1}}$ is a multiplicative arithmetic function. In order to compute f_q for a given q, it is enough to compute f_q for $q = p^l$ where p is a prime number and $l > 0$ is an integer.

If $q = p^l$, $l > 0$ and p a prime number, it can be shown that

$$\delta_p(S, \underline{a}, t) = \lim_{l\to\infty} \frac{A_q(S, \underline{a}, t)}{q^{n-1}}$$

179 exists. In fact, if l is sufficiently large the value of $\dfrac{A_q(S, \underline{a}, t)}{q^{n-1}}$ is independent of l. This shows that $\delta_p(S, \underline{a}, t)$ is really a rational number. Furthermore for all except a finite number of primes (for instance, for $p \nmid 2d$)

$$\delta_p(S, \underline{a}, t) = \frac{A_p(S, \underline{a}, t)}{p^{n-1}}.$$

this enables us to compute $\delta_p(S, \underline{a}, t)$ for almost all primes p. From the fact that $\delta_p(S, \underline{a}, t)$ exists for every p one can construct the product

$$\delta(S, \underline{a}, t) = \prod_p \delta_p(S, \underline{a}, t).$$

It is proved in the analytic theory of quadratic forms that the product above converges and is different from zero only if every factor is different from zero. Moreover

$$\beta_t = \delta(S, \underline{a}, t) \tag{155}$$

This gives an arithmetical meaning for β_t namely that $\beta_t > 0$ if and only if $A_q(S, \underline{a}, t) \neq 0$ for *every integer* $q > 1$. For a proof of these statements see [2].

It should be noticed that since $\psi_{\underline{a}}(z)$ is a linear function in $h_t(z)$ and as $h_t(z)$ is a solution of the equation (137), the function $\psi_{\underline{a}}(z)$ and hence the vector $\underline{\psi}(z)$ defined in (143) is a solution of the differential equation

$$\Delta\underline{\psi}(z) = \underline{0}. \tag{156}$$

The series $\psi_{\underline{a}}(z)$ are called the Eisenstein series. The vector $\underline{\psi}(z)$ of Eisenstein series and the vector $\underline{\varphi}(z)$ satisfy the same differential equation and have the same transformation formula with regard to modular substitutions. 180

9 Main Theorem

We shall now prove the main theorem of the analytic theory of indefinite quadratic forms, namely,

Theorem 11. *If $n > 4$ and S is a rational symmetric matrix which is semi-integral, of signature p, q, $p + q = n$, $pq > 0$ and $\underline{a}$ a rational vector with $2S\underline{a}$ integral, then for $t \equiv S[\underline{a}] (\mathrm{mod}\ 1)$*

$$M(S, \underline{a}, t) = \mu_{\underline{a}}(S) \prod_p \delta_p(S, \underline{a}, t)$$

the product running over all primes p.

Proof. The series

$$\varphi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{\frac{n}{2}} \|S\|^{-1/2}}{\Gamma(p/2)\Gamma(q/2)} \sum_t \frac{M(S, \underline{a}, t)}{\mu_{\underline{a}}(S)} h_t(z)$$

and

$$\psi_{\underline{a}}(z) = \gamma_{\underline{a}} + \frac{\pi^{n/2} \|S\|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_t \beta_t h_t(z)$$

are fourier series in the real part ξ of z. In order to prove theorem 11, it is enough to prove that

$$\varphi_{\underline{a}}(z) - \psi_{\underline{a}}(z) = 0. \tag{157}$$

Then from (155) and the uniqueness theorem of fourier series, it would follow that the coefficients of $\varphi_{\underline{a}}(z) - \psi_{\underline{a}}(z)$ are zero and the theorem is proved. □ 181

We shall therefore prove (157).

Let $\underline{\chi}(z)$ be the vector

$$\underline{\chi}(z) = \begin{pmatrix} \chi_{\underline{a}_1}(z) \\ \vdots \\ \chi_{a_l}(z) \end{pmatrix}$$

where

$$\chi_{\underline{a}}(z) = \varphi_{\underline{a}}(z) - \psi_{\underline{a}}(z).$$

If we put $\alpha_t = \dfrac{M(S, \underline{a}, t)}{\mu_{\underline{a}}(S)}$, then

$$\chi_{\underline{a}}(z) = \frac{\pi^{n/2} \| S \|^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)} \sum_t \quad (\alpha_t - \beta_t) h_t(z) \tag{158}$$

It is to be noticed that $\chi_{\underline{a}}(z)$ lacks the constant term. From theorem 9 and (144) we have

$$\underline{\chi}(z_M) = G(M, z)\underline{\chi}(z). \tag{159}$$

The Unitary matrix $\Lambda(M)$ is defined in § 3 by

$$G(M, z) = \begin{cases} (\gamma z + \delta)^{p/2} (\gamma \bar{z} + \delta)^{q/2} \Lambda(M) & \text{if } \gamma \neq 0 \\ \Lambda(M) & \text{if } \gamma = 0. \end{cases}$$

If we put $z_M = \xi_M + i\eta_M$, then

$$\eta_M = \frac{\eta}{|\gamma z + \delta|^2} \tag{160}$$

Let us now prove some properties of the function $h_t(z)$ introduced in (126). In the first place

$$h_t(z) \sim e^{2\pi it\xi}(4\pi\eta)^{1-n/2}\Gamma(n/2-1) \text{ for } \eta \to 0. \tag{161}$$

This can be proved easily: For, if $t = 0$, then

$$h_0(z) = (4\pi\eta)^{1-n/2}\Gamma(n/2-1)$$

as was seen in (134). Let now $t > 0$. Let us make the substitution $w \to \dfrac{w}{4\pi\eta}$ in the integral for $h_t(z)$. Then 182

$$e^{-2\pi it\xi}h_t(z) = \int\limits_{w>4\pi\eta t} (4\pi\eta)^{1-n/2}w^{p/2-1}(w-4\pi\eta t)^{q/2-1}e^{-w}dw$$

But when $\eta \to 0$

$$\int_{w>4\pi\eta t} w^{p/2-1}(w-4\pi\eta t)^{q/2-1}e^{-w}dw \sim \Gamma(n/2-1).$$

This proves (161). The case $t < 0$ is dealt with in a similar fashion.

In case $\eta \to \infty$ we have

$$\left.\begin{array}{lll} h_t(z) \to 0 & \text{if} & t \neq 0 \\ h_t(z)\eta^{n/2-1} \to 0 & \text{if} & t = 0 \end{array}\right\} \tag{162}$$

This is easily seen from the expression for $h_t(z)$ and (134). In fact, if $t \neq 0$, $h_t(z) \to 0$ exponentially as $\eta \to \infty$.

Let us now consider the equation (158) for $\chi_{\underline{a}}(z)$. The function $h_t(z)$ is monotonic and decreasing in η for fixed ξ. This means that the series for $\chi_{\underline{a}}(z)$ is uniformly bounded in the whole of the fundamental region of the modular group in the z plane. Let $\omega_{\underline{a}}(z) = \eta^{n/4}\chi_{\underline{a}}(z)$. Since $n > 4$ and (162) holds with $h_t(z) \to 0$ exponentially as $\eta \to \infty$, $t \neq 0$, it follows that $\omega_{\underline{a}}(z)$ is bounded, uniformly in ξ, in the fundamental region of the modular group in the upper half z plane.

Let $\omega(z)$ be the vector

$$\underline{\omega}(z) = \begin{pmatrix} \omega_{\underline{a}_1}(z) \\ \vdots \\ \omega_{\underline{a}_l}(z) \end{pmatrix} = \eta^{n/4}\underline{\chi}(z)$$

Then because of (160) and the transformation formula (159) for $\underline{\chi}(z)$ it 183 follows that, if M is a modular matrix,

$$|\omega_{\underline{a}_i}(z_M) \leq \sum_j |\theta_{ij}||\omega_{\underline{a}_j}(z)| \quad (i = 1, \ldots, l) \tag{163}$$

where $\Lambda(M) = (\theta_{ij})$. But $\Lambda(M)$ is a unitary matrix so that $|\theta_{ij}| \leq 1$. This means that

$$|\omega_{\underline{a}_i}(z_M)| \leq \sum_j |\omega_{\underline{a}_j}(z)|$$

From what we have seen above, it follows that $\omega_{\underline{a}}(z)$ is bounded in the whole of the upper half z-plane.

Now $\varphi_{\underline{a}}(z)$ and $\psi_{\underline{a}}(z)$ are fourier series in the real variable ξ and have the period $2d = D$. The fourier coefficient of $\chi_{\underline{a}}(z) = \varphi_{\underline{a}}(z) - \psi_{\underline{a}}(z)$ is

$$\frac{1}{D}\int_0^D \chi_{\underline{a}}(z)e^{-2\pi it\xi}d\xi \tag{164}$$

which clearly equals

$$\frac{\pi^{n/2}||S||^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)}(\alpha_t - \beta_t)h_t(i\eta) \tag{165}$$

Since $n > 4$ and $\eta^{n/4}\chi_{\underline{a}}(z)$ is bounded in the upper half z plane, it follows that

$$\frac{1}{D}\eta^{n/4-1}\int_0^D \eta^{n/4}\chi_{\underline{a}}(z)e^{-2\pi it\xi}d\xi \to 0 \tag{166}$$

184 as $\eta \to 0$. On the other hand the expression on the left of (166) is, in virtue of (164), (165), equal to

$$\frac{\pi^{n/2}||S||^{-\frac{1}{2}}}{\Gamma(p/2)\Gamma(q/2)}(\alpha_t - \beta_t)h_t(i\eta)\eta^{\frac{n}{2}-1}$$

Because of (161)

$$h_t(i\eta)\eta^{n/2-1} \sim (4\pi)^{1-n/2}\Gamma(n/2-1) \neq 0$$

as $\eta \to 0$. Because of (166) therefore, it follows that

$$\alpha_t - \beta_t = 0.$$

Our theorem is thus completely proved.

Going back to the definitions of $\varphi_{\underline{a}}(z)$ in (123) and $\psi_{\underline{a}}(z)$ in (143) we have the partial fraction decomposition

Theorem 12. *If $n > 4$, then*

$$V_{\underline{a}}^{-1} \int\limits_{F_{\underline{a}}} f_{\underline{a}}(z, H) dv =$$

$$= \gamma_{\underline{a}} + \epsilon^{-1} d^{-\frac{1}{2}} \sum_{\substack{(\alpha,\gamma)=1 \\ \gamma>0}} \gamma^{-n} \left(\sum_{\underline{g}((\text{mod})\ \gamma)} e^{2\pi i \frac{\alpha}{\gamma} S[\underline{g}+\underline{a}]} \right) \left(\underline{z} - \frac{\alpha}{\gamma} \right)^{-\frac{p}{2}} \left(\overline{z} - \frac{\alpha}{\gamma} \right)^{-\frac{q}{2}}$$

10 Remarks

Let us consider the main theorem. The right hand side is a product extended over all the primes and is zero if and only if at least one factor is zero. The left hand side is different from zero only if the equation

$$S[\underline{x} + \underline{a}] = t \tag{167}$$

has an integral solution. Thus the main theorem shows that (167) has an integral solution if and only if

$$S[\underline{x} + \underline{a}] \equiv t (\text{mod m})$$

has a solution for every integer $m \succ 1$. Because of the definition of $\delta_p(S, \underline{a}, t)$ we may also say that if S is indefinite and $n > 4$, then (167) has an integral solution if and only if (167) is true in p-adic integers for every p. In the case $t = 0$, this is the Meyer-Hasse theorem. But our main theorem is a quantitative improvement of the Meyer-Hasse theorem, in as much as it gives an expression for the measure of representation of t by $S[\underline{x} + \underline{a}]$. **185**

The method of proof consisted in first obtaining a 'generating function' $f(z)$ for the Diophantine problem (167) and then constructing a function $E(z)$, the Eisenstein series, which behaves like $f(z)$ for all modular substitutions. In other words, we construct a function $E(z)$ which behaves like $f(z)$ when z approaches, in the upper half plane, a rational point on the real axis. This idea was originally used by Hardy and Ramanujan in the problem of representation of integers by sums of k squares. The generating function $f(z)$ here was the theta series

$$f(z) = \left(\sum_{1=-\infty}^{\infty} e^{2\pi i l^2 z} \right)^k$$

The function $E(z)$ is constructed in the same way as here and Hardy and Ramanujan showed that for $k = 5, 6, 7, 8$

$$f(z) = E(z).$$

But for $k = 9$, $f(z) \neq E(z)$. It is remarkable that in the case of indefinite forms, one has equality if $k > 4$. One does not have, in general, for representation of integers by definite forms, a formula like that in theorem 11. One can obtain a modified formula by introducing a genus of forms. If $S > 0$ is an integral matrix, the genus of S consists of all integral matrices $P > 0$ which are such that for each integer $m > 1$ there is an integral matrix U with 186

$$S[U] \equiv P(\text{mod m}), \quad (|\text{U}|, \text{m}) = 1.$$

It is then known that a genus consists of a finite number of classes. Let $S_1, \ldots, S_a$ be representatives of the finitely many classes in the genus of S. If $T > 0$ is any k-rowed integral matrix, we can define for each i, $i = 1, \ldots, a$ the number, $A(S_i, T)$, of representations

$$S_i[X] = T.$$

If $E(S_i)$ denotes the order of the unit group of S_i (this being finite since $S_i > 0$) we can form

$$\overline{A}(S, T) = \frac{\sum\limits_{i=1}^{a} \dfrac{A(S_i, T)}{E(S_i)}}{\sum\limits_{i=1}^{a} \dfrac{1}{E(S_i)}}$$

the average measure of representation of T by a genus of S. Just as in (154) we can define for each p,

$$\delta_p(S,T) = \lim_{1\to\omega} \frac{A_{p^l}(S,T)}{p^{l\lambda}}$$

where $\lambda = nk - \dfrac{k(k+1)}{2}$. This is finite, rational and $\prod\limits_p \delta_p(S,T)$ converges if $k \leq n$. The main theorem would then be

$$\overline{A}(S,T) = c \prod_c \delta_p(S,T) \tag{168}$$

c being a constant depending on n and k. 187

A similar formula, with suitable restrictions, exists if S and T are indefinite also.

One might ask if our theorem 12 could be extended to the cases $n = 2, 3, 4$. In case $n = 4$, and $S[\underline{x}]$ is not a quaternion zero form, then one can prove that $f(z) = E(z)$. The method is slightly more complicated. The differential operator Δ, or slight variants of it, which we had not used in our main theorem, plays an important role here. In case $n = 2$ and 3 it can be proved that our main theorem is false.

Generalizations of the main theorem may be made by considering representations not of numbers, but of rational symmetric matrices. One can generalize the results by considering instead of the domain of rational integers, the ring of integers in an algebraic number field or more generally an order in an involutorial simple algebra over the rational number field. The bibliography gives the sources for these generalizations.

Bibliography

[1] C. L. Siegel : Additive theorie der-Zahlkörper II *Math. Ann*. 88 (1923) P. 184-210. 188

[2] C. L. Siegel : Über die analytische theorie der quadratischen Formen, *Annals of Math*. 36 (1935) P. 527-606; 37 (1936) P. 230-263; 38 (1937) P. 212-291.

[3] C. L. Siegel : Über die zeta funktionen indefiniter quadratischen Formen, *Math. Zeir* 43 (1938) P. 682-708; 44 (1939) P. 398-426.

[4] C. L. Siegel : Einheiten quadratischer Formen *Abhandlungen. Math. Sem. Hansischen. Univ*. 13 (1940) P. 209-239.

[5] C. L. Siegel : Equivalence of quadratic forms Amer. Jour. of Math. 63 (1941) P. 658-680.

[6] C. L. Siegel : On the theory of indefinite quadratic forms, *Annals of Math*. 45 (1944) P. 577-622.

[7] C. L. Siegel : Indefinite quadratische Formen and Modulfunktionen, *Courant. Anniv. Volume* (1948) P. 395-406.

[8] C. L. Siegel : Indefinite quadratische Formen and Funktionentheorie, *Math. Annalen*, 124 (1951) I, P. 17-54; II ibid P. 364-387.

[9] C. L. Siegel : A generalization of the Epstein zeta-function, *Jour. Ind. Math. Soc.* Vol. 20 (1956) P. 1-10. 189

[10] C. L. Siegel : Die Funktionalgleichungen einiger Dirichletscher Reihen, *Math. Zeit* 63 (1956) P. 363-373.

[11] H. Weyl : Fundamental domains for lattice groups in division algebras *Festschrift to A. Speiser,* 1945.