

THE STRUCTURE AND REPRESENTATION
OF
CONTINUOUS GROUPS

HERMANN WEYL

(Notes by Nathan Jacobson)

Delivered at The Institute for Advanced Study

January-April, 1934

THE STRUCTURE AND REPRESENTATION OF CONTINUOUS GROUPS

by

Hermann Weyl

Introduction

1933-4

12-57

In these lectures we shall be concerned chiefly with three important lines in the development of the theory of groups. First, we shall consider certain properties of groups which hold regardless of whether the groups are finite or infinite, discrete or continuous. This part of the theory is of a distinctly algebraic character. In this connection we may mention the Jordan-Hölder-E. Noether theorem and the Burnside theorem. As in these two examples, we shall see that we may distinguish between results which hold for any field and those which hold only for sufficiently extended fields such as the field of all complex numbers. Many of the theorems in this section may be considered from the more general point of view of the theory of structure and representations of associative algebras. The fundamental work in the classification of associative algebras with a finite basis over an arbitrary field is due to Wedderburn. The representation theory was begun by Frobenius and Burnside and was considerably extended and simplified by E. Noether.

Secondly, we shall discuss the classical Lie theory of the generation of a continuous group germ (a neighborhood of the identity) by infinitesimal transformations. We shall be concerned with the relations between the group germ and the algebra of infinitesimal transformations. This algebra is non-associative of a special type and will be called a Lie algebra. The classification of Lie algebras was carried through to the same degree as the classification of associative algebras, mentioned before, by Killing and Cartan. However, the methods used are irrational, and it is as yet an unsolved problem to carry through these results rationally. Cartan has considered also the problem of the representation of these algebras.

Finally, we shall study the properties of the group manifold in its entirety. The considerations will be topological. The chief tool at our disposal is the

integration process introduced by A. Hurwitz. This method was used by I. Schur to obtain all the representations of the orthogonal group and by Weyl to obtain all the representations of any semi-simple Lie group.

The following is a list of references:

A. General theory of groups

- Speiser, Theorie der Gruppen von endlicher Ordnung
- Burnside, Theory of groups
- Miller, Blichfeldt and Dickson, Finite groups
- van der Waerden, Moderne Algebra
- Weyl, The theory of groups and quantum mechanics (translated by H. P. Robertson)

B. Lie theory of continuous groups

- Bianchi, Theoria dei Gruppi Continui
- Kowalewski, Einführung in die Theorie der kontinuierlichen Gruppen
- Buhl, Aperçus Modernes sur la Theorie des Groupes Continus et Finis
- Weyl, Mathematische Analyse der Raumprobleme (5, 6 Vorlesung and Anhang 8)
- Eisenhart, Continuous groups of transformations

C. Integral theory of continuous groups

- Cartan, La Theorie des Groupes Finis et Continus et l'Analysis Situs
- I. Schur, Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie, I, II, III, Sitz. der Preuss. Akad. 1924
- Weyl, Theorie der Darstellung kontinuierlicher half-einfacher Gruppen durch lineare Transformationen, I, II, III, Math. Zeits. 23, 24.

(

Chapter 1.

Fundamental Concepts of Groups and Representations

1. Transformation groups and abstract groups

A mapping or transformation S of a point field (set of points) M upon itself is a correspondence which associates with every element p of M a single element $p' = Sp$. We say that two mappings are equal $S = T$ if $Sp = Tp$ for every point $p \in M$. There is a mapping which we call the identity I having the property that $Ip = p$ for every p . We shall consider only (1-1) mappings; i.e. every element of M is the image of one, and only, element of M . Hence if S is a transformation such that $p' = Sp$ we may define an inverse transformation S^{-1} such that $p' \rightarrow p = S^{-1}p'$. If $p \xrightarrow{S} p' \xrightarrow{T} p''$, then the correspondence directly from p to p'' is called the product of S and T , TS . If M is a finite set of n elements, the mappings are called permutations.

A set $\mathcal{G}$ of mappings of a point field M is called a transformation group if it has the properties:

- (1) The identity transformation $I \in \mathcal{G}$ (belongs to) $\mathcal{G}$
- (2) If $S \in \mathcal{G}$ so does S^{-1}
- (3) If S and $T \in \mathcal{G}$ then so does TS

If p and $p' \in M$, we may consider them as equivalent with respect to $\mathcal{G}$

($p \sim p'$) if there is a transformation $S \in \mathcal{G}$ such that $p \rightarrow p' = Sp$. This definition of equivalence has the properties

- (1) $p \sim p$ (reflexive)
- (2) $p \sim p'$ implies $p' \sim p$ (symmetric)
- (3) $p \sim p'$, $p' \sim p''$ implies $p \sim p''$ (transitive)

These properties are merely a restatement of the group properties of $\mathcal{G}$. Equivalence of figures (sub-sets of M) may be defined in a similar fashion. It is this concept of equivalence which was emphasized by Klein as fundamental in geometry.

Let A be a (1-1) mapping of M into another point field N (not necessarily different from M) such that $p \rightarrow q = Ap$, $p \in M$ and $q \in N$. For every mapping S of $\mathcal{G}$ there corresponds a mapping $\bar{S} = ASA^{-1}$ of N which is defined by

$$q \xrightarrow{A^{-1}} p \xrightarrow{S} p' \xrightarrow{A} q'$$

Evidently the set of transformations $\bar{S}$ of N into itself is a transformation group $\bar{\mathcal{G}}$.

In particular, if $M = N$ and B is a second mapping, then BA is also such a mapping of M upon itself; it changes p into $(BA)p = B(Ap)$ and S into $(BA)S(BA)^{-1} = B(ASA^{-1})B^{-1}$.

An abstract group $\mathcal{G}$ is a set of elements of an unspecified nature for which a law of composition is given which satisfies the following postulates:

- (1) There is an element $i \in \mathcal{G}$ such that $ai = a = ia$ for every $a \in \mathcal{G}$.
- (2) For every a there is an element $a^{-1} \in \mathcal{G}$ such that $aa^{-1} = a^{-1}a = i$.
- (3) If $a, b \in \mathcal{G}$ so does their product ba .
- (4) The associative law holds: $c(ba) = (cb)a$.

Two groups $\mathcal{G}$ and $\bar{\mathcal{G}}$ are simply isomorphic or equivalent if there is a

(1-1) correspondence between their elements which satisfies the conditions

(1) $\bar{i}$ the correspondent of i is the identity of $\bar{\mathcal{G}}$

(2) $\overline{a^{-1}} = \bar{a}^{-1}$

(3) $\overline{ba} = \bar{b} \bar{a}$

More generally if there is a mapping of $\mathcal{G}$ into $\bar{\mathcal{G}}$ which is assumed to be merely single valued and which satisfies these three conditions, then we say that $\mathcal{G}$ is isomorphically mapped into $\bar{\mathcal{G}}$ or is isomorphic with $\bar{\mathcal{G}}$. If we have a mapping of $\mathcal{G}$ into a transformation group $\mathcal{G}$ which has the properties just described where, of course, the product of two transformations is defined as before, then we speak of a realization of $\mathcal{G}$. If the correspondence is (1-1) we call the realization distinct.

The postulates for an abstract group were formulated for the purpose of expressing the formal rules which apply to the composition of transformations. The question arises as to whether or not the postulates are sufficient. Can all the properties of the composition of transformations be deduced from our postulates? We shall show that this is the case by obtaining a distinct realization of any abstract group. For this purpose we use the group manifold $\mathcal{G}$ as the point field. With every element a we associate the left translation γ_a i.e. the transformation

$$s \rightarrow s' = as$$

The set of left translations is easily seen to be a distinct realization of $\mathcal{G}$.

The transformation $s \rightarrow s'' = sa$ of the group manifold is called a right-translation.

2. Sub-groups and factor groups.

A sub-set $\mathcal{H}$ of $\mathcal{G}$ satisfying the group postulates is called a sub-group.

Consider the group $\mathcal{H}$ of left translations of $\mathcal{G}$ which correspond to the elements of $\mathcal{H}$. As in section 1, we can define an equivalence of the elements of $\mathcal{G}$ by means of the transformation group $\mathcal{H}$ i.e. $s \equiv s'$ if $s' = hs$ where $h \in \mathcal{H}$. By abstrac-

tion from this notion of equivalence we obtain a set of elements called the left co-sets of f in G . If the number of co-sets is finite, it is called the index of f in G . In the case of finite groups the sets of mutually equivalent elements of G with respect to f all have the same number of elements and the number of distinct sets obtained in this way is the index of f . Hence, the order of G = order of $f \times \text{index of } f$.

Because of the associative law, the powers (positive, negative, zero) of an element satisfy

$$a^n a^m = a^{n+m}$$

and hence their totality is a sub-group which we call the cyclic group generated by a . If $ba = ab$ for any pair of elements of G , G is abelian or commutative.

The transformation

$$\mathcal{R}_a: \quad s \rightarrow s' = asa^{-1}$$

is a mapping of the group manifold on itself which is called a conjugation by a . The set of conjugations is a realization of G , though not in general distinct. If, for instance, a belongs to the centrum of G , i.e. is commutative with every element of G , then $\mathcal{R}_a = \mathcal{R}_1$. Conjugation has one important advantage over translation in that the correspondence $s \rightarrow s'$ is a simple self-isomorphism or simple automorphism of G . If s and s' are equivalent under the group of conjugations, i.e. $s' = asa^{-1}$, $a \in G$, then they are conjugate. In a similar manner we define conjugate sub-groups. If $f = afa^{-1}$ for every $a \in G$, f is invariant in G .

If f is an invariant sub-group, and $s \equiv s'$, $t \equiv t'$ with respect to f (mod f), then

$$(2.1) \quad ts \equiv t's', \quad s^{-1} \equiv s'^{-1}.$$

Thus, by composing the elements of one co-set with the elements of another, the products obtained are all in the same co-set. We may therefore associate this co-set as the product of the former two. In this way we obtain a group called the factor group of f in G , G/f . Its order is the index of f . By letting an

element of G correspond to its co-set, we obtain by (2.1) an isomorphism of G into G/f . The elements of f correspond to the identity of G/f in this isomorphism.

Conversely, suppose that G is isomorphic with a group $\bar{G}$. The set f of elements of G which correspond to the identity of $\bar{G}$ is an invariant sub-group and the elements which correspond to any other element $\bar{a}$ of $\bar{G}$ have the same co-set with respect to f . It follows easily that $\bar{G}$ and G/f are equivalent. We therefore have the following fundamental

Theorem: If f is an invariant sub-group of G the co-sets of f in G may be taken as the elements of a group G/f to which G is isomorphic. Conversely, if G is isomorphic with $\bar{G}$, then $\bar{G} \approx G/f$ where f is the set of elements of G which correspond to the identity of $\bar{G}$.

As a corollary we have the result that any realization of G is a distinct realization of a certain factor group of G .

3. Vector spaces and linear transformations

The purpose of this section is to recall some of the fundamental concepts of the theory of vectors and linear transformations and to introduce a notation which will be followed without comment in the sequel. The reader is referred to Weyl's "The Theory of groups and quantum mechanics", chapter 1, for a more extensive treatment of this subject. It should be observed, however, that the notation there is different in some particulars from the one developed here.

Summary of the notation:

Vectors: small Latin letters

Vector spaces: German capitals

Linear transformations: Latin capitals underscored

Matrices: plain Latin capitals

Transposed matrix of A denoted by A'

Projection spaces: $\mathcal{R}/\mathcal{R}'$ etc.

Inner product of two vectors: $(x.y)$

Numbers varying within a given number field usually the continuum of complex numbers: small Greek letters

If $\mathcal{R}'$ is a sub-space of $\mathcal{R}$, we define $x \equiv y \pmod{\mathcal{R}'}$ to mean that

$x - y \in \mathcal{R}'$. As in the case of groups and factor groups, we can look upon congruent vectors as the same and obtain a new vector space $\mathcal{R}/\mathcal{R}'$, the projection of $\mathcal{R}$ with respect to $\mathcal{R}'$.

The set of vectors of the form $\sum_{i=1}^n \alpha_i u_i$ is a vector sub-space $\mathcal{R}'$ of $\mathcal{R}$ called the space spanned by $u_1, u_2, \dots, u_n$. We denote it by $(u_1, u_2, \dots, u_n)$.

We assume the space is finite dimensional; i.e. it can be spanned by a finite number of vectors and hence by a finite number, say n , linearly independent vectors. The dimension n is an invariant since it is the maximum number of linearly independent vectors of $\mathcal{R}$. We shall give a proof of this fundamental fact in the next section. If $\mathcal{R} = (o_1, \dots, o_n)$ and the o 's are independent, we call them a coordinate system of $\mathcal{R}$. Any vector $x = \sum \xi_i e_i = (\xi_1, \xi_2, \dots, \xi_n)$ if we denote o_i by $(0 \dots 0, \overset{i}{1}, 0 \dots 0)$. The ξ 's are the coordinates of x . If $o'_1, o'_2, \dots, o'_n$ is another coordinate system, then

$$(3.1) \quad e'_i = \sum_{\kappa} \mu_{\kappa i} e_{\kappa} \quad e_{\kappa} = \sum_j \nu_{j\kappa} e'_j.$$

It follows that $MN = I = NM$ where $M = (\mu_{\kappa i})$ and $N = (\nu_{j\kappa})$ or $N = M^{-1}$. If

$\xi'_1, \xi'_2, \dots, \xi'_n$ are the coordinates of x referred to $o'_1, o'_2, \dots, o'_n$, then

$$x = \sum \xi'_i e'_i = \sum \xi'_i \mu_{\kappa i} e_{\kappa} = \sum \xi_{\kappa} e_{\kappa}$$

and hence

$$(3.2) \quad \xi_{\kappa} = \sum_{i=1}^n \mu_{\kappa i} \xi'_i \quad (\kappa = 1, \dots, n).$$

If $\mathcal{R}'$ is a sub-space of $\mathcal{R}$ and $(u_1, \dots, u_r)$ a coordinate system for $\mathcal{R}'$, we can find vectors $u_{r+1}, u_{r+2}, \dots, u_n$ such that $u_1, u_2, \dots, u_n$ is a coordinate system for $\mathcal{R}$. We call such a coordinate system one adapted to $\mathcal{R}'$.

A transformation $\underline{A}$ of $\mathcal{R}$ upon $\mathcal{R}$ is linear if it satisfies the equations:

$$(3.3) \quad \underline{A}(x + y) = \underline{A}x + \underline{A}y \quad \underline{A}(\alpha x) = \alpha(\underline{A}x).$$

If $o_1, o_2, \dots, o_n$ and $f_1, f_2, \dots, f_m$ are coordinate systems of $\mathcal{R}$ and $\mathcal{R}'$ respectively-

ly, then $\underline{A}$ is fully described by ..

$$(3.4) \quad \underline{A} e_i = \sum_{k=1}^m \alpha_{ki} f_k \quad (i = 1, \dots, n)$$

and conversely any such equations define a linear transformation. The coordinates of the image y of x are given by

$$(3.5) \quad \eta_k = \sum \alpha_{ki} \xi_i .$$

Hence $\underline{A}$ is specified by the matrix $A = (\alpha_{ki})$ and the coordinate systems $e_1, \dots, e_n$ and $f_1, \dots, f_m$. $\underline{A}$ is (1-1) if, and only if, $m = n$ and A is non-singular. If $e'_1, \dots, e'_n$ is a coordinate system defined by (3.1), and $f'_1, \dots, f'_m$ are given by

$$(3.6) \quad f'_j = \sum_k \rho_{kj} f_k \quad f_k = \sum_\ell \sigma_{\ell k} f'_\ell$$

and $R = (\rho_{kj})$, $S = (\sigma_{\ell k})$, then the transformation $\underline{A}$ is represented relative to these new bases by

$$(3.7) \quad A^* = SAM = R^{-1}AM.$$

In particular, if $\mathcal{R} = \mathcal{R}'$, $e_i = f_i$, $e'_i = f'_i$, then (3.7) becomes

$$(3.8) \quad A^* = M^{-1}AM .$$

From (3.8) it follows that the invariants of A under similarity correspondence $A \rightarrow A^* = M^{-1}AM$ are invariants of the transformation $\underline{A}$ of $\mathcal{R}$ upon itself.

The characteristic polynomial

$$(3.9) \quad \varphi(\lambda) = |\lambda E - A| = \lambda^m - \alpha_1 \lambda^{m-1} + \dots + (-1)^m \alpha_m$$

where $E = (\delta_{ij})$ and $| |$ denotes determinant, is such an invariant. The roots ρ_i of $\varphi(\lambda)$ are called the roots of $\underline{A}$. Of particular importance is the trace of $\underline{A}$ or of A ,

$$(3.10) \quad \text{tr}(\underline{A}) = \text{tr}(A) = \alpha_1 = \sum \alpha_{ii} = \sum \rho_i .$$

Because of its invariance or by direct computation, we have

$$(3.11) \quad \text{tr}(AB) = \text{tr}(BA)$$

By (3.10) follows

$$(3.12) \quad \text{tr}(\alpha A + \beta B) = \alpha \text{tr} A + \beta \text{tr} B.$$

If we define the linear transformation $\alpha A + \beta B$ by

$$(\alpha A + \beta B)x = \alpha Ax + \beta Bx$$

then (3.12) holds for the underscored $\alpha A + \beta B$.

The resultant of two linear transformation $B(Ax)$ is linear and may be called the product BA of the two. Its matrix relative to $e_1, \dots, e_n$ is BA . The set of all linear transformations of $\mathcal{P}$ upon itself is an algebra which is essentially the algebra of all matrices of n rows and columns with elements in our number field. The transposed matrix A' of A is defined by $\alpha'_{ik} = \alpha_{ki}$. We have

$$(3.13) \quad (\alpha A + \beta B)' = \alpha A' + \beta B' \quad (AB)' = B'A'.$$

In order to give a geometric meaning to the relation between A and A' we consider the space $\mathcal{Y}$ dual to $\mathcal{P}$. Its elements are the linear forms $l(x) = \sum_i \lambda_i \xi_i$ where $x = (\xi_1, \dots, \xi_m)$ is an arbitrary vector of $\mathcal{P}$. We may take the linear forms ξ_i as a coordinate system in $\mathcal{Y}$ and in this system the coordinates of $l(x)$ are $\lambda_1, \lambda_2, \dots, \lambda_m$. Under the influence of a coordinate transformation of $\mathcal{P}$ defined by (3.2) and (3.1) the coordinates λ_i undergo the contragredient transformation

$$(3.14) \quad \lambda_k = \sum v_{ik} \lambda'_i.$$

The inner product, $(x, y) = \sum \xi_i \eta_i$ of a vector $x \in \mathcal{P}$ and a vector $y = (\eta_1, \dots, \eta_m) \in \mathcal{Y}$, is invariant under a change of coordinate system of $\mathcal{P}$ together with the induced contragredient change of coordinates in $\mathcal{Y}$. If A is a linear transformation of $\mathcal{P}$ upon another space $\mathcal{P}'$ such that $x \rightarrow x' = Ax$, then this defines a linear transformation A' of the dual space $\mathcal{Y}'$ of $\mathcal{P}'$ on $\mathcal{Y}$, $y' \rightarrow y = A'y'$, satisfying the equations

$$(3.15) \quad (Ax, y') = (x, A'y')$$

or

$$(3.16) \quad \eta_j = \sum \alpha'_{ji} \eta'_i = \sum \alpha_{ij} \eta'_i$$

i.e. the matrix of $\underline{\Lambda}'$ is the transpose of the matrix of $\underline{\Lambda}$. If $\underline{\Lambda}$ is (1-1) so is $\underline{\Lambda}'$ and we have defined in this way the contragredient transformation of $\mathcal{R}'$ upon $\mathcal{R}'$ and its matrix is $(\Lambda')^{-1}$.

A sub-space $\mathcal{R}'$ of $\mathcal{R}$ is invariant under $\underline{\Lambda}$ if it is transformed upon itself by $\underline{\Lambda}$. In this case $\underline{\Lambda}$ induces a linear transformation (also denoted by $\underline{\Lambda}$) in $\mathcal{R}/\mathcal{R}'$. If the coordinate system is adapted to $\mathcal{R}' = (u_1, \dots, u_r)$, then

$$(3.17) \quad \underline{\Lambda} u_i = \sum_{j=1}^r \alpha_{ji}^{(1)} u_j \quad (i = 1, \dots, r)$$

$$\underline{\Lambda} u_\ell = \sum_{k=1}^r \beta_{k\ell} u_k + \sum_{m=r+1}^n \alpha_{m\ell}^{(2)} u_m \quad (\ell = r+1, \dots, n)$$

and hence in this coordinate system $\underline{\Lambda}$ is represented by

$$(3.18) \quad \begin{pmatrix} A & B \\ 0 & A_2 \end{pmatrix}.$$

The vectors $u_{r+1} \dots u_n$ are independent mod $\mathcal{R}'$ and may be taken as a coordinate system in $\mathcal{R}/\mathcal{R}'$. A_2 is the matrix of $\underline{\Lambda}$ acting on this coordinate system in $\mathcal{R}/\mathcal{R}'$.

If $\varphi_1(\lambda)$ is the characteristic equation of $\underline{\Lambda}$ in $\mathcal{R}'$ and $\varphi_2(\lambda)$ its characteristic equation in $\mathcal{R}/\mathcal{R}'$, then (3.18) shows that

$$(3.19) \quad \varphi(\lambda) = \varphi_1(\lambda) \varphi_2(\lambda).$$

If $\mathcal{R}$ admits an invariant sub-space $\mathcal{R}' \neq 0$ and $\neq \mathcal{R}$, we call the transformation $\underline{\Lambda}$ of $\mathcal{R}$ reducible. If $\mathcal{R}$ can be decomposed into two invariant sub-spaces $\neq 0$ and having only the 0 vector in common, then $\underline{\Lambda}$ is said to be completely reducible or decomposable. In this case a coordinate system can be found so that the upper corner B as well as the lower corner in (3.18) vanishes.

In general, reducibility does not imply decomposability, but in the important case in which $\underline{\Lambda}$ is unitary it does. We recall that a transformation is unitary if it leaves invariant the form $\sum \xi_i \bar{\xi}_i$, the absolute magnitude of the vector $(\xi_1, \xi_2, \dots, \xi_m)$. It follows easily that the matrix Λ of $\underline{\Lambda}$ referred to the coordinates ξ_i satisfies the equation

$$(3.20) \quad \Lambda \bar{\Lambda}' = E.$$

Now if $\mathcal{R}'$ is invariant the space $\mathcal{R}''$ perpendicular to $\mathcal{R}'$, i.e. consisting of the vectors $(\xi_1'', \xi_2'', \dots, \xi_n'')$ such that $\sum \xi_i' \bar{\xi}_i'' = 0$ for every vector $(\xi_1', \xi_2', \dots, \xi_n')$ of $\mathcal{R}'$, is also invariant and $\mathcal{R} = \mathcal{R}' + \mathcal{R}''$, $\mathcal{R}' \cap \mathcal{R}'' = 0$ ($\mathcal{R}'$ and $\mathcal{R}''$ have only 0 in common). Hence we have that reducibility of $\underline{A}$ implies decomposability.

A set of n mutually perpendicular vectors, each of absolute magnitude 1, is called a normal coordinate system for $\mathcal{R}$. The passage from one normal coordinate system to another is effected by a unitary transformation. Now by choosing normal coordinate systems in $\mathcal{R}'$ and $\mathcal{R}''$ we obtain such a system for $\mathcal{R}$. Hence we have the result that a unitary transformation which is reducible has as its matrix in a suitably chosen normal coordinate system one of the form

$$(3.21) \quad A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

where $A_i \bar{A}_i' = E_i$ ($E_i = (\delta_{ij})$ and of same order as A_i).

A transformation $\underline{A}$ of $\mathcal{R}$ is equivalent to $\underline{A}^*$ of $\mathcal{R}^*$ ($\underline{A} \sim \underline{A}^*$) if $\mathcal{R}$ and $\mathcal{R}^*$ may be put into (1-1) linear correspondence in such a way that if $x \leftrightarrow y$, $x \in \mathcal{R}$, $y \in \mathcal{R}^*$ then $\underline{A} x \leftrightarrow \underline{A}^* y$. If $\underline{A} \sim \underline{A}^*$ a coordinate system may be chosen in $\mathcal{R}$ and $\mathcal{R}^*$ in such a way that the two transformations are represented by the same matrix, and hence in any coordinate systems their matrices are similar. In case the transformations $\underline{A}$ and $\underline{A}^*$ are unitary, we are usually interested in unitary-equivalence, i.e. the correspondence between $\mathcal{R}$ and $\mathcal{R}^*$ preserves the absolute magnitude of vectors. This means that the matrices of $\underline{A}$ and $\underline{A}^*$ referred to normal coordinate systems are similar by means of a unitary matrix,

$$(3.22) \quad \underline{A} = U \underline{A}^* U^{-1} \quad \quad \overline{U U'} = E$$

If $\underline{A}$, $\underline{B}$, $\underline{C}$, ... is a set of transformations, we use the same definitions of reducibility, decomposability, etc., as for a single transformation; o.g. a set of transformations is reducible if by a proper choice of coordinates they may be brought simultaneously to the form (3.18). The above remarks for a single unitary transformation, apply equally well for sets of unitary transformations.

If we compare this section with sections 1 and 2, we observe that a vector space is analogous to an abelian group, sub-space to sub-group, linear transformation to isomorphism, projection space to factor group. The concepts just discussed also have analogues in the theory of groups. Corresponding to a set $\mathcal{A}$ of linear transformations of $\mathcal{R}$ upon itself we have a set α of automorphisms of a group $\mathcal{G}$. To an invariant sub-space under $\mathcal{A}$ there corresponds an allowable invariant sub-group $\mathcal{G}'$ under α , an invariant sub-group which is transformed into itself by the automorphisms of α . Corresponding to invariant sub-spaces of $\mathcal{R}$ which give rise to equivalent components of a set of transformations, we have the concept of operator-isomorphism between allowable sub-groups of $\mathcal{G}$, i.e. an isomorphism between the sub-groups such that if $h \leftrightarrow h^*$, $a(h) \leftrightarrow a(h^*)$ where a is any operator of α .

As a matter of fact, by using this idea of an operator group, we can look upon a vector space as a definite generalization of an abelian group and not merely as an analogue. We may define $\mathcal{R}$ as an abelian group with which there is associated a set of operators $\mathcal{F}$ which are automorphisms of $\mathcal{R}$

$$\alpha(x+y) = \alpha x + \alpha y$$

and such that the set of these automorphisms is a field

$$\alpha(\beta x) = (\alpha\beta)x \quad 1x = x \quad \alpha x + \beta x = (\alpha + \beta)x.$$

Linear transformations are then operator-automorphisms and may be taken as a second set of operators (which are commutative with the set $\mathcal{F}$). This is expressed by (3.3). This generalization of the definition of a group will be very useful for us.

4. Jordan-Hölder argument. Uniqueness of irreducible components.

Let $\mathcal{A}$ be a set of transformations and $\mathcal{R}_1 \neq 0$ a minimal invariant sub-space of $\mathcal{R}$ and $\overline{\mathcal{R}}_2$ a minimal invariant sub-space of $\mathcal{R}/\mathcal{R}_1$. Corresponding to $\overline{\mathcal{R}}_2$ we have an invariant sub-space $\mathcal{R}_2$ of $\mathcal{R}$ which contains $\mathcal{R}_1$ and such that $\overline{\mathcal{R}}_2 \sim \mathcal{R}_2/\mathcal{R}_1$. If $\mathcal{R}_2 \neq \mathcal{R}_1$, we continue in this way and obtain a sequence of invariant sub-spaces $\mathcal{R} = \mathcal{R}_h > \mathcal{R}_{h-1} > \mathcal{R}_{h-2} > \dots > \mathcal{R}_1$ such that $\mathcal{R}_{i+1}/\mathcal{R}_i$ is ~~irreducible~~ ^{minimal}.
~~etc.~~ If we adapt the coordinate system to this reduction by first choosing a basis

of $\mathcal{R}_1$ then supplementing it to obtain a basis for $\mathcal{R}_2$ and so on, then the transformations of $\mathcal{A}$ will have the form

$$(4.1) \quad \begin{pmatrix} A_1 & B_{12} & B_{13} & \cdots \\ 0 & A_2 & B_{21} & \cdots \\ 0 & 0 & A_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Suppose we have a second reduction of the space $\mathcal{R}$, then we will show that the irreducible components of the set of transformations determined by this second reduction are equivalent except for order to the first set. In terms of the matrices this means we may set up a (1-1) correspondence between A_i and the matrices A'_i arising from the second reduction in such a way that the corresponding matrices are similar.

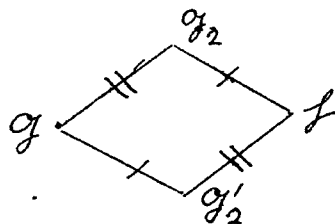
We shall first prove the corresponding theorem for groups. A sequence of sub-groups $\mathcal{G} \supset \mathcal{G}_2 \supset \mathcal{G}_3 \supset \dots$ such that $\mathcal{G}_i$ is invariant in $\mathcal{G}_{i-1}$ will be called a normal sequence for $\mathcal{G}$. If the normal sequence is finite and each $\mathcal{G}_i$ is a maximal invariant sub-group of $\mathcal{G}_{i-1}$ we call it a composition series for $\mathcal{G}$.

Theorem of Jordan-Hölder: If $\mathcal{G}$ has a composition series $\mathcal{G} \supset \mathcal{G}_2 \supset \dots \supset \mathcal{G}_k = 1$, then any normal sequence for $\mathcal{G}$ can be imbedded in a composition series. Any two composition series are equivalent.

By equivalence of composition series we mean that their factor groups $\overline{\mathcal{G}}_i = \mathcal{G}_i / \mathcal{G}_{i+1}$ can be put into (1-1) correspondence in such a way that corresponding pairs are simply isomorphic.

Lemma: If $\mathcal{G}_2$ and $\mathcal{G}_2'$ are invariant sub-groups of $\mathcal{G}$ and $\mathcal{G}_2$ is maximal and does not contain $\mathcal{G}_2'$, then $\mathcal{f} = \mathcal{G}_2 \cap \mathcal{G}_2'$ (intersection) is invariant in $\mathcal{G}_2$ and $\mathcal{G}_2'$ and $\mathcal{G}/\mathcal{G}_2 \simeq \mathcal{G}_2'/\mathcal{f}$, $\mathcal{G}/\mathcal{G}_2' \simeq \mathcal{G}_2/\mathcal{f}$.

The relation between these groups may be indicated by the figure:



Proof: $g'_2 g_2 = g_2 g'_2$ (elements which are products of elements of g'_2 by those of g_2) is an invariant sub-group of g . It contains g_2 and g'_2 and since g_2 was assumed to be maximal and not containing g'_2 it follows that $g'_2 g_2 = g = g_2 g'_2$. Now from

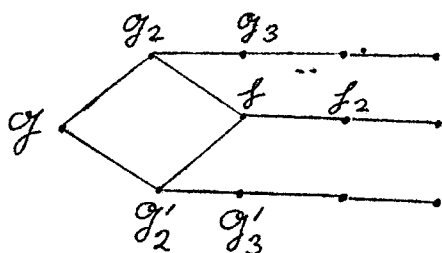
$$s_2 s'_2 \equiv t_2 t'_2 \pmod{g_2}$$

where $s_2, t_2 \in g_2$ and $s'_2, t'_2 \in g'_2$ there follows

$$s'_2 \equiv t'_2 \pmod{f}.$$

Hence we may set up a correspondence between the co-set of any element $s_2 s'_2$ of g with respect to g'_2 and the co-set of s'_2 with respect to f . This correspondence is easily seen to be a simple isomorphism between g/g_2 and g'_2/f . We have made no use of the fact that g_2 is maximal except in showing that $g = g_2 g'_2 = g'_2 g_2$. Hence we may reverse the roles of g_2 and g'_2 and show that $g/g'_2 \cong g_2/f$.

Proof of the Jordan-Hölder theorem: We shall show first that any normal sequence of g has at most k terms. Let $g \supset g'_2 \supset g'_3 \supset \dots$ be a normal sequence. 1) If $g'_2 \supset g'_3$, then we may use induction on the length of the sequence and conclude that $g_2 \supset g'_2 \supset g'_3 \supset \dots$, which is a normal sequence for g_2 , has at most $k-1$ terms and hence $g \supset g'_2 \supset g'_3 \dots$ has at most k terms. 2) If $g'_2 \not\supset g'_3$, we denote $g'_2 \cap g'_3$ by f and construct any normal sequence for f , say $f \supset f_2 \supset f_3 \supset \dots$. Then $g_2 \supset f \supset f_2 \supset \dots$ is a normal sequence for g_2 and by the assumption of induction it can have at most $k-1$ terms. Hence we may obtain a composition series for f . By the lemma, we have that f is a maximal invariant sub-group of g'_2 . Hence we may obtain a composition series for g'_2 consisting of at most $k-1$ terms. It follows then that $g'_2 \supset g'_3 \supset \dots$ has at most $k-1$ terms and hence $g \supset g'_2 \supset g'_3 \supset \dots$ has at most k terms. This process may be illustrated schematically by the accompanying figure:



We use the same type of argument to show that any two composition series are equivalent. Suppose $g \supset g'_2 \supset g'_3 \supset \dots \supset g'_h = i$ is a second composition series. By the first part of the theorem $h=k$. If $g'_2 = g_2$ we use induction. If $g'_2 \neq g_2$ we construct the composition series $g \supset g_2 \supset f \supset f_2$ and $g \supset g'_2 \supset f \supset f_2 \supset \dots$. By the lemma $g/g_2 \simeq g'_2/f$ and $g/g'_2 \simeq g_2/f$ and hence these last two series are equivalent. But by induction $g \supset g_2 \supset g_3 \supset \dots$ is equivalent to $g \supset g_2 \supset f \supset \dots$ and $g \supset g'_2 \supset g'_3 \supset \dots$ is equivalent to $g \supset g'_2 \supset f \supset \dots$. Hence $g \supset g_2 \supset g_3 \supset \dots$ is equivalent to $g \supset g'_2 \supset g'_3 \supset \dots$.

The Jordan-Hölder theorem holds also for groups with operators. We have merely to substitute the phrase "allowable invariant sub-group" for "invariant sub-group" in the above argument. This generalization of the Jordan-Hölder theorem is due to E. Noether.

As a simple application of this theorem we prove the invariance of the dimensionality of a vector space. Let $\mathcal{R} = (o_1 o_2 \dots o_n)$ where the o_i are independent and set $\mathcal{R}_i = (o_i, o_{i+1}, \dots, o_n)$. Then $\mathcal{R} \supset \mathcal{R}_2 \supset \mathcal{R}_3 \supset \dots \supset \mathcal{R}_n \supset 0$ is a composition series for $\mathcal{R}$, the projection spaces $\mathcal{R}_i/\mathcal{R}_{i+1} \simeq (o_i)$ being irreducible. Let $f_1 f_2 \dots f_m$ be a set of independent vectors of $\mathcal{R}$ and denote the spaces $(f_k f_{k+1} \dots f_m)$ by $\mathcal{R}'_k$, then $\mathcal{R} \supset \mathcal{R}'_1 \supset \mathcal{R}'_2 \supset \dots \supset \mathcal{R}'_m \supset 0$ is a normal series for $\mathcal{R}$. Hence by the Jordan-Hölder theorem $m \leq n$ or, the maximum number of independent vectors is n .

We return to the theorem of the invariance of the irreducible components of a set of transformations. If $\mathcal{R} \supset \mathcal{R}_2 \supset \mathcal{R}_3 \supset \dots \supset \mathcal{R}_h = 0$ is a sequence of sub-spaces which are invariant under a set of transformations $\mathcal{A}$ such that $\mathcal{R}_i/\mathcal{R}_{i+1}$ is irreducible, then we may consider $\mathcal{R}$ as an abelian group with the sets of oper-

ators F and α . The sequence of $\mathcal{R}$'s is then a composition series. If we have another series of this type, then the Jordan-Hölder theorem states that they are operator isomorphic or equivalent.

5. Representations. Characters. Formal processes.

A representation of a group $\mathcal{G}$ is a realization of $\mathcal{G}$ by a set $\mathcal{G}$ of linear transformations. If a coordinate system of $\mathcal{R}$ is chosen, we have, to the element S of $\mathcal{G}$ a matrix $U(S)$ such that

$$(5.1) \quad U(1) = 1 \quad U(st) = U(S) U(t).$$

Of course we do not assume that $\mathcal{G}$ is a distinct realization of $\mathcal{G}$. We call the dimensionality of $\mathcal{R}$ the degree of the representation. Equivalence, reducibility, decomposability have the same meaning for representations as for the set of transformations $\mathcal{G}$.

For some considerations it will be convenient to think of the elements s of $\mathcal{G}$ as the linear operators of $\mathcal{R}$ defined by $s(x) = U(s)x$ where $x \in \mathcal{R}$.

The trace $\text{tr } U(s) = \chi(s)$ is called the character of the representation. If two representations are equivalent, then their characters are identical. The character of a representation is a class function, i.e. it is the same for all elements of a conjugate class:

$$(5.2) \quad \chi(asa^{-1}) = \chi(s) \quad \text{or} \quad \chi(st) = \chi(ts)$$

If $\mathcal{G}$ is reducible into $\mathcal{G}_1$ and $\mathcal{G}_2$ and $\chi_i(s)$ is the character of $\mathcal{G}_i$, then by (3.18),

$$(5.3) \quad \chi(s) = \chi_1(s) + \chi_2(s)$$

The identical representation $\mathcal{G}_0$ makes every element of $\mathcal{G}$ correspond to the identity transformation of a one-dimensional space, $\chi_0(s) = 1$.

Before proceeding further in the theory of representations we consider some examples.

Finite cyclic groups. Let $\mathcal{G}$ be the group whose elements are $a, a^2, a^3, \dots, a^h = i$. Suppose that in a one-dimensional representation of $\mathcal{G}$,

$a \rightarrow \zeta$. Then $\zeta^h = 1$. Conversely, if a corresponds to any h^{th} root of unity, we get a representation of degree 1. Let $\mathcal{G}$ be any representation and $a \rightarrow A$. Then $A^h = 1$. To any h^{th} root of 1, ζ we associate the sub-space $\mathcal{R}_\zeta$ of vectors x_ζ such that $Ax_\zeta = \zeta x_\zeta$. The sub-spaces $\mathcal{R}_\zeta$ which are different from 0 are independent. Let $\varphi(\zeta, A) = \frac{1}{h}(1 + \zeta^{h-1}A + \zeta^{h-2}A^2 + \dots + \zeta A^{h-1})$. Any vector $x = \sum_{\zeta} \varphi(\zeta, A)x = \sum_{\zeta} x_\zeta$ where $x_\zeta = \varphi(\zeta, A)x \in \mathcal{R}_\zeta$. Hence $\mathcal{R}$ is decomposed into the sub-spaces $\mathcal{R}_\zeta \neq 0$ and the $\mathcal{R}_\zeta$ may be decomposed into 1-dimensional invariant sub-spaces. (disjoint) Thus any representation of a finite cyclic group can be decomposed into representations of degree 1 and there are as many inequivalent representations as the order of the group.

Finite abelian groups: If a is an element of $\mathcal{G}$ such that $a^h = 1$, and A the corresponding matrix, then by the above argument we may suppose that

$$(5.4) \quad A = \begin{pmatrix} \rho_1 & & & 0 \\ & \rho_2 & & \\ & & \ddots & \\ 0 & & & \rho_m \end{pmatrix}$$

where the ρ 's are h^{th} roots of unity and are arranged so that the first n_1 of them are equal, the next n_2 equal etc. and different from the first n_1 and so on. If B is a matrix which corresponds to b , then $AB = BA$ or $(\rho_i - \rho_k)\beta_{ik} = 0$, $\beta_{ik} = 0$ if $\rho_i \neq \rho_k$. Hence B has the form

$$(5.5) \quad B = \begin{pmatrix} B_1 & 0 & 0 & \dots \\ 0 & B_2 & 0 & \dots \\ 0 & 0 & B_3 & \dots \\ \cdot & \cdot & \cdot & \dots \end{pmatrix}, \quad (B_k \text{ of order } m_k)$$

We can choose a coordinate system in the invariant sub-spaces corresponding to each B_i in such a way that B has a diagonal form. Under this coordinate transformation the form of A is unchanged. If we continue in this manner we can obtain a complete reduction of the representation into 1-dimensional representations. To obtain the 1-dimensional representations we take a basis $a_1, \dots, a_r$ such that every element $a = a_1^{i_1} a_2^{i_2} \dots a_r^{i_r}$ and the order of a_i is h_i ; h , order of $G = \prod h_i$. To each a_i corresponds an h_i th root of 1, ζ_i , and hence to a there corresponds $\zeta_1^{i_1} \zeta_2^{i_2} \dots \zeta_r^{i_r}$. We obtain in this way h inequivalent irreducible representations of G into which every representation can be decomposed.

The argument for the decomposability of the representations of a finite abelian group can be put in such a form that it applies to quite general representations of any abelian group. We assume merely that the transformation which corresponds to any single element of G is decomposable into 1-dimensional parts; then we shall show that all the transformations are simultaneously decomposable in this way. Suppose that $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2 + \dots + \mathcal{R}_k$ is a decomposition of $\mathcal{R}$ into independent sub-spaces invariant with respect to G . If a transformation A is not a mere multiplication (transformation of $x \rightarrow \rho x$) in one of the sub-spaces $\mathcal{R}_i$, then $\mathcal{R}_i = \mathcal{R}_i' + \mathcal{R}_i'' + \dots$ where the operator A is the multiplication by ρ_i' in $\mathcal{R}_i'$, by ρ_i'' in $\mathcal{R}_i''$ etc.: $A x_i' = \rho_i' x_i'$ for $x_i' \in \mathcal{R}_i'$ etc. Then $A B x_i' = B A x_i' = \rho_i' B x_i'$ or $B x_i'$ also belongs to $\mathcal{R}_i'$. Thus $\mathcal{R}_i', \mathcal{R}_i'', \dots$ are invariant sub-spaces under all the transformations and we obtain a further decomposition. We continue in this way until we have a decomposition of $\mathcal{R}$ into sub-spaces such that every transformation A in these sub-spaces is a mere multiplication. But then all the A 's have diagonal form. By choosing a basis for this sub-space we obtain a decomposition of it and then of the whole $\mathcal{R}$ into 1-dimensional invariant spaces.

In particular, if the representation of an abelian group is unitary it can be decomposed into representations of degree one.

For the character of a unitary representation we have

$$(5.6) \quad \overline{\chi(s)} = \chi(s^{-1})$$

for $\text{tr } \bar{U}(s) = \text{tr } U^{-1}(s) = \text{tr } U(s^{-1})$.

We consider next some formal processes by which new representations can be constructed from given ones.

Addition of representations: If $\mathcal{G}_1$ and $\mathcal{G}_2$ are two representations of $\mathcal{G}$ acting on $\mathcal{R}_1$ and $\mathcal{R}_2$ respectively, we form the direct sum $\mathcal{G}_1 + \mathcal{G}_2$. It is the vector space obtained by considering $\mathcal{R}_1$ and $\mathcal{R}_2$ as independent and by "adding" the vectors of $\mathcal{R}_1$ and $\mathcal{R}_2$ to obtain new vectors. This "addition" can be defined abstractly by pairs of vectors of $\mathcal{R}_1$ and $\mathcal{R}_2$. If a coordinate system of $\mathcal{R}$ is determined by coordinate systems for $\mathcal{R}_1$ and $\mathcal{R}_2$, then the matrices of the sum $\mathcal{G}_1 + \mathcal{G}_2$ have the form

$$(5.7) \quad \begin{pmatrix} U_1(s) & 0 \\ 0 & U_2(s) \end{pmatrix}$$

Multiplication of representations: If under a linear transformation $\underline{A}$ of $\mathcal{R}$ an arbitrary vector $(\xi_1 \xi_2 \dots \xi_n) \rightarrow (\xi'_1 \xi'_2 \dots \xi'_n)$ and under $\underline{B}$ of $\mathcal{V}$ the vector $(\eta_1 \dots \eta_m) \rightarrow (\eta'_1 \dots \eta'_m)$, then the products $\xi_i \eta_k$ are transformed linearly into $\xi'_i \eta'_k$. If these products are arranged in some definite order they constitute a vector space and the transformation induced in this vector space by $\underline{A}$ and $\underline{B}$ is called the direct product of $\underline{A}$ and $\underline{B}$, $\underline{A} \times \underline{B}$. The multiplication of direct products satisfies

$$(5.8) \quad (\underline{A} \times \underline{B})(\underline{A}_1 \times \underline{B}_1) = \underline{A} \underline{A}_1 \times \underline{B} \underline{B}_1$$

If $\mathcal{G}_1$ and $\mathcal{G}_2$ are two representations, then the set of transformations $U_1(s) \times U_2(s)$ gives a representation of $\mathcal{G}$ called the direct product $\mathcal{G}_1 \times \mathcal{G}_2$ of $\mathcal{G}_1$ and $\mathcal{G}_2$. By direct computation we have

$$(5.9) \quad \text{tr } (\underline{A} \times \underline{B}) = \text{tr } \underline{A} \text{tr } \underline{B}$$

and hence, if $\mathcal{G}_3 = \mathcal{G}_1 \times \mathcal{G}_2$,

$$(5.10) \quad \gamma_3(s) = \gamma_1(s) \gamma_2(s).$$

Suppose $\mathcal{G}_1$ and $\mathcal{G}_2$ are two groups. Consider the pairs (s_1, s_2) $s_i \in \mathcal{G}_i$ for which multiplication is defined by $(s_1, s_2)(t_1, t_2) = (s_1 t_1, s_2 t_2)$. The system of pairs with this law of composition is a group called the direct product $\mathcal{G}_1 \times \mathcal{G}_2$ of $\mathcal{G}_1$ and $\mathcal{G}_2$. If $\mathcal{G}_i$ is a representation of $\mathcal{G}_i$ then the set of transformations $U_1(s_1) \times U_2(s_2)$ is a representation of $\mathcal{G}_1 \times \mathcal{G}_2$. By letting $U_1(s_1)$ correspond to (s_1, s_2) , $\mathcal{G}_1$ may be taken as a representation of $\mathcal{G}_1 \times \mathcal{G}_2$ and similarly for $\mathcal{G}_2$. The representation whose elements are $U_1(s_1) \times U_2(s_2)$ is then the direct product $\mathcal{G}_1 \times \mathcal{G}_2$.

Specialization to a sub-group. If we have a representation of a group $\mathcal{G}$ we may obtain from it a representation of any sub-group $\mathcal{G}'$ by taking those transformations which correspond to the elements of $\mathcal{G}'$.

Representation of the linear group. If $\mathcal{G}$ is an n-dimensional representation of $\mathcal{G}$: $s \rightarrow u(s)$ and $\mathcal{L}$ a representation in N-dimensions of the full linear group of n-dimensions (all non-singular matrices of n dimensions): $u \rightarrow U$; then $s \rightarrow u(s) \rightarrow U(s)$ gives an N-dimensional representation $s \rightarrow U(s)$ of $\mathcal{G}$. For example, the matrix correspondence $u \rightarrow u^{-1}$ is a representation of the linear group. Hence if $s \rightarrow u(s)$ is a representation, so is $s \rightarrow u^{-1}(s)$, the contragredient representation.

6. Schur's lemma.

Let $\{U(s)\}$ be the matrices of an irreducible set of n-dimensional transformations where s takes on the values of some range not necessarily a group, and $\{V(s)\}$ those of an m-dimensional irreducible set and suppose that A is a matrix of n rows and m columns such that

$$(6.1) \quad U(s) A = A V(s)$$

for all s . Schur's lemma states that there are only two possibilities in such a situation: 1) $A = 0$ 2) $m = n$ and A non-singular. In the second case the two sets of transformations are equivalent. Let a_i denote the vector whose components are the elements of the i th column of A , $a_i = (\alpha_{1i}, \alpha_{2i}, \alpha_{3i}, \dots, \alpha_{ni})$. Equation

(6.1) states that $U(s)a_i$, which is the i th column of AV , is a linear combination of the a_i or that the space $\mathcal{R}' = (a_1 \ a_2 \ \dots \ a_m)$ is invariant with respect to the transformations $U(s)$. Hence, either $\mathcal{R}' = 0$, i.e. $A = 0$ or $\mathcal{R}'$ is the whole space i.e. $m \geq n$ and rank of $A = n$. If $A \neq 0$ we take transposed matrices and obtain from (6.1)

$$V'(s)A' = A'U'(s)$$

and hence $n \geq m$ or $n = m$ and A is non-singular.

Schur's lemma may be given the following equivalent form: If (6.1) holds and 1) $\{U(s)\} \not\sim \{V(s)\}$, then $A = 0$, 2) $U(s) = V(s)$, the set of matrices A satisfying (6.1) is a division algebra. So far we have made no assumption as to the nature of the underlying number field. However, in the case in which it is the complex continuum, the matrices A are multiples of the unit matrix. For if A satisfies (6.1) with $U(s) = V(s)$, so does $A - \alpha E$. If α is a root of the characteristic equation of A , then $A - \alpha E$ is singular and hence by Schur's lemma it must be zero, or $A = \alpha E$. In case the number field $\mathcal{F}$ is not algebraically closed, suppose that

$$\varphi(\lambda) = \varphi_1(\lambda)\varphi_2(\lambda)\dots\varphi_r(\lambda)$$

is a factorization of the characteristic function of A into prime factors in the field. $|\varphi_i(A)| = |A - \alpha_{i_1}| |A - \alpha_{i_2}| \dots = 0$ if α_{i_k} are the roots of φ_i in an extension of $\mathcal{F}$. But then $\varphi_i(A) = 0$ and since the φ_i are prime they must all be equal. Hence the characteristic function of A is a power of an irreducible polynomial in $\mathcal{F}$.

7. The Burnside-Frobenius-Schur theorem.

(1) Burnside's Theorem: If $\{U(s)\}$ is an irreducible multiplicative system of matrices, then the coordinates $u_{ik}(s)$ are linearly independent. (and conversely).

By a multiplicative system we mean one that is closed under multiplication. If we take the linear set of matrices generated by this system we obtain an algebra. The linear independence is expressed by:

$$(7.1) \quad \sum_{i,k} \lambda_{ik} u_{ki}(s) = \text{tr } U(s) L = 0$$

implies $L = (\lambda_{ik}) = 0$. Burnside's theorem states that if the special relations among the u 's which express the reducibility of a system of matrices do not hold, then no relations other than the trivial one $\text{tr } U(s) = 0 = 0$ can hold.

Proof: Consider the set of matrices L which satisfy (7.1). If this set does not consist of the 0 matrix alone, it contains a sub-set having the maximum number of first columns, say $h - 1$ ($< n$), zero, but with the h th column different from zero. We call the matrices of this sub-set special L -matrices. Let $\mathcal{R}'$ denote the sub-space generated by the h th column vectors of the special L -matrices. [The special L -matrices together with the 0 matrix form a vector space of the same number of dimensions as $\mathcal{R}'$.] If L is a special L -matrix, so is $L' = U_0 L$ where U_0 is any matrix of the set $U(s)$. For,

$$\text{tr } (UL') = \text{tr } (UU_0 L) = \text{tr } (U_1 L) = 0$$

where $U_1 = UU_0 \in \{U(s)\}$. Thus if χ is the h th column of L the vector $U_0 \chi \in \mathcal{R}'$, i.e., is the h th column of a special L -matrix. Hence $\mathcal{R}'$ is an invariant sub-space $\neq 0$, and so by the irreducibility of the system $\mathcal{R}' = \mathcal{R}$. Hence there are n special L -matrices $L_1, L_2, \dots, L_n$ such that the h th column χ_i of L_i is $(0 \dots 0 \overset{i}{1} 0 \dots 0)$ and every special L -matrix belongs to the linear set $(L_1, L_2, \dots, L_n)$. It follows that

$$UL_r = \sum_i a_{ir} L_i \quad (r = 1, 2, \dots, n)$$

and comparing h th columns, $a_{ir} = u_{ir}$ or

$$(7.2) \quad UL_r = \sum_i u_{ir} L_i$$

If we compare elements in the j th column of the two sides of (7.2), we obtain

$$(7.3) \quad \sum_k u_{tk} \lambda_{kj}^{(r)} = \sum_i u_{ir} \lambda_{tj}^{(i)} \quad (t = 1 \dots n)$$

where $(\lambda_{1j}^{(i)}, \lambda_{2j}^{(i)}, \dots, \lambda_{nj}^{(i)})$ is the j th column of L_i . If Λ_j is the matrix having $\lambda_{tj}^{(i)}$ as its element in the t th row and i th column, (7.3) may be written as

$$(7.4) \quad U \Lambda_j = \Lambda_j U \quad (j = 1, \dots, n)$$

By Schur's lemma $\lambda_{tj}^{(i)} = \lambda_j \delta_t^i$ and so by (7.1)

$$(7.5) \quad \sum_j \lambda_j u_{ji}(s) = 0$$

or $(\lambda_1, \lambda_2, \dots, \lambda_n)$ is invariant under the set $U'(s)$. Hence $\lambda_i = 0$ (except in the trivial case in which $n = 1$) and all the special L-matrices vanish, contrary to hypothesis.

(2) The generalization of Frobenius-Schur:

If $\{U(s)\}, \{V(s)\} \dots$ are inequivalent irreducible multiplicative systems of transformations, then their coordinates are linearly independent, i.e.

$$(7.6) \quad \text{tr } U(s)L + \text{tr } V(s)M + \dots = 0$$

implies $L = M = \dots = 0$.

We consider sets of matrices $(A, B, C, \dots)$ where A has the same order as U, B the same as V, etc. We define

$$\alpha(A, B, \dots) = (\alpha A, \alpha B, \dots)$$

$$(A, B, \dots) + (A_1, B_1, \dots) = (A + A_1, B + B_1, \dots)$$

$$(A, B, \dots)(A_1, B_1, \dots) = (AA_1, BB_1, \dots).$$

We obtain in this way the direct sum in the terminology of the theory of algebras.

Consider the set of elements $(L, M, \dots)$ which satisfy (7.6). If this set has any elements other than $(0, 0, \dots)$ then there exists a sub-set of elements called special sequences having a maximum number, say $h - 1$, of first columns = 0 while the h th column $\neq 0$. (We consider the columns of $(L, M, \dots)$ in the following order: first the columns of L, then the columns of M etc.) We may suppose that $h \leq n$ where n is the number of rows and columns of $U(s)$. For otherwise we drop a certain number of the systems $U(s), V(s) \dots$ and consider the remaining ones. Now, the proof proceeds as before:

[The special sequences have a basis of n independent elements] say

$(L_k, M_k, \dots)$ where $k = 1, 2, \dots, n$, and the h th column of L_k is $(0 \dots 0 \overset{k}{1} 0 \dots 0)$.

$(U, V, \dots)(L_k, M_k, \dots)$ is a special sequence and hence

$$(7.7) \quad \begin{aligned} U L_k &= \sum u_{ik} L_i \\ V M_k &= \sum v_{ik} M_i \end{aligned}$$

If $\Lambda_j = (\lambda_{kj}^{(i)})$ (k row, i column), $M_p = (\mu_{qp}^r) \dots$, then

(7.8)

$$UA_j = \Lambda_j U$$

$$V M_p = M_p U$$

Hence by Schur's lemma $M_p = 0$, ... and we are brought back to the situation of Burnside's theorem.

(3) Applications:

Inequivalent irreducible representations cannot have the same character. For if $U^{(1)}(s)$ and $U^{(2)}(s)$ determine the representations and $\chi^{(1)}(s) = \chi^{(2)}(s)$, we have a special linear relation between the coordinates of $U^{(1)}$ and $U^{(2)}$ contrary to the above theorem. Thus the character of an irreducible representation determines the representation completely.

By this type of reasoning we can obtain a new proof of the uniqueness of the irreducible components of a representation. For, suppose that $U(s) = n_1 U_1(s) + n_2 U_2(s) + \dots = m_1 U_1(s) + m_2 U_2(s) + \dots$ where the expression indicates that in the first reduction there are n_1 irreducible components equivalent to U_1 , and in the second m_1 , etc. Then $n_1 \chi_1(s) + n_2 \chi_2(s) + \dots = m_1 \chi_1(s) + m_2 \chi_2(s) + \dots$ or $(n_1 - m_1) \chi_1(s) + (n_2 - m_2) \chi_2(s) + \dots = 0$. By the Burnside-Frobenius-Schur theorem $n_k = m_k$. Of course, this proof, as Schur's lemma, holds only for algebraically closed fields. In case the representations are determined by their irreducible components, i.e. if the theorem of decomposability holds, then every representation is determined by its character.

We shall now apply the above results to the representations of the direct product of two groups. Let $\{U(s)\}$ and $\{V(t)\}$ be irreducible representations of $\mathcal{G}$ and $\mathcal{H}$ respectively, then $U(s) \times V(t)$ is an irreducible representation of $\mathcal{G} \times \mathcal{H}$. To prove
The elements of the matrix $U(s) \times V(t)$ are the products $u_{ik}(s) v_{\alpha\beta}(t)$. Suppose

pose
(7.9)

$$\sum \gamma_{ik, \alpha\beta} u_{ik}(s) v_{\alpha\beta}(t) = 0$$

where $\gamma_{ik, \alpha\beta}$ is independent of s and t . For any fixed s (7.9) gives a linear relation between the elements $v_{\alpha\beta}(t)$. Hence by Burnside's theorem

$$\sum_{i,k} \gamma_{ik,\alpha\beta} U_{ik}(s) = 0$$

for all s and $\alpha, \beta = 1, 2, \dots$. Applying the theorem to the irreducible representation $\{U(s)\}$,

$$\gamma_{ik,\alpha\beta} = 0$$

i.e. the elements of $U(s) \times V(t)$ are linearly independent and the representation is irreducible.

Theorem: If every representation of $\mathfrak{g}$ and of $\mathfrak{f}$ is decomposable, then so is every representation of $\mathfrak{g} \times \mathfrak{f}$ and its irreducible components are equivalent to direct products of irreducible components of $\mathfrak{g}$ and $\mathfrak{f}$.

Let $W(s, t)$ be the representing matrix of (s, t) of $\mathfrak{g} \times \mathfrak{f}$ and set $W(s, i) = U(s)$, $W(i, t) = V(t)$. The sub-set of $\mathfrak{g} \times \mathfrak{f}$ of elements of the form (s, i) is a group $\cong \mathfrak{g}$, so that the matrices $U(s)$ give a representation of $\mathfrak{g}$. Similarly $V(t)$ is a representation of $\mathfrak{f}$. Since every representation of $\mathfrak{g}$ is decomposable we may choose a coordinate system in such a way that $U(s)$ has the form

$$(7.10) \quad U = \begin{pmatrix} U_1 & 0 & \dots & & \\ 0 & U_2 & & & \\ \vdots & & \ddots & & \\ & & & U_{n_1} & \\ \hline & 0 & & & U_{n_1+1} & \dots \end{pmatrix}$$

where the blocks off the "diagonal" are all zero and the first n_1 are equal, the next n_2 equal but not equivalent to the first etc. From

$$(s, i)(i, t) = (i, t)(s, i) = (s, t)$$

follows

$$(7.11) \quad U(s)V(t) = V(t)U(s) = W(s, t)$$

Suppose the matrix $V(t)$ is blocked off into sub-matrices $V_{ik}(t)$ in the same way as $U(s)$ is in (7.10). The condition (7.11) in terms of the blocks is

$$U_i(s)V_{ik}(t) = V_{ik}(t)U_k(s)$$

By Schur's lemma, if $U_i(s) \neq U_k(s)$, $V_{ik}(t) = 0$, and if $U_i(s) = U_k(s)$, $V_{ik}(t) =$

$= \beta_{ik}(t) E_{ik}$ where E_{ik} is the identity matrix in the block i, k . Thus

$$(7.12) \quad V = \left(\begin{array}{ccc|ccc} \beta_{11} E_{11} & \beta_{12} E_{12} & \cdots & \beta_{1m_1} E_{1m_1} & & \\ \vdots & \vdots & & \vdots & & \\ \beta_{m_1} E_{m_1} & \beta_{m_2} E_{m_2} & \cdots & \beta_{m_1 m_1} E_{m_1 m_1} & & \\ \hline & & & & \beta_{m_1+1,1} E_{m_1+1,1} & \cdots \\ & & & & \vdots & \\ & & & & & \end{array} \right)$$

We obtain in this way a partial decomposition of $W(s, t)$ where the components are products of matrices of the form

$$(7.13) \quad \begin{pmatrix} U_1 & 0 & \cdots & \cdots \\ 0 & U_1 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & 0 & U_1 \end{pmatrix} \quad \begin{pmatrix} \beta_{11} E_{11} & \beta_{12} E_{12} & \cdots \\ \beta_{21} E_{21} & \beta_{22} E_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

(We have set $U_1 = U_2 = \cdots = U_{n_1}$). We denote the matrices in (7.13) by U and V again and note that

$$U = U_1 \times E_1 \quad V = E_2 \times V_1$$

where E_1 is the identity of n_1 rows and columns, E_2 is the identity of the same number of rows and columns as U_1 and $V_1 = (\beta_{ik})_{n_2 \times n_2}$. The matrices V_1 give a representation of $\mathcal{F}$. If this representation is irreducible we obtain from (7.13)

$$UV = U_1 \times V_1$$

which is an irreducible representation of $\mathcal{G} \times \mathcal{F}$. If $\{V_1(t)\}$ is not irreducible we may decompose it into irreducible components which may be used in place of V_1 in the argument just made.

8. The algebra of a finite group and its reduction

The methods used in the theory of representations of finite groups are fundamental for the representation of continuous groups. We consider them briefly in this and the next sections.

Suppose that G is a finite group and the correspondence $s \rightarrow \underline{U}(s)$ where $\underline{U}(s)$ is a linear transformation of $\mathcal{R}$ is a representation of G . We consider the elements of G as linear operators of $\mathcal{R}$ given through the representation by the equation

$$s(x) = \underline{U}(s)x$$

where x is any vector of $\mathcal{R}$. We may form linear combinations of these operators and obtain an algebra of operators consisting of all the elements of the form $\sum_{s \in G} \alpha_s S$ where

$$(8.1) \quad \begin{aligned} \sum \alpha_s S + \sum \beta_s S &= \sum (\alpha_s + \beta_s) S \\ \alpha \sum \alpha_s S &= \sum \alpha \alpha_s S \\ \sum \alpha_s S \sum \beta_t t &= \sum_s \left(\sum_r \alpha_{sr^{-1}} \beta_r \right) S \end{aligned}$$

It is evident from (8.1) that the addition and multiplication of elements of this algebra depends only on the group G and not on the representation. The linear dependence of the operators, i.e. the representation of the 0 operator or the topology of the algebra does depend on the representation. As we shall see, there exist representations for which the operators s corresponding to the group elements are linearly independent. Apart from this fact, we can define an algebra $\mathcal{A}$ by (8.1) where we suppose that the s 's are independent. This algebra, which is uniquely determined by G and the field of the coefficients, is called the group algebra of G .

A representation of $\mathcal{A}$ is a set $\underline{U}(a)$ of linear transformations of a space

such that $a \rightarrow \underline{U}(a)$ and

$$(8.2) \quad \underline{U}(a + b) = \underline{U}(a) + \underline{U}(b), \quad \underline{U}(\alpha a) = \alpha \underline{U}(a), \quad \underline{U}(ab) = \underline{U}(a)\underline{U}(b).$$

The transformations which correspond to the elements s of G give a representation of

Suppose an algebra $\mathcal{A}$ is represented by linear transformations of a space $\mathcal{V}$ and let $\mathcal{I}$ be a left ideal of $\mathcal{A}$. If y is any vector of $\mathcal{V}$ and $\mathcal{I}y$ denotes the set of transforms of y under the transformations corresponding to the elements of $\mathcal{I}$, then $\mathcal{I}y$ is an invariant sub-space of $\mathcal{V}$. For, if a is an arbitrary element of $\mathcal{A}$ and b an arbitrary element of $\mathcal{I}$ then

$$a(by) = (ab)y = b'y \quad b' \in \mathcal{I}$$

or in symbols

$$\mathcal{A}(\mathcal{I}y) = (\mathcal{A}\mathcal{I})y = \mathcal{I}y.$$

The elements b of $\mathcal{I}$ for which $by = 0$ form an invariant sub-space or left-ideal $\bar{\mathcal{I}}$ of $\mathcal{I}$. Hence if $\mathcal{I}$ is irreducible, $\bar{\mathcal{I}} = 0$ or $= \mathcal{I}$. In the former case

$(o_1y, o_2y \dots o_gy)$ is an independent basis for $\mathcal{I}y$ if $(o_1 \dots o_g)$ is such a basis for $\mathcal{I}$. Then by (8.4),

$$(8.5) \quad a(o_ky) = (ao_k)y = \alpha_{k1}(o_1y) + \alpha_{k2}(o_2y) + \dots + \alpha_{kg}(o_gy) \quad (k = 1 \dots g)$$

and hence the representation determined by $\mathcal{I}y$ is equivalent to the representation determined by $\mathcal{I}$.

Suppose now that the regular representation of $\mathcal{A}$ is decomposable. Then

$$(8.6) \quad \mathcal{A} = \mathcal{I}_1 + \mathcal{I}_2 + \dots + \mathcal{I}_f$$

where the $\mathcal{I}_j$ are irreducible invariant sub-spaces of $\mathcal{A}$ or minimal left ideals such that $\mathcal{I}_j \cap \mathcal{I}_k = 0$. Let $(y_1 y_2 \dots y_n)$ be a basis for the space $\mathcal{V}$ of a representation of $\mathcal{A}$. We shall show that this representation is decomposable. Since $\mathcal{A}$ has an identity, we have

$$(8.7) \quad \mathcal{V} = \mathcal{A}\mathcal{V} = \mathcal{I}_1\mathcal{V} + \dots + \mathcal{I}_f\mathcal{V} = \mathcal{I}_1y_1 + \mathcal{I}_1y_2 + \dots + \mathcal{I}_fy_n.$$

We have seen above that if $\mathcal{I}_ky_j \neq 0$, it is a sub-space which gives a representation of $\mathcal{A}$ equivalent to the representation given by $\mathcal{I}_k$. It follows that if $\mathcal{I}_ky_j \neq 0$, it is an irreducible invariant sub-space of $\mathcal{V}$. Since the intersection of two invariant sub-spaces is also invariant we have $\mathcal{I}_ky_j \cap \mathcal{I}_hy_l = 0$, or

$$\mathcal{I}_ky_j = \mathcal{I}_hy_l \quad \text{Hence we may omit in (8.7) the spaces } \mathcal{I}_ky_j \text{ which are zero, and}$$

select from the remainder a set of independent invariant sub-spaces which determine

γ : Thus, theorem I is established.

In order to prove theorem II we shall show first that every ideal of a group algebra is generated by a single idempotent element, i.e. $\mathcal{I} = \mathcal{A}e$ where $e^2 = e$. For example, from (8.5)

$$i = i_1 + i_2 + \dots + i_f \quad i_k \in \mathcal{I}_k$$

$$a = a_1 + a_2 + \dots + a_f = ai_1 + ai_2 + \dots + ai_f$$

and since $ai_k \in \mathcal{I}_k$ and the $\mathcal{I}_i$'s are independent, $ai_k = a_k$. It follows that

$$\mathcal{I}_k = \mathcal{A}i_k \quad i_k^2 = i_k \quad i_k i_j = 0 \quad (k \neq j).$$

If the ideal $\mathcal{I}$ is generated by a single idempotent quantity e , $\mathcal{I} = \mathcal{A}e$, then the correspondence which sends a into $ae = a'$ is linear and has the properties:

$$(1) \quad a' \in \mathcal{I}$$

$$(2) \quad \text{if } a \in \mathcal{I}, \text{ then } a' = a.$$

Now, consider any linear transformation $\mathcal{A} \xrightarrow{\text{on } \mathcal{A}} \mathcal{A}$ which has the properties (1) and (2). (If $(o_1 \dots o_g)$ is a coordinate system for $\mathcal{A}$ and $(o_1 \dots o_g)$ a coordinate system for $\mathcal{I}$, then we have merely to put $o'_i = o_i$ ($i = 1 \dots g$) and $o'_k =$ linear combination of $o_1 \dots o_g$ ($k = g+1 \dots h$) to obtain such a correspondence). Suppose

$$(8.8) \quad a = \sum \alpha_s s = \sum \alpha(s) s \rightarrow \sum \alpha'(s) s = a'$$

$$\alpha'(s) = \sum_t \gamma(s, t) \alpha(t).$$

The correspondence $a \rightarrow r^{-1}a' = a''$, $r^{-1} \in \mathcal{G}$ is also a mapping satisfying (1) and (2). It sends $a \rightarrow a''$ where

$$\alpha''(s) = \sum_t \gamma(rs, {}^r t) \alpha(t).$$

Finally, the correspondence $a \rightarrow a^*$ where

$$(8.9) \quad \alpha^*(s) = \frac{1}{h} \sum_t \sum_r \gamma(rs, {}^r t) \alpha(t) = \sum_t \mathcal{E}(s, t) \alpha(t)$$

also has the desired properties. From (8.9) we have that $\mathcal{E}(rs, rt) = \mathcal{E}(s, t)$, i.e. $\mathcal{E}(s, t)$ as a function of s and t depends only on $t^{-1}s$ and we may set $\mathcal{E}(s, t) = \mathcal{E}(t^{-1}s)$. The element $e = \sum_s \mathcal{E}(s) s$ is the generating idempotent element of $\mathcal{I}$.

For by (8.9), $a^* = ae$.

We may now complete the proof of II. Suppose that $\mathcal{I}_1$ is a minimal left-ideal of $\mathcal{A}$ (irreducible invariant sub-space). Then $\mathcal{I}_1 = \mathcal{A}i_1$ where i_1 is idempotent. $\mathcal{I}'_1 = \mathcal{A}(i - i_1)$ is a left ideal of $\mathcal{A}$ such that $\mathcal{A} = \mathcal{I}_1 + \mathcal{I}'_1$.

$\mathcal{I}_1 \cap \mathcal{I}'_1 = 0$. For, if $x_1 \in \mathcal{I}_1$ and $x'_1 \in \mathcal{I}'_1$

$$x_1 i_1 = x_1$$

$$x'_1 i_1 = 0$$

so that if $x_1 = x'_1$, then $x_1 = x'_1 = 0$. Further, for any x in $\mathcal{A}$ we have

$$x = x i_1 + (x - x i_1) = x_1 + x'_1. \quad \text{If } \mathcal{I}'_1 \text{ is not a minimal left-}$$

ideal of $\mathcal{A}$, then it contains such an ideal $\mathcal{I}_2$ and $\mathcal{I}_2 \cap \mathcal{I}_1 = 0$. $\mathcal{I}_1 + \mathcal{I}_2$ is a left-ideal of $\mathcal{A}$ and $\mathcal{A} = \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}'_2$. We repeat the argument with $\mathcal{I}'_2$ in place of $\mathcal{I}'_1$, and continuing in this way we obtain a decomposition of $\mathcal{A}$ into minimal left-ideals.

The above proofs are strictly rational. I requires no restriction on the underlying number field, while II requires only the mild restriction that the characteristic of the field shall not be a divisor of the order of the group algebra.

As a simple application of I we mention the theorem that there is only one irreducible representation of a division algebra in the field of its coefficients. For it is easily seen that a division algebra has no left-ideals other than itself and 0.

9. The hermitian argument.

A second proof of the decomposability of the representations of a finite group is the one based on the hermitian argument. It is considerably simpler than the preceding proof, but it has the disadvantage of holding only for special number fields such as the complex continuum.

Let $s \rightarrow \underline{U}(s)$ define a representation of the finite group $\mathcal{G}$. If we perform the transformation $\underline{U}(s)$ on the vector $x = (\xi_1 \xi_2 \dots \xi_n)$ the unit hermitian form $(x, \bar{x}) = \sum \xi_i \bar{\xi}_i$ is transformed into the positive definite hermitian form $(\underline{U}(s)x, \underline{U}(s)\bar{x}) = (x, \underline{U}^*(s)\bar{x}) = H(s)$. Then the form $H_0 = \sum_s H(s)$ is positive definite and invariant under the transformations of the group, i.e. the transformations $x \rightarrow \underline{U}(s)x$. By a change of coordinates H_0 can be transformed into the unit form, and hence the matrices $\underline{U}(s)$ can be transformed into unitary matrices. The theorem of decomposability for the representation of $\mathcal{G}$ then follows from the corresponding theorem for sets of unitary transformations.

Chapter 2.

Survey of Lie's Theory

1. Infinitesimal generation of transformation groups

As an introduction to the Lie theory we consider the group of rotations of a rigid body about a fixed point. In this connection a rotation means a single transformation

$$x' = f(x, y, z) \quad y' = g(x, y, z) \quad z' = h(x, y, z)$$

a displacement of the rigid body from one position to another about a fixed point.

But a "rotation" also calls to mind a continuous set of transformations.

$$(1.1) \quad x' = f(x, y, z; t) \quad y' = g(x, y, z; t) \quad z' = h(x, y, z; t) \quad (t = \text{time}),$$

the displacements obtained by spinning the rigid body about a fixed axis through the fixed point - an axial sub-group. The first sense of rotation is the more technical, the second the more intuitive. The bond between them is the fact that a rotation in the first sense has an axis of fixed points and thus determines (uniquely) a rotation in the second sense.

From the intuitive standpoint a rotation in the second sense breaks up into a succession of infinitesimal steps, iterating an initial one. In this way we are led to the infinitesimal transformations

$$(1.2) \quad dx' = (bx' - cy')dt, \quad dy' = (cx' - az')dt, \quad dz' = (ay' - bx')dt$$

The vectors $(\frac{dx'}{dt}, \frac{dy'}{dt}, \frac{dz'}{dt})$ belong to a 3-dimensional linear family consisting of all vectors of the form $a(o, -\frac{z'}{y'}, \frac{x'}{z'}) + b(\frac{z'}{x'}, o, -x') + c(-y', \frac{x'}{z'}, o)$. It is this idea of the generation of a group of transformations from infinitesimal transformations that is fundamental in the Lie theory.

We consider an r -dimensional continuous group of transformations s of a point field such that $x_o \rightarrow x = sx_o$. Suppose that a coordinate system can be determined for a neighborhood of the identity element of the group manifold in such a way that the identity is given by the coordinates $s^1 = o, s^2 = o, \dots, s^r = o$. If the s^i are sufficiently small, the corresponding transformation will be near the

$$s^\alpha + ds^\alpha = \psi^\alpha(0 + \delta s^\beta; s^\gamma) = \psi^\alpha(0; s^\gamma) + \psi^\alpha_\beta \delta s^\beta,$$

where $\psi^\alpha_\beta \equiv \frac{\partial \psi^\alpha}{\partial s^\beta} \Big|_{\substack{\text{second set of arguments} = s^\gamma \\ \text{wrt first set of arguments; they} = 0^\gamma}}$. But $\psi^\alpha(0; s^\gamma) = s^\alpha$,

so, inverting, $\delta s^\beta = \lambda^\beta_\alpha ds^\alpha$. Etc.

identity and, since the transformation group is continuous, will send x_0 into a nearby point x . We assume that the point field can be parameterized and so we can suppose that x_0 and x can be assigned coordinates in a single coordinate system, say $x_0 = (x_0^1 \dots x_0^n)$, $x = (x^1 \dots x^n)$, and we have

$$(1.3) \quad x^i = \varphi^i(s^1, \dots, s^r; x_0^1 \dots x_0^n) = \varphi^i(s, x_0) \quad (i = 1, \dots, n)$$

where $\varphi^i(0, x_0) = x_0^i$.

The condition $\bar{s} = ss_0$, that the s 's form a group, is expressed for s 's in a suitable neighborhood of the identity by

$$(1.4) \quad \bar{s}^\alpha = \varphi^\alpha(s^1, \dots, s^r; s_0^1, \dots, s_0^r) \quad (\alpha = 1, \dots, r)$$

and since the transformation group is a realization of the abstract group of the s 's we have ${}_0x \xrightarrow{s} x_0 \xrightarrow{\bar{s}} x$ implies ${}_0x \xrightarrow{\bar{s}} x$:

$$(1.5) \quad \varphi^i(\bar{s}; {}_0x) = \varphi^i(s; x_0)$$

If the s^α in (1.3) are infinitesimals δs^α then the displacement of x_0 is given by

$$(1.6) \quad dx_0^i = \left(\frac{\partial \varphi^i}{\partial s^\alpha} \right)_{s=0} \delta s^\alpha = u_\alpha^i(x_0) \delta s^\alpha$$

if we disregard terms of higher order. (In (1.6) and in what follows we use the convention that a repeated index unless otherwise stated indicates summation on that index.) The transformation (1.6) is called an infinitesimal transformation for the identity of the group and corresponding to it we have the transformation

$x_0 \rightarrow (1 + \delta s)x_0 = x_0 + dx_0$. The infinitesimal transformation belongs to the velocity field of the vectors $u_\alpha^i(x_0)$. Since the group is r -dimensional, the velocity vectors u_α are linearly independent.

A general infinitesimal transformation is defined by

$$(1.7) \quad dx^i = \frac{\partial \varphi^i}{\partial s^\alpha} ds^\alpha \quad : \inf T (s + ds) s^{-1}$$

and corresponding to it we have the transformation $x_0 \rightarrow (s + ds)x_0 = (s + ds)s^{-1}x = (1 + \delta s)x = (1 + \delta s)sx_0$. We may write symbolically $s + ds = (1 + \delta s)s$. Similarly if we determine ds from $d\bar{s}$ by $\bar{s} + d\bar{s} = (s + ds)s_0$, and δs from ds as above, then, by the associative law, $\bar{s} + d\bar{s} = (1 + \delta s)\bar{s}$. It follows from this and from (1.4)

see opp page

Know that $(s+ds)s^{-1} : dx^i = \frac{\partial \varphi^i}{\partial s^\alpha} ds^\alpha$ and
 (inf. transf. for the group identity)

$1 + \delta s : dx^i = U_\alpha^i \delta s^\alpha$ are the same T for $ds^\alpha, \delta s^\alpha$

connected by $\delta s^\alpha = \lambda^\alpha_\beta ds^\beta$ (which comes from (1.4) via

$(s+ds)s^{-1} \equiv 1 + \delta s$). Subst. for δs^β , from the lin ind of ds^α
 into the identity (1.9).

that the correspondence $(ds^1 \dots ds^r) \rightarrow (ds^1 \dots ds^r)$ is given by

$$(1.8) \quad \delta s^\beta = \lambda_\rho^\beta(s) ds^\rho = \lambda_\rho^\beta(\bar{s}) d\bar{s}^\rho$$

If we specialize (1.7) by setting $ds^\rho = \delta_\alpha^\rho \epsilon$, we obtain by means of (1.6) and

(1.8),

$$(1.9) \quad \frac{\partial \varphi^i}{\partial s^\alpha} = u_\sigma^i(x) \lambda_\alpha^\sigma(s)$$

i.e. the vectors $\frac{\partial \varphi^i}{\partial s^\alpha}$ belong to our velocity field.

The transformation $s_0 \rightarrow ss_0$ of the group manifold is a realization of the group G , called the parameter group of G . Now, $\lambda_\rho^\beta(0) = \delta_\rho^\beta$ and hence for sufficiently small values of the s 's we can invert (1.8) and obtain

$$(1.10) \quad ds^\alpha = \lambda_\beta^\alpha(s) \delta s^\beta$$

and hence the vectors λ_β^α generate the velocity field of the parameter group. Corresponding to (1.9) we have for the parameter group

$$(1.11) \quad \frac{\partial \varphi^\alpha}{\partial s^\beta} = \lambda_\sigma^\alpha(s) \lambda_\beta^\sigma(s)$$

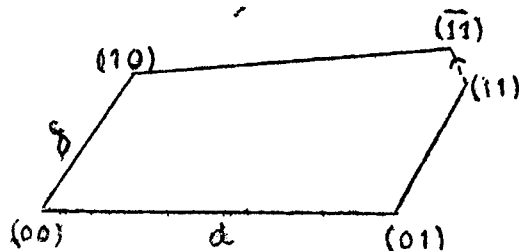
We summarize the above discussion by the following table:

Number of differentiations	Transformation group of group manifold	Transformation group of point field
0	$\bar{s} = ss_0$ $\bar{s}^\alpha = \varphi^\alpha(s^1 \dots s^r; s_0^1 \dots s_0^r)$ $(s_1 s) s_0 = s_1 (ss_0)$	$x = sx_0$ $x^i = \varphi^i(s^1 \dots s^r; x_0^1 \dots x_0^n)$ $(ss_0)_0 x = s(s_0)_0 x$
1	$s + ds = (1 + \delta s)s$ $\delta s^\alpha = \lambda_\beta^\alpha ds^\beta$ $ds^\beta = \lambda_\alpha^\beta(s) \delta s^\alpha$ $\delta s^\alpha = \lambda_\rho^\alpha(s) ds^\rho = \lambda_\rho^\alpha(\bar{s}) d\bar{s}^\rho$ $\bar{s} + d\bar{s} = (s+ds)s_0 = ((1+\delta s)s)s_0$ $= (1 + \delta s)\bar{s}$ $d\varphi^\beta = \lambda_\rho^\beta(s) \lambda_\alpha^\rho(s) ds^\alpha$	$dx^i = u_\rho^i(x) \delta s^\rho$ $= u_\rho^i(x) \lambda_\alpha^\rho(s) ds^\alpha$ $dx^i = u_\rho^i(\bar{x}) \lambda_\alpha^\rho(s) ds^\alpha$ $= u_\rho^i(\bar{x}) \lambda_\alpha^\rho(\bar{s}) d\bar{s}^\alpha$

2. Integrability

We have two fundamental problems: 1) to show that a continuous group (or rather a portion of it in the neighborhood of the identity) can be generated by iteration of its infinitesimal transformations; 2) to determine conditions under which r given linearly independent vectors $u_\alpha(x)$ ($\alpha = 1, \dots, r$) with their linear combinations, form the velocity field of a continuous transformation group.

Suppose that S and T are two transformations in the neighborhood of the identity whose corresponding infinitesimal transformations are $dx^i = u^i(x)\epsilon$ and $\delta x^i = v^i(x)\eta$ respectively. Under the influence of S , suppose that the point $(00) \rightarrow (01) = S(00)$ and $T(00) = (10)$, $S(10) = (\overline{11})$, $T(01) = (11)$. We obtain then $(\overline{11}) = STS^{-1}T^{-1}(11)$. The transformation $STS^{-1}T^{-1}$ belongs to the group. The infinitesimal transformation corresponding to it is $(d\delta - \delta d)x^i$. For we have:



should be by def. here

$$\delta(dx^i) = \{x^i(11) - x^i(01)\} - \{x^i(\overline{11}) - x^i(00)\}$$

$$d(\delta x^i) = \{x^i(\overline{11}) - x^i(10)\} - \{x^i(01) - x^i(00)\}$$

and hence

$$(d\delta - \delta d)x^i = x^i(\overline{11}) - x^i(11).$$

Thus, the velocity vector $[uv]^i = \frac{\partial v^i}{\partial x^k} u^k - \frac{\partial u^i}{\partial x^k} v^k$ belongs to our vector space.

Hence we have a necessary condition that r independent vectors u_α generate a continuous group is that

$$(2.1) \quad [u_\alpha, u_\beta] = c_{\alpha\beta}^\gamma u_\gamma$$

These conditions may be obtained also in the following way. The equations (1.9) are completely integrable since they have a solution which depends on n parameters. Hence the conditions of integrability, $\frac{d}{ds^\beta} \frac{dx^i}{ds^\alpha} = \frac{d}{ds^\alpha} \frac{dx^i}{ds^\beta}$ which

may be written as

$$(2.2) \quad \left(\frac{\partial \lambda_\alpha^\rho}{\partial s^\beta} - \frac{\partial \lambda_\beta^\rho}{\partial s^\alpha} \right) u_\rho^i(x) + \lambda_\alpha^\rho \lambda_\beta^\sigma [\mu_\rho, \mu_\sigma]^i = 0$$

are satisfied identically in x and s . For $s = 0$ we obtain (2.1) and hence from

(2.2) and (2.1) we have

$$(2.3) \quad \left(\frac{\partial \lambda_\alpha^\rho}{\partial s^\beta} - \frac{\partial \lambda_\beta^\rho}{\partial s^\alpha} \right) + c_{\sigma\tau}^\rho \lambda_\alpha^\sigma \lambda_\beta^\tau = 0$$

which are equivalent to

$$(2.4) \quad [\mu_\alpha, \mu_\beta] = c_{\alpha\beta}^\rho \mu_\rho.$$

(2.4) are the conditions of integrability of (1.11). For example, for the full linear group (group of all non-singular linear transformations of a vector space) the infinitesimal transformations are linear. If $dx = Ax$ and $\delta x = Bx$, then $(d\delta - \delta d)x = (BA - AB)x$, and hence the commutator product of the matrices belongs to the set.

The commutator product $[u, v]$ of the velocity vectors u and v is distributive with respect to each term and $\alpha[u, v] = [\alpha u, v] = [u, \alpha v]$. It is anti-commutative:

$$(2.5) \quad [uv] = -[vu]$$

and as is easily verified, it satisfies the Jacobi identity

$$(2.6) \quad [[uv]w] + [[vw]u] + [[wu]v] = 0$$

We may add to the table of sec. 1 the following:

<u>Number of differentiation</u>	<u>Point field</u>	<u>Group manifold</u>
2		$\left(\frac{\partial \lambda_\alpha^\rho}{\partial s^\beta} - \frac{\partial \lambda_\beta^\rho}{\partial s^\alpha} \right) + \lambda_\alpha^\sigma \lambda_\beta^\tau c_{\sigma\tau}^\rho = 0$ $[\mu_\alpha, \mu_\beta] = c_{\alpha\beta}^\rho \mu_\rho$
3		$[u_\alpha, u_\beta] = c_{\alpha\beta}^\rho u_\rho$ $c_{\alpha\beta}^\rho = -c_{\beta\alpha}^\rho$ $c_{\alpha\beta}^\rho c_{\rho\gamma}^\delta + c_{\beta\gamma}^\rho c_{\rho\alpha}^\delta + c_{\gamma\alpha}^\rho c_{\rho\beta}^\delta = 0$

Any manifold of velocity vectors which has the property (2.1) will be called an infinitesimal group. We can abstract from this concept of infinitesimal group the concept of an abstract infinitesimal group or a Lie algebra. We consider any vector space for which there is defined a composition for pairs of elements which satisfies:

$$[a, b + c] = [ab] + [ac]$$

$$\alpha[a, b] = [\alpha a, b] = [a, \alpha b]$$

$$[ab] = -[ba]$$

$$[a[bc]] + [b[ca]] + [c[ab]] = 0$$

A Lie algebra is non-associative but of a rather special type.

We may consider two fundamental problems for Lie algebras. First, we have the problem of realization in which we try to find infinitesimal groups whose velocity vectors can be put into linear correspondence with the elements of a given Lie algebra in such a way that to the product of two elements of the algebra there corresponds the commutator product of the velocity vectors. In particular we may ask if every Lie algebra has a distinct realization. Of special importance is the problem of the representation of a Lie algebra: to find linear sets of linear transformations of a vector space such that to every a of the algebra $\rightarrow \underline{A}$, the correspondence being linear and

$$[ab] \rightarrow \underline{AB} - \underline{BA} = [\underline{A}, \underline{B}].$$

The question of the representations of a Lie algebra is a purely algebraic one.

In the second place we have the algebraic problem of the classification of Lie algebras into non-isomorphic classes. It would be desirable to have a solution of this problem in which no use is made of the nature of the number field. As yet, there is very little known in this connection.

3. Preparations for the proof of Lie's theorems

In the next section we shall prove the fundamental theorems of Lie that (1) the conditions (2.1) are sufficient as well as necessary that a velocity field shall generate a continuous group (or at least a fragment of one); (2) every Lie algebra has a distinct realization. We consider now some preliminaries for the proofs

of these theorems.

One parameter groups. Let $u^i(x)$ be a velocity field and $dx^i = u^i(x)ds$ a corresponding infinitesimal transformation. We wish to show that this infinitesimal transformation generates a 1-parameter group. According to (1.9) we should consider an infinitesimal transformation of the form $dx^i = u^i(x)\lambda(s)ds$. However, we can normalize the parameter as we have done and then by composing two infinitesimal transformations say $dx^i = u^i(x)\epsilon$ and $dx^i = u^i(x)\eta$ we obtain $dx^i = u^i(x)(\epsilon + \eta)$ i.e. the parameters combine additively and hence if we obtain a group by the iteration of this infinitesimal transformation, it will be abelian. This iteration amounts to the integration of

$$(3.1) \quad \frac{dx^i}{ds} = u^i(x)$$

Let $x^i = \varphi^i(s, x_0)$ be a set of solutions of (3.1) which depend on the n independent parameters $x_0^1 \dots x_0^n$ and such that $x_0^i = \varphi^i(0, x_0)$. We must have

$$(3.2) \quad \varphi^i(s, \varphi(s_0, x_0)) = \varphi^i(s+s_0, x_0) \quad s_0 \text{ fixed}$$

For, both sides of this equation satisfy (3.1) and they are equal if $s = 0$ and hence by the uniqueness of the solution of the differential equation, the two sets of functions are identical. (3.2) states that the functions φ^i define a group which is abelian. It should be remarked that we have determined the group only for a neighborhood of the identity. In particular the fact that a 1-parameter group is a group of translations is a local property.

Theory of total differential equations

We consider a system of total differential equations of the form

$$(3.3) \quad \frac{dx^i}{ds^\alpha} = F_\alpha^i(s, x) \quad (i = 1, \dots, n; \alpha = 1, \dots, r)$$

and we seek a set of functions x^i , depending on n parameters x_0^i which satisfy (3.3) and such that for $s = 0$, $x = x_0$. A necessary condition that such a solution exists is that the integrability conditions

$$(3.4) \quad J_{\alpha\beta}^i \equiv [F_\alpha, F_\beta]^i \equiv \left(\frac{\partial F_\alpha^i}{\partial s^\beta} - \frac{\partial F_\beta^i}{\partial s^\alpha} \right) + \left(\frac{\partial F_\alpha^i}{\partial x^\kappa} F_\beta^\kappa - \frac{\partial F_\beta^i}{\partial x^\kappa} F_\alpha^\kappa \right) = 0$$

be satisfied identically in the x 's and the s 's. Conversely, if these integrability conditions are satisfied, there exists a solution of the desired type. We shall show how to integrate (3.3) by a step by step process of "freeing" the variables. First, we consider (3.3) for $\alpha = 1$ and obtain a solution $x^i(s^1, 0, \dots, 0)$ such that $x^i(0, \dots, 0) = x_0^i$. Next, let s_1 be fixed and determine a solution of (3.3) for $\alpha = 2$, say $x^i(s^1 s^2 0 \dots 0)$ such that $x^i(s^1 0 0 \dots 0)$ is the set of functions previously determined. We continue in this way and obtain finally a set of functions $x^i(s^1 \dots s^r)$ which satisfy

$$\frac{dx^i}{ds^\alpha} = F_\alpha^i(s; x)$$

for $(s^1 \dots s^\alpha, 0 \dots 0)$ and such that $x^i(0 \dots 0) = x_0^i$. We shall show that because of the integrability conditions these functions satisfy (3.3) without restriction. Consider

$$D_\alpha^i \equiv \frac{dx^i}{ds^\alpha} - F_\alpha^i(s; x).$$

$D_\alpha^i(s^1 \dots s^\alpha 0 \dots 0) = 0$ and we must show that $D_\alpha^i(s^1 \dots s^r) = 0$. On differentiating we obtain

$$(3.5) \quad \frac{d D_\alpha^i}{d s^\beta} - \frac{d D_\beta^i}{d s^\alpha} = J_{\alpha\beta}^i + \frac{\partial F_\alpha^i}{\partial x^k} D_\alpha^k - \frac{\partial F_\beta^i}{\partial x^k} D_\beta^k$$

By (3.4) $J_{\alpha\beta}^i = 0$. Suppose $\alpha < \beta$ and that $D_\alpha^i(s^1 \dots s^{\beta-1} 0 \dots 0) = 0$. We shall show that $D_\alpha^i(s^1 \dots s^\beta 0 \dots 0) = 0$ and continuing step by step $D_\alpha^i(s^1 \dots s^r) = 0$. Consider $s^1 \dots s^{\beta-1}$ as fixed, s^β as variable and $s^{\beta+1} \dots = 0$. Then (3.5) reduces to (since $D_\beta^k = 0$)

$$(3.6) \quad \frac{d D_\alpha^i}{d s^\beta} = \frac{\partial F_\alpha^i}{\partial x^k} D_\alpha^k$$

The only solution of (3.6) which reduces to 0 for $s^\beta = 0$ is 0. But D_α^i is such a solution and hence $D_\alpha^i(s^1 \dots s^\beta 0 \dots 0) = 0$, and the x 's are the required solution of (3.3).

In proving the existence theorem above, we made use of the integrability conditions $J_{\alpha\beta}^i = 0$ only for $\alpha < \beta$ and for $(s^1 \dots s^\beta 0 \dots 0)$. In other words

these restricted conditions imply the full conditions $J_{\alpha\beta}^i = 0$ for all α, β and $(s^1 \dots s^r)$. The reason for this is that $J_{\alpha\beta}^i$ satisfy the relations

$$(3.7) \quad \frac{dJ_{\alpha\beta}^i}{ds^\gamma} + \frac{dJ_{\beta\gamma}^i}{ds^\alpha} + \frac{dJ_{\gamma\alpha}^i}{ds^\beta} = \frac{\partial F_\alpha^i}{\partial x^\gamma} J_{\beta\gamma}^k + \frac{\partial F_\beta^i}{\partial x^\alpha} J_{\gamma\alpha}^k + \frac{\partial F_\gamma^i}{\partial x^\beta} J_{\alpha\beta}^k$$

Now suppose ^{for any} $\alpha < \beta < \gamma$ and $J_{\alpha\beta}^i (s^1 \dots s^{\gamma-1} 0 \dots 0) = 0$. Let $s^1 \dots s^{\gamma-1}$ be fixed s^i variable and $s^{\gamma+1} \dots = 0$, then (3.7) reduce to

$$\frac{dJ_{\alpha\beta}^i}{ds^\gamma} = \frac{\partial F_\gamma^i}{\partial x^\alpha} J_{\alpha\beta}^k$$

which has the single solution 0, reducing to 0 for $s^\gamma = 0$. Hence

$$J_{\alpha\beta}^i (s^1 \dots s^\gamma 0 \dots 0) = 0 \text{ and finally } J_{\alpha\beta}^i (s^1 \dots s^r) = 0.$$

4. Proof of Lie's main theorems.

Let $u_\alpha^i(x)$ be r independent vectors of an r -dimensional infinitesimal group. We associate the infinitesimal transformation $dx^i = u_\alpha^i(x) ds^\alpha$ with each of these vectors and integrate

$$\frac{dx^i}{ds^1} = u_1^i(x)$$

obtaining a solution $x^i(s^1 0 \dots 0)$ such that $x^i(0 \dots 0) = x_0^i$. Next, we integrate the second infinitesimal transformation $\frac{dx^i}{ds^2} = u_2^i(x)$ obtaining $x^i(s^1 s^2 0 \dots 0)$ where $x^i(s^1 0 \dots 0)$ is the solution of the first set of equations. Continuing, we obtain $x^i(s^1 \dots s^r)$ such that

$$(4.1) \quad \frac{dx^i}{ds^\alpha} (s^1 \dots s^\alpha 0 \dots 0) = u_\alpha^i(x).$$

We shall show that the functions $x^i(s^1 \dots s^r)$ define a group. First, we shall show that functions of the s 's, $\lambda_\alpha^p(s)$ can be determined such that $\lambda_\alpha^p(s^1 \dots s^\alpha 0 \dots 0) = \delta_\alpha^p$ and that the integrability conditions for

$$(4.2) \quad \frac{dx^i}{ds^\alpha} = \lambda_\alpha^p(s) u_p^i(x)$$

are satisfied identically. By (4.1) the functions $x^i(s^1 \dots s^r)$ satisfy (4.2) if we restrict $s^{\alpha+1} \dots s^r = 0$ and hence by the theory of the last section

$x^i(s^1 \dots s^r)$ will satisfy (4.2) if the integrability conditions are satisfied.

These latter conditions are that

$$(4.3) \quad J_{\alpha\beta}^{\gamma} \equiv \left(\frac{\partial \lambda_{\alpha}^{\gamma}}{\partial s^{\beta}} - \frac{\partial \lambda_{\beta}^{\gamma}}{\partial s^{\alpha}} \right) - c_{\rho\sigma}^{\gamma} \lambda_{\alpha}^{\rho} \lambda_{\beta}^{\sigma} = 0$$

since the $u_{\alpha}^i(x)$ are independent. As we saw in the last section (4.3) must hold merely for $\alpha < \beta$ and $(s^1 \dots s^{\beta} 0 \dots 0)$. But under these restrictions (4.3) reduce to

$$(4.4) \quad \frac{d\lambda_{\alpha}^{\gamma}}{ds^{\beta}} - c_{\rho\beta}^{\gamma} \lambda_{\alpha}^{\rho}(s) = 0$$

Suppose that $\lambda_{\alpha}^{\gamma}(s^1 \dots s^{\beta-1} 0 \dots 0)$ are known, then we consider $s^1 \dots s^{\beta-1}$ fixed and determine a solution of (4.4), $\lambda_{\alpha}^{\gamma}(s^1 \dots s^{\beta} 0 \dots 0)$ such that $\lambda_{\alpha}^{\gamma}(s^1 \dots s^{\beta-1} 0 \dots 0)$ are the functions previously determined. The functions obtained at every stage, and hence the $\lambda_{\alpha}^{\gamma}(s^1 \dots s^r)$, are elementary and hence analytic.

If the $u_{\alpha}^i(x)$ are the velocity vectors of a group, then the above argument shows that the parameters may be normalized in such a way that the functions $\lambda_{\alpha}^{\gamma}$ are analytic and hence the Jacobi conditions on the constants c can be derived without assuming the existence of third derivatives.

From (4.3) we can obtain

$$(4.5) \quad \left(\frac{\partial J_{\alpha\beta}^{\rho}}{\partial s^{\gamma}} + \frac{\partial J_{\beta\gamma}^{\rho}}{\partial s^{\alpha}} + \frac{\partial J_{\gamma\alpha}^{\rho}}{\partial s^{\beta}} \right) = c_{\tau\sigma}^{\rho} \lambda_{\alpha}^{\tau} J_{\beta\gamma}^{\sigma} + c_{\tau\sigma}^{\rho} \lambda_{\beta}^{\tau} J_{\gamma\alpha}^{\sigma} + c_{\tau\sigma}^{\rho} \lambda_{\gamma}^{\tau} J_{\alpha\beta}^{\sigma} - \lambda_{\alpha}^{\tau} \lambda_{\beta}^{\mu} \lambda_{\gamma}^{\nu} (c_{\tau\mu}^{\sigma} c_{\sigma\nu}^{\rho} + \dots)$$

By the Jacobi condition the parenthesis on the right-hand side is zero. Suppose $J_{\alpha\beta}^{\rho}(s^1 \dots s^{\gamma-1} 0 \dots 0) = 0$. $\alpha < \beta < \gamma$. Then (4.5) reduces to

$$(4.6) \quad \frac{dJ_{\alpha\beta}^{\rho}}{ds^{\gamma}} = c_{\tau\sigma}^{\rho} \lambda_{\gamma}^{\tau} J_{\alpha\beta}^{\sigma}$$

and this equation has the unique solution 0 subject to the initial condition that

$J_{\alpha\beta}^{\rho}(s^1 \dots s^{\gamma-1} 0 \dots 0) = 0$. Hence $J_{\alpha\beta}^{\rho}(s^1 \dots s^{\gamma} 0 \dots 0) = 0$, and finally,

$J_{\alpha\beta}^{\rho}(S^1 \dots S^r) = 0$. Thus the complete set of integrability conditions of (4.2) is a consequence of the restricted set.

The determination of the functions λ_{α}^{β} as solutions of (4.3) gives a method of obtaining a distinct realization of a Lie algebra. For in the above argument we have used only the Jacobi condition on the constants and shown that the functions λ_{α}^{β} which are solutions of (4.4) in the restricted sense, are solution of (4.3) without restriction, and $\lambda_{\alpha}^{\beta}(0) = \delta_{\alpha}^{\beta}$. In the neighborhood of 0 we can obtain the inverse (μ_{β}^{α}) of the matrix $(\lambda_{\beta}^{\alpha})$. The vectors μ_{β} give a distinct realization of the Lie algebra. They are linearly independent and

$$[\mu_{\alpha}, \mu_{\beta}] = c_{\alpha\beta}^{\rho} \mu_{\rho}$$

We return to the fundamental theorem that the velocity fields u_{α} such that

$$[u_{\alpha}, u_{\beta}] = c_{\alpha\beta}^{\rho} u_{\rho}$$

determine a continuous group. Consider the functions depending on n parameter

$x_0^1 \dots x_0^n$, say $x^i = \varphi^i(S, x_0)$, ($\varphi^i(0, x_0) = x_0^i$), which satisfy

$$\frac{dx^i}{ds^{\alpha}}(S^1 \dots S^{\alpha} 0 \dots 0; x_0) = u_{\alpha}^i(x)$$

We have seen that the function λ_{α}^{β} can be determined to satisfy (4.3) and then x^i will satisfy

$$\frac{dx^i}{ds^{\alpha}} = \lambda_{\alpha}^{\rho}(s) u_{\rho}^i(x)$$

If we make the same argument with the functions μ_{β}^{α} determined from λ_{α}^{β} we can obtain functions $\bar{S}^{\alpha} = \psi^{\alpha}(S; S_0)$ ($\psi^{\alpha}(0, S_0) = S_0^{\alpha}$) such that

$$(4.7) \quad \frac{d\psi^{\alpha}}{dS^{\alpha}} = \lambda_{\alpha}^{\rho}(S) \mu_{\rho}^{\alpha}(\psi)$$

Now the functions φ^i determine a group for which the composition law of the parameters is given by the functions ψ^{α} . For, if S_0 is taken as fixed, the functions $\varphi^i(S; \varphi(S_0; x_0))$ satisfy (4.2) and the functions $\varphi^i(\bar{S}; x_0)$ satisfy

$$dx^i = \lambda_{\alpha}^{\rho}(\bar{S}) u_{\rho}^i(x) d\bar{S}^{\alpha} = u_{\rho}^i(x) \delta \bar{S}^{\rho}$$

But by (4.7)

$$\delta \bar{S}^{\rho} = \lambda_{\alpha}^{\rho}(\bar{S}) d\bar{S}^{\alpha} = \lambda_{\alpha}^{\rho}(S) dS^{\alpha}$$

and hence $\varphi^i(\bar{S}; x_0)$ also satisfy (4.2). Since $\varphi^i(0; \varphi(S_0; x_0)) = \varphi^i(\varphi(0; S_0) x_0)$ we have $\varphi^i(S; \varphi(S_0; x_0)) \equiv \varphi^i(\bar{S}; x_0)$. If we make the same argument for the velocities μ_α we obtain for the parameter group

$$\varphi(S; \varphi(S_0; \circ S)) = \varphi(\bar{S}; \circ S)$$

or $S(S_0 \circ S) = (S S_0) \circ S$ and so the transformation group of the φ 's is associative and the proof is complete.

5. The adjoint group.

In addition to the realization of the continuous group $\mathcal{G}$ by means of the translations $S \rightarrow S' = aS$ (the parameter group), we can consider the group of conjugations $S \rightarrow S' = aS a^{-1}$ as a realization of $\mathcal{G}$. The advantage of this realization is that the elements near the identity are transformed into elements near the identity. Hence a conjugation transforms an infinitesimal element (transformation) of $\mathcal{G}$ into another infinitesimal element. The relation between these is linear and may be expressed by

$$\delta S^\alpha = \lambda_p^\alpha(a) ds^p$$

By the correspondence of $a \rightarrow (\lambda_p^\alpha(a))$ we obtain a representation of $\mathcal{G}$. The linear group of representing matrices is called the adjoint group of $\mathcal{G}$. We wish to determine the infinitesimal elements of this group. Suppose that a and S are near the identity, say $a = (1 + \delta b)$, $S = (1 + \epsilon x)$. Then

$$(1 + \epsilon x') = aS a^{-1} = (aS a^{-1} S^{-1})S = (1 + \epsilon \delta [bx])(1 + \epsilon x)$$

or,

$$dx = x' - x = \delta [bx]$$

is a general infinitesimal element of the adjoint group. It follows that the linear transformation $x \rightarrow dx = [bx] = \underline{B}x$ is a representation of the algebra of infinitesimal transformations. We can verify this also by direct computation. For, by the Jacobi identity

$$[[bc]x] = [b[cx]] - [c[bx]]$$

so that the transformation corresponding to $[bc]$ is $\underline{BC} - \underline{CB}$ if $\underline{B}$ and $\underline{C}$ correspond to

b and c respectively. This representation is analogous to the regular representation of an associative algebra that we considered in Chapter 1.

Chapter 3.

Foundations of a General Theory of Lie Algebras

1. Sub-groups, solvable, derivative; factor groups

In order to emphasize the analogy between the concepts in this theory and in the theory of groups, we adopt the terminology of the latter. We shall call a Lie algebra an infinitesimal group or simply a group. We recall its definition given in Chapter 2: a vector space for which there is defined a composition $[xy]$ for pairs of elements x, y which is linear in both terms and satisfies the conditions

$$(1.1) \quad [xy] = -[yx] \quad [x[yz]] + [y[zx]] + [z[xy]] = 0$$

The dimension of an infinitesimal group is called its order.

Let $\mathcal{G}$ be an infinitesimal group of order g and $e_1 e_2 \dots e_g$ a basis for $\mathcal{G}$. Suppose the multiplication of these basic elements is given by

$$[e_\alpha e_\beta] = c_{\alpha\beta}^\gamma e_\gamma \quad (\text{summed on } \gamma).$$

By (1.1) we have that the constants c satisfy:

$$(1.2) \quad c_{\alpha\beta}^\gamma = -c_{\beta\alpha}^\gamma \quad c_{\alpha\beta}^\rho c_{\rho\gamma}^\delta + c_{\beta\gamma}^\rho c_{\rho\alpha}^\delta + c_{\gamma\alpha}^\rho c_{\rho\beta}^\delta = 0$$

The constants c determine the infinitesimal group since they, together with the distributive law, give the product (or commutator) of any two elements of $\mathcal{G}$. If

$$f_\lambda = a_\lambda^\gamma e_\gamma \quad e_\gamma = b_\gamma^\nu f_\nu$$

define a change of coordinates and $[f_\lambda f_\mu] = c'_{\lambda\mu}^\nu f_\nu$, then the relation between the c' 's and the c 's is given by

$$c'_{\lambda\mu}^\nu = a_\lambda^\alpha a_\mu^\beta b_\gamma^\nu c_{\alpha\beta}^\gamma$$

In other words, the constants c are the components of a tensor which is covariant of order 2 and contravariant of order 1. The condition (1.2) may be expressed in the language of tensor analysis by saying that the tensor is anti-symmetric in its lower indices and that the tensor of order 4 $c_{\alpha\beta}^\rho c_{\rho\gamma}^\delta + \dots = 0$. Our problem is to

classify such tensors. However, it will be advantageous generally to think of the infinitesimal group rather than the tensor defined by it.

We proceed to introduce some definitions which are analogous to those of ordinary group theory.

A group is abelian if the commutator of any pair of elements in it is 0. A sub-group of $\mathcal{G}$ is a sub-space $\mathcal{f}$ which has the closure property $[xy] \in \mathcal{f}$ for every $x, y \in \mathcal{f}$. In order to define the analogue of an invariant sub-group $\mathcal{H}$ of an ordinary group G , we observe that a characterization of such a sub-group can be given by saying that if S is any element of $\mathcal{H}$ its commutator $aSa^{-1}S^{-1}$ with any element a of $\mathcal{G}$ is contained in $\mathcal{H}$. For, $aSa^{-1} = S'S^{-1}$ and hence $aSa^{-1}S^{-1} = S'S^{-1} \in \mathcal{H}$; and the converse follows by retracing the steps. We therefore formulate the definition of an invariant sub-group $\mathcal{f}$ of an infinitesimal group $\mathcal{G}$ as a sub-space $\mathcal{f}$ having the property that it contains the commutator $[ax]$ for any pair $a \in \mathcal{G}$ and $x \in \mathcal{f}$. It follows from this definition that an invariant sub-group is a group. Since $[ax] = -[xa]$, $\mathcal{f}$ is also right-invariant in the sense of associative algebras. We recall that the adjoint representation of $\mathcal{G}$ is the representation given by the linear transformations of $\mathcal{G}$

$$(1.3) \quad x \rightarrow dx = [ax].$$

Thus an invariant sub-group could have been defined as an invariant sub-space under the adjoint representation.

As in the case of ordinary groups, an infinitesimal invariant sub-group determines a factor group. If $\mathcal{f}$ is invariant in $\mathcal{G}$, the projection space $\mathcal{G}/\mathcal{f}$ is a group. For

$$x_1 \equiv x_2 \pmod{\mathcal{f}}, \quad y_1 \equiv y_2 \pmod{\mathcal{f}} \text{ implies } [x_1y_1] \equiv [x_2y_2] \pmod{\mathcal{f}}.$$

The centrum $\mathcal{L}$ of an infinitesimal group is the set of elements c such that $[ca] = 0$ for every a in $\mathcal{G}$. $\mathcal{L}$ is an invariant sub-group of $\mathcal{G}$.

In the case of ordinary groups, the sub-group G' generated by the commutators is invariant in G . For if a' is the commutator $sts^{-1}t^{-1}$ of s and t , $aa^{-1}a'^{-1}$ is the commutator of a and a' . G' is called the first derived group of

$\mathcal{G}$. Correspondingly, if $\mathcal{G}$ is an infinitesimal group, then the linear sub-space $\mathcal{G}'$ generated by the commutators $[xy]$ of the elements of $\mathcal{G}$ is an invariant sub-group. For, if $\alpha[xy] + \beta[zw] + \dots$ is any element of $\mathcal{G}'$, and a any element of $\mathcal{G}$, $\alpha[a[xy]] + \beta[a[zw]] + \dots \in \mathcal{G}'$. We may generalize this result by proving that the derived group $\mathcal{G}'$ of any invariant sub-group $\mathcal{F}$ of $\mathcal{G}$ is itself invariant in $\mathcal{G}$. It is necessary to show merely that if a is any element of $\mathcal{G}$, then $[a[xy]] \in \mathcal{F}'$ if $x, y \in \mathcal{F}$. But this follows from the Jacobi relation:

$$[a[xy]] = -[x[ya]] - [y[ax]] = -[xy'] - [yx']$$

where $y' = [ya]$, $x' = [ax] \in \mathcal{F}$ and hence $[a[xy]] \in \mathcal{F}'$. The analogous result holds also for ordinary groups, as can be shown without difficulty.

Now, suppose $\mathcal{G}''$ is the first derived group of $\mathcal{G}'$. By the result just proved, $\mathcal{G}''$ is invariant in $\mathcal{G}$. We call $\mathcal{G}''$ the second derived group of $\mathcal{G}$ and in general the first derived of the $(m-1)$ th is the m -th derived group of $\mathcal{G}$. It is invariant in $\mathcal{G}$. A group whose sequence of derived groups ends with 0.

$$\mathcal{G} > \mathcal{G}' > \mathcal{G}'' > \dots > \mathcal{G}^{(n)} > \mathcal{G}^{(n+1)} = 0$$

is said to be solvable (integrable). There are analogous definitions for ordinary groups. However, a solvable finite group is usually defined as one having a composition series whose factor groups are all cyclic groups of prime orders. The corresponding concept for infinitesimal groups would be that of a group which has a composition series whose factor groups have the simplest possible structure. As a matter of fact we can show that a solvable infinitesimal group can be defined as a group having a composition series of the type

$$(1.4) \quad \mathcal{G} = \mathcal{G}_g > \mathcal{G}_{g-1} > \mathcal{G}_{g-2} > \dots > \mathcal{G}_1 > 0 \quad \mathcal{G}_k \text{ of order } k.$$

The factor groups, being abelian groups of order 1, have then the simplest possible structure.

Suppose that $\mathcal{G}$ is solvable,

$$\mathcal{G} > \mathcal{G}' > \mathcal{G}'' > \dots > 0.$$

If $\mathcal{G}'$ is of order h , we can interpolate between it and $\mathcal{G}$ the sub-spaces

$$\mathcal{G}_{g-1}, \mathcal{G}_{g-2}, \dots \text{ such that } \mathcal{G} = \mathcal{G}_g > \mathcal{G}_{g-1} > \mathcal{G}_{g-2} > \dots > \mathcal{G}_h = \mathcal{G}'.$$
 But each of

these sub-spaces are invariant sub-groups of $\mathcal{G}$. Similarly we can interpolate invariant sub-groups of $\mathcal{G}^{(k)}$ between $\mathcal{G}^{(k)}$ and $\mathcal{G}^{(k+1)}$. The set of these sub-groups is a composition series of the type (1.4). On the other hand, if $\mathcal{G}$ has a composition series $\mathcal{G} > \mathcal{G}_{g-1} > \mathcal{G}_{g-2} > \dots > 0$, then $\mathcal{G}' < \mathcal{G}_{g-1}$. For, let α be an element of $\mathcal{G}$ which together with $\mathcal{G}_{g-1}$ generates (in the sense of vector spaces) $\mathcal{G}$. If $a = \alpha\alpha + \bar{a}$, $b = \beta\alpha + \bar{b}$, ($\bar{a}, \bar{b} \in \mathcal{G}_{g-1}$) are two arbitrary elements of $\mathcal{G}$,

$$[ab] = \alpha[\alpha\bar{b}] + \beta[\bar{a}\alpha] + [\bar{a}\bar{b}] \in \mathcal{G}_{g-1}$$

because of the invariance of this group. Thus $\mathcal{G}' < \mathcal{G}_{g-1}$ and similarly

$\mathcal{G}'' < \mathcal{G}_{g-1}' < \mathcal{G}_{g-2}$. Hence the sequence of derived groups lead to 0.

A simple infinitesimal group is one which has no invariant sub-group other than itself and 0. A semi-simple group is one which has no invariant solvable sub-group other than 0. If $\mathcal{F} \neq 0$ is a solvable invariant sub-group of $\mathcal{G}$, we have seen that the derived groups of $\mathcal{F}$ are also solvable and invariant in $\mathcal{G}$. Let $\mathcal{F}^{(k)} \neq 0$ and $\mathcal{F}^{(k+1)} = 0$, then $\mathcal{F}^{(k)}$ is an abelian invariant sub-group. It follows that we can define a semi-simple group as one which has no abelian invariant sub-group $\neq 0$. Every sub-space of an abelian group is an abelian invariant sub-group. Hence if $\mathcal{G}$ is simple but not semi-simple, it is an abelian group of order 1.

If $\mathcal{G}$ has a solvable invariant sub-group, it then contains a maximal solvable invariant sub-group $\mathcal{F}$, i.e. one which is contained in no other solvable invariant sub-group. It follows easily that $\mathcal{G}/\mathcal{F}$ is semi-simple. From this it follows that to a large extent the problems of infinitesimal groups can be resolved to these two extreme cases of solvable and semi-simple groups.

2. The characteristic polynomial

We suppose that we have a representation of an infinitesimal group $\mathcal{G}$ such that to the element α of $\mathcal{G}$ there corresponds the linear transformation (or matrix) $\underline{A}$ of an n -dimensional space. The characteristic polynomial of $\underline{A}$

$$f(\lambda; \underline{A}) = |\lambda \underline{E} - \underline{A}| = \lambda^n - \varphi_1(\alpha)\lambda^{n-1} + \varphi_2(\alpha)\lambda^{n-2} - \dots$$

is called the characteristic polynomial of α with respect to the representation.

Of particular importance are the first two coefficients

$$\psi_1(a) = \sum \rho_i$$

$$\psi_2(a) = \sum \rho_i \rho_i$$

where the $\sum$'s indicate symmetric functions and the functions

$$\sigma_1(a) = \sum \rho_i$$

$$\sigma_2(a) = \sum \rho_i^2$$

related to them by the identities

$$(2.1) \quad \sigma_1(a) = \psi_1(a)$$

$$\sigma_2(a) = \psi_1^2(a) - 2\psi_2(a)$$

The second of these identities can be established rationally in the following way:

$$\psi_2 = \sum_{i < k} \begin{vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{vmatrix} = \frac{1}{2} \sum_{i, k} \begin{vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{vmatrix} = \frac{1}{2} (\psi_1^2 - \sigma_2)$$

if we define $\sigma_2(a) = \text{tr}(\underline{A}^2)$. The $\text{tr}(\underline{AB}) = \sigma_2(a, b)$ is a symmetric bilinear form associated with the quadratic form $\sigma_2(a)$. It is called the scalar product of a and b relative to the given representation. Similarly, we denote the symmetric bilinear form associated with $\psi_2(a)$ by $\psi_2(a, b)$. Corresponding to (2.1) we have

$$(2.2) \quad \sigma_2(a, b) = \psi_1(a)\psi_1(b) - 2\psi_2(a, b)$$

If a is any element of $\mathcal{G}$, $\psi_1(a) = 0$. For, a is a linear combination of terms of the type $[xy]$ and $[xy] \rightarrow \underline{XY} - \underline{YX}$ which has 0 trace.

If we are dealing with the adjoint representation we call the characteristic equation of a in it the characteristic equation proper or simply the characteristic equation of a .

Let $\mathcal{f}$ denote a sub-group of $\mathcal{G}$ and $c_1 \dots c_g$ be a basis of $\mathcal{G}$ for which $c_1 \dots c_h$ is a basis of $\mathcal{f}$. If a is any element of $\mathcal{f}$, then

$$(2.3) \quad [ac_1] = \alpha_{11}c_1 + \alpha_{21}c_2 + \dots + \alpha_{h1}c_h$$

$$[ac_h] = \alpha_{1h}c_1 + \alpha_{2h}c_2 + \dots + \alpha_{hh}c_h$$

and so the matrix A representing a in the adjoint representation has the form

$$(2.4) \quad A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1h} & \alpha_{1h+1} & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{h1} & \dots & \alpha_{hh} & \alpha_{hh+1} & \dots & \dots \\ \hline 0 & \dots & 0 & \alpha_{h,h+h} & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \alpha_{g,h+1} & \dots & \dots \end{pmatrix}$$

and $f_g(\lambda | a)$ is divisible by $f_h(\lambda | a)$, the characteristic polynomial of a acting in the group $\mathfrak{f}$ (or the adjoint representation of $\mathfrak{f}$). If $\mathfrak{f}$ is an invariant sub-group we have in addition to (2.3),

$$(2.5) \quad \begin{aligned} [a e_{h+1}] &= \alpha_{1h+1} e_1 + \dots + \alpha_{hh+1} e_h \\ &\vdots \\ [a e_g] &= \alpha_{1g} e_1 + \dots + \alpha_{hg} e_h \end{aligned}$$

or

$$A = \left(\begin{array}{ccc|ccc} \alpha_{11} & \alpha_{12} & \dots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{h1} & \alpha_{h2} & \dots & \vdots & \vdots & \vdots \\ \hline & & & 0 & & 0 \end{array} \right)$$

and hence $f_g(\lambda | a) = \lambda^{g-h} f_h(\lambda | a)$.

If $\mathfrak{f}$ is invariant and a is any element of $\mathfrak{g}$ then (2.3) holds and hence A has the form (2.4). The matrices

$$A_1 = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1h} \\ \vdots & \ddots & \vdots \\ \alpha_{h1} & \dots & \alpha_{hh} \end{pmatrix}$$

give a representation of $\mathfrak{g}$ which is a component of the adjoint representation. If $f_h^*(\lambda | a)$ is the characteristic polynomial of a in this representation, then $f_h^*(\lambda | a)$ is a factor of $f_g(\lambda | a)$.

Suppose that $\xi^1 \dots \xi^{2g}$ are variables, then the element $x = \xi^i e_i$ is a variable element of $\mathfrak{g}$. In a given representation of $\mathfrak{g}$ the matrix X corresponding to x has coordinates which are linear functions of the ξ 's and hence its characteristic polynomial $f(\lambda | x)$ is integral in the ξ 's. As we saw in Chapter 1, $f(\lambda | x)$ is invariant under a coordinate transformation in the representation space. It is also independent of the coordinate system, in the group $\mathfrak{g}$. Finally, it is invariant in the sense that

$$df = \frac{\partial f}{\partial \xi^i} [a \xi]^i dt = 0$$

or,

$$f(\lambda + x + dx) \equiv f(\lambda | x) \pmod{t^2}$$

if $dx = [ax]t$ and a is any element of $\mathcal{O}$. Let $X + t[AX] = X_t$, then

$$(E + tA)X \equiv X_t(E + tA) \pmod{t^2}$$

and hence

$$(E + tA)(\lambda E - X) \equiv (\lambda E - X_t)(E + tA) \pmod{t^2}$$

$$|E + tA| |\lambda E - X| \equiv |\lambda E - X_t| |E + tA| \pmod{t^2}$$

But $|E + tA| \not\equiv 0 \pmod{t}$, so that

$$|\lambda E - X| \equiv |\lambda E - X_t| \pmod{t^2}$$

or,

$$f(\lambda | x + dx) \equiv f(\lambda | x) \pmod{t^2}$$

Suppose $f(\lambda | x) = f_1^{h_1} f_2^{h_2} f_3^{h_3} \dots$ where the f_i are distinct polynomials in the ξ 's and λ which are irreducible and have 1 as coefficients of their highest degree terms in λ . We have

$$df = h_1 f_1^{h_1-1} df_1 f_2^{h_2} f_3^{h_3} \dots + f_1^{h_1} h_2 f_2^{h_2-1} df_2 f_3^{h_3} \dots + \dots = 0.$$

It follows that df_1 is divisible by f_1 , but being of lower degree in λ than f_1 , it must be 0. Hence if g is any factor of f , g is an invariant.

3. Definition of invariant sub-groups by means of invariants

Let $\psi(x) = a_{i_1 k \dots i_l} \xi^{i_1} \xi^k \dots \xi^l$ be an invariant form in the sense of the preceding section:

$$i, k, \dots, l = 0, 1$$

$$\frac{\partial \psi}{\partial \xi^i} [ax]^i = 0$$

We can associate with $\psi(x)$ a symmetric multilinear form $\psi(x, y \dots z) =$

$= a_{i_1 k \dots i_l} \xi^{i_1} \eta^k \dots \zeta^l$ and we shall show that it is an invariant also:

$$\Delta_a \psi = \psi([ax], y \dots z) + \psi(x, [ay], \dots) + \dots = 0$$

For, the form

$$\psi([ax], x \dots x) + \psi(x, [ax] \dots) + \dots \equiv 0$$

and $\Delta_a \psi$ is the corresponding multilinear form and hence is 0.

Now let $\mathcal{f}$ be an invariant sub-group of $\mathcal{O}$ and consider the set $\mathcal{f}_1$ of elements a of $\mathcal{a}$ which satisfy

$$\psi(a, y, \dots, z) = 0$$

where $y \dots z$ are arbitrary elements of $\mathcal{f}$. Evidently $\mathcal{f}_1$ is a linear space. If u is any element of $\mathcal{G}$,

$$\psi([ua], y \dots z) + \psi(a, [uy] \dots z) + \dots = 0$$

and since $[uy] \dots [uz] \in \mathcal{f}$

$$\psi(a, [uy] \dots z) = \dots = \psi(a, y \dots, [uz]) = 0$$

so that

$$\psi([ua], y \dots z) = 0 \quad \text{or} \quad [ua] \in \mathcal{f}_1.$$

Hence $\mathcal{f}_1$ is an invariant sub-group of $\mathcal{G}$.

The special case $\mathcal{f} = \mathcal{G}$ in the above theorem has numerous applications.

For example, if $f(\lambda | x) = \lambda^n - \psi_1(x)\lambda^{n-1} + \psi_2(x)\lambda^{n-2} - \dots$, then $\psi_1(x)$, $\psi_2(x)$ are invariants. The elements a such that $\psi_1(a) = 0$ form an invariant sub-group which contains $\mathcal{G}'$. The elements a such that $\psi_2(ax) = 0$ for all $x \in \mathcal{G}$ form another invariant sub-group. Of course the same remark applies to $\sigma_2(x)$.

We shall show that a necessary and sufficient condition that a group be semi-simple is that the proper scalar product $\sigma_2(x, y)$ shall be a non-degenerate bilinear form. The sufficiency of this condition can be proved very easily. If $\mathcal{G}$ is not semi-simple, it contains an abelian invariant sub-group $\mathcal{f}$. Let $e_1 \dots e_g$ be a basis for $\mathcal{G}$ such that $e_1 \dots e_h$ is a basis for $\mathcal{f}$. A variable element of $\mathcal{G}$, $a = \alpha^i e_i = \alpha^1 e_1 + \dots + \alpha^h e_h + \bar{a}$ and

$$[ae_i] = [\bar{a}e_i] = \alpha_{1i} e_1 + \dots + \alpha_{hi} e_h \quad (i = 1 \dots h)$$

where the α_{ij} ($i, j = 1 \dots h$) are linear functions of the α 's which are independent of the coordinates $\alpha_1 \dots \alpha_h$.

$$[ae_{h+1}] = \alpha_{1, h+1} e_1 + \dots + \alpha_{h+1, g} e_g$$

$$[ae_g] = \alpha_{1g} e_1 + \dots + \alpha_{g, g} e_g$$

and the α_{ij} , ($i, j = h+1, \dots, g$), which are the same as the corresponding functions for $\bar{a}$, are independent of $\alpha_1 \dots \alpha_h$. Thus the matrix which represents a in the adjoint representation has the form

$$A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$$

where the sub-matrices A_1 and A_2 do not involve the coordinates $\alpha^1 \dots \alpha^h$. Hence the characteristic equation of a and in particular $\sigma_2(a)$ is independent of these coordinates. Thus $\sigma_2(a)$ is degenerate. The converse of this theorem will be proved later.

Let $g(x)$ be an invariant non-degenerate quadratic form of $x = \xi^i e_i$, say $g(x) = g_{ik} \xi^i \xi^k$. It is natural to consider the form $g^{ik} e_i e_k$ where the matrix (g^{ik}) is the inverse of (g_{ik}) . However, in $\mathcal{G}$ itself the products $e_i e_k$ are not defined. If we are dealing with a representation of $\mathcal{G}$, then the corresponding form $G = g^{ik} E_i E_k$ where $e_i \rightarrow E_i$ has a meaning. We shall call G the Casimir operator corresponding to $g(x)$ and show that the condition that $g(x)$ is an invariant implies that G is commutative with all the matrices of the representation. For,

$$(3.1) \quad g([e_\alpha x], x) = 0 \quad (\alpha = 1, \dots, g).$$

If $[e_\alpha e_i] = c_{\alpha i}^r e_r$, $[e_\alpha x] = \xi^i c_{\alpha i}^r e_r$ and (3.1) implies that $g_{rk} \xi^i c_{\alpha i}^r \xi^k = 0$ identically in the ξ 's. Hence, if $c_{\alpha i, k}^r$ denotes $g_{rk} c_{\alpha i}^r$, we must have

$$(3.2) \quad c_{\alpha i, k}^r = -c_{\alpha k, i}^r.$$

On the other hand $[E_\alpha, G] = 0$ means

$$(3.3) \quad g^{ik} [E_\alpha, E_i E_k] = g^{rk} [E_\alpha E_r] E_k + g^{ir} E_i [E_\alpha E_r] \\ = (g^{rk} c_{\alpha r}^i + g^{ir} c_{\alpha r}^k) E_i E_k = 0.$$

If $c_{\alpha}^{k, i}$ denotes $g^{rk} c_{\alpha r}^i$, conditions (3.3) are satisfied if $c_{\alpha}^{k, i} = -c_{\alpha}^{i, k}$ and this follows from (3.2).

The infinitesimal groups for the n -dimensional unimodular linear group and for the n -dimensional orthogonal groups can be determined and from them the inner product forms $\sigma_2(x, y)$. It can be shown that these are not degenerate and hence that these groups are semi-simple. As a matter of fact they are all simple, except the orthogonal group of 4 dimensions. Another important class of simple groups are

the complex groups: set of linear transformations which leave a non-degenerate anti-symmetric form invariant.

4. Lie's theorem on solvable groups of linear transformations

Let $\mathcal{G}$ be an infinitesimal group of linear transformations of an n -dimensional space $\mathcal{R}$: a linear set of transformations which contains the commutator of any pair. The theorem of Lie states that if $\mathcal{G}$ is solvable, there exists a vector $e \neq 0$ in $\mathcal{R}$ such that $\underline{A}e = \alpha e$ for every $\underline{A}$ in $\mathcal{G}$.

If $\mathcal{G}$ is solvable, $\mathcal{G} = \mathcal{G}_g > \mathcal{G}_{g-1} > \mathcal{G}_{g-2} > \dots > 0$ where $\mathcal{G}_k$ is of order k and invariant in $\mathcal{G}_{k+1}$. We shall prove the Lie theorem by induction on k , and we may assume just as well that the theorem holds for $\mathcal{F}_g = \mathcal{G}_{g-1}$ and then prove it for $\mathcal{G}$. Suppose then that $e_0 \neq 0$ is a vector of $\mathcal{R}$ such that $\underline{B}e_0 = \beta e_0$ for every $\underline{B}$ in $\mathcal{F}_g$. Let $\underline{A}$ be any transformation of $\mathcal{G}$ not in $\mathcal{F}_g$. The linear space $(\underline{A}, \mathcal{F}_g) = \mathcal{G}$. Let

$$\underline{A}e_0 = e_1, \quad \underline{A}e_1 = \underline{A}^2 e_0 = e_2, \quad \underline{A}e_2 = \underline{A}^3 e_0 = e_3 \dots$$

and suppose that e_χ is the first vector obtained in this way which is dependent on the preceding vectors. We have $\varphi(\underline{A})e_0 = 0$ where $\varphi(\underline{A})$ is a polynomial of degree χ the minimum for which such a relation holds. Let $\underline{B}'$ denote $[\underline{B}\underline{A}]$, $\underline{B}'' = [\underline{B}'\underline{A}]$, ... Then

$$(4.1) \quad \underline{B}e_0 = \beta e_0$$

$$\underline{B}e_1 = \underline{B}\underline{A}e_0 = \underline{B}'e_0 + \underline{A}\underline{B}e_0 = \beta'e_0 + \beta e_1$$

$$\underline{B}e_2 = \underline{B}\underline{A}e_1 = \underline{B}'e_1 + \underline{A}\underline{B}e_1 = \beta''e_0 + 2\beta'e_1 + \beta e_2$$

Consider the χ -dimensional space $\mathcal{Y} = (e_0, e_1, \dots, e_{\chi-1})$. It is invariant with respect to $\mathcal{G}$. By (4.1), $\text{tr } \underline{B} = \chi\beta$ in $\mathcal{Y}$. But $\text{tr } \underline{B}' = \text{tr } (\underline{B}\underline{A} - \underline{A}\underline{B}) = 0$. Hence

$\beta' = 0, \beta'' = 0, \dots$ and (4.1) reduces to

$$(4.2) \quad \underline{B}e_0 = \beta e_0, \underline{B}e_1 = \beta e_1, \underline{B}e_2 = \beta e_2, \dots$$

By (4.2) we have for any vector e of $\mathcal{Y}$, $\underline{B}e = \beta e$. If $\varphi(\lambda) = (\lambda - \alpha)\varphi_1(\lambda)$ and $e = \varphi_1(\underline{A})e_0$, then $e \neq 0$ since $\varphi_1(\lambda)$ is of degree $\chi - 1$. Evidently $\underline{A}e = \alpha e$, and hence e has the required properties.

If $\underline{A} \in \mathfrak{g}'$ the derived group of $\mathfrak{g}$, $\underline{A}e = 0$.

Denote e by e_1 . We have $\underline{A}e_1 = \alpha_1 e_1$ for every $\underline{A}$ in $\mathfrak{g}$. *G solvable ind. of the vector space*
 space $\mathfrak{R}/e_1$ is transformed into itself linearly by $\mathfrak{g}$ and hence we can apply Lie's *The projection on which these*
 theorem to it and determine a vector $e_2 \neq 0 \pmod{e_1}$ such that *(lin.) trans. act.*

$$\underline{A}e_2 = \alpha_2 e_2 \pmod{e_1}.$$

If we consider $\mathfrak{R}/(e_1, e_2)$ we have that there exists an $e_3 \neq 0 \pmod{(e_1, e_2)}$ such that

$$\underline{A}e_3 = \alpha_3 e_3 \pmod{e_1 e_2}$$

and so on. Thus the matrices A of the transformations $\underline{A}$ can be brought simultaneously to the form

$$(4.3) \quad A = \begin{pmatrix} \alpha_1 & \alpha_{12} & \cdots & \alpha_{1n} \\ & \alpha_2 & & \vdots \\ & 0 & \ddots & \vdots \\ & & & \alpha_n \end{pmatrix}$$

If $\underline{A} \in \mathfrak{g}'$, $\alpha_i = 0$. The characteristic polynomial of a variable element $\xi^i \underline{A}_i$ of $\mathfrak{g}$ factors into linear factors $(\lambda - \alpha_1)(\lambda - \alpha_2) \cdots (\lambda - \alpha_n)$ where the α_i are linear functions of the ξ 's. The characteristic polynomial of a variable transformation of $\mathfrak{g}'$ is λ^n , and hence for any element of $\mathfrak{g}'$ we have $\underline{A}^n = 0$.

The theorem of Lie and its consequences for $\mathfrak{g}$ require that the number field be sufficiently extended, say algebraically closed. However, the results for $\mathfrak{g}'$ are rational and hence they hold for any number field. It is obvious that the characteristic equation $f(\lambda | \underline{A})$ of $\underline{A}$ is independent of the field of $\mathfrak{g}$ so that if $\mathfrak{g}$ is solvable and $\underline{A} \in \mathfrak{g}'$, $f(\lambda | \underline{A}) = \lambda^n$. The condition $\underline{A}e = 0$ which can be satisfied in an extended field is merely that $|\underline{A}| = 0$, and hence e can be chosen with components in the original field. It follows that the elements of $\mathfrak{g}'$ can be transformed rationally into the form (4.3) with $\alpha_i = 0$.

5. Applications of Lie's theorems to abstract infinitesimal groups

The adjoint representation

$$x \rightarrow dx = [ax]$$

of a solvable group $\mathfrak{g}$ is a solvable group $\mathcal{G}$ of linear transformations acting on the g -dimensional space $\mathfrak{g}$. We can apply Lie's theorem and determine a basis

$(e_1 \dots e_g)$ of $\mathfrak{g}$ such that $[ae_i]$ is a linear combination of $e_1 \dots e_i$. It follows that $[e_i e_k]$ is a linear combination of $(e_1 \dots e_s)$ where $s = \min(i, k)$ or,

$c_{ik}^r = 0$ if $r > i$ or $> k$ for this basis.

The characteristic equation proper $f(\lambda | a) = \prod (\lambda - \rho_i)$ where the ρ_i are linear forms in the variables α^i where $a = \alpha^i e_i$. For $a \in \mathfrak{g}'$, $f(\lambda | a) = \lambda^g$.

Let g' be the order of $\mathfrak{g}'$. We have seen in sec. 2 that

$f(\lambda | a) = \lambda^{g-g'} f_{g'}(\lambda | a)$ where $f_{g'}$ is the characteristic equation of $a \in \mathfrak{g}'$ in the adjoint representation of $\mathfrak{g}'$. Hence $f_{g'}(\lambda | a) = \lambda^{g'}$. A group for which

the characteristic equation proper of every element is a power of λ will be called

nilpotent. Thus we have that the derived group of a solvable group is nilpotent.

Nilpotent groups, as we shall see, are closely related to nilpotent associative algebras.

The converse of the rational form of Lie's theorem is the

Theorem of Engel: If $\mathfrak{g}$ is nilpotent, then it is solvable.

We require some preliminary remarks for the proof of this theorem:

Lemma 1: Every sub-group of a nilpotent group is nilpotent.

For, the characteristic equation proper $f_h(\lambda | a)$ of an element a of the sub-group $\mathfrak{f}$ is a factor of $f(\lambda | a)$.

Lemma 2: If $\mathfrak{g} > \mathfrak{f}$ a solvable sub-group of order h , then there exists a sub-group $\mathfrak{f}^+$ of order $h + 1$, such that $\mathfrak{g} > \mathfrak{f}^+ > \mathfrak{f}$

For any $b \in \mathfrak{f}$, $dx = [bx]$ is a linear transformation of $\mathfrak{g}$. $\mathfrak{f}$ is invariant under this transformation, and hence so is $\mathfrak{g}/\mathfrak{f}$. Since the set of transformations, i.e. $\mathfrak{f}$, is a solvable group, by Lie's theorem, there exists an element $e \neq 0$

(mod f) such that $[bo] = \beta o$ (mod f) for all $b \in f$. $f^+ = (f, e)$ is the required group.

Lemma 3: Every sub-group $\mathcal{G}_{g-1}$ of a nilpotent group $\mathcal{G}_g$ is invariant.

Let e be any element $\neq 0$ (mod $\mathcal{G}_{g-1}$), then $\mathcal{G}_g = (\mathcal{G}_{g-1}, e)$. Suppose $[be] = \beta e$ (mod $\mathcal{G}_{g-1}$). Then $\lambda - \beta$ is a factor of the characteristic polynomial of b . Hence $\beta = 0$ and $\mathcal{G}_{g-1}$ is invariant. since $\mathcal{G}_{g-1}$ is a subgroup.

Proof of Engel's theorem: Suppose $f(\lambda | a) = \lambda^g$. We shall construct a composition series for $\mathcal{G}$ for which the factor groups are of order 1. We take any 1-dimensional sub-group $\mathcal{G}_1$ and find a $\mathcal{G}_2$ of order 2 such that $\mathcal{G} > \mathcal{G}_2 > \mathcal{G}_1 > 0$. By Lemma 1 $\mathcal{G}_2$ is nilpotent. By Lemma 3 $\mathcal{G}_1$ is invariant in $\mathcal{G}_2$ and hence $\mathcal{G}_2$ is solvable. By Lemma 2 there exists a $\mathcal{G}_3$ such that $\mathcal{G} > \mathcal{G}_3 > \mathcal{G}_2 > \mathcal{G}_1 > 0$. We repeat the argument with $\mathcal{G}_3$. Continuing in this way we obtain a composition series of the desired type.

The proof just given for Engel's theorem is based on the stronger form of Lie's theorem and hence holds only for algebraically closed fields. The theorem can be extended to arbitrary fields by the following argument. Let $\mathcal{G}$ be nilpotent, then it remains nilpotent if we extend the field to be algebraically closed. The extended group $\mathcal{G}_c$ is solvable. If we use the definition of solvability that $\mathcal{G}_c > \mathcal{G}'_c > \mathcal{G}''_c > \dots > 0$, it follows easily that $\mathcal{G} > \mathcal{G}' > \mathcal{G}'' > \dots > 0$ or, $\mathcal{G}$ is solvable.

The two theorems on solvable groups can be summarized in the following criterion: A necessary and sufficient condition that a group be solvable is that its first derived group be nilpotent.

Cartan sharpened this criterion by showing that a sufficient condition that a group be solvable is that $\mathcal{D}_2(x) = 0$ for all x of the derived group.

6. Irreducible transformation groups and semi-simplicity

Let $\mathcal{G}$ be an irreducible infinitesimal group of linear transformations of an n -dimensional space $\mathcal{R}$. Suppose that the number field is algebraically closed, then we shall show that a solvable invariant sub-group of $\mathcal{G}$ is a 1-dimensional group of dilations. For example, consider the set of all linear transformations. These form an irreducible group which contains the sub-group of dilations: transformations whose matrices have the form

$$(6.1) \quad \begin{pmatrix} \gamma & & 0 \\ & \ddots & \\ 0 & & \gamma \end{pmatrix}$$

In the general case, suppose that $\mathcal{L}$ is a solvable invariant sub-group and $\underline{C}$ is any element of $\mathcal{L}$. By Lie's theorem there exists a vector $c \neq 0$ such that $\underline{C}c = \gamma c$. If $\underline{C}' = [\underline{C}\underline{A}]$ where $\underline{A}$ is any transformation of $\mathcal{G}$, it can be shown as in the proof of Lie's theorem that $\gamma' = 0$. It follows that $\underline{A}_1 c, \underline{A}_2 \underline{A}_1 c, \dots$, where $\underline{A}_1, \underline{A}_2 \dots$ are arbitrary transformations of $\mathcal{G}$, are invariant under $\mathcal{L}$: $\underline{C}\underline{A}_r \underline{A}_{r-1} \dots \underline{A}_1 c = \gamma \underline{A}_r \underline{A}_{r-1} \dots \underline{A}_1 c$. The set of vectors obtained in this way generate a vector space $\neq 0$ which is invariant under $\mathcal{G}$. Since $\mathcal{G}$ is irreducible, this space $= \mathcal{R}$. Hence for any vector $c \in \mathcal{R}$ we have $\underline{C}c = \gamma c$ or $\underline{C}$ has the form (6.1). $\mathcal{L}$ is the centrum of $\mathcal{G}$.

If we have a linear set $\mathcal{G}$ of linear transformations which is not closed under multiplication, we can extend it by adding the products of pairs of transformations and their linear combinations. If the new linear set is not closed we repeat the process on it. The dimensionality of each new linear set is greater than that of the preceding, and since it can be at most n^2 this process comes to a stop after a finite number of steps and we obtain an algebra $\mathcal{A}$ which contains $\mathcal{G}$. We call the enveloping algebra of $\mathcal{G}$. If $\mathcal{G}$ is irreducible so is $\mathcal{A}$, and conversely.

Now let $\mathcal{G}$ be an irreducible infinitesimal group of linear transformations defined in a field $\mathcal{F}$ (not necessarily algebraically closed). We form the envelop-

ing algebra $\mathcal{A}$ of $\mathcal{G}$. Schur has shown that there exists an algebraic field $\mathcal{F}(\theta)$ of degree m over $\mathcal{F}$ such that all the matrices of $\mathcal{A}$ can be transformed simultaneously by a matrix in $\mathcal{F}(\theta_1 \theta_2 \dots \theta_m)$ (θ_i the conjugates of $\theta = \theta_1$) into the form

$$\begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_m \end{pmatrix}$$

where the algebra $\mathcal{A}_i$ of the matrices A_i is absolutely irreducible (irreducible in an algebraically closed extension of $\mathcal{F}$) and $A_1, A_2, \dots, A_m$ are conjugates. If $\mathcal{L}$ is a solvable invariant sub-group of $\mathcal{G}$ then the corresponding $\mathcal{L}_i$ is such a sub-group of $\mathcal{G}_i$. By Cartan's result, ^{just given} the elements of $\mathcal{L}_i$ must have the form $\gamma_i E_r$ where E_r is the identity matrix of r rows and columns. Hence the elements of $\mathcal{L}$ have the form

$$\begin{pmatrix} \gamma_1 E_r & & \\ & \gamma_2 E_r & \\ & & \ddots \\ & & & \gamma_m E_r \end{pmatrix}$$

Every element $\neq 0$ is non-singular and $\mathcal{L}$ is the centrum. The enveloping algebra of $\mathcal{L}$ is isomorphic to a sub-field of $\mathcal{F}(\theta)$.

An associative irreducible algebra of linear transformations is simple.

We will show next that this is not the case for infinitesimal groups.

Direct sum of two groups. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two infinitesimal groups.

We define the direct sum $\mathcal{G} = \mathcal{G}_1 \oplus \mathcal{G}_2$ as the group whose elements are the pairs (a_1, a_2) $a_i \in \mathcal{G}_i$ for which scalar multiplication, addition and multiplication are defined by

$$\alpha(a_1, a_2) = (\alpha a_1, \alpha a_2)$$

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$$

$$[(a_1, a_2)(b_1, b_2)] = ([a_1 b_1], [a_2 b_2]).$$

The sub-group of $\mathcal{G}$ consisting of the set of elements of the form $(a_1, 0)$ is isomorphic with $\mathcal{G}_1$ and the sub-group of elements of the form $(0, a_2)$ is isomorphic with $\mathcal{G}_2$. We use $\mathcal{G}_1$ and $\mathcal{G}_2$ again to denote these sub-groups. We have that every element of $\mathcal{G}$ can be written in a unique way in the form $b_1 + b_2$ $b_1 \in \mathcal{G}_1$ and if b_1 and b_2 are arbitrary elements of these sub-groups, $[b_1 b_2] = [b_2 b_1] = 0$.

If $\mathcal{G}_1$ and $\mathcal{G}_2$ are semi-simple, so is $\mathcal{G}$. For, let $\mathcal{f}$ be a solvable invariant sub-group of $\mathcal{G}$. It is easily seen that $\mathcal{f} = \mathcal{f}_1 + \mathcal{f}_2$, $\mathcal{f}' = \mathcal{f}_1' + \mathcal{f}_2' \dots$ where $\mathcal{f}_i \in \mathcal{G}_i$. Hence $\mathcal{f}_i$ is a solvable invariant sub-group of $\mathcal{G}_i$. Since these groups are semi-simple, we must have $\mathcal{f}_i = \mathcal{f} = 0$.

A representation of a simple group $\mathcal{G}$ is either distinct or it sends every element into 0. For, the elements which correspond to the 0 transformation form an invariant sub-group $\mathcal{f}$ of $\mathcal{G}$. $\mathcal{f}$ is either 0 or $\mathcal{G}$. In the first case the representation is distinct and in the second it is 0.

The adjoint representation of a simple group (not of order 1) is a distinct representation. It is also irreducible as follows from the definition of a simple group.

Consider $\mathcal{G}_1$ and $\mathcal{G}_2$ two simple groups and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two distinct representations of $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively. Denote the order of the space $\mathcal{R}_i$ on which $\mathcal{G}_i$ acts by n_i ($i = 1, 2$). Finally let $\mathcal{L}_i$ denote the continuous transformation group generated by the infinitesimal transformations $dx_i = \Lambda_i x_i$ of $\mathcal{G}_i$. It can be shown that the group generated by $\mathcal{G} = \mathcal{G}_1 \oplus \mathcal{G}_2$ is isomorphic with the direct product $\mathcal{L}_1 \times \mathcal{L}_2$. But the infinitesimal group of this direct product is obtained from

$$d(x_1^\alpha x_2^\beta) = x_1^\alpha dx_2^\beta + dx_1^\alpha x_2^\beta$$

and hence its transformations have the form $\underline{E}_1 \times \underline{A}_2 + \underline{A}_1 \times \underline{E}_2$ where $\underline{E}_i$ is the identity of n_i rows or columns. It follows that these transformations give a distinct realization of $\mathcal{G}$. This can also be verified directly. The enveloping algebra $\mathcal{U}$ of $\mathcal{G}$ contains all $\underline{A}_1 \times \underline{A}_2$ and more generally all $\underline{A}_1^{(1)} \underline{A}_1^{(2)} \dots \times \underline{A}_2^{(1)} \underline{A}_2^{(2)} \dots$ and their

linear combinations and hence is the direct product of the enveloping algebras $\mathcal{O}_i$ of $\mathcal{G}_i$. If the field is algebraically closed, we saw in Chapter 1 that this direct product is irreducible whenever the components are. Thus $\mathcal{G}$ is irreducible but not simple.

7. The enveloping algebra. Extension of Engel's theorem

To study the relation between a group $\mathcal{G}$ and its enveloping algebra $\mathcal{U}$ it is necessary to compute the characteristic polynomial or, what is just as good for associative algebras, the trace of an arbitrary element of $\mathcal{U}$ from the characteristic polynomials of the elements of $\mathcal{G}$. We shall show how this can be done. Since the trace is linear, it is necessary to determine only $\text{tr} (\underline{A}_1 \underline{A}_2 \dots \underline{A}_r)$ where the $\underline{A}$'s are arbitrary transformations of the group. $\sum \underline{A}_{i_1} \underline{A}_{i_2} \dots \underline{A}_{i_r}$ the summation being taken over all permutations of $1 \dots r$, is a symmetric multilinear form associated with $\text{tr} \underline{A}^r$. $\text{tr} \underline{A}^r$ can be expressed rationally in terms of $\psi_1(\underline{A}), \psi_2(\underline{A}) \dots$ where $f(\lambda | \underline{A}) = \lambda^n - \psi_1(\underline{A}) \lambda^{n-1} + \psi_2(\underline{A}) \lambda^{n-2} \dots$ is the characteristic equation of $\underline{A}$. Now, $\text{tr} (\underline{A}_1 \underline{A}_2 \dots \underline{A}_{k-1} \underline{A}_{k+1} \underline{A}_{k+2} \dots \underline{A}_r \underline{A}_1 \underline{A}_2 \dots \underline{A}_r) = \text{tr} \underline{A}_1 \dots \underline{A}_{k-1} [\underline{A}_{k+1} \underline{A}_k] \underline{A}_{k+2} \dots \underline{A}_r$. Since $[\underline{A}_{k+1} \underline{A}_k]$ belongs to $\mathcal{G}$, $\underline{A}_1 \dots \underline{A}_{k-1} [\underline{A}_{k+1} \underline{A}_k] \dots \underline{A}_r$ can be expressed linearly in terms of products of $r-1$ terms. Using induction we can suppose that the trace of this element can be determined. Since any permutation is a product of such transpositions, we have $\text{tr} \underline{A}_1 \underline{A}_2 \dots \underline{A}_r = \frac{1}{r!} \text{tr} \sum \underline{A}_{i_1} \dots \underline{A}_{i_r} + \mu$ where μ can be determined. Hence $\text{tr} \underline{A}_1 \dots \underline{A}_r$ can be computed.

Suppose every transformation of $\mathcal{G}$ has λ^n as its characteristic function,
then the trace of every element of $\mathcal{U}$ is 0. For in the above argument we may assume that $\mu = 0$ since it is the trace of a sum of products of $r-1$ terms. Hence $\text{tr} \underline{A}_1 \underline{A}_2 \dots \underline{A}_r = \frac{1}{r!} \text{tr} \sum \underline{A}_{i_1} \underline{A}_{i_2} \dots \underline{A}_{i_r}$. The right hand side is the multilinear form associated with $\text{tr} \underline{A}^r = 0$, and hence $\text{tr} \underline{A}_1 \underline{A}_2 \dots \underline{A}_r = 0$. This remark is the basis of a rational proof of Engel's theorem that we will give below.

Definition 1: An associative algebra $\mathcal{A}$ is nilpotent if there exists an integer N such that $\mathcal{A}^N = 0$, i.e. $a_1 a_2 \dots a_N = 0$ for arbitrary $a_i \in \mathcal{A}$.

To formulate an analogous definition for infinitesimal groups we note that if $\mathcal{F}$ is an invariant sub-group of $\mathcal{G}$ then so is $[\mathcal{G}\mathcal{F}]$ = linear space generated by the elements of the form $[xy]$ where $x \in \mathcal{G}$, $y \in \mathcal{F}$. For, if $a \in \mathcal{G}$

$$[a[xy]] = -[x[ya]] - [y[ax]]$$

Thus the derived group $\mathcal{G}' = [\mathcal{G}\mathcal{G}] = \mathcal{G}^2$ and in general $\mathcal{G}^k = [\mathcal{G}\mathcal{G}^{k-1}]$ is invariant in $\mathcal{G}$. Evidently $\mathcal{G}^2 = \mathcal{G}'$, $\mathcal{G}^3 \supset \mathcal{G}''$, $\mathcal{G}^4 \supset \mathcal{G}''' \dots$

Definition 2: An abstract infinitesimal group $\mathcal{G}$ is nilpotent if there exists an N such that $\mathcal{G}^N = 0$. (We shall see that this definition is equivalent to the one given in sec. 5 for a nilpotent group.)

If $\mathcal{G}$ is nilpotent it is solvable. $\mathcal{G}$ is nilpotent if and only if the enveloping algebra of its adjoint representation is nilpotent. For if $a \rightarrow \underline{A}$, the transformation $x \rightarrow [a_1[a_2 \dots [a_N x] \dots]] = \underline{A}_1 \underline{A}_2 \dots \underline{A}_N x$. Hence $\mathcal{A}^N = 0$ implies $\mathcal{G}^{N+1} = 0$, and conversely

Theorem (Generalization of Engel's theorem). If $\mathcal{G}$ is a λ -distinct representation of $\mathcal{G}$ and the transformations of $\mathcal{G}$ all have λ^n as characteristic functions, then $\mathcal{G}$ is nilpotent.

It should be observed that this is a two-fold generalization of Engel's theorem. In the first place it deals with any representation rather than the regular representation, and secondly the conclusion that $\mathcal{G}$ is nilpotent is stronger than that $\mathcal{G}$ is solvable.

Proof: Let $\mathcal{A}$ be the enveloping algebra of $\mathcal{G}$. We have seen that the trace of every element of $\mathcal{A}$ is 0. $\mathcal{A}$ is nilpotent. For if not, according to a theorem of Wedderburn's, $\mathcal{A}$ contains an idempotent element $\underline{U}$. But from $\underline{U}^2 = \underline{U}$ it follows that the matrix of $\underline{U}$ can be taken in the form

$$U = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

and hence $\text{tr } U = h \neq 0$. Hence $\mathcal{U}$ is nilpotent. Say $\mathcal{U}^N = 0$. Since

$[a_1[a_2[\dots a_N]]\dots] \rightarrow [\underline{a}_1[\underline{a}_2[\dots \underline{a}_N]]\dots] = 0$, and $\mathcal{U}$ is a distinct realization of $\mathcal{g}$, $[a_1[a_2 \dots a_N]\dots] = 0$ or $\mathcal{g}^N = 0$.

The proofs in this section are strictly rational. It would be desirable to obtain such proofs for Lie's theorem and Cartan's theorems. However, this seems to be considerably more difficult.

8. Roots of a single linear transformation. Maximal nilpotent sub-groups

For the proof of Cartan's criterion for solvable groups we shall require a part of the theory of a single linear transformation.

Let $\underline{A}$ be such a transformation acting on an n -dimensional space $\mathcal{R}$ and $f(\lambda | \underline{A}) = f(\lambda) = (\lambda - \alpha_1)^{h_1} (\lambda - \alpha_2)^{h_2} \dots (\lambda - \alpha_r)^{h_r}$ its characteristic polynomial. A 1-dimensional sub-space invariant under $\underline{A}$ is determined by a vector $e \neq 0$ which satisfies $\underline{A}e = \alpha e$. From the theory of homogeneous equations it follows that $|\underline{A} - \alpha \underline{E}| = 0$, or α is a root say $= \alpha_1$ of $f(\lambda)$. Conversely, corresponding to α_1 , there exists such an invariant vector. The projection space $\mathcal{R}/e_1$ is transformed linearly by $\underline{A}$ and the characteristic function of $\underline{A}$ in this space is $\frac{f(\lambda)}{\lambda - \alpha_1}$.

If $h_1 > 1$, we may use α_1 again to determine a vector $e_2 \neq 0 \pmod{e_1}$ and $\underline{A}e_2 = \alpha_1 e_2 \pmod{e_1}$. We then consider the invariant space (e_1, e_2) and its projection space. Continuing, we obtain finally an invariant space $\mathcal{R}_{\alpha_1}$ of dimensionality h_1 such that the characteristic function of $\underline{A}$ in this sub-space is $(\lambda - \alpha_1)^{h_1}$ and such that the matrix of $\underline{A}$ in this space can be taken in the form

$$\begin{pmatrix} \alpha_1 & \alpha_{12} & \dots & \alpha_{1n} \\ & \alpha_2 & \dots & \alpha_{2n} \\ & & \ddots & \\ 0 & & & \alpha_n \end{pmatrix}$$

Similarly we can determine $\mathcal{R}_{\alpha_k} \equiv \mathcal{R}_k$ of dimensionality h_k ($k = 1, 2, \dots, r$).

We wish to show that the spaces $\mathcal{R}_k$ are independent and that $\mathcal{R} = \sum \mathcal{R}_k$.

For this purpose we consider the partial fraction expansion of $\frac{1}{f(\lambda)}$:

$$\frac{1}{f(\lambda)} = \frac{g_1(\lambda)}{(\lambda - \alpha_1)^{h_1}} + \frac{g_2(\lambda)}{(\lambda - \alpha_2)^{h_2}} + \dots$$

where $g_k(\lambda)$ has degree $< h_k$ and is prime to $(\lambda - \alpha_k)$. By multiplying by $f(\lambda)$ we obtain

$$(8.1) \quad 1 = f_1(\lambda) + f_2(\lambda) + \dots$$

where $f_k(\lambda)$ is divisible by $(\lambda - \alpha_1)^{h_1} \dots (\lambda - \alpha_{k-1})^{h_{k-1}} (\lambda - \alpha_{k+1})^{h_{k+1}} \dots (\lambda - \alpha_r)^{h_r}$. For any vector $x_k \in \mathcal{R}_k$ we have $(A - \alpha_k)^{h_k} x_k = 0$, and hence if E_e denotes $f_e(\underline{A})$,

$$(8.2) \quad E_e x_k = 0 \quad (e \neq k)$$

From (8.1) we have

$$(8.3) \quad E = E_1 + E_2 + \dots + E_r$$

and

$$(8.4) \quad x_k = E x_k = E_1 x_k + \dots + E_r x_k = E_k x_k.$$

If the spaces $\mathcal{R}_k$ are not independent there exists a relation of the type

$$x_1 + x_2 + \dots + x_r = 0 \quad x_i \in \mathcal{R}_i$$

where not all the x_i are 0. If we transform this relation by E_k , we obtain

$E_k x_k = x_k = 0$, which contradicts the assumption. Hence the $\mathcal{R}_k$ are independent and since the sum of their dimensions is $\sum h_k = n$, we have $\mathcal{R} = \sum \mathcal{R}_k$.

$f(\underline{A})x_k = 0$ and hence $f(\underline{A})x = 0$. We have then the theorem that a transformation satisfies its characteristic equation. The proof just given is irrational, but the following is a simple rational proof: We denote the unit vectors by

$e_1 e_2 \dots e_n$ and the matrix of $\underline{\Lambda}$ in this coordinate system by $\Lambda = (\alpha_{ij})$. Then

$$\underline{\Lambda} e_1 = \alpha_{11} e_1 + \alpha_{21} e_2 + \dots + \alpha_{n1} e_n$$

$$\underline{\Lambda} e_n = \alpha_{1n} e_1 + \alpha_{2n} e_2 + \dots + \alpha_{nn} e_n$$

or,

$$(\underline{\Lambda} - \alpha_{11} \underline{E}) e_1 - \alpha_{21} e_2 - \dots - \alpha_{n1} e_n = 0$$

$$-\alpha_{1n} e_1 - \alpha_{2n} e_2 - \dots - (\underline{\Lambda} - \alpha_{nn} \underline{E}) e_n = 0$$

The determinant of these homogeneous equations in the e_k is $f(\underline{\Lambda})$. By multiplying by the co-factors of the k th column we obtain as in the ordinary theory $f(\underline{\Lambda}) e_k = 0$ for $k = 1, 2, \dots, n$. Hence $f(\underline{\Lambda}) = 0$.

It is possible to carry through, by the methods used above, a part of the theory of a single transformation in an arbitrary field. Suppose that $f(\lambda) = \prod \varphi_i^{k_i}(\lambda)$ where the $\varphi_i(\lambda)$ are distinct irreducible polynomials. As before, we can obtain

$$1 = f_1(\lambda) + f_2(\lambda) + \dots$$

where $f_k(\lambda)$ is of degree $< n$ and is divisible by $\varphi_1^{k_1} \dots \varphi_{\chi-1}^{k_{\chi-1}} \varphi_{\chi+1}^{k_{\chi+1}} \dots \varphi_n^{k_n}$ but not by φ_χ . Hence

$$\underline{E} = \underline{E}_1 + \underline{E}_2 + \dots \quad \underline{E}_k \underline{E}_\ell = \delta_{k\ell} \underline{E}_\ell$$

$$x = \underline{E}_1 x + \underline{E}_2 x + \dots = x_1 + x_2 + \dots$$

where $\underline{E}_k = f_k(\underline{\Lambda})$. The x_ℓ for a fixed ℓ determine a vector space $\mathcal{R}_\ell$ and $\mathcal{R} = \sum \mathcal{R}_\ell$. The $\mathcal{R}_\ell$ are independent and invariant. The two additional facts that we had before, namely that the characteristic function of $\underline{\Lambda}$ in $\mathcal{R}_\ell$ is $\varphi_\ell^{k_\ell}$ and the dimensionality of $\mathcal{R}_\ell$ is $k_\ell n_\ell$ (n_ℓ the degree of φ_ℓ) are not so immediate.

We return to the algebraically closed case. We note that a vector $x \in \mathcal{R}_\ell$ if and only if

$$(\underline{\Lambda} - \alpha_\ell)^N x = 0 \quad \text{all } k \quad (\text{for some } N).$$

For, suppose that $x = x_1 + x_2 + \dots$, then $(\underline{\Lambda} - \alpha_\ell)^N x_k = 0$ since $\mathcal{R}_k$ is invariant

under $\underline{\Delta}$ and the $\mathcal{R}$'s are independent. But $f_{\chi}(\underline{\Delta})x_k = 0$ if $k \neq \chi$ and $f_{\chi}(\lambda)$ is prime to $(\lambda - \alpha_{\chi})^N$ so that there exist two polynomials $M(\lambda)$, $N(\lambda)$ such that $M(\lambda)(\lambda - \alpha_{\chi})^N + N(\lambda)f_{\chi}(\lambda) = 1$. It follows that $x_k = 0$ and $x = x_{\chi}$.

Application to infinitesimal groups. We consider the adjoint representation of an infinitesimal group $\mathfrak{g}$ and its characteristic function $f(\lambda | a) = \lambda^g - \psi_1(a)\lambda^{g-1} + \psi_2(a)\lambda^{g-2} - \dots \pm \psi_{g-\ell}(a)\lambda^{\ell}$ where $\psi_{g-\ell}(a)$ is the last coefficient not identically 0. $\ell > 0$. For, if a is taken as one of the base elements, $[aa] = 0$ and the matrix of the transformation $x \rightarrow [ax]$ has a column of zeros so that its determinant = 0. (If we assume that the number field is infinite, it follows that $\psi_n(a)$ vanishes algebraically. For the present we assume in fact that the field is algebraically closed.)

Let C_0 be any element of $\mathfrak{g}$ such that $\psi_{g-\ell}(C_0) \neq 0$. C_0 will be called a regular element of $\mathfrak{g}$ and ℓ the efficiency of $\mathfrak{g}$. Consider the transformation

$$x \rightarrow [C_0 x] = \underline{C}_0 x.$$

The above discussion shows that if $f(\lambda | C_0) = \lambda^g (\lambda - \alpha)^{h_{\alpha}} \dots$, $\mathfrak{g} = \sum \mathfrak{g}_{\alpha}$, $\downarrow$ corr. to R and R_i

$\mathfrak{g}_{\alpha}$ has dimensionality h_{α} and is invariant under $c_0(\underline{C}_0)$ and the characteristic polynomial of $c_0(\underline{C}_0)$ in $\mathfrak{g}_{\alpha}$ is $(\lambda - \alpha)^{h_{\alpha}}$. We require the fundamental

Lemma: If $x_{\alpha} \in \mathfrak{g}_{\alpha}$ and $x_{\beta} \in \mathfrak{g}_{\beta}$ then $[x_{\alpha} x_{\beta}] \in \mathfrak{g}_{\gamma}$, $\gamma = \alpha + \beta$

We have to show merely that $(\underline{C}_0 - \gamma)^N [x_{\alpha} x_{\beta}] = 0$. The Jacobi relation and anti-symmetry give

$$[c_0[x_{\alpha} x_{\beta}]] = [x_{\alpha} [c_0 x_{\beta}]] + [[c_0 x_{\alpha}] x_{\beta}]$$

or

$$\begin{aligned} \underline{C}_0 [x_{\alpha}, x_{\beta}] &= [x_{\alpha}, \underline{C}_0 x_{\beta}] + [\underline{C}_0 x_{\alpha}, x_{\beta}] \\ (\underline{C}_0 - \gamma) [x_{\alpha}, x_{\beta}] &= [x_{\alpha}, (\underline{C}_0 - \beta) x_{\beta}] + [(\underline{C}_0 - \alpha) x_{\alpha}, x_{\beta}] \end{aligned}$$

By iteration we obtain

$$(\underline{C}_0 - \gamma)^N [x_{\alpha} x_{\beta}] = \sum_r \binom{N}{r} [(\underline{C}_0 - \alpha)^r x_{\alpha}, (\underline{C}_0 - \beta)^{N-r} x_{\beta}]$$

so that if $N = h_{\alpha} + h_{\beta} - 1$, $(\underline{C}_0 - \gamma)^N [x_{\alpha} x_{\beta}] = 0$.

If $\alpha + \beta$ is not a root it follows that $[x_{\alpha} x_{\beta}] = 0$.

$\mathcal{G}_0$ is a solvable group. If c_1 and $c_2 \in \mathcal{G}_0$, then $[c_1 c_2] \in \mathcal{G}_0$ by the lemma. Then $\mathcal{G}_0$ is a group. Let $c = \gamma^1 c_1 + \gamma^2 c_2 + \dots + \gamma^l c_l$ denote any element of $\mathcal{G}_0$. $[c x_\alpha] \in \mathcal{G}_\alpha$, i.e. $\mathcal{G}_\alpha$ is invariant under the transformation $x \rightarrow [cx]$. Hence $f(\lambda | c) = \prod_\alpha f_\alpha(\lambda | c) = \lambda^l g(\lambda | c)$ where $f_\alpha(\lambda | c)$ is the characteristic polynomial of c acting on $\mathcal{G}_\alpha$. If $\alpha \neq 0$ the last coefficient of $f_\alpha(\lambda | c)$ is not identically 0, and hence there exists a c for which all these last coefficients $\neq 0$. Since $f(\lambda | c)$ is divisible by λ^l it follows that $f_0(\lambda | c) = \lambda^l$. By Engel's theorem $\mathcal{G}_0$ is solvable. $\mathcal{G}_\alpha$ is the representation space of a solvable group $\mathcal{G}_0$, so that according to Lie's theorem the characteristic function of c in $\mathcal{G}_\alpha$

$$f_\alpha(\lambda | c) = (\lambda - \rho_1)(\lambda - \rho_2) \dots$$

where $\rho = \beta_1 \gamma^1 + \beta_2 \gamma^2 + \dots + \beta_l \gamma^l$ (β 's constants)

In the same way we obtain linear forms $\sigma_1, \sigma_2, \dots$ in $\mathcal{G}_\alpha'$, and so on. In particular, for $\mathcal{G}_0$ we have l linear forms $= 0$. Now the forms obtained from two different spaces, say ρ_i, σ_j of $\mathcal{G}_\alpha$ and $\mathcal{G}_\alpha'$ respectively, are not identically equal. It may happen that some of the linear forms belonging to the same space (say, ρ_1, ρ_2 in $\mathcal{G}_\alpha$) are not identically equal. At any rate we may choose an element $c^0 \in \mathcal{G}_0$ such that all the forms which are not identically equal have different values for c^0 . If we repeat the construction with c^0 in place of c_0 and obtain the decomposition of $\mathcal{G} = \sum \mathcal{G}_\alpha$, then the forms ρ in the same space $\mathcal{G}_\alpha$ are identically equal. We change our notation and denote the characteristic polynomial of $c = \gamma^1 c_1 + \dots + \gamma^l c_l$ in $\mathcal{G}_\alpha$ by $(\lambda - \alpha)^{h_\alpha}$, $\alpha = \alpha_1 \gamma^1 + \alpha_2 \gamma^2 + \dots + \alpha_l \gamma^l$ and call the α the roots of $\mathcal{G}_0$.

$\mathcal{G}_0$ is the same as determined for c_0 or for c^0 (in fact for any two regular elements of $\mathcal{G}_0$). The relation between the decomposition of $\mathcal{G}$ relative to c_0 and c^0 will be clearer if we obtain the final decomposition in the following way: We return to the decomposition for c_0 $\mathcal{G} = \sum \mathcal{G}_\alpha$. In $\mathcal{G}_\alpha$ suppose $\rho_1 \neq \rho_2$ and let c_0' be a regular element for which $\rho_1 \neq \rho_2$. $\mathcal{G}_\alpha$ may be decomposed into at least

two different invariant sub-spaces corresponding to its different roots. But these sub-spaces according to the lemma are also invariant with respect to all the elements of $\mathfrak{g}_0$. We thus obtain a finer decomposition of $\mathfrak{g}$. Continuing in this way we obtain finally the final decomposition obtained before by means of c^0 .

In order to bring out the full significance of this result for $\mathfrak{g}_0$ we compare the situation in abelian transformation groups and solvable groups.

1) Abelian set of transformations. We have seen that the matrix of a single linear transformation could be assumed to be in the form:

$$(8.5) \quad A = \begin{pmatrix} \alpha \cdots & 0 & 0 \\ & \beta \cdots & \\ 0 & & \beta \\ 0 & 0 & \ddots \end{pmatrix}$$

By the argument of Chapter 1, sec. 5, it can be shown that if $\underline{A}$, $\underline{B}$, ... are the elements of an abelian set, their matrices can be brought simultaneously to the form (8.5).

2) Solvable group of transformations. According to Lie's theorem the matrices of a solvable group $\mathfrak{g}$ can be taken simultaneously in the form

$$A = \begin{pmatrix} \alpha \cdots & & \\ & \alpha \cdots & \\ & & \beta \cdots \\ 0 & & & \beta \cdots \\ & & & & \ddots \end{pmatrix}$$

If this result can be improved to obtain (8.5) we shall call $\mathfrak{g}$ semi-abelian.

Thus the transformation group determined by the transformations of the solvable group $\mathfrak{g}$, acting on $\mathfrak{g}$ is semi-abelian.

9. Cartan's criterion for semi-simple groups and its applications

We shall prove first Cartan's criterion for solvable groups: if

$$\sigma(a) = \text{tr}(\Lambda^2) = 0, \mathfrak{g} \text{ is solvable.}$$

Consider the decomposition $\mathfrak{g} = \sum \mathfrak{g}_\alpha$ where $\mathfrak{g}_0$ is a maximal solvable sub-group and the characteristic polynomial of $c \in \mathfrak{g}_0$ acting on $\mathfrak{g}_\alpha$ is $(\lambda - \alpha)^{h_\alpha}$, α = linear form in the coordinates of c . Let e_α and $e_{-\alpha}$ be variable elements of $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ respectively, then $[e_\alpha e_{-\alpha}] = c'$ varies in $\mathfrak{g}_0$. Any root ρ taken for c' is a rational multiple of the root α taken for c' . For, suppose that the linear forms $\rho - i\alpha, \rho - (i-1)\alpha, \dots, \rho, \rho + \alpha, \dots, \rho + k\alpha$ are roots but $\rho - (i+1)\alpha$ and $\rho + (k+1)\alpha$ are not. Let $\sum_{j=-i}^k \mathfrak{g}_{\rho+j\alpha} = \mathfrak{g}$. $\mathfrak{g}$ is invariant under the transformations $x \rightarrow [e_\alpha x] = \underline{E}_\alpha x$ and $x \rightarrow [e_{-\alpha} x] = \underline{E}_{-\alpha} x$ and of course under $x \rightarrow [c'x] = [\underline{E}_\alpha \underline{E}_{-\alpha}]x = \underline{C}'x$. Thus $\text{tr } \underline{C}' = 0$ in $\mathfrak{g}$. On the other hand, $\text{tr } c' = \text{tr } \underline{C}'$ in $\mathfrak{g}_\alpha$ is $h_\alpha \alpha'$ (α' the value of α for c') and hence $\text{tr } \underline{C}'$ in $\mathfrak{g} = \sum h_{\rho+j\alpha} (\rho' + j\alpha')$. Thus

$$\sum h_{\rho+j\alpha} (\rho' + j\alpha') = 0$$

and ρ' is a rational multiple of α'

$\sigma(c') = \sum \rho'^2 = 0$ implies that every $\rho' = 0$ since the ρ' lie on a straight line in the complex plane. Now any element in $\mathfrak{g}' \cap \mathfrak{g}_0$ is a linear combination of elements of the form $c' = [e_\alpha e_{-\alpha}]$. Since the roots are linear forms it follows that they are 0 for all elements of $\mathfrak{g}' \cap \mathfrak{g}_0$.

Suppose $\mathfrak{g}' = \mathfrak{g}$, the above argument shows that all the roots of $\mathfrak{g}_0$ are 0 and hence $\mathfrak{g}_0 = \mathfrak{g}$. But $\mathfrak{g}_0$ is solvable so that $\mathfrak{g}' \neq \mathfrak{g}$. Hence $\mathfrak{g}'$ has smaller order than $\mathfrak{g}$.

To complete the proof of the solvability of $\mathfrak{g}$ we repeat the argument with $\mathfrak{g}'$. We may do this since $f_g(\lambda | x) = \lambda^{g-g'} f_{g'}(\lambda | x)$ if $x \in \mathfrak{g}'$ so that $\sigma_2'(x) = \sigma_2(x) = 0$.

Theorem: $\mathfrak{g}$ is semi-simple if and only if σ_2 is non-degenerate.

We have shown the sufficiency of this condition in sec. 3. It remains to show that if $\mathfrak{g}$ is semi-simple, $\sigma = \sigma_2$ is non-degenerate.

Consider the elements a for which $\sigma(ax) = 0$ for all x in $\mathfrak{g}$. By sec. 3 these elements form an invariant sub-group $\mathfrak{f}$ of $\mathfrak{g}$. Since $f_{\mathfrak{g}}(\lambda | a) = \lambda^{g-h} f_{\mathfrak{h}}(\lambda | a)$, $\sigma = 0$ for $\mathfrak{f}$ and $\mathfrak{f}$ is solvable. Since $\mathfrak{g}$ is semi-simple $\mathfrak{f} = 0$, i.e. there exists no $a \neq 0$ such that $\sigma(ax) = 0$ for all x . But this is the definition of non-degeneracy of a quadratic form. Q.E.D.

The two criteria of Cartan hold for groups defined over any (infinite) number field. The extension to this case involves only the same type of reasoning that we have had before.

Applications

1) Independence of semi-simplicity of number field: If $\mathfrak{g}$ is semi-simple in a field K then the group $\mathfrak{g}_1$ obtained from $\mathfrak{g}$ by extending the field to $K_1 \supset K$ is semi-simple. For if $\sigma(xy)$ is the scalar product of $x, y \in \mathfrak{g}$, $\sigma(x_1, y_1)$ is the scalar product of $x_1, y_1 \in \mathfrak{g}_1$. Hence $\sigma(x_1, y_1)$ is non-degenerate and $\mathfrak{g}_1$ is semi-simple.

2) Decomposition of semi-simple groups: A semi-simple group is decomposable in a unique way into a direct sum of simple groups (of orders > 1). Let $\mathfrak{f}$ be an invariant sub-group of $\mathfrak{g}$ and consider the invariant sub-group $\mathfrak{f}^*$ of elements y orthogonal to every element x of $\mathfrak{f}$, i.e. $\sigma(x, y) = 0$. If h, h^*, g denote the orders of $\mathfrak{f}, \mathfrak{f}^*$ and $\mathfrak{g}$ respectively, then $g = h + h^*$. The elements common to $\mathfrak{f}$ and $\mathfrak{f}^*$ form an invariant sub-group of $\mathfrak{g}$ whose scalar product form is identically 0. Thus $\mathfrak{f} \cap \mathfrak{f}^*$ is a solvable invariant sub-group so that $\mathfrak{f} \cap \mathfrak{f}^* = 0$ and $\mathfrak{g} = \mathfrak{f} + \mathfrak{f}^*$. If $x \in \mathfrak{f}, y \in \mathfrak{f}^*, [xy] \in \mathfrak{f} \cap \mathfrak{f}^*$, and hence $[xy] = 0$, $\mathfrak{g} = \mathfrak{f} \oplus \mathfrak{f}^*$. If $\mathfrak{f}$ is a maximal invariant sub-group, $\mathfrak{f}^* = \mathfrak{g}_1 \cong \mathfrak{g}/\mathfrak{f}$ is simple (of order > 1). $\mathfrak{f}$ is semi-simple and of lower order than $\mathfrak{g}$, and hence we may suppose that $\mathfrak{f} = \mathfrak{g}_2 \oplus \mathfrak{g}_3 \oplus \dots, \mathfrak{g}_i$ simple. Then $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_3 \oplus \dots$. To show the uniqueness of this decomposition we shall show that every invariant sub-

group $\mathfrak{g}$ of $\mathfrak{g}$ is a direct sum of certain $\mathfrak{g}_i$. Let x be an arbitrary element of $\mathfrak{g}$, then $x = x_1 + x_2 + \dots, x_i \in \mathfrak{g}_i$. The set of elements x_i (i fixed) form an invariant sub-group $\mathfrak{f}_i$ of $\mathfrak{g}_i$. Hence $\mathfrak{f}_i = 0$ or $\mathfrak{g}_i$, and in either case $\mathfrak{f}_i' = \mathfrak{f}_i$. If a is any element of $\mathfrak{g}$, $[ax_i] = [a_i x_i] \in \mathfrak{f}_i$. But the linear combination of elements of the form $[a_i x_i]$ determine $\mathfrak{f}_i' = \mathfrak{f}_i$. Thus $\mathfrak{f}_i \subset \mathfrak{g}$ and $\mathfrak{g} = \mathfrak{f}_1 \oplus \mathfrak{f}_2 \oplus \dots = \mathfrak{g}_{i_1} \oplus \mathfrak{g}_{i_2} \oplus \dots$ (if we omit the $\mathfrak{f}_i = 0$). If we have any other decomposition of $\mathfrak{g}$ into simple groups, the components are simple invariant subgroups and hence coincide with the $\mathfrak{g}_i$.

3) If $\mathfrak{g}$ is semi-simple, $\mathfrak{g}' = \mathfrak{g}$. For $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \dots$ where the $\mathfrak{g}_i$ are simple. Hence $\mathfrak{g} = \mathfrak{g}_1' \oplus \mathfrak{g}_2' \oplus \dots = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \dots = \mathfrak{g}$. It follows that $\psi_1(x) \equiv 0$ for a semi-simple group, and $\psi_2(x)$ can replace $\sigma_2(x)$ in Cartan's criteria.

4) Invariant bilinear forms. We recall that a symmetric bilinear form $g(x, y)$ is an invariant of $\mathfrak{g}$ if $g([ax] y) + g(x [ay]) = 0$, i.e. the trilinear form $H(x, y, z) = g([xy]z)$ is anti-symmetric in y and z . Suppose now that the number field of $\mathfrak{g}$ is algebraically closed, and 1) $\mathfrak{g}$ is simple. Any invariant bilinear form either vanishes or is non-degenerate. For the elements a such that $g(ax) = 0$ for all $x \in \mathfrak{g}$ form an invariant sub-group. $g(x, y) - \lambda \sigma(x, y)$ is an invariant. There exists a λ such that $g(x, y) - \lambda \sigma(x, y)$ is degenerate. It follows for this λ that $g(x, y) \equiv \lambda \sigma(x, y)$. The only invariant bilinear forms for a simple group over an algebraically closed field are the constant multiples of $\sigma(x, y)$.

2) Let $\mathfrak{g}$ be semi-simple, $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_3 \oplus \dots$, ($x = x_1 + x_2 + \dots$), $\mathfrak{g}_i$ simple. Let $\sigma_i(x_i, y_i)$ be the scalar product of $x_i, y_i \in \mathfrak{g}_i$ in $\mathfrak{g}_i$ and $g(x, y)$ any invariant of $\mathfrak{g}$. $g(x, y)$ is an invariant of $\mathfrak{g}_1$, and hence $g(x_1, y_1) = \lambda_1 \sigma_1(x_1, y_1)$. Set

$$g(x, y) - \lambda_1 \sigma_1(x_1, y_1) = g^*(x, y)$$

We have

$$g^*([a_1 x_1], y) + g^*(x_1, [a, y]) = 0 \quad x_1, a_1 \in \mathfrak{g}_1$$

Since $[a, y] = y_1 \in \mathfrak{g}_1$, $g^*(x_1, y_1) = 0$ so that

$$g^*([a_1 x_1], y) = 0$$

or $g^*(x_1, y) = 0$ for $x_1 \in \mathfrak{g}_1'$. But $\mathfrak{g}_1' = \mathfrak{g}_1$, and hence

$$(9.1) \quad g^*(x_1, y) = 0$$

for $x_1 \in \mathfrak{g}_1$. (9.1) shows that g^* does not involve the variables of x_1 or y_1 , and hence we may use induction and prove that any invariant

$$g(x, y) = \lambda_1 \sigma_1(x_1, y_1) + \lambda_2 \sigma_2(x_2, y_2) + \dots$$

Criteria analogous to those of Cartan hold for associative algebras. The proofs there are considerably simpler.

10. Rational proofs of Lie's theorems and Cartan's theorem on irreducible groups

The following discussion, with slight modification, is taken from a paper by N. Jacobson which will appear in the Annals of Mathematics.

Definition. An associative algebra is semi-simple if it contains no nilpotent invariant sub-algebra $\neq 0$.

Let $\mathcal{G}$ be an infinitesimal group of transformations and $\mathcal{A}$ its enveloping algebra. We propose to investigate the structure of $\mathcal{G}$ when $\mathcal{A}$ is semi-simple. We give first some lemmas (1, 2, 3) which do not make use of this assumption.

Lemma 1. If $\mathcal{H}$ is an invariant sub-group of $\mathcal{G}$ and $\mathcal{A}$ and $\mathcal{L}$ respectively their enveloping algebras, then $\mathcal{A}\mathcal{L}$ and $\mathcal{L}\mathcal{A}$ are invariant sub-algebras of $\mathcal{A}$. If $\mathcal{L}$ is nilpotent, so are $\mathcal{A}\mathcal{L}$ and $\mathcal{L}\mathcal{A}$.

We note first that

$$(10.1) \quad \mathcal{A}\mathcal{L} \subset \mathcal{L}\mathcal{A} + \mathcal{L} \quad \mathcal{L}\mathcal{A} \subset \mathcal{A}\mathcal{L} + \mathcal{L}$$

It is sufficient to show the first of these, and because of the distributive law we need to prove only that

$$G_1 G_2 \dots G_s H_1 H_2 \dots H_t \in \mathcal{L}\mathcal{A} + \mathcal{L}$$

($G_i \in \mathcal{G}$, $H_i \in \mathcal{H}$). Let $[G_s, H_1] = H' \in \mathcal{H}$, then $G_s H_1 = H' + H_1 G_s$ and

$$G_1 \dots G_s H_1 \dots H_t = G_1 \dots G_{s-1} H' H_2 \dots H_t + G_1 \dots G_{s-1} H_1 G_s H_2 \dots H_t$$

If we use induction on s and t , we get the desired result. (10.1) implies that

$\mathcal{A}\mathcal{L}\mathcal{A} \subset \mathcal{A}\mathcal{L} + \mathcal{L}\mathcal{A}$ and hence $\mathcal{A}\mathcal{L}$ and $\mathcal{L}\mathcal{A}$ are invariant. From (10.1) follows also

$$(10.2) \quad \mathcal{A}\mathcal{L}^k \subset \mathcal{L}^k \mathcal{A} + \mathcal{L}^k \quad \mathcal{L}^k \mathcal{A} \subset \mathcal{A}\mathcal{L}^k + \mathcal{L}^k$$

and from (10.2)

$$(10.3) \quad (\mathcal{A}\mathcal{L})^k \subset \mathcal{A}\mathcal{L}^k \quad (\mathcal{L}\mathcal{A})^k \subset \mathcal{L}^k \mathcal{A}$$

This equation shows that if $\mathcal{L}$ is nilpotent, so are $\mathcal{A}\mathcal{L}$ and $\mathcal{L}\mathcal{A}$.

Lemma 2. If $[G, H] = H'$ is commutative with H , then H' is nilpotent.

The operation $(')$ defined by G has the formal properties of a derivative.

It is linear and

$$(H_1 H_2)' = [G, H_1 H_2] = [G, H_1] H_2 + H_1 [G, H_2] = H_1' H_2 + H_1 H_2'$$

From the commutativity of H and H' , we have $\varphi(H)' = \varphi'(H)H'$ where $\varphi(\lambda)$ is a polynomial with coefficients in the field $\mathcal{F}$ and $\varphi'(\lambda)$ is the derivative. Now suppose $\varphi(H) = 0$ ($\varphi(\lambda)$ may be the characteristic polynomial of H). It follows by induction that

$$\varphi^{(k)}(H)(H')^{2k-1} = 0$$

If the degree of $\varphi(\lambda)$ is h , $\varphi^{(h)}(H) = h!$ and $(H')^{2h-1} = 0$.

Any associative algebra $\mathcal{O}$ (of transformations) determines an infinitesimal group in which $[AB] = AB - BA$. $\mathcal{O}$ is its own enveloping algebra. We use the above notation $\mathcal{O}'$ for the derived of $\mathcal{O}$ and denote the centrum of $\mathcal{O}$ by $\mathcal{L}$.

Lemma 3. $\mathcal{L}_1 = \mathcal{O}' \cap \mathcal{L}$ is a nilpotent algebra.

For if C and $C' \in \mathcal{L}_1$ and $C' = [XY] + [ZW] + \dots$, then $CC' = [CX, Y] + [CZ, W] + \dots \in \mathcal{O}'$. But $CC' \in \mathcal{L}$ also, and hence $\mathcal{L}_1$ is closed under multiplication. Since $C, C^2, C^3 \dots \in \mathcal{O}'$, $\text{tr } C = \text{tr } C^2 = \text{tr } C^3 = \dots = 0$ and C is nilpotent. Thus every element of $\mathcal{L}_1$ is nilpotent and so the commutative algebra $\mathcal{L}_1$ is nilpotent.

Lemma 4. If $\mathcal{O}$ is semi-simple, $\mathcal{L}_1 = \mathcal{O}' \cap \mathcal{L} = 0$.

By Lemma 3, $\mathcal{L}_1$ is a nilpotent algebra. $\mathcal{O}\mathcal{L}_1 = \mathcal{L}_1\mathcal{O}$ is nilpotent and invariant in $\mathcal{O}$. Because of the semi-simplicity $\mathcal{O}\mathcal{L}_1 = \mathcal{L}_1\mathcal{O} = 0$. But then $\mathcal{L}_1$ is nilpotent and invariant. Hence $\mathcal{L}_1 = 0$.

Let $\mathcal{F}$ denote the maximal solvable invariant sub-group of $\mathcal{G}$ and $\mathcal{O}$ the enveloping algebra of $\mathcal{G}$. The following is an important generalization of Cartan's theorem on absolutely irreducible groups:

Theorem 1. If $\mathcal{O}$ is semi-simple, then $\mathcal{G} = \mathcal{F} \oplus \mathcal{G}_1$ where $\mathcal{F}$ is Abelian.

We wish to show first that $[\mathcal{F}, \mathcal{G}] = \mathcal{F}' = 0$. $\mathcal{F}' \subset \mathcal{G}' \subset \mathcal{O}'$. If $\mathcal{F}'$ is $\neq 0$, it is a solvable invariant sub-group of $\mathcal{G}$ and it contains $\mathcal{F}_1 \neq 0$ an Abelian invariant sub-group of $\mathcal{G}$. ($\mathcal{F}_1$ may be taken to be one of the derived sequence of $\mathcal{F}'$.) Write $\mathcal{F}_2 = [\mathcal{F}_1, \mathcal{G}]$, $\mathcal{F}_3 = [\mathcal{F}_2, \mathcal{G}]$, ... and suppose

$\mathfrak{r}_n = \mathfrak{r}_{n+1}$. This means that $\mathfrak{r}_n$ has a basis of elements of the form $S' = [S, G]$ where $S \in \mathfrak{r}_n$ and $G \in \mathfrak{g}$. By Lemma 2 these elements are nilpotent. Hence the commutative enveloping algebra $\mathcal{L}$ of $\mathfrak{r}_n$ is nilpotent. By Lemma 1, $\mathcal{O}$ is nilpotent and invariant in $\mathcal{O}$. Hence $\mathcal{L} = 0$ and $\mathfrak{r}_n = 0$. Thus $\mathfrak{r}_1 > \mathfrak{r}_2 > \mathfrak{r}_3 > \dots > \mathfrak{r}_{n-1} > \mathfrak{r}_n = 0$. $\mathfrak{r}_{n-1}$ consists exclusively of elements of the centrum. This is impossible because of Lemma 4. Hence $\mathfrak{r}' = 0$.

On account of Lemma 4, we may obtain a $\mathfrak{g}'$ such that $\mathfrak{g} = \mathfrak{r} + (\mathfrak{g}' + \mathfrak{g}^*) = \mathfrak{r} + \mathfrak{g}_1$, where the spaces $\mathfrak{r}$, $\mathfrak{g}'$ and $\mathfrak{g}^*$ are independent. $\mathfrak{g}_1$ is an infinitesimal group and $[\mathfrak{g}_1, \mathfrak{r}] = 0$. Hence $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{r}$.

$\mathfrak{g}_1 \cong \mathfrak{g}/\mathfrak{r}$ is semi-simple and hence $\mathfrak{g}_1 = \mathfrak{g}'$. Thus $\mathfrak{g}_1 = \mathfrak{g}'$, $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{r}$. This stronger form of Theorem 1 depends on Cartan's theorem on semi-simple groups and will not be used in the sequel.

The following result is fundamental for Lie's theorems on solvable groups:

Theorem 2. If $\mathfrak{g}$ is solvable and $\mathfrak{r}$ is the radical of the enveloping algebra $\mathcal{O}$ of $\mathfrak{g}$, then $\mathcal{O}(\text{mod } \mathfrak{r})$ is Abelian. $\mathfrak{g}'$ is a nilpotent infinitesimal group.

If $G_1 \equiv H_1, G_2 \equiv H_2 \pmod{\mathfrak{r}}$, then $[G_1 G_2] \equiv [H_1 H_2] \pmod{\mathfrak{r}}$. Thus the elements of $\mathfrak{g}$ taken mod $\mathfrak{r}$ define a group which is isomorphic with $\mathfrak{g}$ (not necessarily (1-1)). We denote this group by $\mathfrak{g}(\text{mod } \mathfrak{r})$. It is solvable and its enveloping algebra is $\mathcal{O}(\text{mod } \mathfrak{r})$. By Theorem 1, $\mathfrak{g}(\text{mod } \mathfrak{r})$ and $\mathcal{O}(\text{mod } \mathfrak{r})$ are Abelian. For any pair of elements $G_1, G_2 \in \mathfrak{g}$, $[G_1 G_2] \equiv 0 \pmod{\mathfrak{r}}$ or $\mathfrak{g}' \subset \mathfrak{r}$. Hence $\mathfrak{g}'$ is a nilpotent Lie algebra.

Let $\mathcal{R}$ denote the vector space on which the solvable group $\mathfrak{g}$ of transformations acts. In the notation of Theorem 2, suppose $\mathfrak{r}^s = 0$ but $\mathfrak{r}^{s-1} \neq 0$. If $\mathcal{L}$ is any sub-set of $\mathcal{O}$, we denote the sub-space of $\mathcal{R}$ generated by the transforms of the vectors of $\mathcal{R}$ by the elements of $\mathcal{L}$ by $\mathcal{L}\mathcal{R}$. Denote $\mathfrak{r}^k \mathcal{R}$ by $\mathcal{R}_k$. We have $\mathcal{R}_{s+1} = 0, \mathcal{R}_s \neq 0$. From the associative law $\mathfrak{r}^{\lambda-k} \mathcal{R}_k = \mathcal{R}_\lambda$ ($\lambda > k$).

Hence

$$\mathcal{R} > \mathcal{R}_1 > \mathcal{R}_2 > \dots > \mathcal{R}_s > \mathcal{R}_{s+1} = 0.$$

The spaces $\mathcal{R}_k$ are invariant with respect to $\mathcal{R}$. For,

$$\mathcal{R}\mathcal{R}_k = \mathcal{R}\mathcal{R}^k\mathcal{R} = (\mathcal{R}\mathcal{R})\mathcal{R}^{k-1}\mathcal{R} \subset \mathcal{R}^k\mathcal{R} = \mathcal{R}_k$$

since $\mathcal{R}\mathcal{R} \subset \mathcal{R}$

Consider the transformations of $\mathcal{R}$ in the projection space $\mathcal{R}_k \pmod{\mathcal{R}_{k+1}}$.

If $A \equiv A' \pmod{\mathcal{R}}$ then $Ae \equiv A'e \pmod{\mathcal{R}_{k+1}}$ where $e \in \mathcal{R}_k$. Thus the transformation of $\mathcal{R}$ in $\mathcal{R}_k \pmod{\mathcal{R}_{k+1}}$ is also a transformation of $\mathcal{R} \pmod{\mathcal{R}}$. In particular the transformations of $\mathcal{R}$ are 0 in $\mathcal{R}_k \pmod{\mathcal{R}_{k+1}}$. Let us choose a basis for $\mathcal{R}_s$, supplement it to obtain a basis for $\mathcal{R}_{s-1}$, supplement this to obtain a basis for $\mathcal{R}_{s-2}$, etc. With respect to this basis the matrices of the transformations of $\mathcal{R}$ all have the form

$$(10.4) \quad A = \begin{pmatrix} A_1 & A_{12} & \dots & A_{1r} \\ & A_2 & \dots & A_{2r} \\ & 0 & \dots & \\ & & \ddots & \\ & & & A_r \end{pmatrix}$$

where the A_i are the matrices which represent a commutative semi-simple algebra $\mathcal{R} \pmod{\mathcal{R}}$. In particular for $A \in \mathcal{R}$, $A_1 = A_2 = \dots = A_r = 0$.

It follows from the general theory of representations of algebras that there exists a fixed algebraic field $\mathfrak{F}$ of finite degree over $\mathcal{F}$ such that any representation of the semi-simple commutative algebra $\mathcal{R} \pmod{\mathcal{R}}$ can be decomposed into 1-dimensional representations in $\mathfrak{F}$. It follows that by transforming in $\mathfrak{F}$ the matrices (10.4) may be taken to have such a form that

$$(10.5) \quad A_i = \begin{pmatrix} \alpha_1^{(i)} & & & \\ & \alpha_2^{(i)} & & 0 \\ & 0 & \ddots & \\ & & & \alpha_r^{(i)} \end{pmatrix}$$

We have proved in this way an extension of Lio's theorem:

Theorem 3. If $\mathcal{G}$ is a solvable group of transformations, its matrices may be taken in the form (10.4). The matrices of $\mathcal{G}'$ have $A_1 = A_2 = \dots = A_r = 0$. By transforming in an algebraic field $\mathcal{Z}$ of finite degree over $\mathcal{F}$ we may take the A_i to have the form (10.5).

Theorem 2 may be applied to abstract infinitesimal groups to obtain

Theorem 4. The first derived group of an abstract solvable group is nilpotent.

The adjoint representation of the solvable group $\mathcal{G}$ is a solvable group $\mathcal{G}'$ of transformations. By Theorem 2, $\mathcal{G}'$ is contained in a nilpotent algebra. Hence there exists an integer t such that $G_1' G_2' \dots G_t' = 0$, $G_i' \in \mathcal{G}'$. According to the definition of the adjoint representation, if $g_i \rightarrow G_i'$, then we have $[g_1' [g_2' \dots [g_t' x] \dots]] = 0$ for all $x \in \mathcal{G}$. Using this result for $x \in \mathcal{G}'$, we have $(\mathcal{G}')^{t+1} = 0$.

Chapter 4. Integral-Topological Viewpoint

1. Invariant volume measure on a compact Lie group

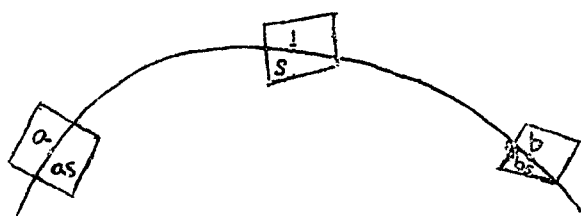
We consider an r -dimensional compact Lie group $\mathcal{G}$ and suppose that it can be parameterized in such a way that the group relation $s_2 s_1 = s_3$ (for s_1 in a suitable neighborhood of the identity) can be given by functions which have continuous first derivatives. Using this assumption we can define the infinitesimal transformations (or the tangent space) of $\mathcal{G}$. If δs is a tangent vector at the identity there is associated with it the vector $a \delta s$ at the point a . Compare Chap. 2, sec. 1.

Let $\delta s \dots ds$ be r independent vectors at the identity. They determine a volume element, a parallelepiped whose volume we assign as

$$\text{absolute value of } \begin{vmatrix} \delta s^1 & \dots & \delta s^r \\ \vdots & & \vdots \\ ds^1 & \dots & ds^r \end{vmatrix}$$

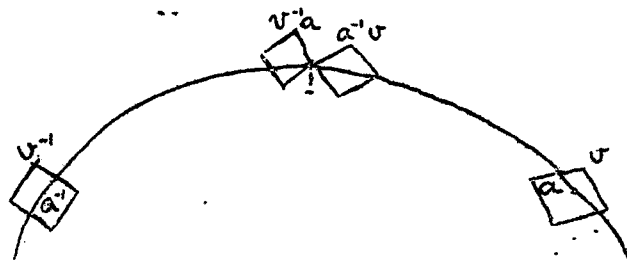
Under a change of coordinates this volume is multiplied by a constant. Consider

r independent vectors at the point a . The parallelepiped determined by them may be obtained by a left translation $s \rightarrow s' = as$ of a parallelepiped at the identity. We define the volume of the element at a to be the volume of the corresponding element at 1 (the identity). This volume measure is invariant under left translation since $(ba^{-1})as = bs$:



The volume measure is also invariant under right translation $s \rightarrow s' = sa$. To prove this it is necessary merely to show that if a volume element at the identity is transported to a by a left translation and returned to 1 by a right translation, then the volume is unchanged. We consider the linear transformation of the tangent space at the identity, which is given by $s \rightarrow asa^{-1}$. Let A be the matrix of this transformation. The correspondence $a \rightarrow A$ gives the adjoint representation of G . We must show that $\text{abs.}|A| = 1$. Suppose $\text{abs.}|A| < 1$, then $\text{abs.}|A|^n \rightarrow 0$. On the other hand, the sequence of powers $a, a^2, a^3, \dots$ have a limit point b and a subsequence a^{n_k} converges to b . Let $b \rightarrow B$ in the adjoint representation. $|B| \neq 0$. By the continuity of the representation $\lim A^{n_k} = B$ and $\lim \text{abs.}|A|^{n_k} = \text{abs.}|B| \neq 0$. From this contradiction we conclude that $\text{abs.}|A| \geq 1$. By considering the negative power $a^{-1}, a^{-2}, \dots$ we find that $\text{abs.}|A| > 1$ is impossible so that $\text{abs.}|A| = 1$.

The measure of volume is invariant under inversion $s \rightarrow s^{-1}$. This inversion sends a volume element at 1 determined by $\delta s, \dots, ds$ into the element defined by $-\delta s, \dots, -ds$. Now consider an element v at a and the inverted element v^{-1} at a^{-1} . Under the left translation $s \rightarrow a^{-1}s$, v is transported into $a^{-1}v$ at 1 and under the right translation $s \xrightarrow{-1} s^{-1}a$, v^{-1} is transported into $v^{-1}a$. But $v^{-1}a$ and $a^{-1}v$ are inverse elements at 1 and hence are equal. Hence $v = v^{-1}$. We indicate this argument by the figure



By means of the volume measure we can define an integral. It has the usual properties of an integral and in addition is invariant under translations and inversion. Every continuous function has a finite integral over the whole of $\mathfrak{g}$.

We have seen that the condition that the volume be left-invariant determines it except for a constant factor. We may determine this factor by supposing that the volume of the whole manifold (the integral of 1 over the manifold) is = 1.

2. Decomposability of representation of a compact group

Theorem: Every representation of a compact Lie group is equivalent to a unitary representation.

The proof is the same as for finite groups except that integration replaces summation. We take any ρ we define hermitian form $A = \sum a_{ik} x_i \overline{x_k}$ and apply $U(s)$ to it obtaining $sA = \sum a_{ik}(s) x_i \overline{x_k}$ the vector x is transformed into $U(s)x$ in passing from A to sA). sA is positive definite and hermitian. Let $H = \int sA ds$. H is positive definite and hermitian and is invariant under $\mathfrak{g}$, i.e., $aH = H$ for every $a \in \mathfrak{g}$. For,

$$aH = \int asA ds = \int tA ds = \int tAdt = H.$$

We have set $t = as$ and made use of the left-invariance of the volume. Now transform H into the unit hermitian form $\sum x_i \overline{x_i}$ by a change of coordinates. The matrices $U(s)$ are transformed into matrices which leave $\sum x_i \overline{x_i}$ invariant and hence are unitary.

It follows that every representation of a compact group is decomposable if reducible.

3. Invariants. Finiteness theorem for compact groups

In this section we shall apply the integration method to prove the finiteness theorem of invariant theory.

Let $\mathcal{G}$ be an arbitrary group and $U(s), V(s), \dots$ representations of $\mathcal{G}$ in the spaces $\mathcal{R}, \mathcal{S}, \dots$ respectively. A vector $x \in \mathcal{R}$ is transformed into $sx \equiv U(s)x, y \in \mathcal{S} \rightarrow sy \equiv V(s)y, \dots$. Let $f(x, y, \dots)$ be a polynomial which is homogeneous of degree μ in the x_i , of degree ν in the $y_i, \dots$. We define the transformed function of f by s to be $sf(x, y, \dots) = f(s^{-1}x, s^{-1}y, \dots)$.

Definition: f is invariant under $\mathcal{G}$ if $sf = f$ for every s .

Classical invariant theory dealt with the following special case: Let $\mathcal{G}$ be the unimodular linear group and $\varphi = \sum_{i,k,\dots} a_{ik\dots} x_i x_k \dots$ a form of degree μ ($a_{ik\dots}$ is a symmetric tensor). Under the influence of the transformation $x_i = \sum \gamma_{ik} x'_k$ of the unimodular group, $\varphi \rightarrow \sum a'_{ik\dots} x'_i x'_k \dots$. The transformation of the coefficients $a_{ik\dots} \rightarrow a'_{ik\dots}$ gives a representation of $\mathcal{G}$. A polynomial in the coefficients $f(a_{ik\dots})$ is transformed in this way into $f'(a'_{ik\dots}) = f(a_{ik\dots})$. f is an invariant in the sense of classical invariant theory if $f'(a'_{ik\dots}) = f(a_{ik\dots})$. Thus the classical theory deals with a special group, the unimodular group, and a special representation given by the transformation of symmetric tensor $a_{ik\dots}$.

Theorem: There exists a finite number of invariants for a compact group such that any invariant is expressible as a polynomial in them.

This holds in particular for the orthogonal and for the unitary groups. It does not apply directly to the unimodular group which is not compact but the theorem can be extended to this case of classical invariant theory by the following argument: The condition that $f(a_{ik\dots})$ be an invariant for the unimodular group is that $df = 0$ for an infinitesimal unimodular transformation

$$dx_i = \sum \gamma_{ik} x'_k \quad \left(\sum \gamma_{ii} = 0 \right). \quad \text{This condition is of the form } \sum A_{ik} \gamma_{ik} = 0$$

where A_{ik} is a polynomial in the $a_{ik\dots}$. If f is an invariant under the uni-

tary group $df = \sum A_{ik} \gamma_{ik} = 0$ for all $a_{ik} \dots \gamma$ and $\gamma_{ik} = -\bar{\gamma}_{ki}$. ($\gamma_{ik} = -\bar{\gamma}_{ki}$ is the condition that $dx_i = \sum \gamma_{ik} x'_k$ be an infinitesimal unitary transformation.) But if $df = 0$ for $\gamma_{ik} = -\bar{\gamma}_{ki}$, $df = 0$ for all γ_{ik} and f is an invariant under the unimodular group. The finiteness theorem for the unimodular group follows from the corresponding theorem for the unitary group.

We require a theorem of Hilbert's on polynomial ideals. Such an ideal is a set of polynomials which contains the sum of any two of its members and the product of any of its members by an arbitrary polynomial. Hilbert proved that any polynomial ideal $\mathcal{J}$ has a finite basis, i.e. $\mathcal{J} \supset F_1, F_2, \dots, F_n$ such that any $F = L_1 F_1 + L_2 F_2 + \dots + L_n F_n$, $F \in \mathcal{J}$, L_i polynomials. It follows from this that any set $\mathcal{Y}$ of polynomials has a finite basis. For we take the ideal generated by $\mathcal{Y}$ and a basis $F_1, F_2, \dots, F_n$ for it. The F_i may not belong to $\mathcal{Y}$ but they are linear combinations of polynomials $J_1, J_2, \dots, J_N$ in $\mathcal{Y}$. Hence every member of $\mathcal{Y}$ is a linear combination of the J 's.

Proof of the theorem: We apply the Hilbert theorem to the set of invariants which are not constants and obtain the basis $J_1, \dots, J_N$ such that any invariant

$$(3.1) \quad J = L_1 J_1 + L_2 J_2 + \dots + L_N J_N.$$

If J is homogeneous of degree μ and the J_i of degree μ_i in the variables x, y and y_i respectively in the y 's ..., then L_i may be taken to be homogeneous of degree $\mu - \mu_i$ in the $x, y - y_i$ in the y If the L_i are invariants we could conclude the theorem by induction on the degree of homogeneity. We may arrange that this is so by the method of integration. We transform (3.1) by s obtaining

$$sJ = J = (sL_1)J_1 + (sL_2)J_2 + \dots + (sL_N)J_N$$

Now integrate with respect to s over and obtain

$$J = I_1 J_1 + I_2 J_2 + \dots + I_N J_N$$

where $I_k = \int s L_k ds$ is an invariant of degree $\mu - \mu_k$. We repeat the process for the I_k obtaining finally J a polynomial in the J_k .

HARMONICS ON HOMOGENEOUS MANIFOLDS

[To appear in The Annals of Mathematics]

by

H. Weyl

Introduction

Our investigation starts from the following data: a compact Lie group $\mathcal{G}$, a point field $\mathcal{M}$ and a realization of $\mathcal{G}$ by means of transformations (one-to-one correspondences) in $\mathcal{M}$, so that the element s of $\mathcal{G}$ carries over the point P of $\mathcal{M}$ into a point sP . $P, Q, \dots$ denote arbitrary points of the point field, $s, t, \dots$ arbitrary elements of the group. We assume that the group of transformations $\mathcal{G}$ is transitive on $\mathcal{M}$ so that all points P are equivalent with respect to $\mathcal{G}$. A function $f(P)$ is carried over by a transformation s of our group into the function sf defined by

$$sf(sP) = f(P) \quad \text{or} \quad sf(P) = f(s^{-1}P)$$

An h -dimensional linear set $\mathcal{R}_h$ of functions $x_1\varphi_1(P) + \dots + x_h\varphi_h(P)$ on with the basis consisting of the linearly independent functions $\varphi_1, \dots, \varphi_h$ is invariant with respect to $\mathcal{G}$, if the transformed functions $s\varphi_\mu$ are linear combinations of the φ_μ :

$$(1) \quad \varphi_{\mu}(P) = \sum_{\nu=1}^h u_{\nu\mu}(S) \varphi_{\nu}(P)$$

$U(s) = \|u_{\mu\nu}(s)\|$ then is a representation U of the group σ ; we say that the invariant set belongs to, or is associated with, the representation U . The set is primitive if the corresponding representation is irreducible.

The classical example of the situation at hand is the sphere $\mathbb{W}$ under the influence of the group σ of all rotations. Here the spherical harmonics of a given order form a primitive invariant set. In accordance with this example one will expect that the primitive invariant sets form together a complete orthogonal system on $\mathbb{W}$. We may stress the analogy by using the term harmonic set of order U instead of invariant set associated with the representation U .

Every abstract group σ allows realization ("regular realization") the point field of which is the group manifold σ itself, and which represents the group element a by the transformation $s \rightarrow as$. In this case the columns of the irreducible representations are the primitive harmonic sets: thus an irreducible representation of degree h yields h linearly independent harmonic sets associated with it. For this case orthogonality was proved by Schur, completeness by Peter and the author*.

* I. Schur, Sitzungsber. Berl. Akad. 1905, p. 406; 1924, p. 199. Peter and Weyl, Math. Ann. 97 (1927), p. 737 (cited as P.W.).

The method can be carried over, practically without any change, to our more general problem which was first propounded and treated by A. Cartan**. It becomes in this

** Rend. Circ. Mat. di Palermo 53 (1929), p. 217.

case even more lucid since the coincidence of the point field with the group manifold in the special case is embarrassing rather than simplifying. Two observations caused me to come back to this problem after Cartan's paper:

- 1) It is not necessary to assume $\mathbb{W}$ to bear a metric, for by means of the invariant volume measure on the group manifold one is able to form the mean value of a continu-

ous function $f(P)$ on π in the following way:

$$m(P) = \int_S f(s^{-1}P)$$

$\int_S$ denotes the average with respect to s . We have

$$m(a^{-1}P) = \int_S f(s^{-1}a^{-1}P) = \int_t f(t^{-1}P) = m(P) \quad [t = as]$$

Since every point Q can be derived from P by means of a transformation a^{-1} of the group, $m(Q)$ is equal to $m(P)$ for any two points P and Q ; hence $m(P)$ is a constant m and this constant may be called the mean value of $f(P)$

$$m = \int_{\pi} f(P)$$

2) Cartan proves completeness in an indirect way, making use of kernels $F(P, Q)$ which are invariant, for the transformations of the group

$$(2) \quad F(sP, sQ) = F(P, Q).$$

But one gets the relation of completeness holding for the function $f(P)$ quite directly if one, in even closer accordance with the procedure of Peter and myself, chooses

$$(3) \quad F(P, Q) = \int_S \bar{f}(s^{-1}P) f(s^{-1}Q)$$

or, what amounts to the same thing, one considers, following E. Schmidt's theory of the eigen-values of asymmetric kernels, the two hermitian conjugate operations

$$(4) \quad \begin{aligned} f : \varphi(P) &\rightarrow \psi(s) = \int_P f(s, P) \varphi(P) \\ f^* : \psi(s) &\rightarrow \varphi(P) = \int_S \bar{f}(s, P) \psi(s) \text{ with } \bar{f}(s, P) = \overline{f(s^{-1}P)} \end{aligned}$$

(The notation of course is not intended to suggest that the two operators are reciprocal.)

Indeed, if $\gamma \neq 0$ is an h -fold eigen-value of $F(P, Q)$ with the h independent eigen-functions $\bar{\varphi}_\mu(P)$, chosen as a normalized, unitary orthogonal system:

$$\int_P \bar{\varphi}_\mu(P) \varphi_\nu(P) = \delta_{\mu\nu}$$

then, on account of the invariance of F , $\bar{\varphi}_\mu(sP)$ is an eigen-function too, and

$\{\varphi_1(P), \dots, \varphi_h(P)\}$ is an invariant set of functions, eq. (1), associated with

a definite representation U . The φ being normalized, this U is unitary.

$$(5) \quad \alpha_\mu = \int_P f(P) \bar{\varphi}_\mu(P)$$

may denote the corresponding Fourier coefficients of f . A simple computation

which is carried through in §1 yields

$$(6) \quad h\gamma = \alpha_1 \bar{\alpha}_1 + \dots + \alpha_h \bar{\alpha}_h$$

Consequently, the theorem that the sum of the eigen-values of the kernel $F(P, Q)$ equals its trace $\text{tr}(F) = \int_P F(P, P)$, leads immediately to the relation of completeness, for we have

$$F(P, P) = \int_S |f(sP)|^2 = \int_P |f(P)|^2, \quad \text{tr}(F) = \int_P |f(P)|^2$$

These preliminary remarks were intended for the expert who is familiar with Cartan's paper. I now embark upon the systematic exposition by which I should like to replace the two papers Peter-Weyl and Cartan. Concerning the group, we make those assumptions which grant the existence of a differential invariant volume measure on the group manifold (cf. P-W). $P' = sP$ may depend uniformly continuously on s and P . In the first part we discuss orthogonality in the same manner as Cartan discussed it; it is shown that at most h independent harmonic sets can be associated with an irreducible representation of degree h . The second part contains the proof of completeness and the expansion and approximation theorem.

I may point out, by the way, that the whole investigation could be carried through within the more general frame of von Neumann's theory of almost periodic functions on arbitrary groups*. Here group and point field need not be compact,

* Which will soon appear in the Trans. of the Amer. Math. Soc.

nor even topological. On the contrary, we topologize the group manifold and the point field relatively to the given function f by agreeing to say that two group elements s and t and two points P and Q , have a distance $\leq \varepsilon$ if in the first case $|f(sP) - f(tP)| \leq \varepsilon$ identically with respect to P , and if in the second case $|f(sP) - f(sQ)| \leq \varepsilon$ identically with respect to s . Following Bochner and von Neumann, f is called almost periodic if, in consequence of this topologizing, the two manifolds σ and π become finite in the sense that, to every positive ε one can choose a finite number of points such that every point has a distance $\leq \varepsilon$ from

one of these chosen points. (A "finite" country can be watched by a "finite" number of policemen, however small the radius of action of the single policeman may be.) Following von Neumann, $s f(P) = f(s^{-1}P)$, considered as a function of s then has a mean value; this is independent of P and may be taken as the mean value of the almost periodic function $f(P)$ on π . When looked upon from this more general viewpoint, the present note should absorb my paper on integral equations and almost periodic functions, Math. Ann. 97 (1926), p. 338, as well as von Neumann's forthcoming paper on almost periodic functions on groups.

§1. Orthogonality

Let $\varphi(P)$ run through a harmonic set $\mathcal{R}_h$ of continuous functions

$$(7) \quad x_1 \varphi_1(P) + \dots + x_h \varphi_h(P)$$

with the unitary basis $\varphi_\mu(P)$:

$$\int_P \bar{\varphi}_\mu(P) \varphi_\nu(P) = \delta_{\mu\nu}$$

According to (1) it belongs to a definite unitary representation U : the function (7) with the parameters x_μ is carried over into the function with the parameters

$$x'_\mu = \sum_\nu u_{\mu\nu}(s) x_\nu$$

by means of the transformation s . If one makes a unitary change of basis, the representation is changed into a unitary equivalent one. The harmonic set is called primitive provided the corresponding representation be irreducible. For well known reasons of elementary unitary geometry it follows that a unitary representation can be completely decomposed into irreducible constituents; by means of the corresponding choice of bases $\mathcal{R}_h$ breaks up into a number of mutually orthogonal primitive harmonic sets.

Theorem 1: The functions of two primitive sets belonging to inequivalent representations are mutually orthogonal.

Let the two sets belonging to the inequivalent irreducible representations U and V consist of the functions

$\varphi(P) = x_1 \varphi_1(P) + \dots + x_h \varphi_h(P)$ and $\varphi'(P) = y_1 \varphi'_1(P) + \dots + y_k \varphi'_k(P)$ respectively.

$$(8) \quad \int_P \bar{\varphi}(P) \varphi'(P) = \sum_{\mu, \nu} a_{\mu\nu} \bar{x}_\mu y_\nu$$

is invariant with respect to the substitutions $U(s)$ and $V(s)$ simultaneously performed on x and y :

$$U^*(s)AV(s) = A \quad \text{or} \quad AV(s) = U(s)A$$

A is the matrix of the $a_{\mu\nu}$, A^* in general the transposed-conjugate of A . In accordance with I. Schur's well-known lemma, this yields:

$$A = 0 \quad a_{\mu\nu} = 0 \quad \text{or} \quad \int \bar{\varphi}_\mu \varphi'_\nu = 0$$

Theorem 2: Two primitive harmonic sets $\varphi_\mu(P)$ and $\varphi'_\nu(P)$ associated with the same representation U satisfy the relations

$$\int_P \bar{\varphi}_\mu(P) \varphi'_\nu(P) = c \delta_{\mu\nu}$$

The constant c is determined by

$$(9) \quad (\varphi, \varphi') = \frac{1}{h} \int_P \sum_{\mu=1}^h \bar{\varphi}_\mu(P) \varphi'_\mu(P) = c$$

Again one forms (8). Its coefficient matrix A commutes with all $U(s)$ and is therefore, according to Schur's lemma, a multiple c of the unit matrix^{*)}.

^{*)} Theorem 2 implies the fact: An invariant set $\{\varphi_\mu(P)\}$ belonging to the unitary irreducible representation U satisfies equations

$$\int \bar{\varphi}_\mu(P) \varphi_\nu(P) = c \delta_{\mu\nu}$$

Therefore only two cases are possible: either all the $\varphi_\mu(P)$ vanish ($c = 0$) or the $\varphi_\mu(P)$ are linearly independent and orthogonal (with a "normalizing" constant $c \neq 0$ however, instead of $c = 1$).

Consequences of Theorem 2: Being given several primitive harmonic sets $\varphi^{(1)}, \varphi^{(2)}, \dots$ associated with the same representation U , every linear combination $a^{(1)} \varphi^{(1)} + a^{(2)} \varphi^{(2)} + \dots$ is such a set again. Within this linear manifold one may define a Hermitian metric by means of the scalar product (11.9) and hence

one may choose the linearly independent among the sets such that $(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}$. Theorem 2 then shows that all the functions of all the sets are orthogonal to each other

$$\int_P \varphi_\mu^{(i)}(P) \varphi_\nu^{(k)}(P) = \delta_{ik} \delta_{\mu\nu} = \begin{cases} 1 & \text{for } i = k, \mu = \nu \\ 0 & \text{in all other cases} \end{cases}$$

This implies the statement, being given several linearly independent harmonic sets associated with the same irreducible representation U --- so that no non-trivial linear relation

$$\sum_i a^{(i)} \varphi_\mu^{(i)}(P) = 0$$

with constant coefficients $a^{(i)}$ exists, all the functions of the sets are linearly independent; that is, there does not exist any non-trivial relation of the form

$$\sum_{i,\mu} a_{i\mu}^{(i)} \varphi_\mu^{(i)}(P) = 0$$

Theorem 3: There exist most h linearly independent primitive harmonic sets associated with a given representation of degree h .

Consider the transformed functions of the given set $s \varphi_\mu^{(i)}(P) = \varphi_\mu^{(i)}(s^{-1}P)$ as functions of s . These functions of s are linearly independent as well as the functions $\varphi_\mu^{(i)}(P)$ of P since P can be carried over into every point by means of a transformation s^{-1} . But these functions of s are, according to the equations

$$s \varphi_\mu^{(i)}(P) = \sum_\nu \mu_{\nu\mu}(s) \varphi_\nu^{(i)}(P)$$

linear combinations of the h^2 functions $\mu_{\nu\mu}(s)$. Hence their number cannot be greater than h^2 and the number of the sets greater than h .

The harmonic sets associated with a given irreducible representation of degree h form the elements of a linear family of $\chi \leq h$ dimensions. The basis $(\varphi^{(i)} (i = 1, \dots, \chi))$ within this family may always be chosen unitary according to the metric (19)

$$(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}$$

By this requirement the basis is determined but for an χ -dimensional unitary transformation: a definite choice of the basis is called a "normalization".

Application to the regular realization: Let U be an irreducible representation of degree h . The equation

$$U(a^{-1}s) = U(a^{-1})U(s) = U^{-1}(a)U(s) = U^*(a)U(s)$$

states that the columns $\varphi^{(1)}(s) \dots \varphi^{(h)}(s)$ of $U(s)$ are invariant sets on the group manifold associated with the representation $\bar{U}$: $\varphi^{(i)}(s) = u_{\mu i}(s)$ so in this case the maximum possible number h of existing linear invariant sets is reached.

Since we have

$$\sum_j \bar{\varphi}_j^{(i)}(s) \varphi_j^{(k)}(s) = \delta_{ik}$$

($U(s)$ being unitary) there follows in using the notation (9)

$$(\varphi^{(i)}, \varphi^{(k)}) = \frac{1}{h} \delta_{ik}$$

and hence the relations of orthogonal

$$\int u_{\mu i}(s) u_{\nu k}(s) = \frac{1}{h} \delta_{\mu\nu} \delta_{ik}$$

first derived by I. Schur, obtain. In accordance with Theorem 1, the elements

$u_{\mu i}(s)$, $v_{\nu k}(s)$ of the matrices $U(s)$, $V(s)$ of two inequivalent irreducible representations are unitary orthogonal to each other.

§ 2. Completeness

Let $f(\pi)$ be a continuous function on π . We call the mean value

$\int_P \bar{f}(P) f(P)$ the magnitude $\|f\|^2$ of f . We form its Fourier coefficients

$$\int_P f(P) \bar{\varphi}_\mu(P) = \alpha_\mu$$

with respect to a given primitive harmonic set

$$\mathcal{R}_h: \varphi_1(P), \dots, \varphi_h(P)$$

or orthogonal functions.

$$(10) \quad e(P) = \alpha_1 \varphi_1(P) + \dots + \alpha_h \varphi_h(P)$$

is the orthogonal projection of f upon $\mathcal{R}_h$; for we have

$$(11) \quad \int_P (f(P) - e(P)) \bar{\varphi}_\mu(P) = 0$$

$$(12) \quad \alpha_1 \bar{\alpha}_1 + \dots + \alpha_h \bar{\alpha}_h$$

the magnitude of e is the contribution of $\mathcal{R}_h$ to the magnitude of f . If one considers several independent primitive harmonic sets (1), (2), (3), ... which may be chosen in orthogonal normalization as far as they belong to the same representation --- the sum of the individual contributions (1), (2), (3), ... will certainly not surpass the total magnitude of f according to Bessel's inequality. The theorem of completeness maintains that here the sign of equality holds, provided the sum be extended over a complete system of normalized primitive sets. This implies the fact that there exist sets occurring in f if f is not identically zero. A set $\mathcal{R}_h$ "occurs" in f if its contribution (11.12) is different from zero; we shall say that $\mathcal{R}_h$ occurs in f with a "weight"

$$= \frac{1}{h} (\alpha_1 \alpha_1 + \dots + \alpha_h \bar{\alpha}_h)$$

One proves the theorem of completeness by constructing the eigen-values and eigen-functions of the couple of Hermitian conjugate operators f, f^* , formula (4) following E. Schmidt. We are dealing with an eigen-function $\varphi(P)$ corresponding to the eigen-value γ , if $\varphi(P) \rightarrow \psi(s) \rightarrow \gamma \varphi(P)$ by performing first f and then f^* . At the same time $\psi(s)$ is then carried over into $\gamma \psi(s)$ by the successive operations f^*, f (in reversed order). It is not quite, but nearly, true that every primitive harmonic set consists of eigen-functions of the couple (f, f^*) so that the eigen-functions of such a couple are "essentially" independent of the nature of the function $f(P)$. The statement becomes correct if one normalizes the basis of harmonic sets associated with the same irreducible representation in an appropriate manner dependent on f . The proof rests upon the following fact.

The harmonic set $\varphi_\mu(P)$ of order U gives rise to a harmonic set $\hat{\varphi}_\mu(P)$ of the same order U by applying the operators f, f^* :

$$\varphi_\mu(P) \rightarrow \psi_\mu(s) \rightarrow \hat{\varphi}_\mu(P)$$

Indeed, we have

$$\hat{\varphi}_\mu(P) = \int_Q F(P, Q) \varphi_\mu(Q)$$

where $F = f^* f$ arises from f and f^* by means of "convolution"

$$F(P, Q) = \int_S f^*(P, s) f(s, Q) = \int_S \bar{f}(s^{-1}P) f(s^{-1}Q) \quad [\text{compare (8)}]$$

Hence

$$\hat{\varphi}_\mu(aP) = \int_Q F(aP, Q) \varphi_\mu(Q) = \int_Q F(aPaQ) \varphi_\mu(aQ) = \int_Q F(P, Q) \varphi_\mu(aQ)$$

Consequently the $\hat{\varphi}_\mu(aP)$ are connected with the functions $\hat{\varphi}_\mu(P)$ by means of the same linear transformation $U(a^{-1})$ by which $\varphi_\mu(P)$ is carried over into $\varphi_\mu(aP)$.

Hence $\varphi^{(i)}(P)$ ($i = 1, \dots, \chi$) being the χ harmonic sets of the primitive order U in normalized form, equations like

$$\hat{\varphi}^{(i)}(P) = \sum_{k=1}^{\ell} c_{ik} \varphi^{(k)}(P)$$

will hold with constant coefficients c_{ik} :

$$c_{ik} = (\varphi^{(k)}, \varphi^{(i)}) = \frac{1}{\chi} \sum_{\mu} \int_Q \bar{\varphi}_\mu^{(k)}(P) F(P, Q) \varphi_\mu^{(i)}(Q)$$

$\|c_{ik}\|$ is a Hermitian matrix: $\overline{c_{ki}} = c_{ik}$ since $F(P, Q)$ is a Hermitian kernel:

$\bar{F}(Q, P) = F(P, Q)$. By appropriate choice of the unitary basis $\|c_{ik}\|$ may be reduced to the form

$$\left\| \begin{array}{ccc} \gamma_1 & & \\ & \gamma_2 & \\ & & \ddots \\ & & & \gamma_\ell \end{array} \right\|$$

under such circumstances the set $\varphi^{(i)}(P)$ consists of eigen-functions of the couple (f, f^*) corresponding to the eigen-value γ_i .

For this reason we look for the primitive harmonic sets occurring in f by constructing according to E. Schmidt the eigen-values of this kernel arranged in decreasing order, and the corresponding eigen-functions. The essential steps may be briefly repeated as follows. We suppose that f is not identically zero. One forms the trace Γ^n of F^n , that is of the n -fold iterated kernel F . In particular

$$\Gamma^1 = \int_P |f(P)|^2 = \|f\|^2, \quad \Gamma^2 = \iint_Q |F(P, Q)|^2.$$

Γ^2 is > 0 since $F(P, P) = \Gamma^1 > 0$. We have the inequalities *)

*) For the proof of the following inequalities cf. Weyl, Math. Ann. 97 (1926),

pp. 342-343. It is immaterial that the arguments of some of the kernels here

considered vary within different fields (σ and π). Of course, a trace like $\text{tr}(F) = \int_P F(P, P)$ can be formed only from a kernel $F(P, Q)$ the arguments of which vary in the same field. In proving the inequalities (12) one has to make use of the fact that the trace $\text{tr}(f^* f \dots f^* f)$ of the $2n$ -term convolution is not changed when one shifts the factors cyclically by one place.

$$(13) \quad \Gamma^{n+n} \leq \Gamma^n \Gamma^n \quad (\Gamma^n)^2 \leq (\Gamma^{n-1})(\Gamma^{n+1})$$

Hence Γ^{n+1}/Γ^n increases monotonely towards a positive limit γ which is less or equal to $\sqrt[n]{\Gamma^n}$, in particular $\leq \Gamma^1 = \|f\|^2$. For later use we note the inequality

$$(14) \quad \Gamma^2 \leq \gamma \Gamma^1$$

$\frac{\Gamma^n}{\gamma^n}$ decreases monotonely towards a limit $h \geq 1$ and $\frac{F^n(P, Q)}{\gamma^n}$ tends uniformly towards a limit $E(P, Q)$. Indeed, in writing $F^{n+1} = f^* G^n f$ one gets from Schwarz's inequality

$$\left| \frac{F^{n+1}}{\gamma^{n+1}} - \frac{F^{n+1}}{\gamma^{n+1}} \right|^2 \leq \frac{1}{\gamma^2} \left(\frac{\Gamma^{2n}}{\gamma^{2n}} - 2 \frac{\Gamma^{n+m}}{\gamma^{n+m}} + \frac{\Gamma^{2m}}{\gamma^{2m}} \right)$$

(Here, for convenience' sake, we supposed $\|f\|^2$ to be ≤ 1 .) $E(P, Q)$ is an invariant Hermitian kernel for which the following equalities hold:

$$(15) \quad FE = \gamma E \quad EE^* = EE = E \quad \text{tr}(E) = h \geq 1$$

On account of the second and third equality, one gets an expression

$$(16) \quad E(P, Q) = \bar{\varphi}_1(P) \varphi_1(Q) + \dots + \bar{\varphi}_h(P) \varphi_h(Q)$$

where the φ form the unitary-orthogonal system of the h eigen-functions of F corresponding to the eigen-value γ (consequently $h \geq 1$ ultimately proves to be an integer!). $\varphi_\mu(s^{-1}P)$ is necessarily a linear combination of the $\varphi_\mu(P)$ themselves: the $\varphi_\mu(P)$ therefore form a harmonic set corresponding to a definite representation U of degree h . For one has

$$\begin{aligned} \varphi_\mu(P) &= \int_P \varphi_\mu(P) E(P, Q) \text{ hence} \\ \varphi_\mu(s^{-1}Q) &= \int_P \varphi_\mu(P) E(P, s^{-1}Q) = \int_P \varphi_\mu(s^{-1}P) E(s^{-1}P, s^{-1}Q) \\ &= \int_P \varphi_\mu(s^{-1}P) E(P, Q) = \sum_\nu \varphi_\nu(Q) \int_P \varphi_\mu(s^{-1}P) \bar{\varphi}_\nu(P). \end{aligned}$$

The set $\{\varphi_\mu(P)\}$ -- if it is not primitive itself -- can be split into primitive mutually orthogonal harmonic sets.

In performing first the operation f , then f^* [compare equ. (4)], we have

$$\bar{\varphi}_\mu(P) \rightarrow \psi_\mu(S) \rightarrow \gamma \varphi_\mu(P)$$

The $\psi(S)$ are orthogonal on σ ; for,

$$(18) \quad \int_S \bar{\psi}_\mu(S) \psi_\nu(S) = \int_P \int_Q \varphi_\mu(P) F(P, Q) \bar{\varphi}_\nu(Q) = \gamma \int_P \varphi_\mu(P) \bar{\varphi}_\nu(P) = \gamma \delta_{\mu\nu}$$

and just as the φ_μ are eigen-functions of $F = f^* f$ corresponding to the eigen-value γ , so are the $\psi_\mu(S)$ eigen-functions of $G = f f^*$ corresponding to the same eigen-value. From (18) there follows

$$(19) \quad \int_S \sum_\mu \bar{\psi}_\mu(S) \psi_\mu(S) = h \gamma$$

One finds

$$\psi_\mu(S) = \int_P f(S^{-1}P) \bar{\varphi}_\mu(P) = \int_P f(P) \bar{\varphi}_\mu(SP) = \sum_\nu U_{\mu\nu}(S^{-1}) \int_P f(P) \bar{\varphi}_\nu(P) = \sum_\nu U_{\mu\nu}(S) \alpha_\nu$$

Since $U(S)$ is unitary, one obtains from this expression:

$$(20) \quad \sum_\mu \bar{\psi}_\mu(S) \psi_\mu(S) = \alpha_1 \bar{\alpha}_1 + \dots + \alpha_h \bar{\alpha}_h$$

Comparison of (19) and (20) yields the relation (6) which is decisive for our whole investigation. γ being > 0 , the invariant set $\{\varphi_\mu(S)\}$ actually occurs in f , namely with the weight γ . The same holds for any of the partial primitive harmonic sets into which the whole set may be decomposed.

The next step in constructing all primitive harmonic sets occurring in f consists in deducting from f the contribution just found out:

$$(21) \quad f_1(P) = f(P) - o(P) = f(P) - \{\alpha_1 \varphi_1(P) + \dots + \alpha_h \varphi_h(P)\}$$

and in applying the same construction on f_1 which we carried through for f (provided f_1 does not vanish identically). In this way we get an invariant set

$\varphi^{(1)}(P)$ the h_1 terms of which are eigen-functions of the kernel $F_1 = f_1^* f_1$ corresponding to an eigen-value $\gamma_1 > 0$ and thus occurring in f_1 with the weight γ_1 .

It is important to realize 1) that the new functions $\varphi^{(1)}$ are orthogonal to the previous φ . Hence they occur in f with the same weight γ_1 as in f_1 and are eigen-functions of the kernel F corresponding to the eigen-value γ_1 . 2) $\gamma_1 < \gamma$.

We have

$$\int_S \int_Q \bar{f}_1(S^{-1}P) f_1(S^{-1}Q) \bar{\varphi}_\nu^{(1)}(Q) = \gamma_1 \varphi_\nu^{(1)}(P)$$

If we multiply the left side $\varphi_\mu(P)$ and average over P , we get 0; for

$$\int_P \bar{f}_1(s^{-1}P) \varphi_\mu(P) = \int_P \bar{f}_1(P) \varphi_\mu(sP) = 0$$

since $\varphi_\mu(sP)$ is a linear combination of the $\varphi_\mu(P)$ and since the equation

$\int_P \bar{f}_1(P) \varphi_\mu(P) = 0$ holds for every $\varphi_\mu(P)$. Consequently, as we maintained before:

$$\gamma_1 \int_P \varphi_\mu(P) \bar{\varphi}_\nu^{(1)}(P) = 0$$

The subtracted function $e(P)$ gives rise to a corresponding kernel

$e(s^{-1}P) = e(s, P)$ which we compute to be

$$(22) \quad e(s, P) = \sum_\nu \alpha_\nu \varphi_\nu(s^{-1}P) = \sum_{\nu, \mu} \alpha_\nu \mu_{\mu\nu}(s) \varphi_\mu(P) = \sum_\mu \psi_\mu(s) \varphi_\mu(P)$$

From this we derive following (17): $f^*e = \gamma^*E$ and since we have, in addition to

that, $e^*e = \gamma^*E$ following (18), we get $f_1^*e = 0$, $e^*f_1 = 0$, and hence

$f^*f = f_1^*f_1 + \gamma^*E$ or $F = \gamma^*E + F_1$. The last equation shows in the simplest

manner that the $\varphi^{(1)}$ being orthogonal to the φ 's are eigen-functions of F as well

as of F_1 . Furthermore, iteration will give

$$F^n = \gamma^n E + F_1^n \quad \Gamma^n = h \gamma^n + \Gamma_1^n$$

Hence $\frac{\Gamma^n}{\gamma^n}$ tends towards zero with increasing n , and consequently the number γ_1

for which $\frac{\Gamma_1^n}{\gamma_1^n}$ tends towards the limit $h_1 > 1$ is necessarily $< \gamma$.

After applying the construction p times, one is left with a certain re-

mainder f_p . We have to prove -- provided the construction does not come to an end

by f_p becoming identically zero for certain finite p -- that

$$\Gamma_p^1 = \int_P |f_p(P)|^2 = \text{tr} (f_p^* f_p)$$

tends towards zero with increasing p . From the relation (Bessel's inequality)

$$(3) \quad (h+h_1+\dots+h_p) \gamma_p \leq h\gamma+h_1\gamma_1+\dots+h_p\gamma_p = \|e\|^2 + \|e_1\|^2 + \dots + \|e_p\|^2 \leq \|f\|^2 \leq 1$$

we first conclude only that γ_p tends to zero with increasing p . The connection

with the result wanted is established by means of the inequality

$$(24) \quad \Gamma_p^2 \leq \gamma_p \Gamma_p^1$$

Let $\mathcal{N}_\epsilon$ denote the domain on the group manifold to which s belongs if

$$(25) \quad \int_P |f(s^{-1}P) - f(P)|^2 \leq \epsilon^2$$

One cannot be sure that $\mathcal{N}_\epsilon$ has a volume in the Jordan sense, but we certainly can

assign a positive lower bound V_ε for its volume. After division by the volume V of the whole group manifold, the fraction $w_\varepsilon = \frac{V_\varepsilon}{V}$ is a lower bound of the "probability" of s satisfying the inequality (25). I maintain: As soon as $\gamma_p \leq \varepsilon^2 w_\varepsilon$

$$\|f_p\|^2 = \|f\|^2 - \|e\|^2 - \dots - \|e_{p-1}\|^2 \leq 4\varepsilon^2$$

Proof:

$$G_p = f_p f_p^* = f f_p^*, \text{ hence}$$

$$G_p(s) = G_p(s, \underline{1}) = \int_p f(s^{-1}P) \bar{f}_p(P)$$

($\underline{1}$ denotes the unit element on σ .) $\Gamma_p^2 = \int_p |f_p(P)|^2 = G_p(\underline{1})$ may be put $= \delta_p^2$.

Suppose it to be $> 4\varepsilon^2$. Then

$$(26) \quad G_p(\underline{1}) > 2\varepsilon \delta_p$$

On the other hand,

$$G_p(s) - G_p(\underline{1}) = \int_p \{f(s^{-1}P) - f(P)\} \bar{f}_p(P)$$

and consequently by Schwarz's inequality,

$$(27) \quad |G_p(s) - G_p(\underline{1})|^2 \leq \int_p |f(s^{-1}P) - f(P)|^2 \cdot \int_p |f_p(P)|^2 \leq \varepsilon^2 \delta_p^2 \text{ in } \mathcal{V}_\varepsilon$$

We then have in $\mathcal{V}_\varepsilon$, according to (26) and (27)

$$|G_p(s)| > \varepsilon \delta_p$$

and therefore

$$\Gamma_p^2 = \int_s |G_p(s)|^2 > \varepsilon^2 \delta_p^2 \varepsilon$$

This is in contradiction with the inequality (24) or $\Gamma_p^2 \leq \gamma_p \delta_p^2$ as soon as

$$\gamma_p \leq \varepsilon^2 w_\varepsilon.$$

Let us number the orthogonal functions in the order in which we got them, without marking the individual sets:

$$\varphi_1(P), \varphi_2(P), \dots$$

and let us now cancel the supplementary condition $\|f\|^2 \leq 1$ again. Then, in taking account of the inequality (23), our final result reads as follows:

$$(28) \quad \|f\|^2 - (\alpha_1 \bar{\alpha}_1 + \alpha_2 \bar{\alpha}_2 + \dots + \alpha_n \bar{\alpha}_n) \leq 4\varepsilon^2 \|f\|^2 \quad \text{for } n \geq \frac{1}{\varepsilon^2 w_\varepsilon}$$

It was our intention to make the estimation as exact as possible.

Our construction will furnish all the primitive invariant sets occurring

in f . This is readily seen in the following manner. It will furnish certain, let us say, χ' of the χ mutually orthogonal primitive harmonic sets of order U . The remaining $\chi - \chi'$ cannot occur in f or they contribute nothing to the total magnitude of f as is obvious from Bessel's inequality according to which the sum of all contributions cannot exceed $\|f\|^2$ whereas the contributions which are furnished by our construction alone give the whole amount $\|f\|^2$. Hence we may formulate our final result thus:

Theorem 4: We gave a construction yielding all the primitive invariant sets occurring in f , in a definite normalization and order depending on f . (The order is such that the weights with which the successive sets occur in f form a monotonely decreasing sequence.) The relation of completeness is valid; it can be supplemented by the explicit estimation of the remainder (28) if one keeps to the normalization and order originating from the construction.

The expansion theorem now is an immediate consequence. We consider a function $g(P)$ of the form

$$g(P) = \int_S f(s^{-1}P)x(s)$$

Writing again

$$f(P) = \{e(P) + e_1(P) + \dots + e_{p-1}(P)\} + f_p(P)$$

one has, according to Schwarz's inequality,

$$\left| \int_S f_p(s^{-1}P)x(s) \right|^2 \leq \int_P |f_p(P)|^2 \cdot \int_S |x(s)|^2$$

and the remainder $\int f_p(s^{-1}P)x(s)$ tends therefore to zero with increasing p in a manner that can be estimated explicitly. For the individual contribution we get from

(22):

$$\int_S e(s^{-1}P)x(s) = \sum_{\mu} \beta_{\mu} \varphi_{\mu}(P) \quad \text{with} \quad \beta_{\mu} = \int_S \psi_{\mu}(s)x(s)$$

β_{μ} coincides with the Fourier coefficient of $g(P)$ with respect to $\varphi_{\mu}(P)$

$$\int_P g(P) \bar{\varphi}_{\mu}(P) = \int_S \int_P f(s^{-1}P) \bar{\varphi}_{\mu}(P)x(s) = \int_S \psi_{\mu}(s)x(s)$$

Theorem 5: The Fourier series

$$\beta_1 \varphi_1(P) + \beta_2 \varphi_2(P) + \dots, \quad \beta_i = \int_P g(P) \bar{\varphi}_i(P)$$

of a function $g(P)$ of the form

$$g(P) = \int_S f(s^{-1}P)x(s)$$

converges uniformly to $g(P)$.

Only the primitive sets of functions actually occurring in f appear in this expansion. They are ordered by decreasing weight in f , and the summation should be performed "by sections", without separating the terms of the individual primitive sets of functions.

One obtains the theorem of approximation in choosing $x(s)$ as a function which is $\neq 0$ only in the immediate neighborhood of the unit element:

Theorem 6: $f(P)$ can be approximated uniformly by a finite sum of harmonics occurring in f .

4. Remarks on the special case of the regular realization

We have seen that in the case of the regular realization ($G = \Pi$) the columns of the representation matrix $\bar{U}(s)$ are an invariant set associated with the representation $U(s)$. If $\varphi_\mu(t)$ ($\mu = 1, \dots, h$) is any such invariant set:

$$\varphi_\mu(s^{-1}t) = \sum \mu_{\nu\mu}(s) \varphi_\nu(t),$$

then

$$(4.1) \quad \varphi_\mu(s) = \sum \bar{\mu}_{\mu\nu}(s) \varphi_\nu(1) = \sum \bar{\mu}_{\mu\nu}(s) c_\nu$$

i.e. $\{\varphi_\mu(s)\}$ is a linear combination of the columns of $\bar{U}$.

If $U(s)$ and $V(s)$ are inequivalent and irreducible, the orthogonality relations read

$$\int_S \bar{\mu}_{\mu i}(s) \nu_{\nu k}(s) = 0$$

$$\int_S \bar{\mu}_{\mu i}(s) \mu_{\nu k}(s) = \frac{1}{h} \delta_{\mu\nu} \delta_{ik}$$

(h = degree of U). These relations imply the Burnside-Frobenius-Schur theorem for the representations of compact groups. They yield also the orthogonality of the primitive characters:

$$\int_S \bar{\chi}(s) \chi'(s) = 0 \quad \chi, \chi' \text{ inequivalent}$$

$$\int_S \bar{\chi}(s) \chi(s) = 1$$

The completeness theorem states that if $f(s)$ is a continuous function on σ , $u_{\mu\nu}(s)$ are the elements of a representative set of irreducible inequivalent unitary representations and $\alpha_{\mu\nu} = \int f(s) \bar{u}_{\mu\nu}(s)$, then

$$(4.2) \quad \|f\|^2 = h \sum |\alpha_{\mu\nu}|^2 + \dots$$

We wish to examine more carefully the construction employed above for the harmonic sets occurring in f .

Let $\int f(s) \bar{U}(s) = \{\alpha_{\mu\nu}\} = A(f)$. The "product" of two group functions $f(s)$ and $g(s)$ is defined as the convolution

$$(4.3) \quad fg(s) = \int_t f(st^{-1})f(t) \text{ or } fg(s^{-1}r) = \int_t f(s^{-1}t)g(t^{-1}r)$$

It is easily verified that

$$(4.4) \quad A(fg) = A(f)A(g)$$

i.e. the matrices $A(f)$ give a representation of the algebra of group functions.

The matrix $\bar{U}(t)$ is an eigen-matrix of the kernel $f(s, t) = f(s^{-1}t)$:

$$\int_t f(s^{-1}t) \bar{U}(t) = \int_t f(t) \bar{U}(st) = \bar{U}(s)A(f).$$

Similarly if $f^*(s, t) = \bar{f}(t, s)$

$$\int_t f^*(s^{-1}t) \bar{U}(t) = \bar{U}(s)A^*$$

where A^* is the hermitian conjugate of A . We may combine these two results and obtain

$$(4.5) \quad \int_t F(s, t) \bar{U}(t) = U(s)\Gamma, \quad \Gamma = A^*A, \quad F = f^*f$$

$\Gamma = A^*A$ is a hermitian semi-definite matrix and we may choose a unitary basis so that

$$\Gamma = \begin{pmatrix} \gamma_1 & & 0 \\ & \gamma_2 & \\ 0 & & \ddots \\ & & & \gamma_h \end{pmatrix}, \quad \gamma_i \geq 0.$$

Equation (4.5) becomes

$$(4.6) \quad \int_t F(s^{-1}t) \bar{u}_{\mu\nu}(t) = \bar{u}_{\mu\nu}(s) \gamma_\nu$$

The matrix elements are eigen functions of F corresponding to the roots γ_ν of Γ .

Let $\varphi_\mu(s)$ be a primitive harmonic set of order U obtained as a set of eigen functions of F corresponding to the eigen value γ . The $\varphi_\mu(s)$ have the form (4.1). By (4.6) it follows that $c_\nu = 0$ for those ν such that $\gamma_\nu \neq \gamma$. The ordering of the eigen functions given by the construction is according to the magnitude of the roots γ_i . The maximum γ_i can be determined as the $\max \sum_\nu |\sum_\mu \alpha_{\mu\nu} x_\nu|^2$ for $\sum_\nu |x_\nu|^2 \leq 1$.

If $x(s)$ is the character of the representation $X(s)$ and $\chi(s), \chi'(s), \dots$ the primitive characters, then

$$x(s) = m\chi(s) + m'\chi'(s) + \dots$$

where $m, m', \dots$ are integers which give the number of times the irreducible representations of characters $\chi(s), \chi'(s), \dots$ occur in $X(s)$. m may be determined by

$$\int_S x(s) \bar{\chi}(s) = m$$

because of the orthogonality relations. From

$$\int_S x(s) \bar{x}(s) = m^2 + m'^2 + \dots$$

follows that $X(s)$ is irreducible if and only if $\int_S x(s) \bar{x}(s) = 1$.

The primitive characters are a complete normalized orthogonal system of functions in the domain of continuous class functions.

For, if $f(st) = f(ts)$,

$$\int_t f(s^{-1}t) \bar{U}(t) = \bar{U}(s)A = \int_t f(ts^{-1}) \bar{U}(t) = A\bar{U}(s)$$

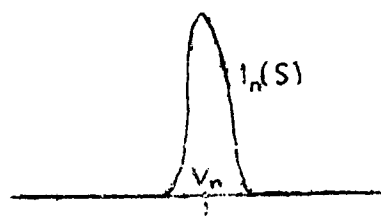
$$A\bar{U}(s) = \bar{U}(s)A$$

By Schur's lemma $A = \alpha E$. Equation (4.2) gives

$$\|f\|^2 = |\alpha|^2 h^2 + \dots = \left| \int_S f(s) \bar{\chi}(s) \right|^2 + \dots$$

which is the completeness relation. The expansion and approximation theorems may be proved as before.

In order to determine all the representations of G we approximate the identity of G by continuous functions $1_n(s)$ of mean value 1, i.e. we let V_n be a sequence of neighborhood of $\underline{1}$ such that $\lim V_n = \underline{1}$ and let $1_n(s) = 0$ outside of V_n :



Since

$$\int_S l_n(s) \bar{U}(s) \rightarrow E$$

for sufficiently large n $\int_S l_n(s) \bar{U}(s) \neq 0$ and the elements of $\bar{U}(s)$ occur in $l_n(s)$.

There are a denumerable number of representations occurring in each $l_n(s)$ and a denumerable number of $l_n(s)$. Hence, the total number of inequivalent irreducible representations of a compact group is denumerable. In the general case of harmonics on a point field π there are at most h primitive sets belonging to a U of degree h . Hence, the number of primitive harmonic sets is denumerable.

The method of Frobenius-Schur for constructing all the representations of a finite group consisted in determining the representations which occur in the function $l(s) = \begin{cases} 0 & s \neq \frac{1}{a} \\ 1 & s = \frac{1}{a} \end{cases}$. This function is replaced in our construction by the approximating sequence $l_n(s)$.

The representations distinguish between the elements of $\mathcal{G}$: if
 $U(s_0) = U(t_0)$ for all U , then $s_0 = t_0$. For, let $t_0 = as_0$, $a \neq \frac{1}{a}$; then

$$U(t_0) = U(s_0) = U(a)U(s_0)$$

Since $U(s_0)$ is non-singular, $U(a) = E$ and

$$U(a)U(s) = U(s)$$

for all s . Hence

$$\int_S f(s) \{ \bar{U}(as) - \bar{U}(s) \} = \int_S \{ f(a^{-1}s) - f(s) \} \bar{U}(s) = 0$$

for all $f(s)$. But if $a \neq \frac{1}{a}$ there is a continuous function $f(s)$ such that

$f(1) \neq f(a)$ or $f(a^{-1}s) - f(s) \neq 0$. By the completeness theorem there exists a $\bar{U}$ such that

$$\int_S \{ f(a^{-1}s) - f(s) \} \bar{U}(s) \neq 0$$

contrary to the above.

If $\chi(s_0) = \chi(t_0)$ for all primitive χ , then s_0 and t_0 are conjugates.

By the approximation theorem it follows for all continuous class functions that $f(s_0) = f(t_0)$. But if s_0 and t_0 are not conjugates there is a continuous class function such that $f(s_0) \neq f(t_0)$.

5. Fourier series. Casimir's theory of spherical harmonics. Bohr's almost-periodic functions

1) Fourier series. The theory of Fourier series is the special case of the above theory for which $\mathcal{G}$ is the group of rotations of a circle. Here the group manifold is a circle. The irreducible unitary representations are 1-dimensional and are the same as their characters. A character $\chi(s)$ is a continuous periodic (of period 1) function of a real variable which satisfies

$$(5.1) \quad \chi(s + t) = \chi(s) \chi(t), \quad |\chi| = 1.$$

The second condition expresses the unitary character of the representation. Let $\chi = e^{2\pi i \lambda(s)}$ where $\lambda(s)$ is ^{real and} a continuous function and $\lambda(0) = 0$. (5.1) gives

$$\lambda(s + t) \equiv \lambda(s) + \lambda(t) \pmod{1}$$

The function $\lambda(s + t) - \lambda(s) - \lambda(t)$ is continuous and has only integral values and so is constant. By the initial condition

$$(5.2) \quad \lambda(s + t) = \lambda(s) + \lambda(t)$$

$\lambda(1)$ is an integer n . From (5.2) we have

$$\lambda\left(\frac{l}{2^m}\right) = \frac{l}{2^m} n \quad \lambda, m \text{ integers}$$

and by the continuity $\lambda(s) = ns$ or $\chi = e^{2\pi i ns}$.

We have proved the completeness, expansion and approximation theorems for the functions $e^{2\pi i ns}$ in the domain of continuous periodic functions.

The general procedure in applying the theory of representations of groups to special cases is illustrated in this example. In all cases we have to determine the irreducible representations or the primitive characters and this is not always as simple as in the present special case.

2) Casimir's theory of spherical harmonics. Casimir considers the theory of harmonics from the infinitesimal rather than the integral point of view. We

return to the situation of the paper Harmonics on homogeneous manifolds. In addition we suppose that the differentiability conditions which enable us to define the infinitesimal group of σ are satisfied. To the element e_α of the infinitesimal group of σ we have corresponding the velocity vector $u_\alpha(x)$ in Π . The transformation $sf(P)$ may be obtained by iterating a suitable infinitesimal transformation, $f \rightarrow f + df$ where

$$df = \sum \frac{\partial f}{\partial x^i} u^i(s),$$

$u^i(x)$ in the velocity field.

A representation U of σ gives a representation (the infinitesimal group of U) of the infinitesimal group of σ .

Suppose that σ is semi-simple. If $\xi = \sum \xi^\alpha e_\alpha$ is a variable element of the infinitesimal group of σ then $\sum g_{\alpha\beta} \xi^\alpha \xi^\beta$ ^{the scalar product} is non-degenerate. If $(g^{\alpha\beta})$ is the inverse of the matrix $(g_{\alpha\beta})$ and $e_\alpha \rightarrow E_\alpha$ in a representation of the infinitesimal group of σ then $H = \sum g_{\alpha\beta} E_\alpha E_\beta$ commutes with all the E_α [see Chap. 3, sec. 3].

Let $\varphi_\mu(x)$ ($\mu = 1, \dots, h$) be a harmonic set corresponding to the representation U .

$$s \varphi_\mu(x) = \sum u_{\nu\mu}(s) \varphi_\nu(x).$$

This condition for the infinitesimal transformations is

$$d\varphi_\mu = \sum \frac{\partial \varphi_\mu}{\partial x^i} u^i(x) = \text{linear combination of the } \varphi_\mu$$

This is a system of partial differential equations satisfied by the harmonic set.

We suppose $\{\varphi_\mu(x)\}$ is primitive. In order to determine a differential equation satisfied by a single $\varphi_\mu(x)$ we consider the operator

$$\Delta f = \sum g^{\alpha\beta} u_\alpha^i(x) \frac{\partial}{\partial x^i} (u_\beta^k(x) \frac{\partial f}{\partial x^k})$$

Let e_α and E_α be corresponding elements of the infinitesimal groups of σ and U . Then

$$\Delta \varphi_\mu = H \varphi_\mu$$

where $H = \sum g^{\alpha\beta} E_\alpha E_\beta$. But U is irreducible and hence $H = \lambda E$ and

$$(5.3) \quad \Delta \varphi_\mu = \lambda \varphi_\mu$$

which is the required equation.

Casimir obtains the orthogonality and completeness of the φ_μ from the theory of eigen functions of a second order partial differential equation. For this purpose it is necessary to show that Δ is self adjoint and elliptic.

Consider $\int \psi \Delta \varphi \rho dx$ where ρ is a scalar density determined by the volume element. The following is the analogue of Green's formula:

$$(5.4) \quad \begin{aligned} \int \psi \Delta \varphi \rho dx &\equiv \int \sum g^{\alpha\beta} u_\alpha^i \frac{\partial}{\partial x^i} \left(u_\beta^k \frac{\partial \varphi}{\partial x^k} \right) \rho dx \\ &= - \int \sum g^{\alpha\beta} u_\alpha^i u_\beta^k \frac{\partial \varphi}{\partial x^k} \frac{\partial \psi}{\partial x^i} \rho dx \equiv -D(\varphi, \psi) \end{aligned}$$

This is obtained by partial integration. The additional term arising from the partial integration is 0 because of the invariance of the volume measure. Since $D(\varphi, \psi) = D(\psi, \varphi)$,

$$\int \psi \Delta \varphi \rho dx = \int \varphi \Delta \psi \rho dx$$

i.e. Δ is self-adjoint.

It can be shown that because of the compactness $\sum g_{\alpha\beta} \xi^\alpha \xi^\beta$ is a definite form. It follows that $D(\varphi, \varphi) \geq 0$, or Δ is elliptic.

3) Bohr's theory of almost-periodic functions. In the case of Fourier series we were dealing with a compact group. The theory of almost periodic functions is the simplest example of the representation theory of an open group. G is the Abelian group of translations of a straight line. To determine the irreducible unitary representations one must solve the functional equations

$$\chi(s+t) = \chi(s)\chi(t) \quad |\chi| = 1$$

where s and t are real variables. One shows readily that the only continuous solutions of these equations are the pure oscillations $\chi(s) = e^{i\lambda s}$, λ real. Thus the number of inequivalent representations is non-denumerable.

Bohr proposed and solved the following problem: to characterize the class of functions which can be uniformly approximated by linear combinations of pure os-

cillations. He showed that those functions are the almost periodic (a.p.) functions which we proceed to define.

Definition 1. An ε -near period (translation number) of a function f is a real number t such that

$$|f(s+t) - f(s)| \leq \varepsilon \text{ for all } s$$

Definition 2. A function is almost periodic if its ε -near periods have a "regular distribution", i.e. there exists a length L_ε such that every interval of length L_ε on the real axis contains an ε -near period.

A fundamental point in the Bohr theory is the existence of a mean value for an a.p. function,

uniformly in s . We denote $\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) dt$ by $M\{f(t)\}$. $M\{f(t)e^{-i\lambda t}\} = a_\lambda$ exists for all λ and because of Bessel's inequality

$$(5.5) \quad \sum |a_\lambda|^2 \leq M\{|f(t)|^2\}$$

$a_\lambda \neq 0$ for only a denumerable set of λ 's. Bohr proved the completeness theorem that the sign of equality holds in (5.5).

The method we used in obtaining the representations which occur in a function on a group can be applied to obtain the terms which actually occur ($a_\lambda \neq 0$) in the Fourier series $\sum a_\lambda e^{i\lambda s}$ of $f(s)$. We shall discuss this more fully in the next section in connection with von Neumann's theory of almost periodic functions.

6. Von Neumann's theory of almost periodic functions. Haar's measure

We consider an arbitrary group $\mathcal{G}$ without topology and a realization of by transformations of a point field Π . If $f(P)$ is a function on Π , we may topologize $\mathcal{G}$ and Π relative to f in the following way: We say that the distance $|st| \leq \varepsilon$ if

$$|f(sP) - f(tP)| \leq \varepsilon \text{ for all } P$$

and $|P_s| \leq \varepsilon$ if

$$|f(sP) - f(sQ)| \leq \varepsilon \text{ for all } s.$$

If, in this topology, the spaces $\mathcal{U}$ and $\mathcal{T}$ can be covered by a finite number of spheres of radii $\leq \varepsilon$ no matter how small ε is, then f is an almost periodic function for $\mathcal{U}$ and $\mathcal{T}$. This definition was first given by Bochner and shown by him to be equivalent to Bohr's definition for the continuous functions on the group of translations of the straight line.

Von Neumann (in a paper to appear in the Transactions) proved the existence of a mean value for a.p. functions which has the properties of Bohr's mean and coincides with it for Bohr's almost periodic functions. Choose n points a_i on $\mathcal{U}$ and form the average $g_n(P) = \frac{1}{n} \sum f(a_i P)$. It can be shown that the points a_i can be chosen in such a way that the functions $g_n(P)$ are "smoothed out", i.e. the functions $g_n(P)$ have a constant c as a limit approached uniformly in P . We denote this limit by $M_a \{f(aP)\}$. In the same way we may form $\frac{1}{n} \sum f(aP_i) = g_n(a)$ and show that the P_i 's can be chosen so that $g_n(a) \rightarrow c \equiv M_P \{f(aP)\} = M_a \{f(aP)\}$. The mean c is independent of the choice of the points a_i or the points P_i . It has the usual properties and is invariant:

$$M_P \{f(aP)\} = M_P \{f(asP)\} \qquad M_a \{f(aP)\} = M_a \{f(asP)\}$$

The sequences $a_i(P_i)$ which lead to the mean may be considered as "uniformly distributed" relative to the function f . In order to prove a relation between the means of several functions it is necessary to show that there exist sequences which are simultaneously uniformly distributed for the functions.

If we replace the continuous functions by the a.p. functions, all the results of the paper Harmonics on homogeneous manifolds may be carried over to the present situation.

Von Neumann has shown that the following conditions on representations are equivalent: 1) bounded, 2) ^{to unitary,} equivalent $\wedge$ 3) almost periodic. Because of the completeness theorem for the elements of the unitary representations in the domain of a.p. functions the question as to the sufficiency of the unitary representations to distinguish group elements is reduced to the same question for a.p. functions. If $f(s_0) \neq f(t_0)$ then there is a unitary representation $U(s)$ such that $U(s_0) \neq U(t_0)$.

In the case of compact groups the a.p. functions coincide with the continuous function and we have the favorable situation that if $s_0 \neq t_0$ there exists a $U(s)$ such that $U(s_0) \neq U(t_0)$. On the other hand there are some important groups for which the only a.p. functions are the constants, and hence the only unitary representation is the identical representation. For these groups the theory is vacuous.

We discuss next the Haar measure for locally compact separable metric groups as presented by Haar in an article in the Annals of Mathematics, 34 (1933).

We recall that a space is metric if it has a distance function defined for pairs of points in it which satisfies the usual condition (symmetry, positive character, triangle inequality); it is separable if it contains a denumerable dense set of points; it is locally compact if every point has a neighborhood whose closure is compact.

Let K_n be a sequence of spheres with centers at 1 and radii tending toward 0 . If G is an arbitrary open set whose closure is compact (an o.c. set) and a a point of G , then the neighborhood aK_n "parallel" to K_n covers the point a .



We obtain in this way a covering of G by neighborhoods parallel to K_n , and because of the compactness a finite number of these are sufficient for the covering. Let $\lambda_n(G)$ be the smallest number of neighborhoods parallel to K_n which are sufficient to cover G . We choose a fixed o.c. set E as a unit and form $\frac{\lambda_n(G)}{\lambda_n(E)}$. This sequence of numbers can be shown to be bounded from above and below. This is of course not sufficient to insure the existence of a limit in the ordinary sense but it can be shown that a generalized limit $\text{Lim } \frac{\lambda_n(G)}{\lambda_n(E)} = \chi(G)$ may be assigned. From this point one pro-

ceeds as in the Lebesgue theory. If s is an arbitrary set, the exterior measure of s is the greatest lower bound of $\sum \chi(A_n)$ where A_n are o.c. and $\sum A_n \supset S$.

The Haar measure leads to a Haar integral. For the most part in the applications so far, the von Neumann mean and the Haar integral may be used interchangeably.

7. Abstract topological groups

In this section we shall systematize some of the elementary concepts of abstract topological groups. This was first done by O. Schreier in two papers in the *Abhandlungen Hamb. Sem.* 4, 5 (1925, 1927).

A topological group $\mathcal{G}$ is a Hausdorff space for which there is defined a composition for pairs of points which has the following continuity properties:

- I. If $a_n \rightarrow a$, $b_n \rightarrow b$, then $a_n b_n \rightarrow ab$
- II. If $a_n \rightarrow a$, then $b_n \rightarrow b^{-1}$.

Generally the "first denumerability axiom of Hausdorff" concerning the existence of a denumerable base for the neighborhoods of a single point is assumed. In addition we make the natural assumptions that if U is a neighborhood of $\underline{1}$, so is U^{-1} (the set of inverse of the elements of U) and if U is a neighborhood of a . One can then show without difficulty that I and II are equivalent with

- I' For every neighborhood V of $\underline{1}$ there exists a U of $\underline{1}$ such that $U^2 \subset V$
- II' For every V of $\underline{1}$ there exists a U of $\underline{1}$ such that $U^{-1} \subset V$.

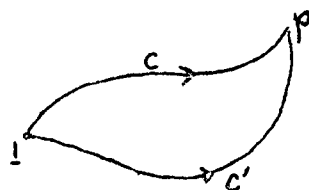
That aU is stated that $\mathcal{G}$ is topologically homogeneous, i.e. there is a (1-1) bi-continuous transformation which carries any point into any other.

We suppose that $\mathcal{G}$ is connected: Any two points a, b can be joined by an arc c ; c is the continuous image of an interval, i.e. c is given by a continuous function $p(t)$, $0 \leq t \leq 1$, $p(0) = a$, $p(1) = b$.

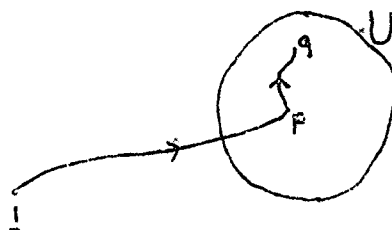
A closed curve ($p(0) = p(1)$) is said to be shrinkable to a point O if there exists a family $p(\lambda, t)$ of closed curves depending continuously on a real parameter λ such that $p(\lambda, t) = p(t)$ for $\lambda = 0$. We suppose that every sufficiently small closed curve is shrinkable to a point O for $\lambda = 1$.

closed curve is shrinkable to a point. A space having this property is called locally simply connected.

Let $\mathcal{G}$ and $\mathcal{G}'$ be two spaces having the topological properties just described (except possibly the homogeneity). Suppose that there is a continuous single valued correspondence of $\mathcal{G}'$ into $\mathcal{G}$ which is locally (1-1), i.e. if $P \rightarrow p$ then there are neighborhoods of P and p whose points are in (1-1) correspondence. Usually we are concerned with the case in which $\mathcal{G}$ is a manifold, a space whose points have neighborhood which are homeomorphic with the interior of a hyper-sphere. In this case we require also that $\mathcal{G}'$ is a manifold; $\mathcal{G}'$ shall have no boundaries, folds or ramification points. A $\mathcal{G}'$ satisfying this condition is called a covering space (c.s.) of $\mathcal{G}$. If $P \rightarrow p$, P is said to lie over p . Among the c.s. of there is a unique maximal c.s. (one which has no c.s. other than itself). This is called the universal covering space (u.c.s.) of $\mathcal{G}$. It is determined by the condition that a curve in it is closed if and only if the corresponding curve in $\mathcal{G}$ is shrinkable to a point. The u.c.s. $\mathcal{G}'$ can be constructed as follows: The points of $\mathcal{G}'$ over p are the arcs $c(p)$ issuing from a fixed point 1 to p where we consider $c(p) = c'(p)$ if $c = c'$ (or cc'^{-1}) is shrinkable to the unit point.

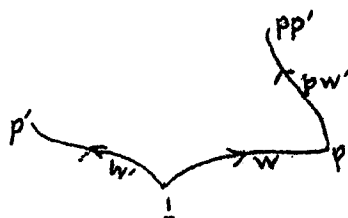


To obtain the neighborhoods of $c(p)$ we take a neighborhood U of p in $\mathcal{G}$ and a point q in it. The curves equivalent to the curves $c(q)$ consisting of $c(p)$ plus a curve in U joining p to q for a neighborhood of $c(p)$.



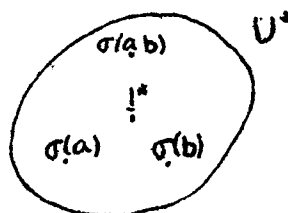
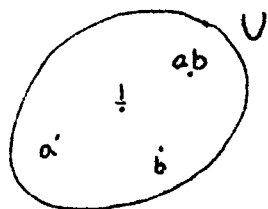
From the assumption that there exists a neighborhood of p such that every closed curve in it is shrinkable to a point, it follows that the correspondence $c(p) \rightarrow p$ is locally (1-1). From the construction we have that a path C in $\mathcal{G}$ is closed if and only if the corresponding path c in $\mathcal{g}$ is shrinkable to a point. Hence every closed curve in $\mathcal{G}$ is shrinkable to a point, or $\mathcal{G}$ is simply connected.

The u.c.s. $\mathcal{G}$ of a topological group $\mathcal{g}$ is also a group. Let w and w' be paths to the point p and p' respectively and let pw' denote the translation of w' by p . pw' has p and pp' for end points. We define $ww' = wpw'$ the path w extended by pw' .



It is necessary to show that if $w = w_1$ (ww_1^{-1} shrinkable to $\underline{1}$) and $w' = w'_1$, then $ww' = w_1w'_1$. This follows since $ww' (w_1w'_1)^{-1} = w(w'_1w_1^{-1})w_1^{-1} = ww_1^{-1} = \underline{1}$.

Two groups $\mathcal{g}$ and $\mathcal{g}^*$ are said to be locally simply isomorphic if there is a (1-1) correspondence σ between a neighborhood U of $\underline{1}$ in $\mathcal{g}$ and U^* in $\mathcal{g}^*$ such that if $a, b, ab \in U$, then $\sigma(ab) = \sigma(a)\sigma(b)$, and if $a^*, b^*, a^*b^* \in U^*$, then $\sigma^{-1}(a^*b^*) = \sigma^{-1}(a^*)\sigma^{-1}(b^*)$.



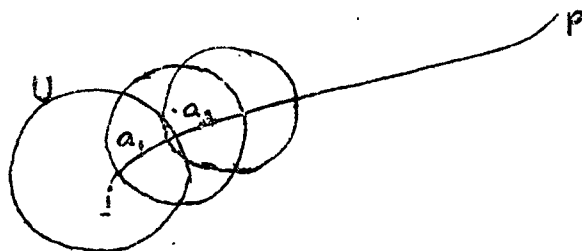
A group $\mathcal{g}$ and its u.c.g. $\mathcal{G}$ are locally simply isomorphic. $\mathcal{G}$ is multiply isomorphic in the large with $\mathcal{g}$. The elements of $\mathcal{G}$ which correspond to $\underline{1}$ in this isomorphism form a discrete invariant sub-group $\mathcal{F}$ of $\mathcal{G}$ which is simply isomorphic with the Poincare group of $\mathcal{g}$.

$\mathcal{F}$ is contained in the centrum of $\mathcal{G}$. We shall show in fact that this is true for any discrete invariant sub-group. For, otherwise there exist a p in such that $b = p^{-1}ap \neq a$. Join $\underline{1}$ and p by an arc. As s varies



along this arc, sas^{-1} varies along an arc which connects a and b and is contained in $\mathcal{F}$. This is impossible since $\mathcal{F}$ is discrete.

The simply connected group (or the u.c.g.) generated by a group germ is unique. We suppose that $\mathcal{G}$ and $\mathcal{G}^*$ are simply connected and locally simply isomorphic. We wish to show that $\mathcal{G}$ and $\mathcal{G}^*$ are simply isomorphic in the large. Let the local correspondence between $\mathcal{G}$ and $\mathcal{G}^*$ be given by $a \rightarrow \sigma(a) = a^*$. The proof is accomplished by "analytic continuation" of the function σ . Let p be any point in $\mathcal{G}$. Join it by an arc to $\underline{1}$ and determine a finite set of nested neighborhood of the form $a_i U$ as in the figure, i.e. $a_i \subset a_{i-1} U$. For a point $a_1 x$ in $a_1 U$ we define



$\sigma(a_1 x) = \sigma(a_1) \sigma(x)$. This determines $\sigma(a_2)$. Then for $a_2 x$ in $a_2 U$ we define $\sigma(a_2 x) = \sigma(a_2) \sigma(x)$ etc. Finally we obtain $\sigma(p)$. One shows as in the theory of analytic continuation that $\sigma(p)$ is independent of the path joining $\underline{1}$ to p . For sufficiently near paths have this independence and since any two paths can be deformed into each other σ is uniquely defined. We treat σ^{-1} in the same way.

σ and σ^{-1} are inverses over the whole of $\mathcal{G}$ and $\mathcal{G}^*$ and hence the extended correspondence is (1-1). Since $\sigma(ab) = \sigma(a) \sigma(b)$ we have an isomorphism.

Any group $\mathfrak{g}$ which is locally isomorphic with a simply connected group G is simply isomorphic in the large with a factor group of G by a discrete sub-group contained in the centrum.

In Chapter 2, in proving Lie's theorems we obtained as solutions of differential equations certain functions which may be composed in the manner of the elements of a group. However, the solutions are valid only for a restricted set of values of the parameters and the composition for a possibly still more restricted set. Abstractly we have two neighborhoods U and V of a space such that $U \supset V$ and

1. If $a, b \in V$, then $ab \in U$
2. If $a, b, c \in V$, and $(ab)c$, $a(bc)$ are defined, then $(ab)c = a(bc)$
3. $V \supset \underline{1}$
4. For every $a \in V$, $a^{-1} \in U$.

We call such a pair (U, V) a group germ. It is necessary to supplement the discussion of Chapter 2 by showing that any group germ generates a group in the large. An outline of a proof of this has been given by Cartan in his *Theorie des groupes finis* etc. (Memorial 42), page 18.

Chapter 5.

The Unimodular Unitary Group

1. Unimodular unitary group is simply connected

With respect to an appropriate coordinate system the unitary group $\check{W}_n$ of degree n may be defined as the group of n by n matrices a such that $aa' = 1$ (1 = unit matrix). We shall usually be concerned with the unimodular sub-group $\check{W}_n$ of $\check{W}_n$, i.e. the sub-group of elements of determinant unity.

Any unitary transformation is the conjugate of a diagonal unitary transformation: $a = u^{-1} \epsilon u$ where u, ϵ are unitary and

$$(1.1) \quad \epsilon = \begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \ddots & \\ & & & \epsilon_n \end{pmatrix} \quad |\epsilon_i| = 1$$

Let ϵ_1 be a root of the characteristic equation $|\lambda I - a| = 0$, and e_1 a vector such that $ae_1 = \epsilon_1 e_1$. From $aa' = 1$ follows $|\epsilon_1| = 1$. Suppose e_1 normalized so that its length $(e_1, e_1) = 1$. e_1 may then be supplemented by vectors $e_2 \dots e_n$ to form a unitary basis for the whole space. It follows that there is a unitary matrix u_1 such that

$$a = u_1^{-1} \begin{pmatrix} \epsilon_1 & \alpha_{12} & \alpha_{13} & \dots \\ 0 & & & \\ & a_1 & & \\ 0 & & & \end{pmatrix} u_1.$$

Since $|\epsilon_1| = 1$ and $|\epsilon_1|^2 + |\alpha_{12}|^2 + \dots + |\alpha_{1n}|^2 = 1$,

$\alpha_{12} = \alpha_{13} = \dots = \alpha_{1n} = 0$. Using induction we obtain the desired result,

The matrix ϵ is unique except for permutations of the roots ϵ_k . Let $\epsilon_k = e^{2\pi i \varphi_k} = e(\varphi_k)$ where the angles φ_k are real and determined mod 1 by ϵ_k . If $a = u^{-1} \epsilon u = v^{-1} \epsilon v$, then $v = cu$ where c is commutative with ϵ . If all the ϵ_k are different we call a regular, otherwise singular. A ma-

trix ϵ commutative with a regular ϵ must be a diagonal matrix.

The dimensionality of $\check{W}_n(W_n^\circ)$ is $n^2(n^2 - 1)$. The conditions that $a = (\alpha_{ik})$ be an infinitesimal unitary matrix are $\alpha_{ik} = -\bar{\alpha}_{ki}$. For $i = k$ this gives $\alpha_{ii} = -\bar{\alpha}_{ii}$ and hence there are n real parameters given by the diagonal terms. If $i \neq k$, $\alpha_{ki} = -\bar{\alpha}_{ik}$ determines the elements below the main diagonal by those above, and hence these terms give in all $\frac{n(n-1)}{2}$ complex or $n(n-1)$ real parameters. The total number of real parameters in $N = n^2$. The condition that a be infinitesimal unimodular is $\text{tr}(a) = \sum \alpha_{ii} = 0$. Hence the dimensionality of $\check{W}_n^\circ$ is $N - 1$.

The manifold of singular elements of $\check{W}_n(W_n^\circ)$ has dimensionality $N - 3(N - 4)$. Let $a = u^{-1}\epsilon u$ be singular. There are $n - 1$ real parameters φ_k involved in ϵ . We write $u \equiv u_1$ if $u = cu_1$ where c commutes with ϵ . By considering congruent elements as identical we obtain a projection space $[\check{W}_n]$. The dimensionality of the c 's is $n + 2$ since if $\epsilon_1 = \epsilon_2 \dots$ such matrix has the form

$$c = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_3 & \dots & \gamma_n \end{pmatrix}$$

Hence the dimensionality of $[\check{W}_n]$ is $N - (n + 2)$ and the dimensionality of the singular elements is $N - (n + 2) + n - 1 = N - 3$. It follows that any curve in $\check{W}_n$ or $\check{W}_n^\circ$ may be deformed continuously in such a way as to avoid the singular elements. We require this in proving

Theorem 1. $\check{W}_n^\circ$ is simply connected.

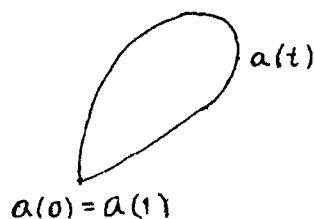
We must show that any closed curve defined by $a(t)$ $0 \leq t \leq 1$

$a(0) = a(1)$ can be shrunk to the point 1.

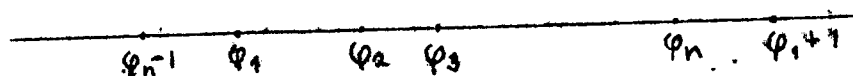
We may take $a(0) = \epsilon_0$ diagonal. Suppose

that $a(t) = u^{-1}(t)\epsilon(t)u(t)$ and that $a(t)$

is regular for all t . Since the roots



$\epsilon_k = e(\varphi_k)$ vary continuously ϵ is continuous in t . We may take $u(t)$ also to be continuous in t . We plot the possible values of the periodic φ_k at $a(0)$ on a line labeling them as in the figure



i.e. $\varphi_1 < \varphi_2 < \dots < \varphi_n < \varphi_1 + 1$. We call $\varphi_1, \varphi_2, \dots, \varphi_n$ a representative set. Since $\det(a) = 1$, $\varphi_1 + \varphi_2 + \dots + \varphi_n = k$ an integer. If we choose the representative set $\varphi'_1, \varphi'_2, \dots, \varphi'_n = \varphi_2, \varphi_3, \dots, \varphi_n, \varphi_1 + 1$, we have $\varphi'_1 + \varphi'_2 + \dots + \varphi'_n = k + 1$. Then by shifting to a suitable sequence on the graph we may suppose that $\varphi_1 + \varphi_2 + \dots + \varphi_n = 0$. $\varphi_1 < \varphi_2 < \dots < \varphi_n < \varphi_1 + 1$ at $a(0)$. Since the φ_i vary continuously with t and none of them become equal for any value of t , we have for all t , $\varphi_1 < \varphi_2 < \dots < \varphi_n < \varphi_1 + 1$. $\varphi_1 + \varphi_2 + \dots + \varphi_n$ is a continuous function of t having only integral values and initial value (for $t = 0$) 0. Hence for all t , $\varphi_1 + \varphi_2 + \dots + \varphi_n = 0$. For $t = 1$ we return to the only 0 representative set and $\varphi_i(0) = \varphi_i(1)$. For all λ , $0 \leq \lambda \leq 1$ we set

$$\epsilon_\lambda(t) = \begin{pmatrix} e(\lambda\varphi_1(t)) & & \\ & \ddots & \\ & & e(\lambda\varphi_n(t)) \end{pmatrix}$$

The $\epsilon_\lambda(t)$ is unitary and unimodular. $a_\lambda(t) = u(t)^{-1} \epsilon_\lambda(t) u(t)$ defines a closed curve in $\check{W}_n^\circ$ which avoids the singular elements. For $\lambda = 1$, $a_\lambda(t) = a(t)$ and for $\lambda = 0$, $a_\lambda(t) = 1$. As $\lambda \rightarrow 0$ we obtain a continuous deformation of the curve $a(t)$ into the point 1 of $\check{W}_n^\circ$.

2. Primitive characters of $\check{W}_n$

The machinery developed in Chapter 4 is powerful enough to enable us to determine the primitive characters of $\check{W}_n$ without actually finding the irreducible representations.

Let $a = u^{-1} \epsilon u$ where ϵ is given by (1.1).

(1) $\chi(a) = \chi(\epsilon) = \chi(\varphi_1 \dots \varphi_n)$ is a continuous periodic function of the φ_i of period 1 and is symmetric in the φ_i .

(2) The diagonal elements such as ϵ form an abelian subgroup of $\check{W}_n$. Suppose that in a given representation of degree m

$$\epsilon \rightarrow E = \begin{pmatrix} \epsilon_1^x & & \\ & \epsilon_2^x & \\ & & \ddots \\ & & & \epsilon_m^x \end{pmatrix}$$

where the root $\epsilon^* = \epsilon^*(\varphi_1 \dots \varphi_n)$, $|\epsilon^*| = 1$ and

$$\epsilon^*(\varphi_1 + \varphi'_1, \dots, \varphi_n + \varphi'_n) = \epsilon^*(\varphi_1, \dots, \varphi_n) \epsilon^*(\varphi'_1, \dots, \varphi'_n).$$

Hence $\epsilon^*(\varphi_1, \varphi_2, \dots, \varphi_n) = \epsilon^*(\varphi_1, 0, \dots, 0) \epsilon^*(0, \varphi_2, 0, \dots, 0)$

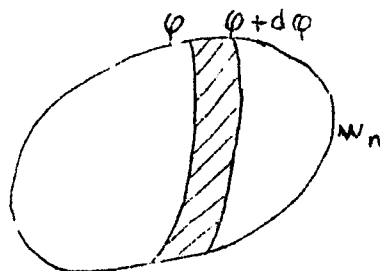
$$\dots \epsilon^*(0, 0, \dots, \varphi_n) = \epsilon_1^*(\varphi_1) \epsilon_2^*(\varphi_2) \dots \epsilon_n^*(\varphi_n)$$

since $\epsilon_k^*(\varphi_k + \varphi'_k) = \epsilon_k^*(\varphi_k) \epsilon_k^*(\varphi'_k)$ by sec. 5 Chapter 4,

$$\epsilon_k(\varphi_k) = e^{2\pi i m_k \varphi_k} = e(m_k \varphi_k) \text{ and } \epsilon^* = e(m_1 \varphi_1 + m_2 \varphi_2 + \dots + m_n \varphi_n).$$

Hence $\chi(\varphi_1, \varphi_2, \dots, \varphi_n)$ is a finite Fourier series with positive integral coefficients.

(3) We apply next the orthogonality relations making use for the first time of the irreducibility of the representation. We have $\int \chi(a) \bar{\chi}(a) da = V$ the volume of $\check{W}_n$. Let da be the volume of the strip of $\check{W}_n$ obtained by letting the angles φ_i vary from φ_i to $\varphi_i + d\varphi_i$. We shall show in the next section



that

$$da = \Delta \bar{\Delta} d\varphi_1 d\varphi_2 \dots d\varphi_n$$

where Δ is the Vandermonde determinant

$$\Delta = |\epsilon^{n-1}, \epsilon^{n-2}, \dots, 1| = \begin{vmatrix} \epsilon_1^{n-1} & \epsilon_1^{n-2} & \dots & 1 \\ \epsilon_2^{n-1} & \epsilon_2^{n-2} & \dots & 1 \\ \dots & \dots & \dots & \dots \\ \epsilon_n^{n-1} & \epsilon_n^{n-2} & \dots & 1 \end{vmatrix}$$

The orthogonality relations may be written as

$$(2.1) \quad \int_0^1 \dots \int_0^1 \chi(\varphi_1, \dots, \varphi_n) \bar{\chi}(\varphi_1, \dots, \varphi_n) \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n = \int_0^1 \dots \int_0^1 \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n$$

$\chi \Delta = \xi$ is skew symmetric in the φ_i since χ is symmetric and

Δ is skew symmetric. Let $\xi(\lambda_1, \lambda_2, \dots, \lambda_n) [\lambda_1 > \lambda_2 > \dots > \lambda_n]$ denote the

elementary skew symmetric polynomial $= \sum \pm e(\lambda_1 \varphi_1 + \lambda_2 \varphi_2 + \dots + \lambda_n \varphi_n)$ ob-

tained from $e(\lambda_1 \varphi_1 + \lambda_2 \varphi_2 + \dots + \lambda_n \varphi_n)$ by performing all the permutations

of the φ 's and putting + or - in front of the term according as the permutation

is even or odd. Thus, for example $\Delta = \xi(n-1, n-2, \dots, 1, 0)$. We order

the ξ 's lexicographically, i.e. $\xi(\lambda_1, \dots, \lambda_n)$ precedes (has higher weight)

$\xi(\lambda'_1, \dots, \lambda'_n)$ if the first of the differences $\lambda_{i_1} - \lambda'_{i_1} \neq 0$ is > 0 and

$\xi(\lambda_1, \dots, \lambda_n) \neq \xi(\lambda'_1, \dots, \lambda'_n)$. In a similar manner we order the sym-

metric polynomials. Let the highest term occurring in χ have the weight

$m_1 \varphi_1 + \dots + m_n \varphi_n$ where $m_1 \geq m_2 \geq \dots \geq m_n$. The highest term in ξ has the

weight $\lambda_1 \varphi_1 + \dots + \lambda_n \varphi_n$ where $\lambda_1 = m_1 + (n-1)$, $\lambda_2 = m_2 + (n-2)$, $\dots$, $\lambda_n = m_n$

and consequently $\lambda_1 > \lambda_2 > \dots > \lambda_n$. If the coefficient of the highest term in

ξ is the integer c , then $\xi - c \xi(\lambda_1, \lambda_2, \dots, \lambda_n)$ has lower weight than ξ .

By repeating this process we obtain an expansion for

$$\xi = c \xi(\lambda_1, \dots, \lambda_n) + c' \xi(\lambda'_1, \dots, \lambda'_n) + \dots \quad c \neq 0,$$

We may verify readily that these elementary polynomials are orthogonal:

$$(2.2) \quad \int_0^1 \dots \int_0^1 \xi(\lambda'_1, \dots, \lambda'_n) \xi(\lambda_1, \dots, \lambda_n) d\varphi_1 \dots d\varphi_n = \begin{cases} n! & \text{if } \lambda_i = \lambda'_i \\ 0 & \text{otherwise} \end{cases}$$

In particular $\int_0^1 \dots \int_0^1 \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n = n!$ and (2.1) becomes

$$c\bar{c} + c'\bar{c}' + \dots = 1$$

Hence $c = 1$, $c' = c'' = \dots = 0$, or

$$(2.3) \quad \chi = \frac{\sum (\chi_1, \dots, \chi_n)}{\sum (n-1, \dots, 1, 0)}.$$

(4) Every function of the form (2.3) occurs as a character. For if not we would have a class function on $\tilde{W}_n$ which is orthogonal to all the characters by (2.2). This contradicts the completeness of the primitive characters.

Applications. The above calculations show that for any set of integers $m_1 \geq m_2 \geq \dots \geq m_n$ we have an irreducible representation $D(m_1 m_2 \dots m_n)$ whose character is

$$(2.4) \quad \chi_{m_1 \dots m_n} = \frac{|e^{\chi_1}, \dots, e^{\chi_n}|}{|e^{n-1}, \dots, 1|} = \frac{\sum \pm o(\chi_1 \varphi_1 + \dots + \chi_n \varphi_n)}{\sum \pm o((n-1)\varphi_1 + \dots + 0\varphi_n)}$$

where $\chi_i = m_i + (n-i)$. The degree of this representation is the value of χ at the unit point, i.e. for $\varphi_i = 0$. (2.4) is indeterminate for $\varphi_i = 0$. To evaluate its limit we set

$$\varphi_1 = (n-1)\varphi = \chi_1^0 \varphi, \quad \varphi_2 = (n-2)\varphi = \chi_2^0 \varphi, \quad \dots \quad \varphi_n = 0 = \chi_n^0 \varphi$$

Then

$$\sum \pm o(\chi_1 \varphi_1 + \dots + \chi_n \varphi_n) = \sum \pm o((\chi_1 \chi_1^0 + \dots + \chi_n \chi_n^0) \varphi)$$

or if we denote $\chi_k \varphi$ by ψ_k

$$\begin{aligned} \sum \pm o(\chi_1 \varphi_1 + \dots + \chi_n \varphi_n) &= |o((n-1)\psi_1), o(n-2)\psi_2, \dots, 1| \\ &= \prod_{i < k} o(\psi_i) - o(\psi_k) \end{aligned}$$

For small φ , $o(\varphi) \sim 2\pi i \varphi$, $o(\chi_i \varphi) - o(\chi_k \varphi) \sim 2\pi i \varphi (\chi_i - \chi_k)$ and hence

$$\chi \sim \frac{(2\pi i \varphi)^{\frac{n(n-1)}{2}} \prod_{i < k} (\chi_i - \chi_k)}{(2\pi i \varphi)^{\frac{n(n-1)}{2}} \prod_{i < k} (\chi_i^0 - \chi_k^0)}$$

or, the degree of the representation

$$N = \frac{\Delta(\chi_1, \dots, \chi_n)}{\Delta(n-1, \dots, 0)}$$

The form of χ in (2.4), though convenient, involves the elements of

the variable element s of $\check{W}_n$ implicitly. Since χ is symmetric in the ϵ 's, it is integral in the elements of s . We proceed to obtain the form of χ as a function of those elements. We use the Cauchy identity

$$\det \left(\frac{1}{1 - x_i y_k} \right) = \frac{\Delta(x_1, \dots, x_n) \Delta(y_1, \dots, y_n)}{\prod (1 - x_i y_k)}$$

for $x_i = \epsilon_i$ and $y_k = \lambda_k$:

$$(2.5) \quad \det \left(\frac{1}{1 - \epsilon_i \lambda_k} \right) = \frac{|\epsilon^{n-1}, \dots, 1| |\lambda^{n-1}, \dots, 1|}{f(\lambda_1) f(\lambda_2) \dots f(\lambda_n)}$$

where $f(\lambda) = \prod (1 - \epsilon_i \lambda) = |1 - \lambda s| = 1 + c_1 \lambda + \dots + c_n \lambda^n$. If we substitute

$$\frac{1}{f(\lambda)} = p_0 + p_1 \lambda + p_2 \lambda^2 + \dots \quad (p_0 = 1)$$

$$\frac{1}{1 - \epsilon \lambda} = 1 + \epsilon \lambda + \epsilon^2 \lambda^2 + \dots$$

in (2.5) and compare the coefficients of $\lambda_1^{l_1} \lambda_2^{l_2} \dots \lambda_n^{l_n}$ on both sides, we obtain

$$|\epsilon_1^{l_1} \epsilon_2^{l_2} \dots \epsilon_n^{l_n}| = \frac{|\epsilon^{n-1}, \dots, 1|}{|p_{l_1-(n-1)}, \dots, p_{l_n}|}$$

and hence

$$\chi_{m_1 \dots m_n}(s) = |p_{l_1-(n-1)}, \dots, p_{l_n}|$$

3. Computation of volume on $\check{W}_n$

We apply the definition given in Chapter 4, sec. 1, of volume on a compact Lie group to the special case of the unitary group. Let ds be a line element or vector from a point s of $\check{W}_n$ to a neighboring point $s + ds$. n^2 independent vectors at s determine a volume element whose volume is defined as the absolute value of the determinant of the components of the translated vector $(ds)s^{-1} = \delta s$.

The components of δs satisfy the condition: $\delta s_{ik} = -\overline{\delta s_{ki}}$. Hence for real coordinates of the infinitesimal elements we may take $\frac{i}{\sqrt{-1}} \delta s_{ii}$ and

$\Re \delta_{s_{ik}}, \Im \delta_{s_{ik}} (i < k)$. However it is more convenient to use as coordinates $\delta_{s_{ii}}, \delta_{s_{ik}},$ and $\delta_{s_{ki}} = -\overline{\delta_{s_{ik}}} (i < k)$ which are related to the former by

$$\delta_{s_{ik}} = \Re \delta_{s_{ik}} + \sqrt{-1} \Im \delta_{s_{ik}} \quad s_{ki} = -\Re \delta_{s_{ik}} + \sqrt{-1} \Im \delta_{s_{ik}}$$

We have seen that a unitary transformation s is similar to a diagonal unitary transformation

$$(3.1) \quad s = u^{-1} \epsilon u$$

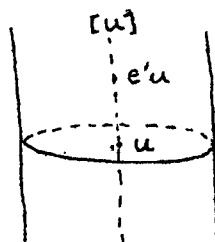
$$\epsilon = \begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \ddots & \\ & & & \epsilon_n \end{pmatrix}$$

ϵ is determined by s except for a permutation of the roots $\epsilon_1, \epsilon_2, \dots, \epsilon_n$.

For a fixed ϵ , u is determined but for a left multiplier belonging to the abelian group of diagonal elements. (We assume that the ϵ 's are all distinct.)

If $u' = \epsilon' u$, ϵ' diagonal the u and u' are said to lie on the same vertical $[u]$.

The projection space of verticals $[u]$ will be denoted by $\check{W}$



Thus any two u 's for which (3.1) holds for a fixed s and ϵ belong to the vertical :

If $[u]$ runs through the whole manifold $\check{W}$ and the φ_i range independently from 0 to 2π , then s describes $\check{W}_n$ $n!$ times.

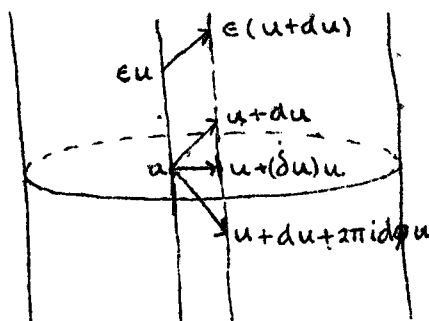
Let $[u + du]$ be a vertical neighboring $[u]$. The points on it which are near to u are of the form

$$(1 + 2\pi i d\varphi)(u + du)$$

where

$$d\varphi = \begin{pmatrix} d\varphi_1 & & \\ & d\varphi_2 & \\ & & \ddots \\ & & & d\varphi_n \end{pmatrix} \quad d\varphi_i: \text{real}$$

The line element determined by this point together with du is $du + 2\pi i(d\varphi)u$:

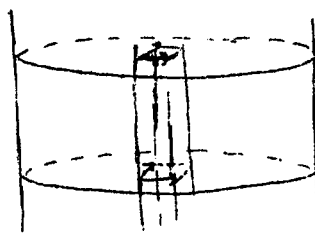


The vector at the identity which is "right parallel" to $du + 2\pi i(d\varphi)u$ is $\delta u + 2\pi i d\varphi$. Thus $d\varphi$ (or the $d\varphi_i$) may be determined uniquely so that $\delta u + 2\pi i d\varphi$ has only 0's on the main diagonal. We denote this particular determination of $u + 2\pi i d\varphi$ by $\dot{\delta}u$ and call $(\dot{\delta}u)u$ the horizontal transition from u on $[u]$ to $[u + du]$ (see figure). The horizontal transition from eu is determined by

$$(3.2) \quad \dot{\delta}'u = e(\dot{\delta}u)e^{-1}.$$

We choose $n^2 - n$ independent horizontal vectors issuing from u .

These determine a parallelogram whose volume is defined to be the absolute value of the components δu_{ik} ($i \neq k$) of the vectors $\dot{\delta}u$. Consider the tube of verticals through this parallelogram. Any horizontal cross-section of it is a par-



allogram whose volume is the same as that of the original parallelogram and the transformation $\dot{\delta}u \rightarrow \dot{\delta}'u$ in (3.2) is unimodular. The volume of one of these parallelograms may be taken as the volume of the tube, the volume element in $[\check{W}]$.

From (3.1) we have

$$(du)s + u(ds) = (d\epsilon)u + \epsilon(du)$$

Multiplying on the right by $(us)^{-1} = (\epsilon u)^{-1}$ we obtain

$$\delta u + u(\delta s)u^{-1} = 2\pi i(d\varphi) + \epsilon(\delta u)\epsilon^{-1}$$

or

$$(3.3) \quad \delta's = u(\delta s)u^{-1} = \{\epsilon(\dot{\delta}u)\epsilon^{-1} - \dot{\delta}u\} + 2\pi i d\varphi$$

where

$$(d\epsilon)\epsilon^{-1} = 2\pi i(d\varphi) = 2\pi i \begin{pmatrix} d\varphi_1 & & \\ & d\varphi_2 & \\ & & \ddots \\ & & & d\varphi_n \end{pmatrix}$$

$$\dot{\delta}u = \delta u + 2\pi i d\varphi$$

We determine a volume element in s : Let $n^2 - n$ of the sides ds be such that the corresponding $d\varphi$ in (3.3) are 0 and $\dot{\delta}u$ give a horizontal transition from u . The components $\delta's$ are then

$$\delta u_{ik} \left(\frac{\epsilon_i}{\epsilon_k} - 1 \right) \quad (i \neq k)$$

and the remaining ones 0. The other n sides of the volume element are chosen so that a single angle $d\varphi_k \neq 0$ ($k = 1, 2, \dots, n$) and $[u]$ remains unchanged. The volume of the parallelepiped is

$$\prod_{i \neq k} \left(\frac{\epsilon_i}{\epsilon_k} - 1 \right) |du| d\varphi_1 d\varphi_2 \dots d\varphi_n$$

where $|du|$ is the magnitude of the volume element of $[\check{W}]$. If we integrate over $[\check{W}]$ and normalize so that the volume of $[\check{W}]$ is 1, we obtain

$$\prod_{i \neq k} \left(\frac{\epsilon_i}{\epsilon_k} - 1 \right) d\varphi_1 d\varphi_2 \dots d\varphi_n$$

as the volume of the strip of u in which the angles vary from φ_k to $\varphi_k + d\varphi_k$.

Since $\bar{\epsilon}_i = \frac{1}{\epsilon_i}$ this expression may be rewritten as

$$\Delta \Delta d\varphi_1 d\varphi_2 \dots d\varphi_n \quad \Delta = \prod_{i < k} (\epsilon_k - \epsilon_i)$$

4. Additional applications of the determination of the characters of $\check{W}_n$

The problem of representations from the infinitesimal viewpoint is a purely algebraic one. For the case of the unitary group $\check{W}_n$ it is the following: to associate with every transformation of degree n , $a = (a_{ik})$, $a_{ik} = -\bar{a}_{ki}$, a transformation A of degree N such that the association is linear and $[ab] \rightarrow [AB]$. The restriction to the unimodular sub-group $\check{W}_n^\circ$ means that only the matrices a such that $\text{tr}(a) = 0$ are considered. In this case $\check{W}_n^\circ$ in the large is simply connected and every representation in the large can be obtained by iteration of a representation of the infinitesimal group. If we allow the real parameters which enter into the infinitesimal group of $\check{W}_n$ to have complex values, we obtain the infinitesimal linear group. From this it follows that the theory of the representations of the linear group can to a large extent be obtained from that of the representations of $\check{W}_n$.

We may construct in a simple way certain representations of $\check{W}_n$. Let $\mathcal{X} = (x_1 \dots x_n)$ and $\mathcal{Y} = (y_1 y_2 \dots y_n)$ be arbitrary vectors in the n -space $\mathcal{R}$ in which $\check{W}_n$ acts. The two-dimensional elements $[\mathcal{X}, \mathcal{Y}]$ whose components are $(x_i y_k - x_k y_i)$ form a space of $\binom{n}{2}$ dimensions $\mathcal{R}^{(2)}$. The transformation u of $\check{W}_n$ induces a transformation $u^{(2)}$ in $\mathcal{R}^{(2)}$ and the correspondence $u \rightarrow u^{(2)}$ is a representation. Similarly the three-dimensional elements

$$\begin{vmatrix} x_i & y_i & z_i \\ x_k & y_k & z_k \\ x_l & y_l & z_l \end{vmatrix}$$

give a representation of degree $\binom{n}{3}$ etc. If $[n_i, n_k]$ ($i \neq k$) form a basis

in $\mathcal{R}^{(2)}$ and $d n_i = a n_i$ is an infinitesimal element of $\check{W}_n$, then

$$(4.1) \quad d[n_i, n_k] = [d n_i, n_k] + [n_i, d n_k] \\ = [a n_i, n_k] + [n_i, a n_k]$$

is the corresponding infinitesimal element in $\mathcal{R}^{(2)}$.

We shall show that the representation of $\check{W}_n$ in $\mathcal{R}^{(2)}$ is irreducible.

To $d n_i = \lambda_i n_i$ corresponds by (4.1)

$$(4.2) \quad d[n_i, n_k] = (\lambda_i + \lambda_k) [n_i, n_k]$$

Now suppose $\sum J_{ik} [n_i, n_k] \neq 0$ is in an invariant sub-space $\mathcal{R}_0^{(2)}$ of $\mathcal{R}^{(2)}$.

According to (4.2) we have the vector

$$J_{ik} (\lambda_i + \lambda_k)^s [n_i, n_k]$$

also in $\mathcal{R}_0^{(2)}$. Choose λ_i so that the $\binom{n}{2}$ number $\lambda_i + \lambda_k$ are distinct. The Vandermonde determinant

$$V = \begin{vmatrix} 1 & 1 & \dots & 1 \\ (\lambda_1 + \lambda_2) & (\lambda_1 + \lambda_3) & \dots & (\lambda_1 + \lambda_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda_1 + \lambda_2)^{\binom{n-1}{2}-1} & (\lambda_1 + \lambda_3)^{\binom{n-1}{2}-1} & \dots & (\lambda_1 + \lambda_n)^{\binom{n-1}{2}-1} \end{vmatrix}$$

is not 0 and hence each $J_{ik} [n_i, n_k]$ belongs to $\mathcal{R}_0^{(2)}$. At least one J , say $J_{12} \neq 0$ and so $[n_1, n_2] \in \mathcal{R}_0^{(2)}$. Suppose $a n_j = \sum_k \alpha_{kj} n_k$ then $[a n_1, n_2] + [n_1, a n_2]$ is in $\mathcal{R}_0^{(2)}$. From this it follows that all $[n_k, n_j] \in \mathcal{R}_0^{(2)}$ and hence $\mathcal{R}_0^{(2)} = \mathcal{R}^{(2)}$. A similar argument may be used to show the irreducibility of $\mathcal{R}^{(3)}$, $\mathcal{R}^{(4)}$, ... By direct multiplication of these representations all irreducible representations of u_n can be obtained.

For the case $\check{W}_2$ the characters are

$$\chi = \frac{\begin{vmatrix} e_1^{l_1} & e_1^{l_2} \\ e_2^{l_1} & e_2^{l_2} \end{vmatrix}}{e_1 - e_2} = \frac{\begin{vmatrix} e_1^{j+1} & 1 \\ e_2^{j+1} & 1 \end{vmatrix}}{e_1 - e_2} \quad (j+1 = l_1 - l_2) \\ = e_1^j + e_1^{j-1} e_2 + \dots + e_2^j$$

Thus there is a single representation D_j of degree j . If

$$(4.3) \quad x' = \alpha x + \beta y \quad y' = \gamma x + \delta y$$

$\in \check{W}_2$, then the transformation induced by (4.3) on the monomials $\frac{x^a y^b}{a! b!}$ ($a + b = j$) defines a j -dimensional unitary representation. It can be shown by the methods just used that this representation is irreducible. Hence all irreducible representations of $\check{W}_2$ are determined in this way.

The question of what irreducible representations occur in the direct product may be answered by means of the characters. The method has been discussed in sec. 4, Chapter 4. For the case $n = 2$ we obtain easily

$$D_j \times D_{j'} = D_{j+j'} + D_{j+j'-2} + \dots + D_{|j-j'|}$$

If we restrict the representation $D(m_1, m_2, \dots, m_n)$ of $\check{W}_n$ to the subgroup of transformations of the form

$$\begin{pmatrix} & & 0 \\ & & \vdots \\ & u_{n-1} & \\ \hline 0 & \dots & 0 & 1 \end{pmatrix}$$

we obtain a representation $\tilde{D}(m_1, m_2, \dots, m_n)$ of $\check{W}_{n-1}$. We may determine by means of the characters the irreducible representations $D(m'_1, m'_2, \dots, m'_{n-1})$ which occur in $\tilde{D}$. It is necessary to expand the character

$$\tilde{\chi} = \frac{|e^{l_1}, \dots, e^{l_n}|}{|e^{n-1}, \dots, 1|} \quad \text{with } e_n = 1$$

in terms of the primitive characters $\chi_{m'_1 m'_2 \dots m'_{n-1}}$

$$\begin{aligned} \tilde{\chi} \prod_{\substack{i < k = \\ 1, \dots, n-1}} (e_i - e_k) &= \frac{1}{(e_1 - 1)(e_2 - 1) \dots (e_{n-1} - 1)} \begin{vmatrix} e_1^{l_1} - e_1^{l_2} & e_1^{l_2} - e_1^{l_3} & \dots & e_1^{l_{n-1}} - 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ e_{n-1}^{l_1} - e_{n-1}^{l_2} & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{vmatrix} \\ &= \sum \begin{vmatrix} e_1^{l'_1} & e_1^{l'_2} & \dots & e_1^{l'_{n-1}} \\ \vdots & \vdots & & \vdots \\ e_{n-1}^{l'_1} & e_{n-1}^{l'_2} & \dots & e_{n-1}^{l'_{n-1}} \end{vmatrix} \end{aligned}$$

where $\sum$ is extended over all $\lambda'_1, \lambda'_2, \dots$ such that

$$\lambda'_1 > \lambda'_1 \geq \lambda'_2 > \lambda'_2 \geq \lambda'_3 \dots > \lambda'_{n-1} \geq \lambda'_n.$$

Hence if

$$\begin{aligned} \lambda'_1 &= m'_1 + (n-2) \dots \lambda'_{n-1} = m'_{n-1} \\ m_1 &\geq m'_1 \geq m_2 \geq m'_2 \dots \end{aligned}$$

and

$$\widetilde{D}(m_1 \dots m_n) = \sum D(m'_1 \dots m'_{n-1})$$

summed over all m'_i for which the above inequalities hold.

The Institute for Advanced Study

THE STRUCTURE AND REPRESENTATION OF CONTINUOUS GROUPS

by

Hermann Weyl

1934-35

Notes by Dr. Richard Brauer

TABLE OF CONTENTS

	Page
SURVEY OF LAST TERM'S COURSE ON CONTINUOUS GROUPS -----	1
CHAPTER 1. FULL LINEAR, COMPLEX AND ORTHOGONAL GROUPS FROM THE INFINI- TESIMAL STANDPOINT -----	11
1. Structure of the full linear group -----	11
2. The weights of a representation -----	12
3. The first main theorem -----	16
4. The second main theorem -----	18
5. Construction of the irreducible representations -----	21
6. Young's symmetry operators -----	26
7. Full reducibility -----	32
8. Transcendental determination of the characters of the irreduci- ble representations of the full linear group -----	34
9. Structure of the complex group -----	35
10. The unitary restriction. The topology of the unitary complex group -----	38
11. Construction of the irreducible representations of the complex- group -----	43
12. Algebraic supplements to the above construction -----	46
13. Primitive characters of the complex group -----	54
14. Orthogonal group -----	63
Literature and Related Topics -----	67
CHAPTER 2. STRUCTURE OF SEMISIMPLE GROUPS -----	69
1. The maximal nilpotent subgroup of an infinitesimal group -----	69
2. Representations. Weights -----	74
3. Cartan's criterion for semisimple groups -----	80
4. The finite group (S) and the metric of the λ -space -----	89

CHAPTER 2 (Continued)	Page
5. Determination of the infinitesimal group by the vector diagram	94
6. The unitary restriction -----	100
7. Automorphisms of an infinitesimal group -----	102
8. Adjoint group and true group -----	107
9. The characters of the irreducible representations -----	113
CHAPTER 3. SEMINAR ON VARIOUS TOPICS IN GROUP THEORY -----	121
1. Connection between the Betti numbers of the manifold of a semi-simple group and the number of integral-invariants -----	121
2. Construction of the invariant integrals for the manifold of a compact simple group -----	132
3. Direct determination of the Betti numbers of the group manifolds of the compact simple groups -----	148
4. The fundamental region of finite groups generated by reflections	155
5. Direct characterization of the finite group (S) associated with a semi-simple continuous group -----	164
6. Discrete groups generated by reflections (Notes by Dr. H.S.M.Coxeter) -----	186

THE STRUCTURE AND REPRESENTATION OF CONTINUOUS GROUPS

by

Hermann Weyl

1934-35

Survey of Last Term's Course on Continuous Groups

Chapter 1. Fundamental concepts of groups and representations*

We assume that the reader is familiar with the concept of an abstract group. We speak of a realization of an abstract group $\mathcal{G}$ as a group of transformations of a given point field Π , if to every element g of $\mathcal{G}$ corresponds a transformation of Π (i.e. a one-to-one mapping of Π into itself) in such a way that to the product of elements g_1 and g_2 of $\mathcal{G}$ corresponds the product of the associated transformations. A realization is called distinct if to different group elements there always corresponds different transformations.

There arises the question whether every group has distinct realizations. This can be answered affirmatively in the following way: Let Π be the set of elements of $\mathcal{G}$ itself. If a is a given element of $\mathcal{G}$, we associate the transformation $s \rightarrow as$ of $\mathcal{G}$ into itself where a ranges over the elements of $\mathcal{G}$. It is easily seen that this is a distinct realization. We call this transformation $s \rightarrow as$ the (left-hand) translation of $\mathcal{G}$ by a . Similarly, a right-hand-translation is a transformation $s \rightarrow sa$ of $\mathcal{G}$ into itself.

There is another way to obtain a realization for which the point field is identical with $\mathcal{G}$. We associate to a in $\mathcal{G}$ the transformation $s \rightarrow asa^{-1}$, the so-called conjugation by a . This realization is not always distinct, since the elements a of the centrum (i.e. the elements commutative with every element of $\mathcal{G}$) give the identity transformation $s \rightarrow s$.

* Page numbering in brackets refers to the notes of last term's course.

A realization of $\mathcal{G}$ by linear transformations of a vector space $\mathcal{R}$ is called a representation of $\mathcal{G}$. The dimension of $\mathcal{R}$ is called the degree of the representation.

To every element s of $\mathcal{G}$ corresponds here a matrix $U(s)$ of degree n . If we change the coordinate system by a linear transformation with the matrix A , the matrix $U(s)$ is changed into $AU(s)A^{-1}$ and naturally these matrices form a new representation which is said to be equivalent to the former representation.

A representation is reducible if there is a proper subspace $\mathcal{R}_1$ of $\mathcal{R}$ which is invariant under its linear transformations. If n_1 is the dimension of $\mathcal{R}_1$ and the coordinate-system in $\mathcal{R}$ is chosen in such a way that the first n_1 basis-vectors belong to $\mathcal{R}_1$ the matrices $U(s)$ have the form

$$U(s) = \begin{pmatrix} A_1(s) & B(s) \\ 0 & A_2(s) \end{pmatrix} \begin{matrix} n_1 \text{ rows} \\ n - n_1 \text{ rows} \end{matrix}$$

$$\underbrace{\hspace{10em}}_{n_1 \text{ columns}} \quad \underbrace{\hspace{10em}}_{n-n_1 \text{ columns}}$$

Here $A_1(s)$ gives the transformation of $\mathcal{R}_1$ under $U(s)$ and $A_2(s)$ gives the transformation of the projection space $\mathcal{R} \pmod{\mathcal{R}_1}$ under $U(s)$. $A_1(s)$ and $A_2(s)$ are representations of $\mathcal{G}$ themselves.

If it is possible to find a second subspace $\mathcal{R}_2$ which is invariant under $U(s)$ such that $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$ (every vector of $\mathcal{R}$ can be written uniquely as the sum of a vector of $\mathcal{R}_1$ and of a vector of $\mathcal{R}_2$), the matrices $U(s)$ in a properly chosen coordinate-system will have the form

$$(1) \quad U(s) = \begin{pmatrix} A_1(s) & 0 \\ 0 & A_2(s) \end{pmatrix}$$

In this case the representation is said to be decomposable.

If two representations $A_1(s)$ and $A_2(s)$ of a group are given, the matrices (1) form a new representation, the sum of the given representations. If $A_1(s)$ is performed in a space with the coordinates $x_1, x_2, \dots, x_m$ and $A_2(s)$ in a space with the coordinates

$y_1, \dots, y_n$, the mn products $x_i y_k$ are transformed linearly by applying $A_1(s)$ on the x_i and $A_2(s)$ on the y_k . The matrix of this new transformation is denoted by $A_1(s) \times A_2(s)$. Obviously, we thus obtain a representation of degree mn , the product of $A_1(s)$ and $A_2(s)$.

The function $\chi(s)$ obtained by forming the trace of $U(s)$ for every s in $\mathcal{G}$ is called the character of the representation $U(s)$.

If all matrices $U(s)$ are unitary matrices and the representation is reducible, it is decomposable. For if $\mathcal{R}_1$ is an invariant subspace and $\mathcal{R}_2$ is the space of all vectors unitary orthogonal to every vector of $\mathcal{R}_1$, then $\mathcal{R}_2$ is also invariant and $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$.

Fundamental theorems:

I. Every representation $U(s)$ can be written in the form

$$U(s) = \begin{pmatrix} A_1 & B_{21} & \dots & B_{K1} \\ \sigma & A_2 & \dots & B_{K2} \\ \dots & \dots & \dots & \dots \\ \sigma & \sigma & \dots & A_K \end{pmatrix}$$

if the coordinate-system is properly chosen, where the A_k form an irreducible representation. These constituents of $U(s)$ are uniquely determined, if equivalent representations are considered as equal. [Proof pp. 12-16]

II. Theorems of Burnside and Frobenius-Schur: If $U(s) = \|u_{ik}(s)\|$ is an irreducible representation of a group, there exists no linear relation $\sum_{i,k} a_{ik} u_{ik}(s) = 0$ with coefficients a_{ik} which are independent of s and not all zero (Burnside). If $U_1(s) = \|u_{ik}^{(1)}(s)\|$, $U_2(s) = \|u_{ik}^{(2)}(s)\|$, ... are irreducible representations of degrees $n_1, n_2, \dots$, no two of which are equivalent, the $n_1^2 + n_2^2 + \dots$ coefficients $u_{ik}^{(1)}(s), u_{ik}^{(2)}(s), \dots$ are linearly independent in the analogous sense (Frobenius-Schur). [Proof pp. 21-24]

Whereas theorem I holds for representations by matrices $U(s)$ whose coefficients belong to any field, in the case of theorem II the field must be as-

sumed to be algebraically closed. From theorem II it follows that in the case of a non-modular and algebraically closed field two representations have the same character if and only if the constituents are the same in both representations [p. 24].

In the case of a finite group, every reducible representation is decomposable (except in the case where the characteristic of the field is different from zero and divides the order of the group) [pp. 27-31].

Chapter 2. Survey of Lie's Theory

Chapter 3. Foundations of a General Theory of Lie Algebras

We consider a vector space for which there is defined a multiplication $[xy]$ of any two vectors x and y , which is linear homogeneous in both factors and satisfies the conditions

$$[xy] = -[yx]; \quad [x[yz]] + [y[zx]] + [z[xy]] = 0.$$

Such a vector space is called a Lie algebra or an infinitesimal group, the dimension of the space is the order of the algebra.

A realization of a Lie algebra $\mathfrak{g}$ is given, if to every x of $\mathfrak{g}$ corresponds a velocity-field $U^i(\xi_1, \dots, \xi_m)$ in a fixed pointfield Π where $\xi_1, \dots, \xi_m$ are the coordinates in Π such that the following conditions are fulfilled: If $x \rightarrow U^i(\xi_1, \dots, \xi_m)$ and $y \rightarrow V^i(\xi_1, \dots, \xi_m)$ then

$$x + y \rightarrow U^i(\xi_1, \dots, \xi_m) + V^i(\xi_1, \dots, \xi_m)$$

$$\lambda x \rightarrow \lambda U^i(\xi_1, \dots, \xi_m) \quad (\lambda \text{ scalar})$$

$$[xy] \rightarrow \sum_k \left(\frac{\partial U^i}{\partial \xi_k} V^k - \frac{\partial V^i}{\partial \xi_k} U^k \right) = [UV]^i$$

Of course, assumptions on the differentiability of the occurring functions are necessary.

Every continuous group of transformations of a point field Π (with assumptions about differentiability of the occurring functions) gives rise to a

definite Lie algebra $\mathfrak{g}$ and a realization; the U^i are the infinitesimal elements of the group [Chapter 2].

A realization of a Lie algebra $\mathfrak{g}$ is said to be a representation, if Π is a vector space and the U^i depend linearly and homogeneously on the ξ^i 's. Then to every element a of $\mathfrak{g}$ corresponds a matrix A (non-singular or not), such that from $a \rightarrow A$, $b \rightarrow B$ follows $a + b \rightarrow A + B$, $\lambda a \rightarrow \lambda A$ (λ scalar),
 $[ab] \rightarrow AB - BA = [AB]$

We obtain a representation with $\Pi = \mathfrak{g}$ by associating to a the linear transformation $x \rightarrow [ax]$ of the vector space $\mathfrak{g}$, the so-called adjoint representation. In the case of a continuous group it can be obtained from the conjugation $s \rightarrow asa^{-1}$ by application on infinitesimal elements [p. 43].

If $a \rightarrow A$ and $a \rightarrow A'$ are two representations of a Lie algebra, the product of the two representations is here the representation
 $a \rightarrow (I_m \otimes A') + (A \otimes I_n)$ where I_m and I_n denote the unit matrices of degree m and n .

A subspace $\mathfrak{f}$ of a Lie algebra $\mathfrak{g}$ which is closed under the multiplication $[xy]$ is called a subgroup of $\mathfrak{g}$; $\mathfrak{f}$ is invariant subgroup if $\mathfrak{f}$ contains with x the element $[ax]$ for all a in $\mathfrak{g}$; i.e. if the subspace $\mathfrak{f}$ is invariant under all transformation of the adjoint representation. If $\mathfrak{f}$ is invariant and $x = x'(\mathfrak{f})$, $y = y'(\mathfrak{f}) \rightarrow$ i.e. if $x = x'$ and $y = y'$ belong to $\mathfrak{f}$ -- it follows that

$$x + y = x' + y', \quad \lambda x = \lambda x', \quad [xy] = [x'y'] \quad (\mathfrak{f}).$$

Therefore, the projection space of $\mathfrak{g}$ with respect to $\mathfrak{f}$ may be considered as a Lie algebra, which is called the factor group. The order is the difference of the orders of $\mathfrak{g}$ and of $\mathfrak{f}$.

An infinitesimal group in which $[ab] = 0$ holds for all a and b will be called an Abelian group. In any case, the elements a for which $[ax] = 0$ for all x in $\mathfrak{g}$, form an invariant Abelian subgroup, the centrum.

If $\mathfrak{f}$ is invariant in $\mathfrak{g}$, all elements $[gh]$, g in $\mathfrak{g}$, h in $\mathfrak{f}$, and all their linear combinations form a new invariant subgroup $\mathfrak{f}'$ of $\mathfrak{g}$, we have

$\mathfrak{f}' < \mathfrak{f}$ We denote $\mathfrak{g}'$ by $\mathfrak{g}^2$, $\mathfrak{g}''$ by $\mathfrak{g}^3$ and so on, $\mathfrak{g} = \mathfrak{g}^1$. $\mathfrak{g}$ is nilpotent, if $\mathfrak{g}^{\nu} = 0$ for a properly chosen ν , it is sufficient to set $\nu = r + 1$, r order of $\mathfrak{g}$.

The nilpotency of an infinitesimal group is the same as the nilpotency of the adjoint representation, where a set of matrices is called nilpotent, if there exists a positive integer N such that the product of any N matrices of the set vanishes [p. 61].

If in the series

$$\mathfrak{g} = \mathfrak{g}_0 > \mathfrak{g}_1 > \mathfrak{g}_2 > \dots > \mathfrak{g}_n = 0$$

every $\mathfrak{g}_i$ is an invariant subgroup of $\mathfrak{g}_{i-1}$, we speak of a composition series of $\mathfrak{g}$. Groups $\mathfrak{g}$ with composition series where all the factor groups $\mathfrak{g}_{i-1}/\mathfrak{g}_i$ are Abelian of order 1, are called solvable infinitesimal groups.

In the case of a nilpotent group,

$$\mathfrak{g}, \mathfrak{g}^2, \mathfrak{g}^3, \dots, \mathfrak{g}^{\nu} = 0$$

form a composition series. We can find subspaces such that

$$\mathfrak{f}_0 = \mathfrak{g}^{i-1} > \mathfrak{f}_1 > \mathfrak{f}_2 > \dots > \mathfrak{f}_{\lambda-1} > \mathfrak{g}^i = \mathfrak{f}_{\lambda}$$

such that the dimension of $\mathfrak{f}_{\mu-1}$ is larger by 1 than the dimension of $\mathfrak{f}_{\mu}$. Obviously, all $\mathfrak{f}_{\mu}$ are invariant subgroups of $\mathfrak{g}^{i-1}$. The factor groups are Abelian groups since the order is 1 and $[a, a] = 0$ holds in every Lie algebra. If we repeat this process for $i = 1, 2, \dots, h$, we obtain: Every nilpotent group is solvable.

Conversely, if $\mathfrak{g}$ is solvable and $\mathcal{O}$ is a representation, the matrices $\mathcal{O}'$ corresponding to the elements of $\mathfrak{g}' = \mathfrak{g}^2$ form a nilpotent set (Lie's theorem). This follows [p. 54] from the Lie theorem proper: If $\mathfrak{g}$ is solvable and $\mathcal{O}$ is a representation in the space $\mathcal{R}$, then there exists a vector $e \neq 0$ in $\mathcal{R}$ such that $\underline{A}e = \alpha e$ for every $\underline{A}$ in $\mathcal{O}$, α is a scalar depending on $\underline{A}$ [proof of the Lie the-

orem proper: p. 53].

Connection between nilpotency and characteristic equations: To every

linear transformation with the matrix A , there belongs a characteristic equation

$$\det (\omega I - A) = f(\omega) = \omega^n - c_1 \omega^{n-1} + \dots$$

where ω is a variable and I denotes the unit matrix. This polynomial $f(\omega)$ is invariant if we change the coordinate-system of the vector space in which the linear transformation takes place. In particular, c_1 is invariant. c_1 is the trace $\text{tr}(A)$ of A , the sum of all elements in the main diagonal of A . We have

$$\text{tr}(AB) = \text{tr}(BA)$$

for any two matrices A, B , where the products are defined. This is the case, if the number of rows in A is equal to the number of columns in B and the number of rows in B is equal to the number of columns in A .

The roots of the characteristic equation of A^k are the k -th powers of the roots of the characteristic equation of A . If a power of A vanishes, all roots of the characteristic equation $f(\omega) = 0$ of A must vanish, and consequently $f(\omega) = \omega^n$ holds. For a nilpotent set of matrices, all characteristic polynomials have the form ω^n .

If $\mathfrak{g}$ is a Lie algebra and $\mathcal{U}$ is a representation we form the set of all matrices which can be written as a linear combination of products of a finite number of elements of $\mathcal{U}$. These elements form an associative algebra, the enveloping algebra of $\mathcal{U}$.

An associative algebra is called nilpotent if there is a positive integer N such that in the algebra every product with N factors vanishes.

If $\mathfrak{g}$ is nilpotent, the matrices of the adjoint representation $\mathcal{U}$ form a nilpotent set as mentioned above. Consequently, the enveloping algebra is nilpotent and all characteristic polynomials have the form ω^n .

The inversion is also true in the following form: (Generalized Engel's theorem): If $\mathcal{U}$ is a distinct representation of a Lie algebra $\mathfrak{g}$ and all charac-

teristic polynomials have the form ω^n , $\mathfrak{g}$ is nilpotent. [Proof, p. 61-62, for this and Lie's theorem compare also pp. 72-76].

An infinitesimal group is called simple, if it does not possess an invariant subgroup except the whole group and the subgroup 0 consisting only of the vector 0. An infinitesimal group is called semi-simple if there is no solvable invariant subgroup except 0. Except for the Abelian group of order 1, every simple group is semi-simple.

Every Lie group $\mathfrak{g}$ contains a maximum invariant solvable group $\mathfrak{f}$ which is not contained in a larger invariant solvable subgroup. The factor group $\mathfrak{g}/\mathfrak{f}$ is semi-simple [p. 47].

Chapter 4: Integral-Topological Viewpoint

Let σ be a continuous group (with assumptions about differentiability of the occurring functions) which forms a compact manifold. Then, it is possible to define a volume element ds in the manifold σ which is invariant under left-hand and right-hand translation [pp. 76-78]. If $\varphi(s)$ is a function defined in σ , we can define $\int_{\sigma} \varphi(s) ds$ formed over σ . The quotient of the integral by the volume $\int ds$ of σ will be denoted by $\int_{\sigma} \varphi(s)$. From the invariance of the volume element under translations follows

$$\int_{\sigma} \varphi(s) = \int_{\sigma} \varphi(as) = \int_{\sigma} \varphi(sa)$$

for every a in σ .

If $\mathcal{T}$ is a representation of σ of degree n , we transform the Hermitian form $\sum_{v=1}^n x_v \bar{x}_v = H_0$ by the element of $\mathcal{T}$ corresponding to s . We obtain a positive definite Hermitian form which will be denoted by $s(H_0)$. Then $\int_{\sigma} s(H_0)$ is a positive definite Hermitian form invariant under all transformation of $\mathcal{T}$. Consequently, $\mathcal{T}$ is equivalent to a unitary group. (Hermitian argument)

Let σ be a group of transformations of a point field Π which is transitive, i.e. every point of Π can be transformed into every other point of Π by a properly chosen transformation of σ . We assume that σ is compact, but do not

make any assumptions about Π . If $f(P)$ is a function defined in Π , we set

$$sf(P) = f(s^{-1}P)$$

Now, we can define a mean value $m(P)$ of $f(P)$ by

$$m(f(P)) = \int_S sf(P) = \int_S f(s^{-1}P) = \int f(P)$$

It is easily seen that $m(f(P))$ does not depend on P ; we can denote it by $m(f)$.

The functions $f(P)$ form a space, to every s in σ corresponds a linear transformation of the elements of this space. A metric can be introduced by the Hermitian form $\int \overline{f(P)}f(P)$ and, correspondingly, two elements $f(P)$ and $g(P)$ are called orthogonal functions, if $\int \overline{f(P)}g(P) = 0$.

Let $\varphi_1(P), \varphi_2(P), \dots, \varphi_n(P)$ generate an invariant subspace of n dimensions, the $\varphi_i(\)$ being linearly independent. The subspace consists of all functions

$$\varphi(P) = x_1 \varphi_1(P) + \dots + x_n \varphi_n(P) \quad (x_k \text{ constant}).$$

$(x_1, \dots, x_n)$ are the coordinates of $\varphi(P)$. By a transformation s of σ , $\varphi(P)$ is transformed into an element of the same subspace. If

$$s\varphi_i(P) = \sum_{k=1}^n u_{ki}(s) \varphi_k(P)$$

the coordinates of $s\varphi(P)$ are

$$x'_i = \sum_{k=1}^n u_{ik}(s) x_k$$

Obviously, the matrices $\|u_{ik}(s)\|$ form a representation $\mathcal{U}$ of σ . We use the term harmonic set of order $\mathcal{U}$ for $\varphi_1(P), \dots, \varphi_n(P)$. The set is primitive if $\mathcal{U}$ is irreducible. By choosing another basis in the subspace, we obtain a set belonging to an equivalent representation. The basis can be chosen in such a manner that the $\varphi_i(P)$ form a unitary system of functions.

$$\int \overline{\varphi_i(P)} \varphi_k(P) = \delta_{ik} \quad \delta_{ik} = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$$

Then the matrix is unitary.

Primitive sets belonging to non-equivalent representations are orthogonal [pp. 85-86]. To the same representation $\mathcal{U}$ of degree n there belongs at

most n linearly independent primitive harmonic sets. These can be chosen in such a manner that the functions belonging to the independent sets form an orthogonal system [86-87]. If we do this for every irreducible representation, all the harmonic sets together form an orthogonal system for which the theorem of completeness holds [pp. 88-96].

In the case of the regular realization (realization by left-hand translations, $\sigma = \pi$), we obtain a harmonic set of order $\mathcal{U}$ by taking a fixed column of the matrix $\bar{U}(\sigma)$ which represents s . If n is the degree, we have n linearly independent sets corresponding to the n columns [p. 88].

A class function $f(s)$ is a function which has the same value for conjugate group elements s and asa^{-1} . The primitive characters (i.e. the characters belonging to irreducible representations) form a complete normalized orthogonal system in the domain of continuous class functions.

Chapter 1.

Full linear, complex and orthogonal groups from the infinitesimal standpoint

1. Structure of the full linear group

The full linear group of degree n consists of all non-singular linear transformations of a vector space of n dimensions. The transformations of determinant 1 form a subgroup, the unimodular group. The infinitesimal group $\mathfrak{g}$ of the full linear group is formed by all matrices (singular or non-singular) of degree n , the order being n^2 . In the case of the unimodular group, the infinitesimal group contains only the matrices of trace 0, the order being $n^2 - 1$.

Let h_i be the matrix whose coefficient in the i -th row and column is 1 and all other coefficients are zero. Then the matrices

$$(1) \quad h = \lambda_1 h_1 + \lambda_2 h_2 + \dots + \lambda_n h_n = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{vmatrix}$$

form an Abelian subgroup $\mathfrak{f}$ of $\mathfrak{g}$ of order n . If $a = \|\alpha_{ik}\|$ is another element of $\mathfrak{g}$, we have

$$(2) \quad [ha] = ha - ah = \|(\lambda_i - \lambda_k) \alpha_{ik}\|$$

Consequently a is commutative with every h , $[ha] = 0$, if and only if $\alpha_{ik} = 0$ for $i \neq k$ and therefore a belongs to $\mathfrak{f}$. $\mathfrak{f}$ is a maximum Abelian subgroup of $\mathfrak{g}$.

The matrix with a coefficient 1 in the i -th row and the k -th column and all other coefficients 0, will be denoted by e_{ik} ($i \neq k$). The elements h_i and e_{ik} form a basis of $\mathfrak{g}$. From (2) follows

$$(3) \quad [he_{ik}] = (\lambda_i - \lambda_k)e_{ik}, \quad [hh_i] = 0.$$

The element of the adjoint group corresponding to h is given by $s \rightarrow [hs]$ where s ranges over $\mathfrak{g}$. Therefore, the significance of (3) is that e_{ik} is an eigenvector of this transformation of the adjoint group belonging to the

eigenvalue (or characteristic root) $\lambda_i - \lambda_k$. Furthermore, h_i is an eigenvector belonging to the eigenvalue 0. The $n^2 - n$ linear forms $\lambda_i - \lambda_k$ ($i, k = 1, 2, \dots, n; i \neq k$) will be called the roots.

If $\alpha = \lambda_i - \lambda_k$ is a root, so is $-\alpha = \lambda_k - \lambda_i$. We also write e_α for e_{ik} . As is easily seen,

$$(4) \quad [e_\alpha, e_{-\alpha}] = h_i - h_k$$

belongs to $\mathfrak{f}$, ($\lambda_i = 1, \lambda_k = -1, \lambda_j = 0$ ($j \neq i, k$) in (1)). We set

$h_i - h_k = h_\alpha$. We obtain by a simple computation for $\alpha = \lambda_i - \lambda_k$,

$$\beta = \lambda_j - \lambda_l$$

$$(5) \quad \begin{aligned} [e_\alpha, e_\beta] &= 0 && \text{if } \alpha + \beta \text{ is not a root and } \alpha \neq -\beta \\ [e_\alpha, e_\beta] &= \pm e_{\alpha+\beta} && \text{if } \alpha + \beta \text{ is a root} \end{aligned}$$

(The + sign corresponds to the case $k = j$, the - sign to $i = l$)

The formulas (3), (4) and (5) give the structure of $\mathfrak{g}$

2. The weights of a representation

We have a representation $\mathcal{G}$ of $\mathfrak{g}$ if to every element a of $\mathfrak{g}$ corresponds a linear transformation A in a vector space of N dimensions with the following properties. From $a \rightarrow A$ and $b \rightarrow B$ follows

$$a + b \rightarrow A + B, \quad \lambda a \rightarrow \lambda A, \quad [ab] \rightarrow AB - BA.$$

If $e_\alpha \rightarrow E_\alpha$, $h_i \rightarrow H_i$, then

$$h = \lambda_1 h_1 + \lambda_2 h_2 + \dots + \lambda_n h_n \rightarrow \lambda_1 H_1 + \lambda_2 H_2 + \dots + \lambda_n H_n = H$$

and from (3), (4), (5) follows ($\alpha = \lambda_i - \lambda_k$)

$$(6) \quad \begin{aligned} [HE_\alpha] &= \alpha E_\alpha, & [HH_i] &= 0 \\ [E_\alpha, E_{-\alpha}] &= H_i - H_k = H_\alpha \\ E_\alpha E_\beta &= 0, & (\alpha + \beta \neq 0 \text{ and not a root}) \\ [E_\alpha, E_\beta] &= \pm E_{\alpha+\beta} & (\alpha + \beta \text{ a root}) \end{aligned}$$

The problem of finding all representations is that of finding all matrices fulfilling (6); it is a purely algebraic one.

The vectors in the space of N dimension will be denoted by small German letters. We say that a vector $\neq 0$ has a weight Λ , if for each $H = \lambda_1 H_1 + \dots + \lambda_n H_n$ a formula

$$H\mathcal{M} = \Lambda \mathcal{M}$$

holds where Λ is a linear form in $\lambda_1, \lambda_2, \dots, \lambda_n$

There exist weights always. This can be seen from Lie's theorem or directly in the following manner. Let m_1 be an eigenvalue (root of the characteristic equation) of H_1 . Consider all vectors $\mathcal{P}$ such that

$$H_1 \mathcal{P} = m_1 \mathcal{P}$$

There are such eigenvectors $\mathcal{P} \neq 0$. All these $\mathcal{P}$ form a subspace $\mathcal{R}_1$. From the commutativity of the H_k follows easily that $\mathcal{R}_1$ is invariant under the H_k .

Let m_2 be an eigenvalue of the transformation which is induced by H_2 in $\mathcal{R}_1$.

Let $\mathcal{R}_2$ be the subspace of $\mathcal{R}_1$ consisting of all vectors $\mathcal{P}$ of $\mathcal{R}_1$ for which $H_2 \mathcal{P} = m_2 \mathcal{P}$ and so on. Finally, we find vectors $\neq 0$ such that

$$H_i \mathcal{M} = m_i \mathcal{M} \quad (i = 1, 2, \dots, n)$$

$$H\mathcal{M} = (\lambda_1 H_1 + \dots + \lambda_n H_n) \mathcal{M} = (\lambda_1 m_1 + \dots + \lambda_n m_n) \mathcal{M}$$

Consequently, $\Lambda = \lambda_1 m_1 + \dots + \lambda_n m_n$ is a weight.

Lemma 1: If a vector $\mathcal{P}$ of weight Λ is a sum of vectors $\mathcal{P}_i$ of weight Λ_i and all Λ_i are different from Λ , then $\mathcal{P} = 0$

Proof. Let $f(\omega)$ be the polynomial $(\omega - \Lambda_1)(\omega - \Lambda_2) \dots (\omega - \Lambda_r)$ in a variable ω . Then apply the operator

$f(H) = (H - \Lambda_1)(H - \Lambda_2) \dots (H - \Lambda_r)$ on the equation

$$\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2 + \dots + \mathcal{P}_r$$

Since the factors in $f(H)$ are commutative and $(H - \Lambda_i) \mathcal{P}_i = 0$, we obtain 0 on the right side. We find

$$f(\Lambda) \mathcal{P} = 0$$

and since $\Lambda \neq \Lambda_i$, $f(\Lambda) \neq 0$, we have $\mathcal{P} = 0$, q.e.d.

It follows that there are not more than N different weights. In the case of the adjoint representation the weights different from zero are the roots.

Lemma 2: If $\mathcal{M}$ is a vector of the weight Λ , $H\mathcal{M}$ belongs to the weight Λ too and $E_\alpha \mathcal{M}$ belongs to the weight $\Lambda + \alpha$.

Proof. The first remark follows immediately from the commutativity of all the H 's. Furthermore,

$$\begin{aligned} H \cdot E_\alpha \mathcal{M} &= E_\alpha H \mathcal{M} + [H, E_\alpha] \mathcal{M} = \Lambda E_\alpha \mathcal{M} + \alpha E_\alpha \mathcal{M} \\ &= (\Lambda + \alpha) E_\alpha \mathcal{M}. \end{aligned}$$

Lemma 3: If the representation under consideration is irreducible, the basis vectors of the space $\mathcal{R}$ can be chosen in such a manner that each has a definite weight.

Proof. Let $\mathcal{M} \neq 0$ be a vector which has a weight and form the vectors

$$(7) \quad \mathcal{M}, E_\alpha \mathcal{M}, E_\alpha E_{\alpha'} \mathcal{M}, \dots$$

for all α , all possible combinations α, α' and so on. They generate a subspace $\mathcal{R}_1$ which, obviously, is invariant under all E_α . By lemma 2, $\mathcal{R}_1$ is invariant under all H , and thus $\mathcal{R}_1$ is invariant under all elements of the group. From the irreducibility of the representation follows $\mathcal{R}_1 = \mathcal{R}$. A maximum system of linearly independent vectors among the vectors (7) forms a basis of $\mathcal{R}$; all its vectors have a definite weight. The weights which we obtain in this way are all possible weights as can be seen from lemma 1. In this particular basis, H is a diagonal matrix containing all the N weights in the main diagonal.

For the moment, we will take the integral standpoint. By iteration of the infinitesimal transformation h we obtain the finite transformation

$$(\mathcal{E}) = \begin{vmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{\lambda_n} \end{vmatrix}$$

and by performing the corresponding process in the irreducible representation

of we get

$$(E) = \begin{vmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{\lambda_N} \end{vmatrix}$$

An element s of the full linear group is, in the general case, conjugate to an element (E) , where $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_N}$ are the roots of the characteristic polynomial of s

$$s = u(E)u^{-1}$$

If S and U are the elements which correspond to s and u by the representation, we have

$$S = U(E)U^{-1}$$

$$\text{tr}(S) = \text{tr}(E) = e^{\lambda_1} + e^{\lambda_2} + \dots + e^{\lambda_N}$$

Thus, the knowledge of the trace of a representation is equivalent to knowledge of the weights.

There are two remarks which are quite obvious from the integral standpoint and which will be proved in the next section by infinitesimal methods:

- 1) If we permute $\lambda_1, \lambda_2, \dots, \lambda_n$ in (E) we get a conjugate element. Therefore, the corresponding elements in the representation also are conjugate. By a permutation of the λ_i , every weight λ_μ is carried over into another weight.
- 2) If we confine ourselves to the unimodular unitary subgroup of the full linear group, we have to impose upon the λ 's the condition that they are pure imaginary and fulfil

$$(8) \quad \lambda_1 + \lambda_2 + \dots + \lambda_n = 0$$

By using (8) every weight λ can be written such that λ_n does not occur.

Now, the unimodular unitary group is simply connected [p. 183]. Therefore, there

exist only one-valued continuous functions in the group manifold. By increasing λ_i by $2\pi i$ and decreasing λ_k by $2\pi i$ in (ε) , (ε) is not changed and so (E) cannot be changed either. Therefore, each Λ_μ can only be increased by an integral multiple of $2\pi i$ and hence we see that the coefficient of λ_i in each Λ is an integer. If we drop (8) again, we can say that in every weight Λ the difference of the coefficients of any two λ 's is an integer.

3. The first main theorem

If $\alpha = \lambda_i - \lambda_k$ is a root, we have, according to (4)

$$\begin{bmatrix} 0 & 0 \\ \alpha & -\alpha \end{bmatrix} = h_\alpha = \alpha_1 h_1 + \dots + \alpha_n h_n$$

with $\alpha_i = 1$, $\alpha_k = -1$, $\alpha_j = 0$, $j \neq i, k$.

The value which an arbitrary linear function Λ of $\lambda_1, \lambda_2, \dots, \lambda_n$ has for $\lambda_i = \alpha_i$ will be denoted by Λ_α . In particular

$$(9) \quad \alpha_\alpha = (\lambda_i - \lambda_k)_\alpha = 2.$$

Theorem 1. If Λ is a weight, $2 \frac{\Lambda_\alpha}{\alpha_\alpha}$ is an integer and $\Lambda - 2 \frac{\Lambda_\alpha}{\alpha_\alpha} \alpha$ also is a weight for each root $\alpha = \lambda_i - \lambda_k$.

Remarks: Because of (9) we have $2 \frac{\Lambda_\alpha}{\alpha_\alpha} = \Lambda_\alpha$. The formulation

above is chosen with a view of a later generalization. If

$\Lambda = m_1 \lambda_1 + \dots + m_n \lambda_n$, then $\Lambda_\alpha = m_i - m_k$ and $\Lambda - \Lambda_\alpha \alpha$ is obtained from Λ by interchanging λ_i and λ_k . Therefore, the statement of theorem 1 is identical with the results 1) and 2) obtained at the end of section 2 by taking the integral standpoint.

Proof. First of all, let us assume that the weight Λ is extreme in the sense that $\Lambda + \alpha$ is not a weight. If $m_0 \neq 0$ belongs to Λ , we form the vectors

$$m_0, E_{-\alpha} m_0 = m_1, E_{-\alpha} m_1 = m_2, \dots$$

which (by lemma 2, section 2) belong to the weights

$$\Lambda, \Lambda - \alpha, \Lambda - 2\alpha, \dots$$

respectively. There must be an integer h such that $m_h \neq 0$

$$n_{h+1} = 0.$$

We will prove that there exist constants μ_i such that

$$(9) \quad E_\alpha n_i = \mu_i n_{i-1} \quad i = 0, 1, 2, \dots, h+1$$

For $i = 0$, (9) is valid whatever n_{-1} is if we set $\mu_0 = 0$. For $E_\alpha n_0$ belongs to the weight $\Lambda + \alpha$ by section 2, lemma 2, and from the assumption that Λ is extemo follows $E_\alpha n_0 = 0$. If we assume that (9) is proved for i , we have

$$\begin{aligned} E_\alpha n_{i+1} &= E_\alpha E_{-\alpha} n_i = (E_{-\alpha} E_\alpha + [E_\alpha, E_{-\alpha}]) n_i \\ &= E_{-\alpha} E_\alpha n_i + H_\alpha n_i = E_{-\alpha} \mu_i n_{i-1} + (\Lambda - i\alpha)_\alpha n_i. \end{aligned}$$

Therefore (9) holds for $i+1$ and

$$\mu_{i+1} = \mu_i + \Lambda_\alpha - i\alpha_\alpha \quad (i = 1, 2, \dots)$$

Hence by using $\mu_0 = 0$ we find

$$\mu_{i+1} = (i+1)\Lambda_\alpha - \frac{(i+1)i}{2} \alpha_\alpha$$

From (9) and $n_{h+1} = 0$ follows $\mu_{h+1} = 0$. Therefore $\Lambda_\alpha = \frac{h}{2} \alpha_\alpha$.

$$2 \frac{\Lambda_\alpha}{\alpha_\alpha} = h$$

is an integer, $\Lambda - 2 \frac{\Lambda_\alpha}{\alpha_\alpha} = \Lambda - h\alpha$ is a weight, $n_h \neq 0$ belonging to it.

Thus, if Λ is extemo, theorem 1 is true.

If Λ is any weight, we form

$$\dots, \Lambda - 2\alpha, \Lambda - \alpha, \Lambda, \Lambda + \alpha, \Lambda + 2\alpha, \dots$$

If $\Lambda^* = \Lambda + k\alpha$ is the last expression in this sequence which is a weight, Λ^* is extemo and, by what was shown above, it follows that

$$(10) \quad \Lambda^*, \Lambda^* - \alpha, \Lambda^* - 2\alpha, \dots, \Lambda^* - 2 \frac{\Lambda_\alpha^*}{\alpha_\alpha} \alpha$$

also is a weight. In a geometrical representation

$$\begin{array}{ccccccc} | & & | & & | & & | \\ \Lambda^{**} & & \Lambda - \alpha & & \Lambda & & \Lambda + \alpha & & \Lambda + 2\alpha & & \Lambda^* \end{array}$$

these weights lie symmetrically with respect to the "point"

$$(11) \quad \Lambda^* - \frac{\Lambda_\alpha^*}{\alpha_\alpha} \alpha = \Lambda + k\alpha - \frac{(\Lambda + k\alpha)_\alpha}{\alpha_\alpha} \alpha = \Lambda - \frac{\Lambda_\alpha}{\alpha_\alpha} \alpha.$$

On the other hand, if $\Lambda^{**} = \Lambda - k'\alpha$ is the first weight in the sequence (10), we obtain in the same way by using the root $-\alpha$ (with respect to which Λ^{**} is extreme) a set of weights

$$(12) \quad \Lambda^{**}, \Lambda^{**} + \alpha, \Lambda^{**} + 2\alpha, \dots, \Lambda^{**} + 2 \frac{-\Lambda_{\alpha}^{**}}{\alpha_{\alpha}} \alpha$$

which lie symmetrically with respect to the same center (11). Therefore (10) and (12) are identical and Λ belongs to (10). Then we have

$$2 \frac{\Lambda_{\alpha}}{\alpha_{\alpha}} = 2 \frac{(\Lambda^{*} - k\alpha)_{\alpha}}{\alpha_{\alpha}} = 2 \frac{\Lambda_{\alpha}^{*}}{\alpha_{\alpha}} - k = \text{integer}$$

and $\Lambda - 2 \frac{\Lambda_{\alpha}^{*}}{\alpha_{\alpha}} \alpha$ is symmetric to Λ with respect to the center (11). Therefore it belongs to (10) and is a weight, q.e.d.

4. The second main theorem

We arrange the weights of a representation according to the lexicographical order. A linear form $m_1 \lambda_1 + m_2 \lambda_2 + \dots + m_n \lambda_n$ is positive if the first non-vanishing coefficient is positive. A linear form is larger than another form, if the difference is positive.

Lemma: If Λ is the highest weight of an irreducible representation and $\mu \neq 0$ belongs to Λ , then every vector

$$\mu^{*} = E_{\alpha_m} E_{\alpha_{m-1}} \dots E_{\alpha_1} \mu \quad (\alpha_1, \alpha_2, \dots, \alpha_m, \text{ roots})$$

of weight Λ is of the form $k\mu$ (k scalar). k can be computed independently of the special representation, if $\alpha_m, \alpha_{m-1}, \dots, \alpha_1$ are given.

Proof. The vector μ^{*} belongs to $\Lambda + \alpha_1 + \alpha_2 + \dots + \alpha_m$ according to lemma 2, section 2. It is of weight Λ , if $\alpha_1 + \alpha_2 + \dots + \alpha_m = 0$. At least one α_{μ} must be positive. We choose here μ as small as possible. We assume that the lemma is proved for all smaller values of m and in the case of the same m for all smaller values of μ . For $\mu = 1$ it is trivial because $E_{\alpha_1} \mu$ belonging to $\Lambda + \alpha_1$ must vanish since $\Lambda + \alpha_1$ is higher than Λ .

In the general case, we have

$$(13) \quad \begin{aligned} n^* &= E_{\alpha_m} \dots (E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}) \dots E_{\alpha_1} n = E_{\alpha_m} \dots (E_{\alpha_{\mu-1} + \alpha_{\mu}} + [E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}]) \dots E_{\alpha_1} n \\ n^* &= E_{\alpha_m} \dots (E_{\alpha_{\mu-1}} E_{\alpha_{\mu}}) \dots E_{\alpha_1} n + E_{\alpha_m} \dots [E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}] \dots E_{\alpha_1} n \end{aligned}$$

The first summand on the right hand side is a multiple of n because the sum of the indices $\alpha_1 + \alpha_2 + \dots + \alpha_{\mu} + \alpha_{\mu-1} + \dots + \alpha_m$ vanishes and at the $(\mu-1)$ th place stands a positive root α_{μ} . To prove the same for the second summand we distinguish three cases

- 1) $\alpha_{\mu} + \alpha_{\mu-1}$ neither 0 nor a root. Then $[E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}] = 0$ (section 2, (6)) and the statement is trivial.
- 2) $\alpha_{\mu} + \alpha_{\mu-1} = 0$. Then (section 2 (6)) $[E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}] = H_{\alpha_{\mu}}$ and the second summand is

$$E_{\alpha_m} \dots E_{\alpha_{\mu+1}} H_{\alpha_{\mu}} E_{\alpha_{\mu-2}} \dots E_{\alpha_1} n = (\Lambda + \alpha_1 + \dots + \alpha_{\mu-2}) E_{\alpha_m} \dots E_{\alpha_{\mu+2}} E_{\alpha_{\mu-1}} \dots E_{\alpha_1} n$$

because from section 2, lemma 2 follows that $E_{\alpha_{\mu-2}} \dots E_{\alpha_1} n$ belongs to the weight $\Lambda + \alpha_1 + \dots + \alpha_{\mu-2}$. On the right-hand side here, we have a smaller number of factors E_{α} than m and $\alpha_1 + \dots + \alpha_{\mu-2} + \alpha_{\mu+1} + \dots + \alpha_m = 0$ in this case. Therefore, this vector is a multiple of n .

- 3) $\alpha_{\mu} + \alpha_{\mu-1}$ is a root. Then (section 2 (6)) $[E_{\alpha_{\mu}} E_{\alpha_{\mu-1}}] = \pm E_{\alpha_{\mu} + \alpha_{\mu-1}}$, and by replacing this in the second summand in (13) we

find

$$\pm E_{\alpha_m} \dots E_{\alpha_{\mu+1}} E_{\alpha_{\mu} + \alpha_{\mu-1}} E_{\alpha_{\mu-2}} \dots E_{\alpha_1} n.$$

Here the number of factors E_{α} is smaller than m ; the sum of the indices vanishes. Thus this vector is a multiple of n .

Theorem 2. An irreducible representation is uniquely determined by its highest weight Λ . The highest weight occurs only once,

Proof. If $n \neq 0$ belongs to Λ , we form

$$(14) \quad n, E_{\alpha} n, E_{\alpha'} E_{\alpha} n, \dots \quad (\alpha, \alpha', \dots \text{ all roots})$$

Obviously, these vectors generate an invariant subspace and, because of the irreducibility, this is the whole space. All vectors (14) have a definite weight but by the preceding lemma, Λ occurs only in the case of vectors $k n$. Conso-

quently, there are no other vectors of the weight Λ in the space.

If there exist a second irreducible representation $\mathcal{G}^*$ with the highest weight Λ , let E_α^* , $\mathcal{N}^*$ have the same significance for it as E_α , $\mathcal{N}$ have in the case of the original representation. We form

$$(14^*) \quad \mathcal{N}^*, E_\alpha^* \mathcal{N}^*, E_{\alpha'}^* E_\alpha^* \mathcal{N}^*, \dots$$

Thus, to every term in (14) corresponds a term in (14*). If there exist a linear relation between some of the vectors in (14), say $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_m$

$$(15) \quad \gamma_1 \mathcal{N}_1 + \gamma_2 \mathcal{N}_2 + \dots + \gamma_m \mathcal{N}_m = 0$$

we form the vector

$$(16) \quad \gamma_1 \mathcal{N}_1^* + \gamma_2 \mathcal{N}_2^* + \dots + \gamma_m \mathcal{N}_m^* = \mathcal{N}^*$$

where $\mathcal{N}_\mu^*$ is the vector in (14*) corresponding to $\mathcal{N}_\mu$ in (14).

These vectors $\mathcal{N}^*$ for all possible relations (15) form a subspace $\mathcal{R}_1^*$ ^{of the space $\mathcal{R}^*$} in which the group $\mathcal{G}^*$ takes place. Obviously, $\mathcal{R}_1^*$ is invariant under $\mathcal{G}^*$ and therefore, because of the irreducibility of $\mathcal{G}^*$, we have $\mathcal{R}_1^* = \mathcal{R}^*$ or $\mathcal{R}_1^* = 0$. But $\mathcal{N}^*$ does not belong to $\mathcal{R}_1^*$. If, namely $\mathcal{N}^* = \mathcal{N}_\mu^*$, all $\mathcal{N}_\mu^*$ on the left-hand side in (16) with $\gamma_\mu \neq 0$ ^(can be assumed to be) are of weight Λ according to section 2, lemma 1, and therefore $\mathcal{N}_\mu^* = \mathcal{R}_\mu \mathcal{N}^*$ by lemma 4. Furthermore, by lemma 4, we have $\mathcal{N}_\mu = \mathcal{R}_\mu \mathcal{N}$ with the same $\mathcal{R}_\mu$ because the computation of $\mathcal{R}_\mu$ does not depend on the representation. Therefore, from (15) with $\mathcal{N}^* = \mathcal{N}$ follows

$$\gamma_1 \mathcal{N}_1 + \dots + \gamma_m \mathcal{N}_m = \mathcal{N}$$

in contradiction to (15). Thus, $\mathcal{R}_1^*$ being different from $\mathcal{R}^*$ we have

$\mathcal{R}_1^* = 0$. In other words, from (15) follows

$$\gamma_1 \mathcal{N}_1^* + \gamma_2 \mathcal{N}_2^* + \dots + \gamma_m \mathcal{N}_m^* = 0$$

The correspondence

$$E_{\alpha_n} E_{\alpha_{n-1}} \dots E_{\alpha_1} \mathcal{N} \rightarrow E_{\alpha_n}^* E_{\alpha_{n-1}}^* \dots E_{\alpha_1}^* \mathcal{N}^*$$

gives a non-singular linear transformation from the space $\mathcal{R}$ of $\mathcal{G}$ to $\mathcal{R}^*$ and by using corresponding basis vectors selected from (14) and (14*), $\mathcal{G}$ and $\mathcal{G}^*$

5. Construction of the irreducible representations

We obtain special representations in the following way. We take all skew symmetric tensors of rank m for a fixed $m, 0 < m \leq n$. These are functions $F(k_1, k_2, \dots, k_m)$ of the integral variables k_μ ranging over $1, 2, \dots, n$, which depend on the coordinate system and possess the properties

1) (tensor property) by performing a coordinate transformation

$$(17) \quad x'_i = \sum_k a(i, k) x_k$$

the tensor is transformed into

$$(18) \quad F'(i_1, i_2, \dots, i_m) = \sum_{k_1, \dots, k_m} a(i_1, k_1) a(i_2, k_2) \dots a(i_m, k_m) F(k_1, \dots, k_m)$$

2) (skew-symmetry). By interchanging any two variables i_ρ and i_σ $F(i_1, \dots, i_m)$ changes its sign.

We can obtain such tensors in the following way: If $\mathcal{P}$ and $\mathcal{Y}$ are two vectors, with the components $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$ respectively, the products $x_{i_1} y_{i_2}$ form a tensor $\mathcal{P} \times \mathcal{Y}$ of rank 2, and

$$[\mathcal{P}, \mathcal{Y}] = \mathcal{P} \times \mathcal{Y} - \mathcal{Y} \times \mathcal{P}$$

is a skew-symmetric tensor of rank 2.

If $\mathcal{Z} = (z_1, \dots, z_n)$, is a third vector, the products $x_{i_1} y_{i_2} z_{i_3}$ form a tensor $\mathcal{P} \times \mathcal{Y} \times \mathcal{Z}$ of rank 3, we can derive a skew-symmetric tensor by forming the alternating sum

$$[\mathcal{P}, \mathcal{Y}, \mathcal{Z}] = \mathcal{P} \times \mathcal{Y} \times \mathcal{Z} + \mathcal{Y} \times \mathcal{Z} \times \mathcal{P} + \mathcal{Z} \times \mathcal{P} \times \mathcal{Y} - \mathcal{P} \times \mathcal{Z} \times \mathcal{Y} - \mathcal{Z} \times \mathcal{Y} \times \mathcal{P} - \mathcal{Y} \times \mathcal{P} \times \mathcal{Z}$$

(sum over all permutations of $\mathcal{P}, \mathcal{Y}, \mathcal{Z}$, + sign in the case of an even permutation, - sign in the case of an odd permutation).

Correspondingly, we can define $\mathcal{P}_1 \times \mathcal{P}_2 \times \dots \times \mathcal{P}_m$ and

$$[\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_m] \quad \text{and obtain skew-symmetric tensors of rank } m.$$

Skew-symmetric tensors of rank m are described completely by the $\binom{n}{m}$ components $F(i_1, i_2, \dots, i_m)$ with $i_1 < i_2 < \dots < i_m$.

By performing (17) they undergo a linear transformation which can be

obtained from (18) by expressing all $F(k_1, \dots, k_m)$ on the right-hand side by those with $k_1 < k_2 < \dots < k_m$.

We thus obtain a representation Γ_m of the full linear group. Γ_m has the degree $\binom{n}{m}$, $m = 1, 2, \dots, n$.

We have, of course, if $n_1, n_2, \dots, n_n$ is a basis of the vector space

$$[p, q] = \sum x_i y_k [n_i, n_k].$$

The most general skew-symmetric tensor of rank 2 can be written in the form

$$\frac{1}{2} \sum_{i < k} \alpha(ik) [n_i, n_k]$$

where $\alpha(ik)$ is skew symmetric. Corresponding facts hold in the case of rank

m . The tensors $[n_{i_1}, n_{i_2}, \dots, n_{i_m}]$, $i_1 < i_2 < \dots < i_m$ can be used as a basis in the $\binom{n}{m}$ -dimensional "tensor"-space.

By performing the infinitesimal transformation H :

$$dn_i = \lambda_i n_i$$

we obtain

$$d[n_{i_1}, n_{i_2}, \dots, n_{i_m}] = (\lambda_{i_1} + \lambda_{i_2} + \dots + \lambda_{i_m}) [n_{i_1}, \dots, n_{i_m}]$$

Thus, our basis vectors in the $\binom{n}{m}$ -dimensional space have a weight and the weights of the representation Γ_m are

$$(19) \quad \lambda_{i_1} + \lambda_{i_2} + \dots + \lambda_{i_m}$$

In particular, the highest weight is $\lambda_1 + \lambda_2 + \dots + \lambda_m$. The representation

Γ_m is irreducible since every weight (19) can be obtained from the highest weight by a permutation of the λ 's and, consequently, is a weight of the irreducible representation to which $\lambda_1 + \lambda_2 + \dots + \lambda_m$ belongs.

A second proof for the irreducibility of the Γ_m , which does not use the first main theorem, will be given at the end of section 5.

If the vectors p of an n -dimensional space undergo a linear transformation A and the vectors q of an m -dimensional space undergo a second linear

transformation B , the mn products $x_i y_k$ also undergo a linear transformation which is denoted by $A \times B$ (Kronecker product). If $A(s)$ and $B(s)$ are two representations $\mathcal{A}$ and $\mathcal{B}$ of a group, $A(s) \times B(s)$ also is a representation $\mathcal{A} \times \mathcal{B}$.

In the special case that A and B belong to diagonal matrices, to $A \times B$ belongs a diagonal matrix containing in the main diagonal the mn products of diagonal elements of A by diagonal elements of B . Hence it is easily seen that the weights of $\mathcal{A} \times \mathcal{B}$ are the sums of one weight of $\mathcal{A}$ and one weight of $\mathcal{B}$, $\mathcal{A}$ and $\mathcal{B}$ being representations of the full linear group.

In the same way, Kronecker products of more than two factors can be defined. For the product of m equal factors $\mathcal{A}$ will be written $\mathcal{A}^m$.

If a representation $\mathcal{G}$ has the highest weight

$$(20) \quad \Lambda = m_1 \lambda_1 + m_2 \lambda_2 + \dots + m_n \lambda_n$$

we have $m_1 \geq m_2 \geq \dots \geq m_n$, since otherwise we should be able to obtain a higher weight by a permutation of the λ 's.

Γ_n is a representation of degree $\binom{n}{n} = 1$, to each transformation s corresponds its determinant. The only weight is $\lambda_1 + \lambda_2 + \dots + \lambda_n$ according to (19). Therefore, the weights of $\mathcal{G} \times \Gamma_n^k$ can be obtained by increasing each coefficient in a weight of $\mathcal{G}$ by k . The transformations of $\mathcal{G} \times \Gamma_n^k$ are obtained from those of $\mathcal{G}$ by multiplying each coefficient by the k -th power of the determinant of the transformation. If k is sufficiently large, we see that we can assume $m_n \geq 0$ without essential restriction. We have then

$$(21) \quad m_1 \geq m_2 \geq \dots \geq m_n \geq 0$$

By multiplying the Γ_n 's we are able to form another representation with the same highest weight (20). Obviously,

$$\Gamma_n^{p_1} \times \Gamma_n^{p_2} \times \dots \times \Gamma_n^{p_n}$$

has the highest weight (according to (15) and the rule for the weights of a prod-

uct)

$$p_1 \lambda_1 + p_2 (\lambda_1 + \lambda_2) + \dots + p_n (\lambda_1 + \lambda_2 + \dots + \lambda_n) = \\ = (p_1 + p_2 + \dots + p_n) \lambda_1 + (p_2 + p_3 + \dots + p_n) \lambda_2 + \dots + p_n \lambda_n$$

In order to obtain (20) we have to take

$$p_1 = m_1 - m_2, \quad p_2 = m_2 - m_3, \quad \dots, \quad p_{n-1} = m_{n-1} - m_n, \quad p_n = m_n$$

Thus the representation

$$\Gamma = \Gamma_1^{m_1 - m_2} \times \Gamma_2^{m_2 - m_3} \times \dots \times \Gamma_{n-1}^{m_{n-1} - m_n} \times \Gamma_n^{m_n}$$

has the required highest weight. Since there is only one irreducible $\mathcal{G}$ with this highest weight (Theorem 2), $\mathcal{G}$ must be an irreducible constituent of Γ .

In order to separate this irreducible representation in Γ we take a vector λ^* of the highest weight (20) in Γ . Obviously, each vector of highest weight in Γ is a multiple of λ , because the analogous fact is true for $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ and therefore for all Kronecker products formed of the Γ_v . We perform all linear transformations of the representation upon λ^* . The vectors obtained in this way generate an invariant subspace $\mathcal{R}^* \neq 0$ of the total space $\mathcal{R}$. We want to show that $\mathcal{R}^*$ is irreducible. Then we will have that the linear transformations of $\mathcal{R}^*$ which are induced by the transformations of Γ in $\mathcal{R}$ form the irreducible representation $\mathcal{G}$ with the given highest weight (20).

For the proof, we need the fact that Γ is completely decomposable into irreducible constituents. This will be seen in 7.

If $\mathcal{R}^*$ is reducible it is the direct sum of two invariant subspaces $\mathcal{R}^* = \mathcal{R}_1^* + \mathcal{R}_2^*$. If $\lambda^* = \lambda_1^* + \lambda_2^*$, λ_1^* in $\mathcal{R}_1^*$, λ_2^* in $\mathcal{R}_2^*$, we find from $H\lambda^* = \Lambda \lambda^*$ that $H\lambda_1^* + H\lambda_2^* = \Lambda \lambda_1^* + \Lambda \lambda_2^*$. Hence, we have $H\lambda_1^* = \Lambda \lambda_1^*$ and $H\lambda_2^* = \Lambda \lambda_2^*$. But there is just one linearly independent vector of weight Λ and so we have either $\lambda_1^* = 0$ or $\lambda_2^* = 0$, say $\lambda_2^* = 0$, $\lambda^* = \lambda_1^*$. The invariant subspace

$\mathcal{R}_1^*$ containing $\mathcal{M}^*$ contains all vectors generated from $\mathcal{M}^*$ by performing the transformations of the group. Thus we have $\mathcal{R}_1^* = \mathcal{R}^*$ and $\mathcal{G}$ is irreducible.

Second proof for the irreducibility of Γ_m : If $m_1, \dots, m_n$ constitute a basis of the space in which the transformations of $\mathcal{G}$ take place, a basis of the space $\mathcal{R}$ in which the transformations of Γ_m are performed is formed of

$$(22) \quad [m_{i_1}, m_{i_2}, \dots, m_{i_m}] \quad (i_1 < i_2 < \dots < i_m),$$

and this element belongs to the weight $\lambda_{i_1} + \lambda_{i_2} + \dots + \lambda_{i_m}$. We denote the $\binom{n}{m} = k$ elements (22) in any arrangement by $m^1, m^2, \dots, m^k$, and the corresponding weights by $\lambda^1, \lambda^2, \dots, \lambda^k$.

Let $\mathcal{R}' \neq 0$ be an invariant subspace of $\mathcal{R}$, and $\mathcal{M} \neq 0$ an element of

$\mathcal{R}'$

$$\mathcal{M} = \sum \alpha_k m^k$$

At least one coefficient, say α_v is different from 0. We perform H and obtain

$$H\mathcal{M} = \sum_k \alpha_k \lambda^k m^k$$

Hence

$$(H - \lambda^\mu) \mathcal{M} = \sum_k \alpha_k (\lambda^k - \lambda^\mu) m^k.$$

where m^μ has the coefficient 0 on the right-hand side. Then we have

$$\prod_{\mu \neq v} (H - \lambda^\mu) \mathcal{M} = \alpha_v \prod_{\mu \neq v} (\lambda^v - \lambda^\mu) m^v$$

With $\mathcal{M}$, this element belongs to $\mathcal{R}'$ and so does m^v itself, since the factor of m^v is different from 0, $\lambda^1, \lambda^2, \dots, \lambda^k$ being all distinct. By transformations of Γ_m , we can carry over m^v into every other element (22). Therefore $m^1, m^2, \dots, m^k$ belong to $\mathcal{R}'$ and we have $\mathcal{R} = \mathcal{R}'$. This proves the irreducibility of $\mathcal{G}$.

6. Young's symmetry operators

In order to construct the irreducible representation $\mathcal{G}$ with the highest weight

$$m_1 \lambda_1 + m_2 \lambda_2 + \dots + m_n \lambda_n \quad (m_1 \geq m_2 \geq \dots \geq m_n \geq 0)$$

we have to set (according to 5)

$$p_1 = m_1 - m_2, \quad p_2 = m_2 - m_3, \quad \dots, \quad p_{n-1} = m_{n-1}, \quad p_n = m_n$$

and to form

$$(23) \quad \Gamma = \Gamma_1^{p_1} \times \Gamma_2^{p_2} \times \dots \times \Gamma_n^{p_n}.$$

If λ is a vector of highest possible weight of this representation, we have to apply all transformations on λ . All vectors obtained from λ in this manner generate a space $\mathcal{R}$. Then $\mathcal{G}$ consists of the transformations of $\mathcal{R}$ which are induced by the original transformations.

Let $n_1, n_2, \dots, n_m$ be a vector-basis for the original space. The "vector"

$$\lambda = \underbrace{[n_1 n_2 \dots n_n] \times \dots \times [n_1 n_2 \dots n_n]}_{p_n \text{ factors}} \times \underbrace{[n_1 n_2 \dots n_{n-1}] \times \dots \times [n_1 n_2 \dots n_{n-1}]}_{p_{n-1} \text{ factors}} \times \dots \times \underbrace{[n_1] \times \dots \times [n_1]}_{p_1 \text{ factors}}$$

is of highest possible weight for (23). By performing a linear transformation T in the original space, λ is carried over into

$$(24) \quad \underbrace{[\varrho \eta \dots \zeta] \times \dots \times [\varrho \eta \dots \zeta]}_{p_n \text{ factors}} \times \underbrace{[\varrho \eta] \times \dots \times [\varrho \eta]}_{p_2 \text{ factors}} \times \underbrace{\varrho \times \dots \times \varrho}_{p_1 \text{ factors}} = \psi$$

where $\varrho, \eta, \dots, \zeta$ are any n linearly independent vectors in the original space. The rank of this tensor (24) is

$$m = np_n + (n-1)p_{n-1} + \dots + 2p_2 + p_1 = m_1 + m_2 + \dots + m_n$$

since this is the number of vectors occurring in (24). More exactly, ϱ occurs m_1 times, η occurs m_2 times, $\dots$, ζ occurs m_n times.

We replace ψ by the tensor

$$(25) \underbrace{(\rho \times \eta \times \dots \times \zeta)}_{p_1 \text{ parentheses}} \times \dots \times \underbrace{(\rho \times \eta \times \dots \times \zeta)}_{p_2 \text{ parentheses}} \times \dots \times \underbrace{(\rho \times \eta)}_{p_2 \text{ parentheses}} \times \dots \times \underbrace{(\rho \times \eta)}_{p_2 \text{ parentheses}} \times \underbrace{\rho \times \dots \times \rho}_{p_1 \text{ factors}}$$

Let $x_v, y_v, \dots, z_v$ be the components of $\rho, \eta, \dots, \zeta$ respectively.

The components of φ are given by

$$(26) (x_{i_1} y_{j_1} \dots z_{l_1}) \dots (x_{i_m} y_{j_m} \dots z_{l_m}) \dots (x_{i_{m_2+1}} y_{j_{m_2+1}}) \dots (x_{i_{m_2}} y_{j_{m_2}}) x_{i_{m_2+1}} \dots x_{i_{m_1}}$$

We denote this expression by

$$(26^*) \quad \varphi \begin{pmatrix} i_1 \dots i_{m_1} & \dots & i_{m_2} \dots i_{m_1} \\ j_1 \dots j_{m_1} & \dots & j_{m_2} \dots j_{m_1} \\ \dots & \dots & \dots \\ l_1 \dots l_{m_1} & \dots & l_{m_2} \dots l_{m_1} \end{pmatrix}$$

We write the components of ψ , (24), in the same form. We obtain ^{them} from the components (26*) of φ in the following way: We permute the indices in the columns of (26*) in all possible ways, multiply by +1 or -1 corresponding to whether the permutation is even or odd, and add all these expressions. We denote this operator which transforms φ into ψ by A

$$\psi = A \varphi$$

The tensor (26*) is symmetric in the indices in every row. We consider the most general tensor Φ of rank m which, if written in the form (26*) has the same properties of symmetry. We set $\Psi = A \Phi$. Of course every linear combination of tensors φ , (25), (26), is such a Φ . We are going to prove that, conversely, every Φ is a linear combination of tensors φ . It will follow then that the tensors Ψ are linear combinations of tensors ψ (24). Vice versa, every such linear combination is a tensor Ψ . Hence the space $\mathcal{R}$ is identical with the space of all tensors Ψ .

In order to prove that the special tensors φ (25), (26), span the whole tensor space Φ , we have to show that no proper linear subspace of the Φ -space

contains all the φ 's. In other words, we have to prove that the expressions (26) for all systems

$$i_1 \leq i_2 \leq \dots \leq i_{m_1}, \quad j_1 \leq j_2 \leq \dots \leq j_{m_2}, \dots$$

are linearly independent, when we consider $x_1, x_2, \dots, x_{m_1}, y_1, y_2, \dots, y_{m_2}, \dots$ as independent variables. But this is obvious from (26).

Thereby we have proved that the space $\mathcal{R}$ consists of all the tensors Ψ . Every linear transformation in the original space induces a transformation in $\mathcal{R}$ and these induced transformations form $\mathcal{G}$. The tensors Ψ form the substratum of the representation $\mathcal{G}$.

We will now construct all tensors Ψ . We start from an absolutely arbitrary tensor χ of rank m . The $m = m_1 + m_2 + \dots + m_n$ components may be arranged in a scheme of the following form (as before)

$$(27) \quad \begin{array}{|c|c|c|c|c|c|c|} \hline i_1 & i_2 & \dots & i_{m_3} & \dots & i_{m_2} & \dots & i_{m_1} \\ \hline j_1 & j_2 & \dots & j_{m_3} & \dots & j_{m_2} & & \\ \hline k_1 & k_2 & \dots & k_{m_3} & & & & \\ \hline \dots & \dots & \dots & \dots & & & & \\ \hline \end{array}$$

where all the indices run from 1 to n . We denote by p all permutations of the indices which leave every index in its row. We denote by q all permutations which leave every index in its column, and we set $\delta_q = +1$ if q is even, and $\delta_q = -1$ if q is odd.

Then it is clear what is meant by

$$(28) \quad \Psi = \sum_{p, q} \delta_q q p \cdot \chi$$

where the sum is to be extended over all p and q . Obviously, we obtain tensors

Ψ with the required properties of symmetry, and all such tensors, indeed, because Ψ in (28) coincides with χ apart from a positive scalar factor, if χ itself has the properties of symmetry.

The operator $\sum_{\mu, \nu} \delta_{\mu \nu} q^{\mu \nu}$ determined by a scheme (27) is called a Young operator. There is a representation of the full linear group whose substratum consists of the general tensors of rank m , i.e. the representation $\Gamma_1 \times \Gamma_1 \times \dots \times \Gamma_1$ (m factors). It is possible to investigate the splitting of this representation in irreducible constituents in purely algebraic manner by deriving the properties of these operators. One obtains a close connection with the theory of representations of the symmetric permutation group.

In the case $n = 2$, the possible highest weights are $m_1 \lambda_1 + m_2 \lambda_2$ with $m_1 \geq m_2 \geq 0$. We can assume $m_2 = 0$, if we take the unimodular group instead of the full linear group, then $m = m_1$. We write $\mathcal{G}_m$ for the irreducible group $\mathcal{G}$ belonging to $m_1 \lambda_1$ as highest weight. Equation (26) has the form

$$\varphi(i_1, \dots, i_m) = x_1^{\mu} x_2^{m-\mu}$$

if μ of the i_ν 's have the value 1 and $m-\mu$ the value 2. The tensor φ is given by

$$(29) \quad x_1^m, x_1^{m-1} x_2, \dots, x_2^m$$

If x_1, x_2 undergo a unimodular transformation s , the quantities (29) undergo the corresponding transformation S_m in $\mathcal{G}_m$. We can replace (29) by any set of linearly independent linear combinations, i.e. any linearly independent homogeneous polynomials of degree m

$$c_0 x_1^m + c_1 x_1^{m-1} x_2 + \dots + c_m x_2^m$$

if we replace $\mathcal{G}_m$ by an equivalent representation.

Let y_1, y_2 be two variables which are transformed cogrediently with x_1, x_2 . We form

$$(30) \quad y_1^{\tau}, y_1^{\tau-1} y_2, \dots, y_2^{\tau}$$

which yield the representation $\mathcal{G}_{\tau}$. The $(m+1)(\tau+1)$ products $x_1^{\mu} x_2^{m-\mu} y_1^{\rho} y_2^{\tau-\rho}$ or rather any $(m+1)(\tau+1)$ linearly independent polynomials $\mathcal{P}(x_1 x_2; y_1 y_2)$ which are homogeneous of degree m in x_1, x_2 and of degree τ in y_1, y_2 can be

considered as the substratum of the representation $\mathcal{G}_m \times \mathcal{G}_r$. We define $m + r + 1$ special polynomials $\varphi_i(x_1, x_2; y_1, y_2)$ of the required degrees by

$$(31) \quad (\mu x_1 + \nu x_2)^m (\mu y_1 + \nu y_2)^r = \sum_{i=0}^{m+r} \binom{m+r}{i} \mu^i \nu^{m+r-i} \varphi_i(x_1, x_2; y_1, y_2)$$

For $x_1 = y_1, x_2 = y_2$, we have evidently

$$(32) \quad \varphi_i(x_1, x_2; x_1, x_2) = x_1^i x_2^{m+r-i}$$

We set

$$(33) \quad \begin{aligned} x_1 &= \alpha x'_1 + \beta x'_2, & y_1 &= \alpha y'_1 + \beta y'_2 \\ x_2 &= \gamma x'_1 + \delta x'_2, & y_2 &= \gamma y'_1 + \delta y'_2 \end{aligned}$$

Then $\mu x_1 + \nu x_2 = \mu' x'_1 + \nu' x'_2$ with $\mu' = \mu\alpha + \nu\gamma, \nu' = \mu\beta + \nu\delta$ and $\mu y_1 + \nu y_2 = \mu' y'_1 + \nu' y'_2$; furthermore

$$(34) \quad (\mu x_1 + \nu x_2)^m (\mu y_1 + \nu y_2)^r = (\mu' x'_1 + \nu' x'_2)^m (\mu' y'_1 + \nu' y'_2)^r = \sum_i \binom{m+r}{i} \mu'^i \nu'^{m+r-i} \varphi_i(x'_1, x'_2; y'_1, y'_2)$$

On comparing (31) and (33) we see that $\varphi_i(x_1, x_2, y_1, y_2)$ are connected with the $\varphi_i(x'_1, x'_2, y'_1, y'_2)$ by a linear transformation $\mathcal{T}$ with coefficients depending on $\alpha, \beta, \gamma, \delta$ only. We put $x_1 = y_1, x_2 = y_2$. We see from (32) that $\mathcal{T}$ is the transformation which

$$x_1^{m+r}, x_1^{m+r-1} x_2, \dots, x_2^{m+r-1}$$

undergo if we apply (33). Consequently, $\mathcal{T}$ is the transformation corresponding to $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ in $\mathcal{G}_{m+r}$.

If $\mathcal{P}(x_1 x_2; y_1 y_2)$ is a polynomial homogeneous of degree m in x_1, x_2 , and of degree r in y_1, y_2 , there exists a formula like

$$(35) \quad \mathcal{P}(x_1, x_2; y_1, y_2) = \sum_{i=0}^{m+r} a_i \varphi_i(x_1, x_2; y_1, y_2) + (x_1 y_2 - x_2 y_1) Q(x_1, x_2; y_1, y_2)$$

where Q is of degree $m-1$ in $x_1 x_2$ and of degree $r-1$ in y_1, y_2 , and the a 's are constants (Clebsch-Gordan formula).

In order to prove this we remark first that (35) represents $(m+1)(r+1)$ linear equations for the $(m+r+1)$ coefficients a_i and the $m \cdot r$ coefficients of Q . The number of equations and of unknown ^{numbers} being equal, we have only to prove that the corresponding homogeneous equations

$$(36) \quad 0 = \sum a_i \varphi_i(x_1, x_2; y_1, y_2) + (x_1 y_2 - x_2 y_1) Q(x_1 x_2; y_1, y_2)$$

are satisfied only for $a_i = 0$, $Q = 0$. We set $x_1 = y_1$, $x_2 = y_2$. Then (36) has the form $\sum a_i x_1^i x_2^{m+r-i} = 0$. Hence $a_i = 0$. Then (36) shows, after suppressing the factor $x_1 y_2 - x_2 y_1$, that $Q(x_1, x_2; y_1, y_2) = 0$.

We apply (35) to the $(m+1)(r+1)$ linearly independent polynomials P .

We obtain then at least $m \cdot r = (m+1)(r+1) - (m+r+1)$ linearly independent polynomials Q on the right-hand side, hence exactly $m \cdot r$ linearly independent Q . Therefore

$$(37) \quad \mathcal{G}_m \times \mathcal{G}_r = \mathcal{G}_{m+r} + \mathcal{G}_{m-1} \times \mathcal{G}_{r-1}$$

We express $\mathcal{G}_{m-1} \times \mathcal{G}_{r-1}$ in the corresponding form and so on. We get finally for $m \geq r$

$$(38) \quad \mathcal{G}_m \times \mathcal{G}_r = \mathcal{G}_{m+r} + \mathcal{G}_{m+r-2} + \mathcal{G}_{m+r-4} + \dots + \mathcal{G}_{m-r}, \quad (m \geq r).$$

7. Full reducibility

We consider the group $\mathcal{U}$ of all unitary unimodular transformations (i.e. the transformations $x'_i = \sum_k a_{ik} x_k$, $\det a_{ik} = 1$ which leave $x_1 \bar{x}_1 + x_2 \bar{x}_2 + \dots + x_n \bar{x}_n$ ^(invariant)). Since $\mathcal{U}$ forms a compact manifold, the Hermitian argument (compare the introduction, pp. 8 and 3) yields the fact that every representation of $\mathcal{U}$ is completely decomposable into irreducible constituents.

Let $\tilde{\mathcal{M}}$ be the infinitesimal group belonging to $\mathcal{U}$. If a representation A of $\tilde{\mathcal{M}}$ is given, we obtain a representation of $\mathcal{U}$ by integration according to Lie, first for a neighborhood of the identity only, and then for the whole

group $\mathcal{U}$ by continuation. No difficulty arises since $\mathcal{U}$ is a simply connected manifold [p. 114]. (Compare also the corresponding proof in section 10)

We see now that every representation of $\tilde{\mathcal{U}}$ is completely reducible, because it can be extended to a representation of $\mathcal{U}$ and here the theorem of the full reducibility holds.

Next we consider the infinitesimal group $\mathcal{G}_0$ belonging to the group of all unimodular transformations. $\mathcal{G}_0$ consists of all matrices $\|s_{ik}\|$ whose trace vanishes. The subgroup $\tilde{\mathcal{U}}$ is characterized by the conditions $s_{ki} = -\bar{s}_{ik}$.

Let $\mathcal{G}$ be a reducible representation of $\mathcal{G}_0$, we have in a properly chosen coordinate system

$$\mathcal{G} = \begin{pmatrix} \alpha & \sigma \\ \mathcal{L} & \mathcal{V} \end{pmatrix}.$$

If we apply this representation to the subgroup $\tilde{\mathcal{U}}$ of $\mathcal{G}_0$ only, we obtain a representation of $\tilde{\mathcal{U}}$. The theorem of complete reducibility holds for $\tilde{\mathcal{U}}$ and consequently we may assume that $\mathcal{L} = 0$ if $s_{ki} = -\bar{s}_{ik}$. But every coefficient in $\mathcal{L}$ is a linear form in the s_{ik} and hence $\mathcal{L} = 0$ for all elements of $\mathcal{G}_0$.

This shows that the theorem of complete reducibility holds for the representations of the infinitesimal unimodular group.

The above proof used "integral" methods. A purely infinitesimal proof for the complete reducibility of the representation of a semi-simple Lie algebra was given recently by Casimir and van der Waerden (Math. Annalen 111, p. 1)

We made use of the complete reducibility of the constituent $\mathcal{G}$ of the representation Γ , (23) in 5. We can prove this special fact without using the existence of an invariant volume element in the group manifold $\mathcal{U}$. We may confine ourselves to unitary transformations in the original group $\mathcal{G} = \Gamma_1$. The Kronecker product of any two representations which consist of unitary matrices only has the same property again. Every constituent of a group of unitary

matrices is unitary as is seen from the argument on page 3, lines 7-10. Therefore Γ_ν as constituent of $\Gamma_1 \chi \dots \Gamma_1$ (ν factors) is unitary and hence the same is true for Γ and its constituent $\mathcal{G}$.

8. Transcendental determination of the characters of the irreducible representations of the full linear group

We have seen in § that to every system

$$(39) \quad m_1 \geq m_2 \geq \dots \geq m_n$$

of n integers belongs a representation $\mathcal{G}$ of the full linear group. Let A be any non-singular matrix of degree n with the characteristic roots

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. We set

$$\chi_1 = m_1 + (n-1), \chi_2 = m_2 + (n-2), \dots, \chi_n = m_n$$

Then the character of the matrix $\mathcal{G}$ corresponding to A in the representation is given by

$$(40) \quad \chi = \frac{|\varepsilon^{\ell_1}, \varepsilon^{\ell_2}, \dots, \varepsilon^{\ell_n}|}{|\varepsilon^{n-1}, \varepsilon^{n-2}, \dots, 1|} \quad \ell_i = \chi_i$$

where the n rows of the determinants are obtained by replacing ε by ε_1 , by ε_2 , ..., by ε_n .

The proof of this theorem was given in the notes of last term's course, Chapter 5. We do not repeat the proof here because an absolutely analogous proof will be given in Section 13 for the characters of the representations of the complex group.

One obtains from (40) for the degree of the representation (39),

[p. 116],

$$N = \frac{\Delta(\ell_1, \dots, \ell_n)}{\Delta(n-1, n-2, \dots, 0)} \quad \text{where} \quad \Delta(x_1, x_2, \dots, x_n) = \prod_{k>j} (x_k - x_j)$$

We set

$$f(\lambda) = \prod_i (1 - \varepsilon_i \lambda)$$

$$\frac{1}{f(\lambda)} = p_0 + p_1 \lambda + p_2 \lambda^2 + \dots$$

Then $p_0, p_1, \dots$ are rational integral symmetric functions of $\varepsilon_1, \dots, \varepsilon_n$.

The character χ , (40), can be expressed in the form [p. 117]

$$\chi = | p_{\ell-(n-1)}, p_{\ell-(n-2)}, \dots, p_{\ell} |$$

Here the n rows of the determinants are obtained by replacing λ by $\lambda_1, \lambda_2, \dots, \lambda_n$.

The product of two characters can be written as a sum of characters (40) again. This is a consequence of the fact that the Kronecker product of two representations is a new representation whose character is the product of the original ones and hence a sum of irreducible characters (40).

In the case $n = 2$, we have

$$(41) \quad \mathcal{G}_i \times \mathcal{G}_k = \mathcal{G}_{i+k} + \mathcal{G}_{i+k-2} + \dots + \mathcal{G}_{|i-k|}$$

if $\mathcal{G}_j$ is the irreducible representation belonging to the highest weight $j\lambda_j$

(41) can be proved by comparing the characters on both sides.

We gave a direct proof for (41) in 5, which did not make use of the explicit formulas for the characters.

9. Structure of the complex group

We consider a $(2n)$ -dimensional vector space and denote the components of an arbitrary vector x by $x_1, x_2, \dots, x_n, x_{1'}, x_{2'}, \dots, x_{n'}$. Let

$$S(x, y) = \sum a_{ik} x_i y_k \quad (i, k = 1, 2, \dots, n, 1', 2', \dots, n')$$

be an anti-symmetric bilinear form

$$(42) \quad S(x, y) = -S(y, x)$$

which is not degenerate (i.e. $S(x^0, y) = 0$ identical in y holds only if $x^0 = 0$).

Lemma: If the coordinate system is properly chosen, $S(x, y)$ has the

form

$$(43) \quad S(x, y) = \sum_{\nu=1}^n (x_{\nu} y_{\nu'} - x_{\nu'} y_{\nu})$$

Proof: We start from an arbitrary vector $e_1 \neq 0$. Then $S(e_1, y)$ is not identically 0. Therefore we can find a vector $e_{1'}$ such that

$$(44a) \quad S(e_1, e_{1'}) = 1, \quad S(e_{1'}, e_1) = -1$$

Because of (42), we have

$$(44b) \quad S(e_1, e_1) = 0, \quad S(e_{1'}, e_{1'}) = 0.$$

The linear space x defined by the two equations $S(e_1, x) = 0, S(e_{1'}, x) = 0$

has at least $2n-2$ dimensions. If $\lambda e_1 + \lambda' e_{1'}$ belongs to this linear space

we have

$$0 = S(e_1, \lambda e_1 + \lambda' e_{1'}) = \lambda S(e_1, e_1) + \lambda' S(e_1, e_{1'}) = \lambda'; \quad 0 = S(e_{1'}, \lambda e_1 + \lambda' e_{1'}) = -\lambda$$

hence $\lambda = \lambda' = 0$. We can therefore choose the $2n-2$ other basis vectors

$e_2, e_{2'}, \dots, e_n, e_{n'}$ of the coordinate system such that they belong to this linear space. Then

$$(44c) \quad S(e_1, e_\mu) = 0, \quad S(e_1, e_{\mu'}) = 0, \quad S(e_{1'}, e_\mu) = 0, \quad S(e_{1'}, e_{\mu'}) = 0 \text{ for } \mu \neq 1.$$

Then, from (44a), (44b), (44c) follows $(\mu, \nu = 1, 2, \dots, n, 1', 2', \dots, n')$

$$S(x, y) = S(\sum x_\mu e_\mu, \sum x_{\nu'} e_{\nu'}) = \sum x_\mu y_{\nu'} S(e_\mu, e_{\nu'}) = (x_1 y_{1'} - x_{1'} y_1) + S^*$$

where S^* depends on $x_2, \dots, x_n, x_{2'}, \dots, x_{n'}$; $y_2, \dots, y_n, y_{2'}, \dots, y_{n'}$ only.

If the lemma is proved for forms in $n-1$ variables, we obtain it for the case of n variables.

The basis vectors $e_1, e_2, \dots, e_n, e_{1'}, e_{2'}, \dots, e_{n'}$ of the coordinate-system for which S has the normal form (43) are characterized by

$$S(e_i, e_k) = 0, \quad S(e_{i'}, e_{k'}) = 0, \quad S(e_i, e_{k'}) = \delta_{ik} = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$$

Such a basis will be called a normal coordinate system.

The complex-group $\mathcal{L}$ is the group of all linear transformations which applied simultaneously on x and y leave $S(x, y)$, (43), invariant. The infinitesimal complex group $\mathcal{A}$ consists of the transformations

$$(45) \quad C \quad \begin{cases} dx_i = \sum_{\mu} \rho_{i\mu} x_\mu + \sum_{\mu} \rho'_{i\mu} x_{\mu'} \\ dx_{i'} = \sum_{\mu} \sigma_{i\mu} x_\mu + \sum_{\mu} \sigma'_{i\mu} x_{\mu'} \end{cases} \quad \text{with} \quad \begin{cases} \rho_{ik} + \sigma_{ki} = 0 \\ \rho'_{ik} = \rho'_{ki}, \sigma'_{ik} = \sigma'_{ki} \end{cases}$$

as is easily seen from $dS(xy) = 0$.

The order of $\mathcal{V}$ is $n^2 + 2 \frac{n(n+1)}{2} = n(2n+1)$. The matrices

$$(46) \quad h = \begin{pmatrix} \lambda^1 & & & \\ & \ddots & & \\ & & \lambda^n & \\ & & & -\lambda^1 \\ & & & & \ddots \\ & & & & & -\lambda^n \end{pmatrix}$$

form an abelian subgroup. We set $h = h_\nu$ for $\lambda^i = 0, i \neq \nu$ and $\lambda^\nu = 1$.

From (45) follows

$$[h, c] = \left(\begin{array}{c|c} \rho_{ik}(\lambda^i - \lambda^k) & \rho'_{ik}(\lambda^i + \lambda^k) \\ \hline \sigma'_{ik}(-\lambda^i - \lambda^k) & \sigma_{ik}(-\lambda^i + \lambda^k) \end{array} \right)$$

If, in particular, $\rho_{ik} = -\sigma_{ki} = 1$ for one fixed pair $i < k$, and all other coefficients in (45) vanish, we have $[h, c] = (\lambda^i - \lambda^k) \cdot c$. We denote $\lambda^i - \lambda^k$ by α , the particular c by e_α , and have

$$(47) \quad [h e_\alpha] = \alpha e_\alpha$$

We say again as in section 1 that α is a root. We find in a similar manner that all expressions

$$\alpha = \pm \lambda^i \pm \lambda^k \quad i < k$$

are roots. To each one belongs an element e_α of $\mathcal{V}$ for which (47) holds.

We have to set (i, k fixed)

$$\begin{aligned} \rho_{ik} &= \rho_{ki} = 1 \text{ for } \lambda^i + \lambda^k \text{ (all other coefficients} = 0) \\ \sigma_{ik} &= \sigma_{ki} = 1 \text{ for } -\lambda^i - \lambda^k \text{ (all other coefficients} = 0) \\ \rho_{ki} &= -\sigma_{ik} = 1 \text{ for } -\lambda^i + \lambda^k \text{ (all other coefficients} = 0) \end{aligned}$$

Furthermore, (47) holds for $\alpha = 2\lambda^i$ if e_α is given by $\rho_{ii} = 1$, all other coefficients 0. Finally (47) holds for $\alpha = -2\lambda^i$, if e_α is given by $\sigma_{ii} = 1$, all other coefficients = 0. Thus we have $4 \frac{n(n-1)}{2} + 2n = 2n^2$ roots

$$\alpha = \pm \lambda^i \pm \lambda^k \quad (i < k), \quad \alpha = \pm 2\lambda^i$$

The e_α belonging to these α together with $h_1, h_2, \dots, h_n$, form a basis of $\mathfrak{g}$ as is easily seen. A simple computation gives the structure of $\mathfrak{g}$

$$[h, h'] = 0 \quad [h, e_\alpha] = \alpha e_\alpha$$

$$(48) \quad [e_\alpha, e_\beta] = \begin{cases} 0 & , \alpha + \beta \text{ no root}, \alpha + \beta \neq 0 \\ M e_{\alpha+\beta} & , \alpha + \beta \text{ root} \\ h_\alpha = \alpha^1 h_1 + \dots + \alpha^n h_n & , \alpha + \beta = 0 \end{cases}$$

where $M \neq 0$ is a numerical factor and $\alpha^1, \dots, \alpha^n$ are defined by

$$(a) \quad \alpha^i = \pm 1, \alpha^k = \pm 1, \alpha^j = 0 \text{ for } j \neq i, k, \text{ if } \alpha = \pm \lambda^i \pm \lambda^k$$

$$(b) \quad \alpha^i = \pm 1, \quad \alpha^j = 0 \text{ for } j \neq i, \text{ if } \alpha = \pm 2\lambda^i$$

If $\Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n$ is an arbitrary linear form, we set again

$$\Lambda_\alpha = \Lambda(h_\alpha) = m_1 \alpha^1 + \dots + m_n \alpha^n = \begin{cases} \pm m_i \pm m_k & \text{in the case (a)} \\ \pm m_i & \text{in the case (b)} \end{cases}$$

10. The unitary restriction. The topology of the unitary complex group

We denote the coordinates by $x_1, x_2, \dots, x_l$ in this section. We define a scalar product of two vectors x, y by the Hermitian unit form

$$(xy) = x_1 \bar{y}_1 + \dots + x_l \bar{y}_l$$

A coordinate system $e_1, e_2, \dots, e_l$ is orthogonal if $(e_i, e_k) = \delta_{ik}$, ($\delta_{ik} = 0$ for $i \neq k$, $\delta_{ii} = 1$).

Lemma: A unitary transformation T has, in a properly chosen coordinate system, the form $e_i \rightarrow \varepsilon_i e_i$ where $|\varepsilon_i| = 1$.

Proof (by induction): Let ε be a root of the characteristic equation of T , and e an eigenvector belonging to ε . We can assume $(e, e) = 1$. We can find an orthogonal coordinate system $e_1, e_2, \dots, e_l$ with $e_1 = e$. Then T has the form

$$(49) \quad T = \begin{vmatrix} \varepsilon & 0 & \dots & 0 \\ c_2 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ c_\ell & * & \dots & * \end{vmatrix}$$

From the conditions for unitary matrices follows $|\varepsilon|^2 = 1$ and

$$|\varepsilon|^2 + |c_2|^2 + \dots + |c_\ell|^2 = 1,$$

$c_2 = \dots = c_\ell = 0$. The matrix of degree $\ell-1$ in the lower right corner in (49) is unitary again. If for matrices of degree $\ell-1$ the lemma is proved, we obtain a proof for the degree ℓ .

The ε_i are, of course, the characteristic roots of T . The vectors x satisfying $Tx = \varepsilon_i x$ form a linear space $\mathcal{R}_i$. The number of dimensions of $\mathcal{R}_i$ is the multiplicity of the root ε_i . If x belongs to $\mathcal{R}_i$ and y belongs to $\mathcal{R}_j$, $i \neq j$, we have $(xy) = 0$. This can be seen as follows: We consider x and y as columns with n rows. We denote by a prime the transposed of any matrix. In particular y' is a row with n columns. From $Ty = \varepsilon_j y$ follows $\overline{y'}T' = \overline{\varepsilon_j y'}$, and hence $\overline{y'}T'T = \overline{\varepsilon_j y'T}$, $\overline{y'} = \overline{\varepsilon_j y'T}$, because T is unitary. Hence $(xy) = \overline{y'}x = \overline{\varepsilon_j y'T}x = \overline{\varepsilon_j y'}(\varepsilon_i x) = (\varepsilon_i \overline{\varepsilon_j})(\overline{y'}x) = (\varepsilon_i \overline{\varepsilon_j})(xy)$. Now $|\varepsilon_j| = 1$ and $\varepsilon_i \neq \varepsilon_j$. Therefore $\varepsilon_i \overline{\varepsilon_j} \neq 1$, $(xy) = 0$.

Therefore, we can obtain an orthogonal coordinate system in which T has the normal form of the lemma by taking an arbitrary orthogonal coordinate system in every $\mathcal{R}_i$. They form together an orthogonal system for the whole space.

We take now $\chi = 2n$ again. We consider the elements of the complex-group, which are unitary at the same time. This means for the infinitesimal group $\mathfrak{u}$, that we require in (45)

$$(50) \quad \overline{\rho_{ik}} = -\rho_{ki}, \quad \overline{\sigma_{ki}} = -\sigma_{ik}, \quad \overline{\rho'_{ik}} = -\sigma'_{ki}$$

It is important that we impose in (50) only reality conditions on the parameters.

The unitary complex group is a group in $n(2n+1)$ real parameters instead of the same number of complex parameters in $\mathcal{U}$.

Any linear form in $\rho_{ik}, \sigma_{ik}, \rho_{ik}^*, \sigma_{ik}^*$ which vanishes under the restriction (50), vanishes identically.

Lemma: Let T be a unitary element of the complex group $\mathcal{U}$. There exists a normal orthogonal coordinate system in which T has the form $\rho_i \rightarrow \varepsilon_i \rho_i$ ($i = 1, 2, \dots, n, 1', 2', \dots, n'$), $\varepsilon_{i'} = 1/\varepsilon_i$.

In other words: To every T of the unitary restricted complex group $\mathcal{U}_n$ belongs an element U of $\mathcal{U}_n$ such that $U^{-1}TU$ is a diagonal matrix.

Proof: Every vector x determines a vector x^* with

$$x_{i'}^* = \bar{x}_i, \quad x_i^* = -\bar{x}_{i'}$$

We have the following properties,

- (a) $(ax) = S(ax^*)$ for every vector a . This can be chosen as definition of x^*
- (b) $(x+y)^* = x^* + y^*$, $(\alpha x)^* = \bar{\alpha} x^*$ for any scalar α
- (c) $x^{**} = -x$
- (d) $(xx^*) = 0$
- (e) $(x^* y^*) = (yx)$

Let ε be an eigenvalue of T , $|\varepsilon| = 1$. Let $\mathcal{R}_\varepsilon$ be the eigenspace belonging to ε . We choose an orthogonal basis $\rho_1, \dots, \rho_\mu$ in $\mathcal{R}_\varepsilon$. Then T transforms $\rho_j \rightarrow \varepsilon \rho_j$, $j = 1, 2, \dots, \mu$.

If T transforms $x \rightarrow y$, it transforms $x^* \rightarrow y^*$ because the defining property (a) of x^* is invariant under transformations T of $\mathcal{U}_n$. Hence T transforms $\rho_j^* \rightarrow \bar{\varepsilon} \rho_j^*$. Therefore $\bar{\varepsilon}$ is eigenvalue of at least the multiplicity μ of T . The same argument applied to $\bar{\varepsilon}$ instead of ε shows that ε and $\bar{\varepsilon}$ have the same multiplicity μ .

1st case: ε real. We start with a vector e_1 of length 1, $(e_1, e_1) = 1$. We have $(e_1, e_1^*) = 0$ (d); $(e_1^*, e_1^*) = 1$, (e). Let e_2 be a vector of $\mathcal{R}_\varepsilon$ which is orthogonal to e_1 and e_1^* , $(e_1, e_2) = 0$, $(e_1^*, e_2) = 0$ and has the length 1, $(e_2, e_2) = 1$. Then $(e_1, e_2^*) = (e_1^*, e_2^*) = (e_2, e_2^*) = 0$, $(e_2^*, e_2^*) = 1$. We continue in this manner and get finally an orthogonal basis

$e_1, e_1^*, \dots, e_\nu, e_\nu^*$ of $\mathcal{R}_\varepsilon$, μ necessarily being even, $\mu = 2\nu$.

2d case: ε not real. If $e_1, \dots, e_\mu$ is an orthogonal basis of $\mathcal{R}_\varepsilon$, we take $e_1^*, \dots, e_\mu^*$ as basis of $\mathcal{R}_{\bar{\varepsilon}}$

We obtain finally an orthogonal basis $e_1, e_2, \dots, e_n, e_1^*, e_2^*, \dots, e_n^*$ of the whole space, which consists of eigenvectors of T : $e_i \rightarrow \varepsilon_i e_i$, $e_i^* \rightarrow \bar{\varepsilon}_i e_i^*$, $\bar{\varepsilon}_i = 1/\varepsilon_i$. The properties (a) and (c) of the operation $*$ show that this basis is normal at the same time if we set $e_{i+1} = e_i^*$. This proves our lemma.

Next, we prove a lemma about the topological structure of the manifold

Lemma: The group $\mathcal{L}_u$ is simply connected.

Proof. An element T of $\mathcal{L}_u$ will be called singular, if two eigenvalues coincide. We are going to show first that these singular elements form a point set of at least 3 dimensions less. We distinguish two cases: Case 1)

ε_i and ε_j or ε_i and ε'_j , with $i \neq j$, coincide. We may assume $\varepsilon_1 = \varepsilon_2$. The most general T of this type can be written in the form $T = \mathcal{U}(\varepsilon)\mathcal{U}^{-1}$,

with

$$(\varepsilon) = \begin{pmatrix} \varepsilon & & & \\ & \varepsilon & & \\ & & \varepsilon_3 & \\ & & & \ddots \\ & & & & \varepsilon'_1 \\ & & & & & \ddots \end{pmatrix}$$

Here, U is an arbitrary element of the $N = n(2n+1)$ dimensional manifold $\mathcal{L}_n$. But U and UA give the same T if A is commutative with (6). All matrices A

$$A = \begin{pmatrix} B & & \\ & \eta_2 & \\ & & \ddots & \\ & & & \eta_n & \\ & & & & B^{-1} & \\ & & & & & \overline{\eta_3} & \\ & & & & & & \ddots & \\ & & & & & & & \overline{\eta_n} \end{pmatrix}$$

where B is unitary of degree 2, have this property. These A form a manifold of $n+2$ dimensions. Hence the T belonging to a fixed (6) form an $N - (n+2)$ dimensional set. Now (6) depends on $n-1$ parameters. Consequently the singular T form a set of at most $N-3$ dimensions. Case 2) An e_i and an e'_i coincide, say $e_1 = e'_1$. Then $\varepsilon_1 = \pm 1$. A similar argument holds.

We have to prove that any closed curve γ in $\mathcal{L}_n$ can be shrunk to a point. We may assume that γ does not pass through any singular T because of the fact just proved. We mark the points $\varepsilon_\nu = e^{\pi i \varphi_\nu}$, $\varepsilon'_\nu = e^{-\pi i \varphi_\nu}$ on the unit circle. We may assume

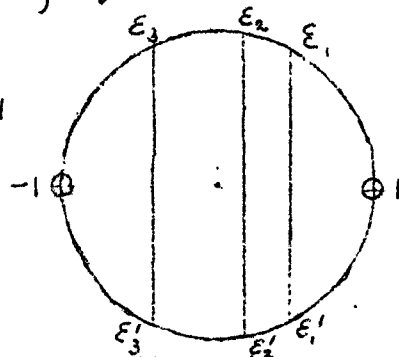
$$(51) \quad 0 < \varphi_1 < \varphi_2 < \dots < \varphi_n < 1$$

If T runs over γ , $\varphi_1, \varphi_2, \dots, \varphi_n$ will vary and so will U in $T = U(\varepsilon)U^{-1}$. But (51) will hold always, since $\varepsilon_1, \dots, \varepsilon_n, \varepsilon'_1, \dots, \varepsilon'_n$ are always distinct. When T will have returned to its initial position, then

$\varepsilon_1, \dots, \varepsilon_n, \varepsilon'_1, \dots, \varepsilon'_n$ will have the same position as at the beginning.

It is conceivable that the single $\varepsilon_i, \varepsilon'_j$ underwent a permutation after a complete circuit of T on γ . But this cannot happen for otherwise two

$\varepsilon_i, \varepsilon'_j$ would coincide at some stage. Hence every $\varepsilon_i, \varepsilon'_j$ returns to its initial position.



We replace every φ_i by $\lambda \varphi_i$. If (ε) ^{then} is replaced by (ε_λ) , we get

$$T_\lambda = U(\varepsilon_\lambda) U^{-1}$$

for every single T on γ . Then T runs over a closed curve γ_λ as T runs over γ , as is easily seen. This would not be so if the ε_ν underwent a permutation during a circuit of T on γ . If λ shrinks from 1 to 0, γ_λ shrinks from γ to γ_0 , which consists of the unit element only. Consequently, γ can be shrunk to a point as stated above; K_u is simply connected.

11. Construction of the irreducible representations of the complex-group

Let $\mathcal{G}$ be an arbitrary representation of the ^(infinite) complex-group $\mathcal{K}$

$$e_\alpha \rightarrow E_\alpha, \quad h = \lambda^1 h_1 + \dots + \lambda^n h_n \rightarrow H = \lambda^1 H_1 + \dots + \lambda^n H_n$$

The necessary and sufficient conditions for these matrices E_α , H_ν of degree N are that the equations (48) on p. 38 hold, if we replace e_α by E_α , h by H .

$$(52) \quad \begin{aligned} [H, H'] &= 0, \quad [H, E_\alpha] = \alpha E_\alpha \\ [E_\alpha, E_\beta] &= \begin{cases} 0 & , \quad \alpha + \beta \text{ no root, } \alpha \neq -\beta \\ E_{\alpha+\beta} & , \quad \alpha + \beta \text{ root} \\ H_\alpha = \alpha^1 H_1 + \dots + \alpha^n H_n, & \alpha = -\beta \end{cases} \end{aligned}$$

The vectors in the space of N dimensions in which $\mathcal{G}$ operates will be denoted by German letters. As on p. 13 for the case of the full linear group, we say that a vector $\mathcal{N}$ has a weight Λ if for each $H = \lambda^1 H_1 + \dots + \lambda^n H_n$ a formula

$$(53) \quad H \mathcal{N} = \Lambda \mathcal{N}$$

holds where Λ is a linear form in $\lambda^1, \dots, \lambda^n$. We call Λ a weight of the representation if a vector $\mathcal{N} \neq 0$ has the weight Λ .

We prove in the same manner as on pp. 13-20

Theorem 1. If Λ is a weight, $2 \frac{\Lambda_\alpha}{\alpha_\alpha}$ is an integer and $\Lambda - 2 \frac{\Lambda_\alpha}{\alpha_\alpha} \alpha$ also is a weight for each root α .

Theorem 2. An irreducible representation is uniquely determined by its highest weight. The highest weight occurs only once.

Remarks: If $\alpha = \pm \lambda^i \pm \lambda^k$, $i < k$, we have $\alpha^i = \pm 1$, $\alpha^k = \pm 1$, $\alpha^j = 0$ for $j \neq i, k$. Hence $\Lambda_\alpha = \pm m_i \pm m_k$ for any linear form $\Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n$. In particular $\alpha_\alpha = 2, \tilde{\Lambda} = \Lambda - 2 \frac{\Lambda_\alpha}{\alpha} \alpha = \Lambda - (\pm m_i \pm m_k)(\pm \lambda^i \pm \lambda^k)$. One sees that $\tilde{\Lambda}$ is obtained by interchanging λ^i and λ^k in the case $\pm \alpha = \lambda^i - \lambda^k$ and by replacing λ^i by $-\lambda^k$, λ^k by $-\lambda^i$ in the case $\pm \alpha = \lambda^i + \lambda^k$.

$$\begin{aligned} \text{If } \alpha = \pm 2 \lambda^i, \quad \Lambda_\alpha = \pm m_i, \quad \alpha_\alpha = 2, \quad 2 \frac{\Lambda_\alpha}{\alpha} = \pm m_i \\ \tilde{\Lambda} = \Lambda - \Lambda_\alpha \alpha = m_1 \lambda^1 + \dots + m_{i-1} \lambda^{i-1} - m_i \lambda^i + \\ + m_{i+1} \lambda^{i+1} + \dots + m_n \lambda^n. \end{aligned}$$

Hence, the coefficients of the weights are integers. If we permute the λ^i in a weight Λ or replace some λ^i by $-\lambda^i$ we obtain a weight again.

If

$$(54) \quad \Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n$$

is the highest weight of a representation, we have

$$(55) \quad m_1 \geq m_2 \geq \dots \geq m_n$$

automatically. We are going to show that to (54) belongs an irreducible representation with this highest weight Λ , if (55) is fulfilled.

We start from a normal coordinate system $e_1, e_2, \dots, e_n, e_1, e_2, \dots, e_n$. These vectors have, for the original group π itself, the weights $\lambda_1, \lambda_2, \dots, \lambda_n, -\lambda_1, \dots, -\lambda_n$ respectively. Every transformation of π induces in the space $\mathcal{K}_V$ of skew-symmetric tensors of rank V a linear transformation and we obtain a representation $\tilde{\Gamma}_V$ of π in this way,

$V = 1, 2, \dots, n$. Here, we have a slight difference compared with the case of the full linear group. In this case we had to take $\tilde{\Gamma}_V$ for all V which do not exceed the number of dimensions of the space of the original group, that

is $2n$ now.

The "vector"

$$[e_{i_1} e_{i_2} \dots e_{i_v}] , \quad i_1, i_2, \dots, i_v = 1, 2, \dots, n, 1', 2', \dots, n'$$

of $\mathcal{K}_v$ has a weight for Γ_v , namely

$$(56) \quad \lambda_{i_1} + \lambda_{i_2} + \dots + \lambda_{i_v} , \quad (\lambda_{p'} = -\lambda_s)$$

The expressions (55) are all the weights, as can be seen from Lemma 1, page 18, which holds for the complex group also. The highest possible weight is

$$\lambda_1 + \lambda_2 + \dots + \lambda_v$$

It occurs only once.

As in the case of the full linear group, we set

$$(57) \quad p_1 = m_1 - m_2, p_2 = m_2 - m_3, \dots, p_{n-1} = m_{n-1} - m_n, p_n = m_n$$

$$(58) \quad \Gamma = \Gamma_1^{p_1} \times \Gamma_2^{p_2} \times \dots \times \Gamma_n^{p_n}$$

The highest weight of Γ is

$$p_1 \lambda^1 + p_2 (\lambda^1 + \lambda^2) + \dots + p_n (\lambda^1 + \lambda^2 + \dots + \lambda^n) = \sum m_i \lambda^i$$

It occurs only once. Therefore, one irreducible constituent of Γ has the prescribed highest weight Λ

The "vector"

$$(59) \quad v = \underbrace{[u_1, u_2, \dots, u_n]}_{p_n \text{ factors}} \times \dots \times \underbrace{[u_1, u_2]}_{p_2 \text{ factors}} \times \dots \times \underbrace{[u_1, u_2]}_{p_1 \text{ factors}} \times u_1 \times u_1$$

is the only vector in the space of Γ with the weight Λ , (54).

We apply all the transformations of Γ to v and construct the linear space $\mathcal{K}$ spanned by all these vectors. Every transformation of κ induces a linear transformation in $\mathcal{K}$ and these form the representation $\mathcal{G}$ of κ with the highest weight Λ . The proof is analogous to the proof in the case of the full linear group. Again, we need the fact that every representation of

$\mathcal{K}$ is completely reducible. We prove this as in 7: We replace the complex-group $\mathcal{L}$ by the unitary restricted complex-group $\mathcal{L}_u$. This group is compact and we can make use of the Hermitian argument: Every representation of $\mathcal{L}_u$ is completely reducible.

Now we take the infinitesimal group $\mathcal{K}_u$ of $\mathcal{L}_u$. Every representation of $\mathcal{K}_u$ can be extended to a representation of $\mathcal{L}_u$ by integration. We obtain a representation of $\mathcal{L}_u$ (and not of a covering manifold) because $\mathcal{L}_u$ is simply connected (cf. 10, page 41).

Finally, we derive the complete reducibility of the representations of $\mathcal{K}$ from the analogous fact for $\mathcal{K}_u$. Every linear relation between the coefficients of an arbitrary element of $\mathcal{K}_u$ holds for the elements of an arbitrary element of $\mathcal{K}$ (cf. 10, page 40).

For the details, compare the proof on pp. 32-33 for the case of the full linear group.

12. Algebraic supplements to the above construction

We had to apply all elements of $\mathcal{L}$ on $\mathcal{V}$, (59), and to determine the smallest linear space $\mathcal{R}$ which contains all these transformed vectors. The representation $\mathcal{V}$ consists of all linear transformations of $\mathcal{R}$ which are induced by the transformations of $\mathcal{L}$.

A transformation C carries $\pi_1, \pi_2, \dots, \pi_n$ over into n vectors

$\rho_1, \rho_2, \dots, \rho_n$ with the property

$$(60) \quad S(\rho_i, \rho_k) = 0 \quad (i, k = 1, 2, \dots, n)$$

because $S(\pi_i, \pi_k) = 0$ and S is invariant under $\mathcal{L}$.

Then i is transformed into

$$(61) \quad \underbrace{[\rho_1, \rho_2, \dots, \rho_n]}_{p_1 \text{ factors}} \times \dots \times \underbrace{[\rho_1, \rho_2, \dots, \rho_n]}_{p_2 \text{ factors}} \times \dots \times \underbrace{[\rho_1, \rho_2]}_{p_3 \text{ factors}} \times \rho_1 \times \dots \times \rho_n$$

This is a tensor of rank $m = m_1 + m_2 + \dots + m_n$ whose components may be denoted by

$$(62) \quad \psi \begin{pmatrix} i_1, \dots, i_{m_3}, \dots, i_{m_2}, \dots, i_{m_1} \\ j_1, \dots, j_{m_3}, \dots, j_{m_2}, \dots, j_{m_1} \\ k_1, \dots, k_{m_3} \\ \dots \end{pmatrix}$$

analogously to (25), p. 27-28.

We replace every bracket $[p_1, p_2, \dots, p_e]$ in (61) by

$p_1 \times p_2 \times \dots \times p_e$ and get

$$(p_1 \times \dots \times p_n) \times \dots \times (p_1 \times \dots \times p_n) \times \dots \times (p_1 \times p_2) \times \dots \times (p_1 \times p_2)$$

On analogy to (26), p. 27-28, we denote the components by

$$(63a) \quad \psi \begin{pmatrix} i_1, \dots, i_{m_3}, \dots, i_{m_2}, \dots, i_{m_1} \\ j_1, \dots, j_{m_3}, \dots, j_{m_2} \\ k_1, \dots, k_{m_3} \\ \dots \end{pmatrix} \times p_1 \times \dots \times p_1$$

If $p_1 = (x_1, x_2, \dots, x_n)$, $p_2 = (y_1, y_2, \dots, y_n)$, $p_3 = (z_1, z_2, \dots, z_n)$, we find for (63) the value

$$(63b) \quad x_{i_1} x_{i_2} \dots x_{i_{m_1}} \cdot y_{j_1} y_{j_2} \dots y_{j_{m_2}} \cdot z_{k_1} z_{k_2} \dots z_{k_{m_3}} \dots$$

Again, ψ is obtained from φ by an operation A which consists of permuting the elements in the columns in (63a) and forming alternating sums

$$(64) \quad \psi = A \varphi$$

We determine the space which is spanned by all tensors ψ . We denote the $m = m_1 + m_2 + \dots + m_n$ indices $i_1, \dots, i_{m_1}, j_1, \dots, j_{m_2}, \dots$ by $t_1, \dots, t_m$ also and write (63) in the form

$$(63c) \quad \varphi(t_1, t_2, \dots, t_m)$$

We put

$$(65) \quad S(\tilde{M}, \Lambda) = \sum_{i, k} \epsilon(i, k) u_i v_k^{(i, k = 1, 2, \dots, n, 1', 2', \dots, n')}$$

for any two vectors $\vec{u} = (u_1, \dots, u_n)$ and $\vec{v} = (v_1, \dots, v_n)$,

$$\varepsilon(i, i') = 1, \quad \varepsilon(i', i) = -1 \quad (i = 1, 2, \dots, n)$$

$$\varepsilon(ik) = 0 \text{ in all other cases}$$

Then it follows from (60) that

$$\sum_{\substack{t_\mu, t_\nu = 1, 2, \dots, n, 1', 2', \dots, n'}} \varepsilon(t_\mu, t_\nu) \varphi(t_1, t_2, \dots, t_m) = 0$$

for any fixed μ, ν . We call the expressions on the left hand side the contractions of the tensor φ .

We denote by $\bar{\Phi}$ the most general tensor of rank m which 1) if written in the form (63a) is symmetric with respect to permutation of indices in every fixed row and 2) whose contractions all vanish. Every linear combination of tensors φ is such a $\bar{\Phi}$. We want to prove that, conversely, every $\bar{\Phi}$ can be represented as a linear combination of special tensors φ , (63). We need two lemmas ^{cf} the proof which will be given at the end of this section.

Lemma 1: Any n linearly independent vectors $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n$ with the properties $S(\mathcal{P}_i, \mathcal{P}_k) = 0$ can be obtained from $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_n$ by an appropriate complex-transformation C .

Hence the vectors $\mathcal{P}_1 = (x_1, \dots, x_n), \mathcal{P}_2 = (y_1, \dots, y_n), \dots$ underlie only the restriction (60) in (63).

Lemma 2: Let $H(\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n)$ be a homogeneous polynomial in the components of $\mathcal{P}_i$ for $i = 1, 2, \dots, n$, which vanishes if all

$S(\mathcal{P}_i, \mathcal{P}_k) = 0$. Then H can be represented in the form

$$(66) \quad H = \sum_{\rho, \sigma} S(\mathcal{P}_\rho, \mathcal{P}_\sigma) \mathcal{P}_{\rho\sigma}$$

where the $\mathcal{P}_{ik}$ are polynomials in the components of $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n$.

We assume now that the special tensors φ (63) all lie in a linear subspace $\mathcal{T}$ of the space $\mathcal{X}$ of all tensors $\bar{\Phi}$. Then the components of every φ satisfy a linear relation

$$(67) \quad \sum a \begin{pmatrix} i_1 & \dots & i_{m_1} \\ j_1 & \dots & j_{m_2} \\ \dots & \dots & \dots \end{pmatrix} \varphi \begin{pmatrix} i_1 & \dots & i_{m_1} \\ j_1 & \dots & j_{m_2} \\ \dots & \dots & \dots \end{pmatrix} = 0$$

where the constant coefficients can be assumed to be symmetric with respect to permutation of the indices in each fixed row. By using (63b) we see that the left-hand side is a polynomial H as in Lemma 2. Equation (67) shows then that $H = 0$ if $\rho_1, \dots, \rho_n$ are n linearly independent vectors with the property $S(\rho_i, \rho_k) = 0$, ($i, k = 1, 2, \dots, n$). According to Lemma 2, H can be written in the form (66). On comparing (66), (67) and taking (63b) and (65) into account, we obtain a formula

$$(68) \quad a \begin{pmatrix} i_1 & \dots & i_{m_1} \\ j_1 & \dots & j_{m_2} \\ \dots & \dots & \dots \end{pmatrix} = \varepsilon(i_1 j_1) b_1 + \varepsilon(i_1 j_2) b_2 + \dots$$

where b_1 does not depend on i_1, j_1 ; b_2 does not depend on $i_1, j_2, \dots$. If we form

$$(69) \quad \sum a \begin{pmatrix} i_1 & \dots & i_{m_1} \\ j_1 & \dots & j_{m_2} \\ \dots & \dots & \dots \end{pmatrix} \Phi \begin{pmatrix} i_1 & \dots & i_{m_1} \\ j_1 & \dots & j_{m_2} \\ \dots & \dots & \dots \end{pmatrix}$$

for any tensor Φ and use (68), we see that (69) vanishes because all contractions of Φ vanish. Hence (67) is true for Φ instead of φ . The subspace $\mathcal{T}$ of the space $\mathcal{X}$ is identical with $\mathcal{X}$, as we wished to show. Hence every tensor Φ is a linear combination of special tensors φ .

On performing the operation A , (64), on Φ , we obtain tensors $\Psi = A \Phi$ which are linear combinations of tensors ψ , (62). Conversely, every linear combination of ψ 's is such a Ψ .

The tensors Ψ form, therefore, the space $\mathcal{R}$. Every transformation $\mathcal{C}$ of the complex-group induces a linear transformation $\mathcal{G}(\mathcal{C})$ in $\mathcal{R}$. All

these $\mathcal{G}(\mathcal{C})$ form the irreducible representation $\mathcal{G}$ of $\mathcal{L}$ belonging to the given highest weight $m_1 \lambda^1 + \dots + m_n \lambda^n$.

Proof of Lemma 1: Let $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_k$ be any linearly independent vectors. The space $\mathcal{T}$ of the vectors A satisfying the k linear homogeneous equations $S(\mathcal{P}_i, A) = 0, i = 1, 2, \dots, k$, has at least $2n-k$ dimensions. But if the number of dimensions were greater than $2n-k$, we could find an $A \neq 0$ in $\mathcal{T}$ which satisfies

$$S(\mathcal{P}_{k+1}, A) = 0, \dots, S(\mathcal{P}_{2n}, A) = 0$$

where $\mathcal{P}_{k+1}, \dots, \mathcal{P}_{2n}$ complete $\mathcal{P}_1, \dots, \mathcal{P}_k$ to a basis of the whole space. Then A would satisfy $S(\mathcal{P}, A) = 0$ for all $\mathcal{P}$ whereas $\mathcal{X}$ is not degenerate.

We assume now that n linearly independent vectors $\mathcal{P}_1, \dots, \mathcal{P}_n$ fulfill the conditions $S(\mathcal{P}_i, \mathcal{P}_k) = 0$ for all i, k . We are going to show that we can find n vectors $\mathcal{P}_{1'}, \dots, \mathcal{P}_{n'}$ which, together with $\mathcal{P}_1, \dots, \mathcal{P}_n$ form a normal basis. Then the linear transformation C which transforms

$$\mu_1 \rightarrow \mathcal{P}_1, \dots, \mu_n \rightarrow \mathcal{P}_n, \mu_{1'} \rightarrow \mathcal{P}_{1'}, \dots, \mu_{n'} \rightarrow \mathcal{P}_{n'}$$

necessarily leaves S invariant; C is a complex-transformation. We assume that we have already found $\mathcal{P}_{1'}, \dots, \mathcal{P}_{r'}$ such that

$\mathcal{P}_1, \dots, \mathcal{P}_n, \mathcal{P}_{1'}, \dots, \mathcal{P}_{r'}$ are linearly independent and that S vanishes for any two of them except $S(\mathcal{P}_p, \mathcal{P}_{p'}) = 1, S(\mathcal{P}_{p'}, \mathcal{P}_p) = -1,$
($p = 1, 2, \dots, r$). The equations

$$(70) \quad S(\mathcal{P}_1, A) = 0, \dots, S(\mathcal{P}_r, A) = 0, S(\mathcal{P}_{r+2}, A) = 0, \dots, S(\mathcal{P}_n, A) = 0, \\ S(\mathcal{P}_{1'}, A) = 0, \dots, S(\mathcal{P}_{r'}, A) = 0$$

determine a linear space of $2n - (n-1+r) = n+1-r$ dimensions, but if we add the further equations $S(\mathcal{P}_{r+1}, A) = 0$, we have one dimension less. Consequently there exists a solution of (70) for which $S(\mathcal{P}_{r+1}, A) \neq 0$. We can

assume $S(p_{r+1}, A) = 1$ and set $p_{(r+1)} = A$. Continuing in this manner we obtain finally a normal basis $p_1, \dots, p_n, p_1, \dots, p_n$.

Proof of Lemma 2: We consider polynomials $H(p_1, \dots, p_n)$ which are homogeneous in the components of p_i of the given degree m_i . These form a linear "vector"-space F . The polynomials H which vanish if $S(p_i, p_k) = 0$, ($i, k = 1, 2, \dots, n$), constitute a linear subspace $\mathcal{H}$. Every transformation $\mathcal{C}$ of $\mathcal{L}$ induces a linear transformation $\mathcal{L}_{\mathcal{C}}$ in F and these $\mathcal{L}_{\mathcal{C}}$ form a representation $\mathcal{L}$ of $\mathcal{L}$. The subspace $\mathcal{H}$ is invariant under $\mathcal{L}$. Then F is the direct sum of $\mathcal{H}$ and another invariant subspace $\mathcal{H}'$, because the theorem of complete reducibility holds for the representations of $\mathcal{L}$.

$$(71) \quad F = \mathcal{H} + \mathcal{H}'$$

Let $y^1, y^2, \dots, y^n$ be n vectors which are transformed contragrediently to $p_1, p_2, \dots, p_n$. Then,

$$f = (p_1 y^1)^{m_1} (p_2 y^2)^{m_2} \dots (p_n y^n)^{m_n}$$

(where $(p y)$ is the inner product of p and y) is invariant under $\mathcal{L}$.

It is of the given degrees in the components of $p_1, \dots, p_n$. We may write f in the form

$$(72) \quad f = \sum_J n_J \mathcal{X}_J y_J$$

where $\mathcal{X}_J$ is a monomial in the components of $p_1, \dots, p_n$ of the prescribed degrees, where y_J is the corresponding monomial in the components of $y^1, \dots, y^n$ and n_J is a positive integer.

If we represent $\mathcal{X}_J$ according to (71) we obtain a formula

$$(73) \quad \mathcal{X}_J \equiv P_J \pmod{\mathcal{H}'}, \quad P_J \text{ in } \mathcal{H}$$

Hence

$$(74) \quad f \equiv \sum_J n_J P_J y_J \pmod{\mathcal{H}'}$$

On applying a transformation $\mathcal{C}$ of $\mathcal{L}$ in the original space and consequently the induced $\mathcal{L}_{\mathcal{C}}$ in F , we obtain

$$(75) \quad L_{\rho}(f) \equiv \sum_J n_J L_{\rho}(P_J y_J) \pmod{\mathfrak{g}'} \quad (\text{mod } \mathfrak{g}')$$

since $\mathfrak{g}'$ is invariant. Now f is invariant, $L_{\rho}(f) = f$. Then comparison of (74) and (75) gives

$$\sum_J n_J P_J y_J \equiv \sum_J n_J L_{\rho}(P_J y_J) \pmod{\mathfrak{g}'}$$

The left side belongs to $\mathfrak{g}$ and so does the right side, since $\mathfrak{g}$ also is invariant under L_{ρ} . The difference of both sides in this congruence belongs to $\mathfrak{g}$ and $\mathfrak{g}'$ at the same time. But $\mathfrak{g}$ and $\mathfrak{g}'$ have only 0 in common, the sum in (71) being direct. Hence we may replace the $\equiv$ sign by an equality sign and derive $P_J y_J = L_{\rho}(P_J y_J)$, $P_J y_J$ invariant under all transformations of L .

The fundamental theorem of the theory of invariants* shows that every

* For a simple proof for the case of the complex group, cf. H. Weyl, Mathematische Zeitschrift 20, p. 140. It is sufficient to prove this theorem for the case that only cogredient vectors p_i occur in the invariant, in which case the $S(p_i, p_k)$ alone are sufficient to express the invariants.

A contragredient vector $y = (y_1, \dots, y_n, y_1', \dots, y_n')$ is invariantly connected with a cogredient vector

$$y_* = (-y_1', \dots, -y_n', y_1, \dots, y_n) \text{ and } S(y^1, y^2) = S(y_*^1, y_*^2), \quad (p, y) = S(p, y_*).$$

such invariant can be expressed as an integral rational function of

$$S(p_i, p_k), \quad S(y_i^1, y_k^2), \quad (p_i, y_k^r)$$

Since $P_J y_J$ is of the same degree in $p_1, \dots, p_n$ as in

$y_1^1, \dots, y_n^1$, terms of the second type always appear multiplied by at least one factor of the first kind. We denote the set of all expressions

$$\sum S(p_i, p_k) u_{ik} \quad (u_{ik} \text{ polynomials in } p_i, y^r)$$

by $\mathfrak{Q}$. Here $\mathfrak{Q}$ is, by the way, an ideal, the ideal with the basis

$S(p_i, p_k), \quad i, k = 1, 2, \dots, n$. It is contained in $\mathfrak{g}$. Then $P_J y_J$ is congruent (mod $\mathfrak{Q}$) to an expression in which only terms of the third kind

occur.

$$(76) \quad P_J Y_J \equiv G(P_1 y^1, P_1 y^2, \dots, P_n y^n) \pmod{\mathcal{Q}}$$

We put $P_1 = n_1, \dots, P_n = n_n$. We have $S(n_1, n_n) = 0$, and hence all polynomials of $\mathcal{Q}$ and of $\mathcal{H}$ vanish. In particular, $P_J = 0$, since

$$P_J \text{ belongs to } \mathcal{H}. \text{ If } y^r = (y_1^r, y_2^r, \dots, y_n^r), \text{ we obtain from (76)}$$

$$0 = G(y_1^1, y_1^2, \dots, y_1^n, y_2^1, \dots, y_2^n, \dots, y_n^1, \dots, y_n^n)$$

Here the arguments are absolutely arbitrary and consequently G vanishes identically. Then (75) shows that

$$P_J Y_J \equiv 0 \pmod{\mathcal{Q}}$$

We choose $y_1^1, \dots, y_n^n$ such that $Y_J \neq 0$ and obtain $P_J \equiv 0 \pmod{\mathcal{Q}}$; every P_J belongs to $\mathcal{Q}$.

The elements H of $\mathcal{F}$ can be written in the form $\sum a_J X_J$ with constant a_J and we obtain from (73)

$$H = \sum a_J X_J \equiv \sum a_J P_J \pmod{\mathcal{H}'}$$

If H belongs to $\mathcal{H}$, so does $H - \sum a_J P_J$. But the latter belongs to $\mathcal{H}'$ and vanishes, as a common element of $\mathcal{H}$ and $\mathcal{H}'$.

$$H = \sum a_J P_J$$

With P_J , H belongs to $\mathcal{Q}$, q.e.d.

We used the fact that the representations of $\mathcal{L}$ are completely reducible. We can avoid this in the following manner: If $F = \sum a_J X_J$, $G = \sum b_J X_J$ are two elements of $\mathcal{F}$, we define a positive definite metric by $(F, G) = \sum \bar{a}_J b_J$. This metric is invariant for the unitary complex-group and then for the whole group $\mathcal{L}$ as can be seen by going back to infinitesimal transformations or directly.

13. Primitive characters of the complex group.

The invariant volume element in the manifold $\mathcal{L}_u$ (unitary complex group) is defined in the following way. We start with N linearly independent infinitesimal vectors, issuing from the unit element. These are N infinitesimal elements δu of the group, and $N = n(2n+1)$ is the number of (real) dimensions. The volume of the parallelopiped spanned by these vectors is the absolute value of the determinant of the parameters. The infinitesimal elements being of the form

$$(77) \quad \delta u = \left\| \begin{array}{c|c} p_{ik} & p'_{ik} \\ \hline \sigma'_{ik} & \sigma_{ik} \end{array} \right\| \quad \left\{ \begin{array}{l} p_{ik} + \bar{\sigma}_{ki} = 0, p'_{ik} = \bar{p}_{ki} \\ \sigma'_{ik} = \bar{\sigma}_{ki} \\ p_{ik} = -\bar{p}_{ki}, \sigma_{ik} = -\bar{\sigma}_{ki}, \sigma'_{ik} = \bar{\sigma}'_{ki} \end{array} \right.$$

we can choose as parameters the real and imaginary parts

$$(78) \quad R(p_{ik}), i < k; \quad I(p_{ik}), i < k; \quad R(p'_{ik}), i \leq k; \quad I(p'_{ik}), i \leq k$$

together with

$$(79) \quad I(p_{ii})$$

If u is any element of $\mathcal{L}_u$, δu determines a neighboring element of the group

$$(80) \quad u + du = (1 + \delta u)u, \quad du = \delta u \cdot u$$

If, conversely, $u + du$ is a neighboring element of u , $\delta u = du \cdot u^{-1}$ is an infinitesimal element of the group.

If a parallelopiped in u is given by n independent du , we define the volume as the volume of the corresponding δu , (80), issuing from the unit element. It is clear from this definition that this volume measure is invariant under right-hand translation in $\mathcal{L}_u$. It can be easily seen, that it is invariant under left-hand translation also (cf. p.77 of the notes of last year's course).

The group $\mathcal{L}_u$ contains an Abelian subgroup consisting of all

$$(80) \quad (\varepsilon) = \begin{pmatrix} \varepsilon_1 & & \\ & \ddots & \\ & & \varepsilon_n \\ & & & \varepsilon_1^{-1} \\ & & & & \ddots \\ & & & & & \varepsilon_n^{-1} \end{pmatrix}, \quad \varepsilon_v = e^{2\pi i \varphi_v} \quad \text{of absolute value one.}$$

The corresponding infinitesimal elements are

$$(81) \quad 2\pi i (d\varphi) = 2\pi i \begin{pmatrix} d\varphi_1 & & \\ & \ddots & \\ & & d\varphi_n \\ & & & -d\varphi_1 \\ & & & & \ddots \\ & & & & & -d\varphi_n \end{pmatrix}.$$

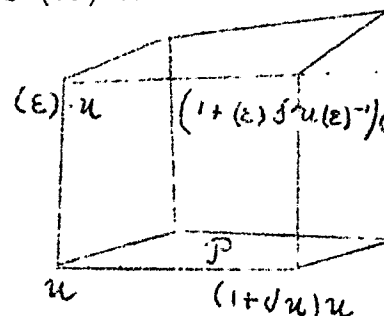
Here only the parameters (79) are different from 0.

We will say that two elements u and v of $\mathcal{L}_n$ are equivalent if $v = (\varepsilon)u$, (ε) of form (80). We say that all elements equivalent to u form the straight line or the vertical $[u]$. All straight lines $[u]$ constitute a manifold $[\mathcal{L}_n]$ of $N - n = 2n^2$ dimensions.

Consider a point u and the vertical through $u + du$. Then $(1 + 2\pi i (d\varphi))(u + du) = (1 + 2\pi i (d\varphi))(1 + \delta u)u = (1 + \delta u + 2\pi i (d\varphi))u$ lies also on $[u + du]$ in the neighborhood of u . From (77) follows that the elements in the main diagonal of δu are purely imaginary and the last n elements are the negatives of the first n elements. Therefore we may choose $(d\varphi)$ uniquely such that $\delta u + 2\pi i (d\varphi)$ has 0's in the main diagonal. We denote this $\delta u + 2\pi i (d\varphi)$ by $\dot{\delta}u$ and call the transition from u to $(1 + \dot{\delta}u)u$ the horizontal transition from u to the vertical $[u + du]$. The $\dot{\delta}u$ are characterized by the fact that the parameters (79) vanish.

We choose $N - n$ linearly independent $\dot{\delta}u$. They determine a "horizontal" parallelogram $\mathcal{P}$ in u .

The verticals passing through points of $\mathcal{P}$ form a "tube". We will define a volume measure in



$[L_u]$ by defining the volume of such a tube. We take (78) as the parameters for the δu ; the volume of the tube is the determinant of the parameters of the defining δu . But we have to show that this definition is independent of the special point u on $[u]$. Let $(\varepsilon)u$ be another point on $[u]$. Then

$$(\varepsilon)(1 + \delta u)u = (1 + (\varepsilon)\delta u (\varepsilon)^{-1}) \cdot (\varepsilon)u$$

lies on $[(1 + \delta u)u]$ in the neighborhood of $(\varepsilon)u$. The main diagonal of $(\varepsilon)\delta u (\varepsilon)^{-1}$ contains only 0 and we have therefore the horizontal transition from $(\varepsilon)u$ to $[(1 + \delta u)u]$. We see from (77) and (80) that in $(\varepsilon)\delta u (\varepsilon)^{-1}$ the ρ_{ik} and ρ'_{ik} are to be replaced.

$$(82) \quad \rho_{ik} \text{ by } \frac{\varepsilon_i}{\varepsilon_k} \rho_{ik}, \quad \rho'_{ik} \text{ by } \rho'_{ik} \cdot \varepsilon_i \varepsilon_k.$$

One sees easily that the determinant of the parameters (78) for the $N = n$ vectors

δu is multiplied by the squares of the absolute values of all $\varepsilon_i \varepsilon_k$ and $\varepsilon_i / \varepsilon_k$, which is 1. Hence the volume measure is independent of the special choice of u on $[u]$, and we denote the measure by $d[u]$.

Every element t of L_u can be written in the form

$$(83) \quad t = u^{-1}(\varepsilon)u$$

where u belongs to L_u and (ε) is of the form (80) (cf. p. 40).

Here u is not uniquely determined. We may replace u by any element $(\varepsilon^*)u$ on $[u]$.

The ε_i are the characteristic roots of t . We call the φ_i the angles of t . We investigate how u and (ε) are transformed if t is replaced by a neighboring element $t + dt$. We derive from $u t = (\varepsilon)u$ that

$$du \cdot t + u \cdot dt = d(\varepsilon)u + (\varepsilon) \cdot du$$

On multiplying by $t^{-1} u^{-1} = u^{-1}(\varepsilon)^{-1}$ on the right, we derive

$$du u^{-1} + u (dt \cdot t^{-1}) u^{-1} = d(\varepsilon) \cdot (\varepsilon)^{-1} + (\varepsilon) (du u^{-1}) (\varepsilon)^{-1}$$

Here, $d(\varepsilon) \cdot \varepsilon^{-1}$ is an infinitesimal element $2\pi i (\delta\varphi)$, (81), and $\delta\varphi_i$ is the increment of the angle φ_i in going over from t to $t + dt$

$$\begin{aligned}\delta u + u \cdot \delta t \cdot u^{-1} &= 2\pi i (\delta \varphi) + (\varepsilon) \delta u (\varepsilon)^{-1} \\ u \cdot \delta t \cdot u^{-1} &= ((\varepsilon) \delta u (\varepsilon)^{-1} - \delta u) + 2\pi i (\delta \varphi).\end{aligned}$$

We can replace δu by $\dot{\delta} u$ since the left side is not changed by adding elements in the main diagonal of δu . We denote $u \cdot \delta t \cdot u^{-1}$ by $\underline{\delta t}$

$$(84) \quad \underline{\delta t} = ((\varepsilon) \dot{\delta} u \cdot (\varepsilon)^{-1} - \dot{\delta} u) + 2\pi i (\delta \varphi)$$

We are interested in the volume of the set $\mathcal{V}$ for all $\tilde{u} = \tilde{u}^{-1}(\tilde{\varepsilon})\tilde{u}$ for which $\tilde{u}$ lies in a horizontal parallelogram $\mathcal{P}$ with the vertex u , and the angles $\tilde{\varphi}_i$ lie between φ_i and $\varphi_i + d\varphi_i$. We see from (84) that $\mathcal{V}$ is a parallelepiped which is spanned by the following vectors: a) We take the $N - n$ vectors $\dot{\delta} u$ which define $\mathcal{P}$, determine $\underline{\delta t}$ for every one from (84) with $\delta \varphi = 0$, and put $\delta t = u^{-1} \cdot \underline{\delta t} \cdot u$; b) We set $\delta \varphi_i = d\varphi_i$ for a fixed i , $\delta \varphi_k = 0$ for $i \neq k$ and $\dot{\delta} u = 0$. We obtain n vectors $\underline{\delta t}$ and $\delta t = u^{-1} \cdot \underline{\delta t} \cdot u$ for $i = 1, 2, \dots, n$. We have to calculate the determinant of the parameters (78) and (79) of these N vectors δt . But we may take the N vectors $\underline{\delta t} = u \cdot \delta t \cdot u^{-1}$ because of the invariance of the volume element for left and right hand translation: The volume of the parallelepiped spanned by the δt at the unit element does not change if we perform first a left-hand translation u and then a right-hand translation u^{-1} ; its vertex after these operations is the identity again.

The determinant of the parameters of the $\underline{\delta t}$ has the form

$$d\omega = \begin{vmatrix} \mathbb{D} & 0 & \dots & 0 \\ 0 & d\varphi_1 & & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d\varphi_n \end{vmatrix}$$

where $\mathbb{D}$ is the matrix of degree $N - n$ formed by the $N - n$ parameters (78) of $(\varepsilon) \dot{\delta} u (\varepsilon)^{-1} - \dot{\delta} u$ for the $N - n$ vectors which define the horizontal parallelogram $\mathcal{P}$. If $\dot{\delta} u$ belongs to $\rho_{ik}, \rho_{ik}^{-1} (\varepsilon) \dot{\delta} u (\varepsilon)^{-1} - \dot{\delta} u$ be-

longs to $\rho_{ik}(\varepsilon_i/\varepsilon_k - 1)$, $\rho'_{ik}(\varepsilon_i \varepsilon_k - 1)$ (cf. (82)), and by a similar argument to that used in connection with (82), one sees that the determinant $\mathcal{D}$ is equal to the determinant $\mathcal{D}^*$ of the parameters (78) for the $\hat{\delta}u$ multiplied by the squares of the absolute values of all the factors $\varepsilon_i \varepsilon_k - 1$, $i \leq k$ and $\varepsilon_i/\varepsilon_k - 1$, $i \neq k$. But $\mathcal{D}^*$ is the volume $d[u]$ of the horizontal parallelogram $\mathcal{P}$. Hence

$$d\omega = \prod_{i < k} |\varepsilon_i/\varepsilon_k - 1|^2 \cdot \prod_{i \leq k} |\varepsilon_i \varepsilon_k - 1|^2 \cdot d[u] d\varphi_1 \dots d\varphi_n.$$

After integrating over $[L_u]$ we obtain for the volume of the set of elements whose angles lie between φ_i and $\varphi_i + d\varphi_i$:

$$d\mathcal{Q} = \prod_{i < k} |\varepsilon_i/\varepsilon_k - 1|^2 \prod_{i \leq k} |\varepsilon_i \varepsilon_k - 1|^2 d\varphi_1 \dots d\varphi_n$$

apart from a constant factor, the volume of $[L_u]$ which we may put equal to 1.

The factor of $d\varphi_1 \dots d\varphi_n$ is the square of the absolute value of

$$\begin{aligned} \prod_{i < k} (\varepsilon_i/\varepsilon_k - 1) \prod_{i \leq k} (\varepsilon_i \varepsilon_k - 1) &= \prod_{i < k} (\varepsilon_i/\varepsilon_k - 1)(\varepsilon_i \varepsilon_k - 1) \prod_i (\varepsilon_i^2 - 1) \\ &= \prod_{i < k} \varepsilon_i (\varepsilon_i + \varepsilon_i^{-1} - \varepsilon_k - \varepsilon_k^{-1}) \prod_i \varepsilon_i (\varepsilon_i - \varepsilon_i^{-1}) \end{aligned}$$

or the square of the absolute value of

$$\Delta = \prod_{i < k} ((\varepsilon_i + \varepsilon_i^{-1}) - (\varepsilon_k + \varepsilon_k^{-1})) \prod_i (\varepsilon_i - \varepsilon_i^{-1})$$

We set $e(\varphi) = e^{2\pi i \varphi}$

$$\begin{aligned} S(\varphi) &= e(\varphi) - e(-\varphi) = 2i \sin 2\pi \varphi, \quad c(\varphi) = e(\varphi) + e(-\varphi) = \\ &= 2 \cos 2\pi \varphi \end{aligned}$$

A simple computation yields

$$\begin{aligned} \Delta &= \prod_{i < k} (c(\varphi_i) - c(\varphi_k)) \prod_i S(\varphi_i) = |c^{n-1}(\varphi), c^{n-2}(\varphi), \dots, 1| \cdot \prod_i S(\varphi_i) \quad (*) \\ &= |c((n-1)\varphi), c((n-2)\varphi), \dots, 1| \cdot \prod_i S(\varphi_i) = |S(n\varphi), S((n-1)\varphi), \dots, S(\varphi)| \end{aligned}$$

* One obtains the n rows of the determinant by replacing φ by

$$\varphi_1, \varphi_2, \dots, \varphi_n.$$

Thus we have the result

$$(85) \quad d\Omega = \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n, \quad \Delta = |s(n\varphi), s((n-1)\varphi), \dots, s(\varphi)|$$

Let $\mathcal{G}$ be any representation of $\mathcal{L}_n$. We may assume that all elements of $\mathcal{G}$ are unitary. The elements $(\mathcal{E})$, (80), form an Abelian subgroup of $\mathcal{L}_n$. We can assume that the corresponding elements (E) in $\mathcal{G}$ are diagonal matrices also, because they form a unitary Abelian group

$$(E) = \begin{pmatrix} E_1 & & \\ & \ddots & \\ & & E_n \end{pmatrix}$$

Every E_s is a function $f(\varphi_1, \dots, \varphi_n)$ of $(\mathcal{E})$ which, because of the property of representation satisfies the conditions

$$f(\varphi_1 + \varphi'_1, \dots, \varphi_n + \varphi'_n) = f(\varphi_1, \dots, \varphi_n) f(\varphi'_1, \dots, \varphi'_n).$$

f periodic in φ_p with the period $2\pi i$

Hence, $f(\varphi_1, \dots, \varphi_n) = f(\varphi_1, 0, \dots, 0) \cdot f(0, \varphi_2, 0, \dots, 0) \dots f(0, \dots, 0, \varphi_n)$. Every $g(\varphi) = (0, \dots, 0, \varphi, 0, \dots, 0)$ satisfies $g(\varphi) \cdot g(\varphi') = g(\varphi + \varphi')$ and is consequently of the form $g(\varphi) = e^{k\varphi}$, k constant. The periodicity yields that $k = 2m\pi i$, m integer. Thus we see that

$$f(\varphi_1, \dots, \varphi_n) = e^{(m_1\varphi_1 + \dots + m_n\varphi_n)}, \quad m_i \text{ integers.}$$

The character of $\mathcal{G}$, that is the trace of (E) , is a sum

$$\chi(\varphi) = \sum e^{(m_1\varphi_1 + \dots + m_n\varphi_n)}$$

ranging over a definite set of lattice points $(m_1, \dots, m_n)$. It is sufficient to consider diagonal matrices $(\mathcal{E})$ in $\mathcal{L}_n$. Every other element can be written in the form $t = u^{-1}(\mathcal{E})u$, u in $\mathcal{L}_n$. If in $\mathcal{G} : u \rightarrow \mathcal{U}$, we have $u^{-1}(\mathcal{E})u \rightarrow \mathcal{U}^{-1}(E)\mathcal{U}$ and the trace is the same as in (E) . The φ_i in $(\mathcal{E})$ are the angles of the element t .

A permutation of the φ_i or a change of sign for some φ_i means replacing (ε) by a $u^{-1}(\varepsilon)u$. Therefore $\chi(\varphi)$ is not altered. We may derive this also from the results of 11, p. 44.

Let $\chi(\varphi)$ be an irreducible character. Then we have (cf. the introduction, p. 10)

$$\int_{\mathcal{L}_u} \bar{\chi}(\varphi) \chi(\varphi) = 1$$

if the volume measure is normalized so that $\mathcal{L}_u$ has the volume 1. Otherwise 1 on the right side is to be replaced by the volume V of $\mathcal{L}_u$.

The subset $\mathcal{T}$ of $\mathcal{L}_u$ of all t where the angles lie between φ_i and $d\varphi_i$ gives $\bar{\chi} \chi d\Omega$ to the integral, χ depending only on the φ_i . Here $d\Omega$ may be expressed by (85). We determine the volume V by putting

$\chi = 1$ (character of the representation $u \rightarrow 1$) and obtain

$$(86) \quad \int_0^1 \dots \int_0^1 \bar{\chi}(\varphi) \chi(\varphi) \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n = \int_0^1 \dots \int_0^1 \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n$$

The product

$$(87) \quad \chi(\varphi) \Delta = \chi(\varphi) \cdot |S(n\varphi), \dots, S(2\varphi), S(\varphi)| = \xi(\varphi)$$

is a finite Fourier series with integral coefficients. Let (S) denote the finite group of linear transformations in a $\varphi_1, \dots, \varphi_n$ -space which is generated by a) permutations of $\varphi_1, \dots, \varphi_n$, b) changes of sign in some of the φ_i 's. The character $\chi(\varphi)$ is invariant under (S) as we have seen. The determinant Δ is alternating, and invariant under a subgroup (S_1) of half the order of (S) . The other transformations of (S) change the sign of Δ , and we denote them by (S_2) . We call a finite Fourier series antisymmetric if it behaves like Δ . For instance, $\xi(\varphi)$ is antisymmetric. Furthermore, for any system $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$,

$$(88) \quad \xi(l_1, \dots, l_n) = \sum_{\sigma \in (S_1)} \sigma e(l_1 \varphi_1 + \dots + l_n \varphi_n) - \sum_{\sigma \in (S_2)} \sigma e(l_1 \varphi_1 + \dots + l_n \varphi_n)$$

is antisymmetric. Any antisymmetric finite Fourier series, for instance $\xi(\varphi)$, is a linear combination of such $\xi(\lambda_1, \dots, \lambda_n)$.

$$(89) \quad \xi(\varphi) = c \xi(\lambda_1, \dots, \lambda_n) + c' \xi(\lambda'_1, \dots, \lambda'_n) + \dots$$

If $c e(\lambda_1 \varphi_1 + \dots + \lambda_n \varphi_n)$ is the highest term in (87), we see easily that

$c > 0$ and $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$. Exactly the same c and the same

$\lambda_1, \dots, \lambda_n$ occur in the "highest" term on the right side in (89). A simple computation gives

$$(90) \quad \int_0^1 \dots \int_0^1 \xi(k_1, \dots, k_n) \xi(k'_1, \dots, k'_n) d\varphi_1 \dots d\varphi_n = \begin{cases} 0, & (k_1, \dots, k_n) \neq (k'_1, \dots, k'_n) \\ 2^n n!, & (k_1, \dots, k_n) = (k'_1, \dots, k'_n) \end{cases}$$

From (88) we derive easily

$$(91) \quad \xi(\lambda_1, \dots, \lambda_n) = \sum^+ s(\lambda_1 \varphi_1), \dots, s(\lambda_n \varphi_n) = |s(\lambda_1 \varphi), \dots, s(\lambda_n \varphi)|$$

where, in the middle term, the sum is to be extended over all permutations of

$\varphi_1, \dots, \varphi_n$. In particular, $\Delta = \xi(n, n-1, \dots, 1)$. Then (90) yields

for $k_1 = n, \dots, k_n = 1, k'_1 = n, \dots, k'_n = 1$

$$(92) \quad \int_0^1 \dots \int_0^1 \Delta \bar{\Delta} d\varphi_1 \dots d\varphi_n = 2^n n!$$

We put the value for $\xi(\varphi) = \chi(\varphi) \Delta$ from (89) into (86). From (90) and

(92) follows

$$(93) \quad 2^n n! (c \bar{c} + c' \bar{c}' + \dots) = 2^n n!$$

Now c as the highest coefficient in $\xi(\varphi)$ is an integer and we saw that $c > 0$.

Hence (93) gives $c = 1, c' = c = \dots = 0$ and

$$\xi(\varphi) = \xi(\lambda_1, \dots, \lambda_n)$$

and consequently, according to (87) and (91)

$$(94) \quad \chi(\varphi) = \frac{|s(l_1 \varphi), \dots, s(l_n \varphi)|}{|s(n \varphi), \dots, s(1 \varphi)|}$$

If $e(m_1 \varphi_1 + \dots + m_n \varphi_n)$ is the highest term in $\chi(\varphi)$, we have

$$(m_1 + n) \varphi_1 + (m_2 + (n-1)) \varphi_2 + \dots + (m_n + 1) \varphi_n = \lambda_1 \varphi_1 + \dots + \lambda_n \varphi_n$$

because $|S(k_1\varphi), S(k_2\varphi), \dots, S(k_n\varphi)|$ has the highest term $c(k_1\varphi_1 + k_2\varphi_2 + \dots + k_n\varphi_n)$ for $k_1 > k_2 > \dots > k_n$. Hence (94) has the highest term $c(m_1\varphi_1 + \dots + m_n\varphi_n)$ for

$$(95) \quad \chi_1 = m_1 + n, \chi_2 = m_2 + n - 1, \dots, \chi_n = m_n + 1$$

We know that there is an irreducible character $\chi(\varphi)$ with the highest term $c(m_1\varphi_1 + \dots + m_n\varphi_n)$; therefore it must be (94) where the χ_i are given by (95).

Theorem. The characters of the irreducible representations of L_n are given by

$$\frac{|S((m_1+n)\varphi), S((m_2+n-1)\varphi), \dots, S((m_n+1)\varphi)|}{|S(n\varphi), S((n-1)\varphi), \dots, S(\varphi)|}$$

where φ is to be replaced by the angles of an arbitrary element of L_n

In order to obtain a second form of this theorem, we set

$$f(\lambda) = |1 - \lambda t|, \quad t \text{ in } L_n$$

$$\frac{1}{f(\lambda)} = p_0 + p_1\lambda + p_2\lambda^2 + \dots$$

The p_j are polynomials in the coefficients of t , and we define $p_j = 0$ for $j < 0$. Then (94) can be expressed as the determinant of degree n

$$(96) \quad |p_{\chi-n}, p_{\chi-n+1} + p_{\chi-n-1}, p_{\chi-n+2} + p_{\chi-n-2}, \dots|$$

The computation is given in H. Weyl, Math. Zeitschrift 24, p. 341. We can drop the unitary restriction now because of the integral rational character of (96).

Theorem. The character of the representation of L , which belongs to the highest weight $m_1\lambda^1 + \dots + m_n\lambda^n$, $m_1 \geq m_2 \geq \dots \geq m_n \geq 0$ is given by

$$|p_{\chi-n}, p_{\chi-n+1} + p_{\chi-n-1}, p_{\chi-n+2} + p_{\chi-n-2}, \dots|$$

where $\chi_i = m_i + n + 1 - i$.

We determine the degree R of the representation. This is the value of the character for $t = 1$. We put $\varphi_1 = n\varphi$, $\varphi_2 = (n-1)\varphi$, $\dots$, $\varphi_n = \varphi$. Then we have, according to (94)

$$\xi(l_1, \dots, l_n) = \left| \begin{array}{ccc} S(n l_1, \varphi), & \dots, & S(n l_n, \varphi) \\ S((n-1) l_1, \varphi), & \dots, & S((n-1) l_n, \varphi) \\ \vdots & & \vdots \\ S(l_1, \varphi), & \dots, & S(l_n, \varphi) \end{array} \right| = \prod_i S(l_i, \varphi) \left| \begin{array}{ccc} c(l_1, \varphi)^{n-1} & \dots & c(l_n, \varphi)^{n-1} \\ \vdots & & \vdots \\ c(l_1, \varphi) & \dots & c(l_n, \varphi) \end{array} \right|$$

The power series in φ for this expression begins with

$$A \prod_j l_j \prod_{j < k} (l_j^2 - l_k^2) \cdot \varphi^{n^2}$$

where A is a number independent of φ and the l_j . If we express numerator and denominator in (94) in this fashion, we find for $\varphi = 0$ that

$$(97) \quad R = \frac{P(l_1, l_2, \dots, l_n)}{P(n, n-1, \dots, 1)} \quad \text{with} \quad P(h_1, \dots, h_n) = \prod_j h_j \prod_{j < k} (h_j^2 - h_k^2)$$

is the degree of the representation in question.

14. Orthogonal group.

The orthogonal group $\mathcal{O}$ is defined as the group of all linear transformations of a vector space which leave a non-degenerate quadratic form Q invariant. We may assume that Q has the form

$$Q = 2(x_1 x_1' + x_2 x_2' + \dots + x_n x_n') \quad \text{for an even number of dimensions } \mathcal{V} = 2n$$

$$Q = -x_0^2 + 2(x_1 x_1' + x_2 x_2' + \dots + x_n x_n') \quad \text{for an odd number of dimensions}$$

$$\mathcal{V} = 2n+1$$

The same theory as given above for the full linear group and the orthogonal group can be developed for the orthogonal group also. We give only the formulas and the results and some additional remarks concerning changes which are necessary in this case.

The elements of the infinitesimal group σ have the form

$$\left\| \begin{array}{cc|cc} \sigma & -\rho_i & -\sigma_i & \\ \hline \sigma_i & \rho_{ik} & \rho'_{ik} & \\ \hline \rho_i & \sigma'_{ik} & \sigma_{ik} & \end{array} \right\|$$

$$i, k = 1, 2, \dots, n$$

$$\rho_{ik} + \sigma_{ki} = 0$$

$$\rho'_{ik} + \rho'_{ki} = 0, \quad \sigma'_{ik} + \sigma'_{ki} = 0$$

The first row and the first column occur only in the case of odd $\nu = 2n+1$.

The diagonal elements

$$h = \begin{pmatrix} 0 & & & \\ & \lambda^1 & & \\ & & \ddots & \\ & & & \lambda^n \\ & & & & -\lambda^1 \\ & & & & & \ddots \\ & & & & & & -\lambda^n \end{pmatrix} = \lambda^1 h_1 + \dots + \lambda^n h_n$$

form an Abelian subgroup $\mathfrak{g}$

The "roots" are here

$$\alpha = \pm \lambda^i \quad (\text{for odd } \nu = 2n+1 \text{ only}) \text{ and } \alpha = \pm \lambda^i \pm \lambda^k, \quad i < k.$$

There exists an element e_α for every α such that

$$[h e_\alpha] = \alpha \cdot e_\alpha$$

Again, $h_1, \dots, h_n$ and the e_α form a basis. $[e_\alpha, e_{-\alpha}]$ is an element h_α of $\mathfrak{g}$. If $\Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n$ is a linear form, we have here

$$\Lambda_\alpha = \begin{cases} \pm 2m_1 & (\text{for } \nu = 2n+1 \text{ only}) \\ \pm m_i \pm m_k \end{cases}$$

We have $\alpha_\alpha = 2$.

Let $\mathfrak{g}$ be a representation of $\mathfrak{v}$. The theorem concerning the weights holds again: Λ_α is an integer for every weight Λ and $\Lambda - \Lambda_\alpha \cdot \alpha$ is a weight again. We derive from this that the coefficients of every weight $m_1 \lambda^1 + \dots + m_n \lambda^n$ are either all integers or all half-integers (these are the numbers $\pm \frac{1}{2}, \pm \frac{3}{2}, \pm \frac{5}{2}, \dots$)*. Let (S) denote the group of linear transformations

* This is not true in the case $2n = 2$. Here $\mathfrak{v}$ is a one-parameter group. We omit this case, which is without interest.

in a $\lambda^1 - \lambda^2 - \dots - \lambda^n$ space consisting of all permutations of the λ_ν

combined with a change of sign in $\frac{\text{an even}}{\text{any}}$ number of λ_i^i , $\frac{\nu = 2n}{\nu = 2n+1}$. The

order of (S) is $\frac{2^{n-1}n!}{2^n n!}$. The operations of (S) transform weights into weights.

The highest weight $m_1 \lambda^1 + \dots + m_n \lambda^n$ satisfies the conditions

$$\begin{aligned} \nu = 2n: & \quad m_1 \geq m_2 \geq \dots \geq m_{n-1} \geq |m_n| \\ \nu = 2n+1: & \quad m_1 \geq m_2 \geq \dots \geq m_{n-1} \geq m_n \geq 0 \end{aligned} \quad m_i = m_k = 0 \text{ or } \frac{1}{2} \pmod{1}$$

There is at most one irreducible representation $\mathcal{G}$ with a given highest weight Λ and Λ occurs only once as a weight of $\mathcal{G}$.

We introduce the unitary restriction. If the parameters are properly chosen this restriction implies only a condition of reality for the parameters.

Every element t of $\mathcal{O}_u$ is equivalent to a diagonal element

$$(98) \quad u t u^{-1} = (\varepsilon) = \begin{pmatrix} 1 & & & \\ & \varepsilon_1 & & \\ & & \ddots & \\ & & & \varepsilon_n \\ & & & & \varepsilon'_1 \\ & & & & & \ddots \\ & & & & & & \varepsilon'_n \end{pmatrix}, \quad \varepsilon'_i = 1/\varepsilon_i$$

(The first row and column occur for $\nu = 2n+1$ only.) u belongs to $\mathcal{O}_u$.

The elements in the main diagonal of (ε) are, of course, the characteristic roots of t . They are all of absolute value 1.

The manifold $\mathcal{O}_u$ is not simply connected. The reason that the old proof fails is that the matrices t with a multiple characteristic root form a manifold of $N-1$ dimensions, N being the number of dimensions of $\mathcal{O}_u$.

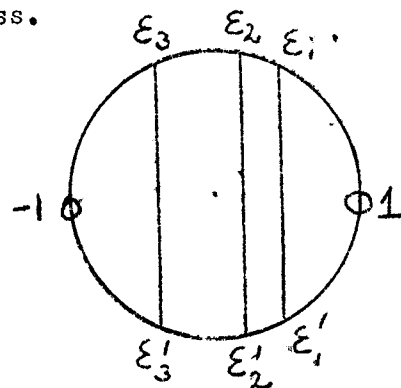
* If $\varepsilon_1 = \varepsilon'_1 = -1$ and ± 1 , $\varepsilon_2, \dots, \varepsilon_n$ are all distinct, the elements of $\mathcal{O}_u$ which are commutative with (ε) are diagonal matrices (ε^*) again. They depend on n parameters. Then $u^{-1}(\varepsilon)u$, u in $\mathcal{O}_u$ form a manifold of $N - n + (n-1) = N - 1$ dimensions, since (ε) depends on $n-1$ parameters.

We are going to prove: Every closed curve on $\mathcal{O}_u$ can be shrunk either to a point or to a fixed closed curve γ . The double of γ , but not γ it-

self can be shrunk to a point. ..

Proof: We call the element t singular if in (98) either $\varepsilon_i = \varepsilon_k$ or $\varepsilon_i = \varepsilon'_k$ for a pair i, k with $i \neq k$, or if in the case $\gamma = 2n+1$, one of the ε_i has the value 1. The same proof as in the case of the complex group (p. 42) shows that the singular t form a manifold of $N-3$ dimensions. Let δ be a closed curve in $\mathcal{O}_u$. It can be assumed without loss of generality that δ does not pass a singular t during the shrinking process.

Let t run over δ . The characteristic roots will move on the unit circle but never will roots on the upper semicircle coincide. After t has returned to its initial point the ε_i must have their original position apart from a permutation. But the only possible permutation is that



ε_1 and ε'_1 are interchanged or ε_n and ε'_n are interchanged, ε_1 being the root nearest to 1 and ε_n the root nearest to -1. If ε_n and ε'_n are not interchanged we can shrink δ into a point as we did in 10 pp. 42/43. If γ and δ are two curves (we may assume with the same starting point), for which ε_n and ε'_n are interchanged, then for the composed curve $\gamma + \delta$, ε_n and ε'_n return to their initial position. Hence $\gamma + \delta$ can be shrunk into a point. This holds in particular for $\gamma + \gamma$. Therefore δ can be shrunk into $-\gamma$ and into γ itself.

It is easy to give examples of curves γ which cannot be shrunk to a point, for instance

$$t = (\varepsilon), \quad \varepsilon_1 = e(i\varphi), \quad 0 \leq \varphi \leq 1, \quad \varepsilon_2 = \varepsilon_3 = \dots = \varepsilon_n = 1$$

We can formulate these facts in the form: The Poincaré group of $\mathcal{O}_u$ has the order 2. Every function on $\mathcal{O}_u$ which is given in the neighborhood of the identity and permits a continuous continuation is either single-valued or double-valued.

In order to give a construction of the irreducible representation analogously to that in 5 and in // we need the representation with the highest weight

$$\frac{1}{2}(\lambda^1 + \lambda^2 + \dots + \lambda^n) \quad \text{for } \gamma = 2n+1$$

and the two representations with the highest weights

$$\frac{1}{2}(\lambda^1 + \dots + \lambda^{n-1} + \lambda^n) \quad \text{and} \quad \frac{1}{2}(\lambda^1 + \dots + \lambda^{n-1} - \lambda^n)$$

A construction of such representation of the infinitesimal group σ was given by Cartan; for an algebraic construction of these representations for $\mathcal{O}$ cf. Brauer and Weyl, American Journal of Math. 57, p. 425.

It is possible to obtain the irreducible representations by Kronecker multiplication of this "Spinor" representation and the tensor representation in quite the same manner as used in 5, p. 24, and //, p. 45.

Furthermore, our methods permit also, in the case of the orthogonal group, determination of the characters of the irreducible representations. The formulas are given in Weyl, Mathematische Zeitschrift 24, pp. 348-353.

Literature and Related Topics

For the theory given in this Chapter I, cf. H. Weyl, Mathematische Zeitschrift 23, p. 271, and 24, pp. 328-353.

For an algebraic method of splitting up the representation of the full linear group which belongs to the general tensor of rank f , cf. I. Schur, Dissertation, Berlin 1901, and I. Schur, Sitzungsberichte Berliner Akademie 1927; H. Weyl, Group theory and quantum mechanics; van der Waerden, Math. Annalen 104, pp. 92, 800; Young, On the quantitative substitutional analysis, I - VII, Proc. London Math. Soc. 1900-1933.

The connection with the representations of the symmetric group given in these papers is of importance. A corresponding treatment for the complex-group or the orthogonal group has not yet been given. Cf. for the case of the complex-group: H. Weyl, *Mathematische Zeitschrift* 35, p. 305 where the enveloping algebra of the tensor representation was determined in the case of the complex-group.

It was necessary to assume that the functions defining the representation are differentiable. In a paper by von Neumann in the *Sitzungsberichte d. Berliner Akademie* 1927 (cf. also von Neumann, *Mathematische Zeitschrift*) it was shown that continuity is sufficient.

It is possible to give the theory of representations in an essentially algebraic manner without using Lie's theory or the integration over the group manifold. One gets all continuous representation in this manner. Cf. for the case of the full linear group, I. Schur, *Sitzungsberichte der Berliner Akademie* 1928; for the case of the orthogonal group, R. Brauer, *Dissertation*, Berlin, 1925, and *Sitzungsberichte Berliner Akademie*, 1929; for the complex-group, H. Mohr, *Dissertation*, Göttingen, 1933.

The characters of the representations of the orthogonal group were determined first by Schur in the *Sitzungsberichte Berliner Akademie*, 1924. The case of the complex group was treated first by H. Weyl in the paper mentioned above.

Chapter 2.

Structure of semisimple groups

1. The maximal nilpotent subgroup of an infinitesimal group.

Let $\mathfrak{g}$ be an infinitesimal group of order r . We consider the adjoint representation of $\mathfrak{g}$. The vector space in which this representation takes place is $\mathfrak{g}$ itself. To an element c of $\mathfrak{g}$ corresponds the linear transformation $x \rightarrow Cx$ with $Cx = [c, x]$. We form the characteristic equation of the matrix C of this transformation

$$(1) \quad f(\omega) = |\omega \cdot 1 - C| = \omega^r - \psi_1 \omega^{r-1} + \psi_2 \omega^{r-2} - \dots = 0.$$

Here 1 denotes the unit matrix and ω is a variable. The roots of (1) will be called the characteristic roots of c .

We say that a vector e belongs to a root α of (1) if

$$(2) \quad (\alpha \cdot 1 - C)^h e = 0$$

for at least one integer h . The following facts are proved in the theory of linear transformations:

1) The vectors belonging to a fixed root α of (1) of multiplicity a form a subspace $\mathcal{R}_\alpha$ of a dimensions. $\mathcal{R}_\alpha$ is invariant under C . A vector e belongs to $\mathcal{R}_\alpha$ if (2) holds for $h = a$.

2) If $\alpha, \beta, \dots$ are the different roots of (1) and $\mathcal{R}_\alpha, \mathcal{R}_\beta, \dots$ the corresponding subspaces, we have

$$\mathfrak{g} = \mathcal{R}_\alpha + \mathcal{R}_\beta + \dots$$

(i.e. every element of $\mathfrak{g}$ can be represented as a sum of an element of $\mathcal{R}_\alpha$, an element of $\mathcal{R}_\beta, \dots$ in a unique manner.

3) We can find a basis $e_1, e_2, \dots, e_a$ of $\mathcal{R}_\alpha$ such that

$$Ce_1 = e_1$$

$$Ce_2 \equiv e_2 \pmod{e_1}$$

$$Ce_a \equiv e_a \pmod{e_1, e_2, \dots, e_{a-1}}$$

The transformation induced by C in the space $\mathcal{R}_\alpha$ has the characteristic polynomial $(\omega - \alpha)^a$.

4) If a linear transformation T leaves all the spaces $\mathcal{R}_\alpha, \mathcal{R}_\beta, \dots$ invariant, the characteristic polynomial of T is of the form $\varphi_\alpha \cdot \varphi_\beta \cdot \dots$ where $\varphi_\alpha, \varphi_\beta, \dots$ are the characteristic polynomials of the linear transformations induced in $\mathcal{R}_\alpha, \mathcal{R}_\beta, \dots$ by T .

Lemma 1: If α and α' are roots of (1) and e belongs to α and

e' belongs to α' , $[e, e']$ belongs to $\alpha + \alpha'$.

$$\text{Proof. } [C[ee']] = -[e[e'c]] - [e'[ce]] = [e[ce']] + [e'e']$$

$$C[ee'] = [e, Ce'] + [Ce, e']$$

We subtract $(\alpha + \alpha')[ee']$ and obtain

$$(C - (\alpha + \alpha')1)[ee'] = [(C - \alpha 1)e, e'] + [e, (C - \alpha' 1)e']$$

on repeating the same argument we get

$$(3) \quad (C - (\alpha + \alpha')1)^k [ee'] = \sum_{k=0}^K \binom{K}{k} [(C - \alpha 1)^k e, (C - \alpha' 1)^{K-k} e']$$

If we take for k the sum of the multiplicities a and a' of α and α' , we have either $K \geq a$ or $k - K \geq a'$. It follows that either $(C - \alpha 1)^K e = 0$ or $(C - \alpha' 1)^{K-k} e' = 0$ since e belongs to α and e' belongs to α' . Hence the expression in (3) vanishes.

The matrix C has the characteristic root 0 under our assumption. If, namely, $c \neq 0$, we have $Cc = [cc] = 0$, and consequently c is a vector $\neq 0$ belonging to the root 0. Furthermore $C = 0$ for $c = 0$. If h and h' are two elements of the corresponding subspace $\mathcal{R}_0$, $[hh']$ belongs to $\mathcal{R}_0$ too according to Lemma 1. Therefore $\mathcal{R}_0$ is not only a subspace of $\mathfrak{g}$ but also a subgroup $\mathfrak{f}$. We denote the roots of (1) now by $0, \alpha, \beta, \dots$ and their multiplicities by $\lambda, a, b, \dots$.

Let $h_1, h_2, \dots, h_e$ form a basis of $\mathcal{H}$. Every element of $\mathcal{H}$ can be written in the form

$$(4) \quad h = \lambda^1 h_1 + \lambda^2 h_2 + \dots + \lambda^e h_e$$

Let $\mathcal{H}^*$ be the corresponding transformation of the adjoint representation:

$\mathcal{H}^* x = [h x]$. If x belongs to $\mathcal{R}_\alpha$, so does $[h x]$ according to Lemma 1.

Hence $\mathcal{R}_0, \mathcal{R}_\alpha, \mathcal{R}_\beta, \dots$ are invariant under $\mathcal{H}$. It follows that the characteristic polynomial of $\mathcal{H}$

$$(5) \quad f(\omega|h) = |\omega \cdot 1 - \mathcal{H}| = \omega^r - \psi_1(h) \omega^{r-1} + \psi_2(h) \omega^{r-2} - \dots$$

splits up into a product

$$(6) \quad f(\omega|h) = f_0(\omega|h) f_\alpha(\omega|h) f_\beta(\omega|h) \dots$$

* Cf. the properties of linear transformations mentioned on p. 70.

where $f_v(\omega|h)$ is the characteristic polynomial of the transformation $H^{(v)}$ induced by H in $\mathcal{R}_v$, ($v = 0, \alpha, \beta, \dots$)

For the special element $h = c$ of $\mathcal{H}$, we have*

* See note above.

$$(7) \quad f_0(\omega|c) = \omega^e, f_\alpha(\omega|c) = (\omega - \alpha)^a, f_\beta(\omega|c) = (\omega - \beta)^b,$$

The coefficients of $H = \lambda^1 H_1 + \dots + \lambda^e H_e$ are linear functions of

$\lambda^1, \dots, \lambda^e$. It follows that ψ_k , (5), is homogeneous of degree k in

$\lambda^1, \dots, \lambda^e$. Furthermore, the coefficients of $f_v(\omega|h)$ are polynomials in $\lambda^1, \dots, \lambda^e$ (for $v = 0, \alpha, \beta, \dots$). We consider $\lambda^1, \dots, \lambda^e$

as variables. The polynomials $f_\alpha(\omega|h), f_\beta(\omega|h), \dots$ are not divisible by ω because this is not true for $\omega = c$ as is seen from (7).

Until now, c was an arbitrarily fixed element of $\mathcal{H}$. We now choose c such that the following conditions are fulfilled:

- A) The multiplicity λ of the root 0 of (1) has the smallest possible value.
 B) There are no elements c^* in $\mathfrak{g}$ fulfilling the condition A) which have a larger number of distinct characteristic roots than c has.

Then the multiplicity of the root 0 of $f(\omega|h) = 0$ is at least λ .
 Hence $f(\omega|h)$ is divisible by ω^λ for variable $\lambda^1, \dots, \lambda^l$. But only the factor f_0 in (6) can be divisible by ω . Its degree being λ we have

$$(8) \quad f_0(\omega|h) = \omega^\lambda$$

We consider only the group f for the moment. The linear transformation $H^0: x \rightarrow [hx]$, (x ranging over f) gives the element corresponding to h in the adjoint representation of f . According to (8), the characteristic polynomial is ω^λ . It follows from Engel's theorem (cf. p. 7) that f is nilpotent. It is a maximum nilpotent subgroup of $\mathfrak{g}$, i.e. f cannot be embedded in a larger nilpotent subgroup of $\mathfrak{g}$, because otherwise $f(\omega|h)$ would be divisible by a higher power of ω than ω^λ as is easily seen.

We consider now the linear transformation $H^\alpha: x \rightarrow [hx]$ where x ranges over $\mathcal{R}_\alpha$. They form a representation of f and therefore a solvable group of linear transformations. According to Lie's theorem (p. 6), we can find a basis $e_\alpha, e'_\alpha, \dots, e_\alpha^{(a-1)}$ such that

$$(9) \quad \begin{aligned} [he_\alpha] &= \eta_1 e_\alpha \\ [he'_\alpha] &= \eta_2 e'_\alpha \pmod{e_\alpha} \\ &\vdots \\ [he_\alpha^{(a-1)}] &= \eta_a e_\alpha^{(a-1)} \pmod{e_\alpha, e'_\alpha, \dots, e_\alpha^{(a-2)}} \end{aligned}$$

for every $h = \lambda^1 h_1 + \dots + \lambda^l h_l$ in f . Each η_μ is a linear function

$$\eta_\mu = \eta_{\mu 1} \lambda^1 + \dots + \eta_{\mu l} \lambda^l \quad \text{where } \eta_{\mu p} \text{ is given by}$$

$$[h_p e_\alpha^{(\mu-1)}] = \eta_{\mu p} e_\alpha^{(\mu-1)} \pmod{e_\alpha, e'_\alpha, \dots, e_\alpha^{(\mu-2)}}. \quad \text{The}$$

characteristic polynomial of H^α is

$$f_\alpha(\omega|h) = (\omega - \eta_1)(\omega - \eta_2) \dots (\omega - \eta_a)$$

as is seen from (9) by using $\alpha_1, \dots, \alpha_a$ as the basis in $\mathcal{R}_\alpha$

The same consideration is possible for $\mathcal{R}_\beta, \mathcal{R}_\gamma, \dots$. If

$$f_\beta(\omega|h) = (\omega - \gamma_1)(\omega - \gamma_2) \dots (\omega - \gamma_b), \dots$$

the linear functions $\gamma_1, \gamma_2, \dots, \gamma_b$ are different from $\eta_1, \eta_2, \dots, \eta_a$, because we have for $h = c$

$$\eta_1 = \eta_2 = \dots = \eta_a = \alpha, \quad \gamma_1 = \gamma_2 = \dots = \gamma_b = \beta, \dots$$

We choose now such values for $\lambda^1, \dots, \lambda^l$ that 1) all

$\eta_1, \dots, \eta_a, \gamma_1, \dots, \gamma_b, \dots$ which are not identical have different

values, 2) $\psi_{r-l}(h) \neq 0$. The element $h = c^*$ we get in this way has 0 as a

characteristic root of multiplicity χ . Furthermore it has a greater number of different characteristic roots than c unless

$$(10) \quad \eta_1 = \eta_2 = \dots = \eta_a, \quad \gamma_1 = \gamma_2 = \dots = \gamma_b, \dots \text{ identically in } \lambda^1, \dots, \lambda^l.$$

But the conditions A) and B) satisfied by c give a contradiction, unless (10) holds. It follows that

$$f_\alpha(\omega|h) = (\omega - \eta_1)^a, \quad f_\beta(\omega|h) = (\omega - \gamma_1)^b, \dots$$

We change the notations and write α for η_1 , β for $\gamma_1, \dots$. They are linear functions $\alpha = \alpha_1 \lambda^1 + \dots + \alpha_l \lambda^l$, $\beta = \beta_1 \lambda^1 + \dots, \dots$

We summarize our results: We have a maximum nilpotent subgroup $\mathcal{f}$ of order χ

$$(11) \quad h = \lambda^1 h_1 + \dots + \lambda^l h_l$$

The characteristic equation of the matrix H corresponding to h in the adjoint representation of $\mathcal{g}$ has the form

$$(12) \quad f(\omega|h) = \omega^l (\omega - \alpha)^a (\omega - \beta)^b \dots$$

$$(13) \quad \alpha = \alpha_1 \lambda^1 + \dots + \alpha_l \lambda^l, \quad \beta = \beta_1 \lambda^1 + \dots + \beta_l \lambda^l, \dots$$

These linear forms are called the roots of $\mathcal{g}$. We have

$$(14) \quad \mathcal{g} = \mathcal{R}_0 + \mathcal{R}_\alpha + \mathcal{R}_\beta + \dots, \quad \mathcal{f} = \mathcal{R}_0$$

Here $\mathcal{K}_\alpha$ is a subspace of a dimensions. There exists a basis

$e_\alpha, e'_\alpha, \dots, e_\alpha^{(a-1)}$ of $\mathcal{K}_\alpha$ such that

$$(15) [he_\alpha] = \alpha e_\alpha, [he'_\alpha] \equiv \alpha e'_\alpha \pmod{e_\alpha}, [he''_\alpha] \equiv \alpha e''_\alpha \pmod{e_\alpha e'_\alpha}$$

Corresponding facts hold for $\mathcal{K}_\beta, \dots$. If ρ belongs to $\mathcal{K}_\rho$ and ρ^* belongs to $\mathcal{K}_{\rho^*}$ then $[\rho, \rho^*]$ belongs to $\mathcal{K}_{\rho+\rho^*}$, ($\rho, \rho^* = 0, \alpha, \beta, \dots$)

In particular: $[\rho, \rho^*] = 0$ if $\rho + \rho^*$ is neither 0 or a root.

2. Representations. Weights.

We consider a representation $\mathcal{G}$ of the infinitesimal group $\mathfrak{g}$. The vectors of the space $\mathcal{R}$ in which the transformations of $\mathcal{G}$ take place will be denoted by German letters. The elements H corresponding in $\mathcal{G}$ to elements of the maximum nilpotent subgroup $\mathfrak{f}$ of $\mathfrak{g}$ form a solvable group of linear transformations. Lie's theorem shows that there is a vector $\mu \neq 0$ in $\mathcal{R}$ such that μ is eigenvector for all H . If $h_\mu \rightarrow H_\mu$,

$$h = \lambda^1 h_1 + \dots + \lambda^l h_l \rightarrow \lambda^1 H_1 + \dots + \lambda^l H_l$$

We have $H_\mu \mu = m_\mu \mu$ with scalar factors m_μ . Hence

$$(16) H \mu = \Lambda \mu$$

where $\Lambda = m_1 \lambda^1 + \dots + m_l \lambda^l$. We say that μ belongs to a linear form

Λ if (16) holds. A linear form to which vectors $\mu \neq 0$ belong is called a

weight. In the case of the adjoint representation, the weights different from

0 are the roots.

From (15) follows that the matrix E_α corresponding to e_α satisfies the equation

$$[HE_\alpha] = \alpha E_\alpha$$

Lemma 2: If μ is of weight Λ , $E_\alpha \mu$ belongs to the weight $\Lambda + \alpha$

Proof. $HE_\alpha \mu = E_\alpha H \mu + [HE_\alpha] \mu = \Lambda E_\alpha \mu + \alpha E_\alpha \mu =$

$$= (\Lambda + \alpha) E_\alpha \mu$$

Lemma 3: If a vector $\mathcal{C}$ of weight Λ is a sum of vectors $\mathcal{C}_i$ of weight Λ_i and all Λ_i are different from Λ then $\mathcal{C} = 0$.

The proof is exactly the same as the proof of Lemma 1 in Chapter 1, 2 p. 13. It follows again that there are not more than N different weights if $\mathcal{R}$ has N dimensions.

If h^* is an element of $\mathfrak{f}$ we denote by $\Lambda(h^*)$ the value of Λ for the values $\lambda^1, \dots, \lambda^l$ belonging to h^* .

Lemma 4: Every weight Λ vanishes for the elements of the commutator group $\mathfrak{f}'$ of $\mathfrak{f}$.

Proof. Let h^* and h^{**} be two elements of $\mathfrak{f}$. We have, if $n \neq 0$ is a vector belonging to Λ

$$H^* n = \Lambda(h^*) n, \quad H^{**} n = \Lambda(h^{**}) n$$

$$\begin{aligned} [H^* H^{**}] n &= (H^* H^{**} - H^{**} H^*) n = H^* \Lambda(h^{**}) n - H^{**} \Lambda(h^*) n = \\ &= \Lambda(h^*) \Lambda(h^{**}) n - \Lambda(h^{**}) \Lambda(h^*) n = 0 \end{aligned}$$

on the other hand, we have $[H^* H^{**}] n = \Lambda([h^* h^{**}]) n$ Hence $\Lambda([h^* h^{**}]) = 0$ and consequently $\Lambda = 0$ for all linear combinations of brackets $[h^* h^{**}]$.

We now assume that $-\alpha$ also is a root. We shall show later that

this is always true in the case of a semi-simple group. If e_α and $E_{-\alpha}$ have the same significance for $-\alpha$ as have e_α, E_α for α , $\begin{bmatrix} e_\alpha & e_{-\alpha} \end{bmatrix}$ is an element

of $\mathcal{R}_{\alpha+(-\alpha)} = \mathcal{R}_0 = \mathfrak{f}$ We set

$$(17) \quad \begin{bmatrix} e_\alpha & e_{-\alpha} \end{bmatrix} = h_\alpha = \alpha^1 h_1 + \dots + \alpha^l h_l \quad *)$$

* The use of the notation $\alpha_1, \dots, \alpha_l$ in (13) and of $\alpha^1, \dots, \alpha^l$ here will be justified later.

At the same time we have

$$(18) \quad [E_\alpha, E_{-\alpha}] = H_\alpha = \alpha^1 H_1 + \dots + \alpha^l H_l$$

If $\Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n$, we set

$$\Lambda_\alpha = \Lambda(h_\alpha) = m_1 \alpha^1 + \dots + m_n \alpha^n$$

Lemma 5: Let $\alpha \neq 0$ and $-\alpha$ be roots of a group $\mathfrak{g}$. Let Λ be a weight of a representation $\mathfrak{g}$. If $\alpha_\alpha = \alpha_1 \alpha^1 + \dots + \alpha_k \alpha^k \neq 0$; then $g = \frac{2\Lambda_\alpha}{\alpha_\alpha}$ is an integer and $\tilde{\Lambda} = \Lambda - g\alpha$ is a weight too. If, however, $\alpha_\alpha = 0$, we have $\Lambda_\alpha = 0$ for every Λ .

Proof. (Cf. Chapter 1, pp. 16-18.) First of all, we assume that

Λ is an extreme weight in the sense that $\Lambda + \alpha$ is not a weight. If n_0 belongs to Λ , we form the vectors

$$(19) \quad n_0, E_{-\alpha} n_0 = n_1, E_{-\alpha} n_1 = n_2, \dots$$

which, according to Lemma 2, belong to the weights $\Lambda, \Lambda - \alpha, \Lambda - 2\alpha, \dots$

$$(20) \quad H n_i = (\Lambda - i\alpha) n_i$$

There exists an integer g such that $n_g \neq 0, n_{g+1} = 0$ since $\Lambda - i\alpha$ cannot be a weight for $i = 1, 2, 3, \dots$. Then we prove

$$(21) \quad \begin{cases} E_\alpha n_i = \mu_i n_{i-1} \\ \mu_i = i\Lambda_\alpha - \frac{i(i-1)}{2} \alpha_\alpha \end{cases} \quad (i = 0, 1, 2, \dots)$$

in exactly the same manner as on page 17 (cf. also the proof of the more general relation (24) below). It follows from $\mu_{g+1} = 0$ that

$$(22) \quad \Lambda_\alpha = \frac{g}{2} \alpha_\alpha$$

If, therefore, $\alpha_\alpha = 0$, then $\Lambda_\alpha = 0, (\Lambda - k\alpha)_\alpha = \Lambda_\alpha - k\alpha_\alpha = 0$.

The statement holds in this case also, if the weight is not extreme.

We assume: $\alpha_\alpha \neq 0$. Then $g = \frac{2\Lambda_\alpha}{\alpha_\alpha}$ is an integer, and $\Lambda - g\alpha$ is a weight. Our statement is true for extreme weight Λ . The consideration on pages 17-18 shows that the theorem is true for all weights Λ . Furthermore, we see at the same time

Lemma 6: All the weights whose differences from a fixed weight are integral multiples of α form an uninterrupted "string of weights"

$$(23) \quad \Lambda - R'\alpha, \Lambda - (R'-1)\alpha, \dots, \Lambda, \dots, \Lambda + R\alpha$$

We need some additional remarks: 1) Let t_α be any element of $\mathcal{R}_\alpha$ (14), i.e. a linear combination of $e_\alpha, e'_\alpha, e''_\alpha, \dots$. If $t_\alpha \rightarrow T_\alpha$, T_α is the same linear combination of $E_\alpha, E'_\alpha, E''_\alpha, \dots$. According to Lemma 1,

$[t_\alpha e_{-\alpha}] = h_\alpha^* = \alpha_1^* h_1 + \dots + \alpha_n^* h_n$ belongs to $\mathfrak{f}$. We set

$$\Lambda_\alpha^* = m_1 \alpha_1^* + \dots + m_n \alpha_n^* \text{ for any linear function } \Lambda = m_1 \lambda^1 + \dots + m_n \lambda^n.$$

We state that we have for an extreme weight Λ :

$$(24) \quad \begin{cases} T_\alpha n_i = v_i n_{i-1} \\ v_i = i \Lambda_\alpha^* - \frac{i(i-1)}{2} \alpha_\alpha^* \end{cases}, \quad (i = 0, 1, 2, \dots)$$

$$(25) \quad \Lambda_\alpha^* = \frac{g}{2} \alpha_\alpha^*$$

in generalization of (21) and (22).

Proof. The vector $E_\alpha n_0$ belongs to $\Lambda + \alpha$. But $\Lambda + \alpha$ is not a weight when Λ is extreme. Hence $E_\alpha n_0 = 0$. We form $E'_\alpha n_0$. If, according to (15), $[h e'_\alpha] = e'_\alpha + \rho e_\alpha$, we have

$$\begin{aligned} H E'_\alpha n_0 &= E'_\alpha H n_0 + [H E'_\alpha] n_0 = \Lambda E'_\alpha n_0 + \alpha E'_\alpha n_0 + \rho E_\alpha n_0 \\ &= (\Lambda + \alpha) E'_\alpha n_0 \end{aligned}$$

because of $E_\alpha n_0 = 0$. Thus $E'_\alpha n_0 = 0$ since it belongs to $\Lambda + \alpha$. We obtain in the same way successively: $E''_\alpha n_0 = E'''_\alpha n_0 = \dots = 0$. Hence

$T_\alpha n_0 = 0$. Therefore (24) is true for $i = 0$, whatsoever n_{-1} is. If (24) is true for i , we have

$$\begin{aligned} T_\alpha n_{i+1} &= T_\alpha E_{-\alpha} n_i = E_{-\alpha} T_\alpha n_i + [T_\alpha E_{-\alpha}] n_i \\ &= E_{-\alpha} v_i n_{i-1} + H_\alpha^* n_i = v_i n_i + (\Lambda - i \alpha)_\alpha^* n_i \\ &= (v_i + \Lambda_\alpha^* - i \alpha_\alpha^*) n_i = v_{i+1} n_i \end{aligned}$$

Hence we obtain (24). We derive (25) from $n_{g+1} = 0, v_{g+1} = 0$.

2) The elements $h_1, \dots, h_l, e_\alpha, e'_\alpha, \dots, e_{(a-1)\alpha}, 0$ generate a subgroup $\mathfrak{g}_1$ as is easily seen from the results of 2, 1, pp. 73-74. We need

not assume in the argument above that a representation $\mathfrak{g}$ of $\mathfrak{g}$ is given. We can derive the same results if we have merely a representation $\mathfrak{g}_1$ of $\mathfrak{g}_1$.

The vectors $n_0, n_1, \dots, n_g$ span a subspace $\mathcal{V}$ of $\mathcal{R}$. It follows from (20), (24) and (19) that $\mathcal{V}$ is invariant under H, T_α and $E_{-\alpha}$ and hence under $\mathcal{G}_1$. The group $\mathcal{G}_1$ induces, therefore, a representation $\mathcal{G}'_1$ in the projection space $\mathcal{R}'$ of $\mathcal{R} \pmod{\mathcal{V}}$. Consequently, there exists a linear form Λ' in $\lambda^1, \dots, \lambda^n$ and vectors $n'_0, n'_1, \dots, n'_g$ such that

$$(19') \quad E_{-\alpha} n'_i \equiv n'_{i+1} \pmod{\mathcal{V}}$$

$$(20') \quad H n'_i \equiv (\Lambda' - i\alpha) n'_i \pmod{\mathcal{V}}$$

$$(24') \quad T_\alpha n'_i \equiv v'_i n'_{i-1} \pmod{\mathcal{V}}, \quad v'_i = i\Lambda'^*_{\alpha} - \frac{i(i-1)}{2} \alpha^*_{\alpha}$$

$$(25') \quad \Lambda'^*_{\alpha} = \frac{g'}{2} \alpha^*_{\alpha}$$

Let $\mathcal{V}'$ denote the subspace spanned by $\mathcal{V}, n'_0, \dots, n'_g$. Then $\mathcal{V}'$ is invariant under $\mathcal{G}_1$. We repeat our argument and continue in this manner.

Finally, we get a basis $n_0, \dots, n_g, n'_0, \dots, n'_g, \dots$ for which formulas as (19), (20), (24), (25), (19'), (20'), (24'), (25'), ... hold.

We write all the matrices with respect to this coordinate system. $H, T_\alpha, E_{-\alpha}$ have 0 above the main diagonal. We have in the main diagonal of H ,

$$(26) \quad \Lambda, \Lambda - \alpha, \dots, \Lambda - g\alpha, \Lambda', \Lambda' - \alpha, \dots, \Lambda' - g'\alpha, \dots$$

according to (20), (20'), ... Consequently, (26) are the characteristic roots of H .

From (19), (24) and (25) it follows that

$$T_\alpha E_{-\alpha} n_i = T_\alpha n_{i+1} = v_{i+1} n_i = (i+1) \left(\Lambda^*_{\alpha} - \frac{i}{2} \alpha^*_{\alpha} \right) n_i = \frac{(i+1)(g-i)}{2} \alpha^*_{\alpha} n_i$$

Similarly, we obtain $T_\alpha E_{-\alpha} n'_i = \frac{(i+1)(g'-i)}{2} \alpha^*_{\alpha} n'_i \pmod{\mathcal{V}}$. Thus the coefficients in the main diagonal of $T_\alpha E_{-\alpha}$ are

$$\frac{1 \cdot g}{2} \alpha^*_{\alpha}, \frac{2(g-1)}{2} \alpha^*_{\alpha}, \dots, \frac{(g+1) \cdot 0}{2} \alpha^*_{\alpha}, \frac{1 \cdot g'}{2} \alpha^*_{\alpha}, \frac{2(g'-1)}{2} \alpha^*_{\alpha}, \dots, \frac{(g'+1) \cdot 0}{2} \alpha^*_{\alpha}, \dots$$

The sum of these numbers is the trace

$$(27) \quad \text{tr}(T_\alpha E_{-\alpha}) = \frac{1}{2} \alpha^*_{\alpha} \left(\frac{g(g+1)(g+2)}{1 \cdot 2 \cdot 3} + \frac{g'(g'+1)(g'+2)}{1 \cdot 2 \cdot 3} + \dots \right)$$

We will apply our results to the case of the adjoint representation.

Here the characteristic roots of H , (26), are 0 (l -times), α (a -times),

β (b -times), ..., according to (12).

Let t_α and $t_{-\alpha}$ be elements of $\mathcal{R}_\alpha$ and $\mathcal{R}_{-\alpha}$ respectively. Then $[t_\alpha t_{-\alpha}] = \tilde{h}_\alpha$ belongs to $\mathfrak{f}$. We set $\rho(\tilde{h}_\alpha) = \tilde{\rho}_\alpha$ for any root ρ of and state that

$$(28) \quad \tilde{\rho}_\alpha = r_{\rho, \alpha} \tilde{\alpha}_\alpha, \quad r_{\rho, \alpha} \text{ rational number}$$

Proof. We consider the string of roots of η to which ρ belongs

$$\rho - R'\alpha, \rho - (R'-1)\alpha, \dots, \rho, \dots, \rho + R\alpha$$

(cf. (23)) and the corresponding eigenspaces $\mathcal{R}_{\rho - R'\alpha}, \dots, \mathcal{R}_{\rho + R\alpha}$

Let $\mathcal{V}$ be the sum of these spaces. Then $\mathcal{V}$ is invariant under $T_\alpha: x \rightarrow [t_\alpha x]$

and $T_{-\alpha}: x \rightarrow [t_{-\alpha} x]$. If $T_\alpha^*, T_{-\alpha}^*, \tilde{H}_\alpha^*$ are the transformations induced by $T_\alpha, T_{-\alpha}, \tilde{H}_\alpha$ in $\mathcal{V}$, we have

$$\text{tr}(\tilde{H}_\alpha^*) = \text{tr}[T_\alpha^* T_{-\alpha}^*] = \text{tr}(T_\alpha^* T_{-\alpha}^* - T_{-\alpha}^* T_\alpha^*) = 0$$

Performance of $\tilde{H}_\alpha^*$ means merely multiplication of the vectors of $\mathcal{R}_{\rho - i\alpha}$ by

$$(\rho - i\alpha)(\tilde{h}_\alpha) = \tilde{\rho}_\alpha - i\tilde{\alpha}_\alpha. \quad \text{If } m_i \text{ is the number of dimensions of } \mathcal{R}_{\rho - i\alpha}$$

the trace of $\tilde{H}_\alpha^*$ equals

$$\text{tr}(\tilde{H}_\alpha^*) = \sum_{i=-R'}^R m_i (\tilde{\rho}_\alpha - i\tilde{\alpha}_\alpha) = \tilde{\rho}_\alpha \sum m_i - \tilde{\alpha}_\alpha \sum i m_i$$

The vanishing of this expression yields (28).

We apply a similar argument to the case $\rho = 0$. The subspace

$$\mathcal{V}^* = \{e_{-\alpha}\} + \mathcal{R}_0 + \mathcal{R}_\alpha + \dots + \mathcal{R}_{R\alpha}$$

is invariant under $H, T_\alpha, E_{-\alpha}$ as is easily seen. Hence the matrix induced

by $H_\alpha^* = [T_\alpha E_{-\alpha}]$ in the projection space of $\mathcal{V}$ (mod $\mathcal{V}^*$) has a vanishing trace. On the

other hand, this trace has the value $((m_{-1} - 1) + 2m_{-2} + \dots + Rm_{-R}) \alpha_\alpha^*$ if m_i

is the number of dimensions of $\mathcal{R}_{i\alpha}$. This expression vanishes in the case

$\alpha_\alpha^* \neq 0$ only if $m_{-1} = 1, m_{-2} = 0, m_{-3} = 0, \dots$. We obtain

Lemma 7: Let α and $-\alpha$ both be roots of $\mathfrak{g}$. If there is an element A_α in $\mathcal{R}_\alpha$ such that $\alpha([A_\alpha, H_{-\alpha}]) = \alpha_\alpha^* \neq 0$, then $2\alpha, 3\alpha, \dots$ are not roots of $\mathfrak{g}$. Furthermore, $\mathcal{R}_{-\alpha}$ is of dimension 1. The root $-\alpha$ of $\mathfrak{g}$ is simple.

3. Cartan's criterion for semisimple groups

Let a and b be two elements of $\mathfrak{g}$ and A and B the corresponding elements of the adjoint representation. We set

$$(29) \quad \sigma(a, b) = \text{tr}(AB) = \text{tr}(BA)$$

If we express a and b by means of a basis of $\mathfrak{g}$, we see that $\sigma(a, b)$ is a bilinear form in the parameters belonging to a and b . This form is symmetric in a and b . The corresponding quadratic form is

$$(30) \quad \sigma(a) = \sigma(a, a) = \text{tr}(A^2)$$

Hence $\sigma(a)$ is the sum of the squares of the characteristic roots

$$(31) \quad \sigma(a) = \sum \rho^2 \quad (\rho \text{ characteristic root of } A)$$

Let

$$(32) \quad f(\omega|a) = |\omega \cdot 1 - A| = \omega^r - \psi_1(a) \omega^{r-1} + \psi_2(a) \omega^{r-2} \dots$$

be the characteristic polynomial of A . Then

$$(33) \quad \sigma(a) = \psi_1^2 - 2\psi_2 = \left(\sum \rho\right)^2 - \sum_{\substack{\rho, \rho' = 0, \alpha, \beta \\ \rho \neq \rho'}} \rho \rho'$$

1st Criterion

The group $\mathfrak{g}$ is solvable if and only if $\sigma(a') = 0$ for all elements a' of the commutator group $\mathfrak{g}'$ of $\mathfrak{g}$.

Proof. We assume that $\sigma(a') = 0$ for all a' in $\mathfrak{g}'$. Let α and $-\alpha$ be roots of $\mathfrak{g}$ and t_α and $t_{-\alpha}$ elements of $\mathcal{R}_\alpha$ and $\mathcal{R}_{-\alpha}$ respectively. Then $[t_\alpha t_{-\alpha}] = \tilde{h}_\alpha$ belongs to $\mathfrak{f}$. We make use of (31) and (28) and have

$$\sigma(\tilde{h}_\alpha) = \sum \rho(\tilde{h}_\alpha)^2 = \tilde{\alpha}_\alpha^2 \sum_{\rho} \rho^2$$

The $\tau_{\rho\alpha}$ ^{are} rational numbers and $\tau_{\alpha\alpha} = 1$. Therefore $\sigma(\tilde{h}_\alpha) = 0$ implies $\tilde{z}_\alpha = 0$. Consequently we have $\rho(\tilde{h}_\alpha) = \rho([t_\alpha t_{-\alpha}]) = 0$ for every root ρ .

Let x and y be two elements of $\mathfrak{g}$, and let

$$x = \sum x_\mu, \quad y = \sum y_\nu \quad (\mu, \nu = 0, \alpha, \beta, \dots)$$

be their representation according to (14). We have

$$[xy] = \sum [x_\mu y_\nu]$$

Every element u of the commutator group $\mathfrak{g}'$, i.e. any linear combination of brackets $[xy]$, can be written in the form

$$(34) \quad u = \sum [v w]$$

where v belongs to some ρ and w to some σ , ($\rho, \sigma = 0, \alpha, \beta, \dots$).

Then $[vw]$ belongs to $\rho + \sigma$ (Lemma 1, p. 70). If u is an element of the intersection of $\mathfrak{g}'$ and $\mathfrak{f}$, then u belongs to 0. It follows from Lemma 3, p. 75, that we may cancel all terms on the right-hand side in (34) which

belong to an $\rho + \sigma \neq 0$. Therefore we get a formula

$$u = \sum [t_\alpha, t_{-\alpha}] \quad (t_\alpha \text{ belongs to } \mathcal{R}_\alpha, t_{-\alpha} \text{ to } \mathcal{R}_{-\alpha})$$

where α is a root such that $-\alpha$ is also a root (or $\alpha = 0$). Hence

$$\rho(u) = \rho(\sum [t_\alpha t_{-\alpha}]) = \rho(\sum \tilde{h}_\alpha) = \sum \rho(\tilde{h}_\alpha) = 0$$

We will assume now that $\mathfrak{g}' = \mathfrak{g}$. Then u may be any element of $\mathfrak{f}$.

Every root vanishes identically. According to (14), we have $\mathfrak{g} = \mathcal{R}_0 = \mathfrak{f}$.

But $\mathfrak{f}$ is solvable, $\mathfrak{f}' \neq \mathfrak{f}$ and the assumption $\mathfrak{g} = \mathfrak{g}'$ leads to a contradiction. Hence the order r' of $\mathfrak{g}'$ is smaller than the order of $\mathfrak{g}$.

Let a' be any element of $\mathfrak{g}'$. If $f'(\omega/a')$ is the characteristic polynomial belonging to the adjoint representation of $\mathfrak{g}'$, we have

$$(35) \quad f(\omega/a') = \omega^{r-r'} f'(\omega/a') \quad *$$

* This can be seen by choosing the basis of $\mathfrak{g}$ such that it consists of a basis of $\mathfrak{g}'$ and $r - r'$ other vectors.

Hence the coefficients ψ_1 and ψ_2 of $f' = \omega^{r-1} - \psi_1 \omega^{r-2} + \psi_2 \omega^{r-3} - \dots$, $\psi_1(a') = \psi_1^2 - 2\psi_2$ vanishes. The group $\mathfrak{g}'$ satisfies itself the assumption of criterion 1.

If this criterion is proved for groups of smaller order, we obtain that $\mathfrak{g}'$ is solvable. Hence $\mathfrak{g}$ is solvable also.

If $\mathfrak{g}$ is solvable, $\mathfrak{g}'$ is nilpotent. Hence all the coefficients in $f'(\omega/a')$ vanish. The same is true for the coefficients of $f(\omega/a')$ (cf. (35)). In particular, $\sigma(a') = \psi_1^2(a') - 2\psi_2(a') = 0$.

2d Criterion

A group $\mathfrak{g}$ is semi-simple if and only if $\sigma(a)$ is not degenerate.

Proof:

a) If $\mathfrak{g}$ contains a solvable invariant subgroup $\mathfrak{A}$, we form the successive commutator groups $\mathfrak{A}' = \mathfrak{A}_1$, $\mathfrak{A}_1' = \mathfrak{A}_2$, $\mathfrak{A}_2' = \mathfrak{A}_3$, ... All these groups are invariant subgroups of $\mathfrak{g}$ as can easily be seen. There is an r such that $\mathfrak{A}_r \neq 0$, $\mathfrak{A}_{r+1} = 0$. Then $\mathfrak{A}_r$ is an invariant Abelian subgroup of $\mathfrak{g}$. We choose a basis of $\mathfrak{g}$ which contains a basis of $\mathfrak{A}_r$. The characteristic equations of the matrices of the adjoint representation do not contain the parameters which refer to basis elements in $\mathfrak{A}_r$, as is easily seen. Hence $\sigma(a)$ is certainly degenerate because it does not contain all the parameters.

b) Let a and b be two elements of $\mathfrak{g}$ belonging to the values y^r and η^r of the parameters in $\mathfrak{g}$ ($r = 1, 2, \dots, r$). We set $\sigma(a, b) = Q(\xi, \eta)$ where Q is a bilinear symmetric form in $y^1, \dots, y^r, \eta^1, \dots, \eta^r$. If, in the adjoint representation, $a \rightarrow A$ and $b \rightarrow B$, we have

$$a' = a + \varepsilon[ca] \rightarrow A + \varepsilon(CA - AC), \quad b' = b + \varepsilon[cb] \rightarrow B + \varepsilon(CB - BC)$$

for an arbitrary c in $\mathfrak{g}$ and a variable ε . Then

$$\sigma(a', b') = \text{tr}((A + \varepsilon(CA - AC))(B + \varepsilon(CB - BC)))$$

$$= \text{tr}(AB) + \varepsilon \text{tr}(ACB) - \varepsilon \text{tr}(ABC) + \varepsilon \text{tr}(CAB) - \varepsilon \text{tr}(ACB) \pmod{\varepsilon^2}$$

$$(3b) \quad \sigma(a', b') = \text{tr}(AB) = \sigma(a, b) \pmod{\varepsilon^2}$$

since $\text{tr}(AB \cdot C) = \text{tr}(C \cdot AB)$. If $[ca]$ and $[cb]$ belong to the values $\tilde{y}'$ and $\tilde{\eta}'$ of the parameters in $\mathcal{Y}$, a' and b' belong to $y' + \varepsilon \tilde{y}'$ and $\eta' + \varepsilon \tilde{\eta}'$ respectively. Hence

$$\begin{aligned}\sigma(a', b') &= Q(y + \varepsilon \tilde{y}, \eta + \varepsilon \tilde{\eta}) = Q(y, \eta) + \varepsilon Q(\tilde{y}, \eta) + \varepsilon Q(y, \tilde{\eta}) + \varepsilon^2 Q(\tilde{y}, \tilde{\eta}) \\ &= \sigma(a, b) + \varepsilon \sigma([ca], b) + \varepsilon \sigma(a, [cb]) + \varepsilon^2 \sigma([ca], [cb])\end{aligned}$$

On comparing with (36) we obtain

$$(37) \quad \sigma([ca], b) + \sigma(a, [cb]) = 0 \quad *$$

* More generally: All coefficients of the characteristic polynomial of a are invariant under the infinitesimal transformation $da = [ca]$, c fixed in $\mathcal{Y}$. Cf. p. 50. of last year's notes.

We consider the subset $\mathcal{M}$ of elements a for which $\sigma(a, x) = 0$ for all x in $\mathcal{Y}$. If σ is degenerate, we have $\mathcal{M} \neq 0$. Obviously, $\mathcal{M}$ is a linear subspace. We replace b in (37) by a variable element x of $\mathcal{Y}$. If a belongs to $\mathcal{M}$ and c is any element of $\mathcal{Y}$, (37) shows that $\sigma([ca], x) = 0$. Hence $[ca]$ belongs to $\mathcal{M}$, $\mathcal{M}$ is an invariant subgroup. In particular $\sigma'(a) = \sigma(a, a) = 0$.

If $f^*(\omega | a)$ is the characteristic polynomial of a in the adjoint representation, $f(\omega | a)$ and $f^*(\omega | a)$ are equal apart from a factor ω^m (cf. (35), where only the fact that $\mathcal{Y}'$ is an invariant subgroup was used. If σ^* , ψ_1^* , ψ_2^* have the same significance for $\mathcal{M}^*$ as σ , ψ_1 , ψ_2 have for $\mathcal{Y}$ we have

$$\sigma^*(a) = \psi_1^{*2} - 2\psi_2^* = \psi_1^2 - 2\psi_2 = \sigma(a) = 0.$$

It follows from Criterion 1 that $\mathcal{M}$ is solvable.

If, therefore, σ is degenerate, $\mathcal{Y}$ has a solvable invariant subgroup $\mathcal{M}$. Then $\mathcal{Y}$ is not semi-simple, q.e.d.

Let $\mathfrak{g}$ be semi-simple. We choose

$$(38) \quad h_1, h_2, \dots, h_\ell, e_\alpha, e'_\alpha, \dots, e_\alpha^{(a-1)}, e_\beta, e'_\beta, \dots, e_\beta^{(b-1)}, \dots$$

as a basis of $\mathfrak{g}$ (cf. pp. 73-74). Let

$$(39) \quad a = \sum_{i=1}^{\ell} \lambda^i h_i + \sum_{\rho=\alpha, \beta, \dots} (\sigma_\rho e_\rho + \sigma'_\rho e'_\rho + \dots + \sigma_\rho^{(s-1)} e_\rho^{(s-1)})$$

$$(40) \quad b = \sum_{i=1}^{\ell} \mu^j h_j + \sum_{\nu=\alpha, \beta, \dots} (\tau_\nu e_\nu + \tau'_\nu e'_\nu + \dots + \tau_\nu^{(n-1)} e_\nu^{(n-1)})$$

be two elements of $\mathfrak{g}$. The associated elements in the adjoint representation will always be denoted by the corresponding capitals. We have

$$\begin{aligned} \sigma(a, b) = \text{tr}(AB) &= \sum_{i,j} \lambda^i \mu^j \text{tr}(H_i H_j) + \sum_{i, \nu, \kappa} \lambda^i \tau_\nu^{(\kappa)} \text{tr}(H_i E_\nu^{(\kappa)}) \\ &\quad + \sum_{\rho, j, i} \sigma_\rho^{(j)} \mu^j \text{tr}(E_\rho^{(j)} H_j) + \sum_{\rho, j, \nu, \kappa} \sigma_\rho^{(j)} \tau_\nu^{(\kappa)} \text{tr}(E_\rho^{(j)} E_\nu^{(\kappa)}) \end{aligned}$$

The transformation $E_\rho^{(j)} E_\nu^{(\kappa)}$ transforms $\mathcal{R}_\mu$ into $\mathcal{R}_{\mu+\nu+\rho}$

($\mu = 0, \alpha, \beta, \dots$). If we use the basis (38), all the coefficients in

the main diagonal of $E_\rho^{(j)} E_\nu^{(\kappa)}$ vanish unless $\rho + \nu = 0$. Hence

$\text{tr}(E_\rho^{(j)} E_\nu^{(\kappa)}) = 0$ for $\rho + \nu \neq 0$. Similarly, $\text{tr}(H_i E_\nu^{(\kappa)}) = 0$, $\text{tr}(E_\rho^{(j)} H_j) = 0$.

$$(41) \quad \sigma(a, b) = \sum_{i,j} \lambda^i \mu^j \text{tr}(H_i H_j) + \sum_{\rho, j, \kappa} \sigma_\rho^{(j)} \tau_{-\rho}^{(\kappa)} \text{tr}(E_\rho^{(j)} E_{-\rho}^{(\kappa)})$$

The second sum extends over those roots ρ for which $-\rho$ is also a root. But

this must hold for every ρ because otherwise the parameter σ_ρ would not appear in $\sigma(a, b)$ and this is not possible in the case of a non-degenerate σ .

If ρ is a root of a semi-simple group, $-\rho$ is a root also. Therefore we can

apply (27) which was proved under the assumption that the negative of a root is a root again.

We gather the terms in (41) which contain a fixed parameter $\tau_{-\alpha}$

These are

$$\tau_{-\alpha} \text{tr} \left((\sigma_\alpha E_\alpha + \sigma'_\alpha E'_\alpha + \dots + \sigma_\alpha^{(a-1)} E_\alpha^{(a-1)}) E_{-\alpha} \right) = \text{tr}(\tau_\alpha, E_{-\alpha}) \cdot \tau_{-\alpha}$$

where T_α is an element of $\mathcal{R}_\alpha$. According to (27), the factor of $T_{-\alpha}$ is

$$\frac{1}{2} \alpha_\alpha^* \left(\frac{g(g+1)(g+2)}{1 \cdot 2 \cdot 3} + \frac{g'(g'+1)(g'+2)}{1 \cdot 2 \cdot 3} + \dots \right)$$

where $g, g', \dots$ are positive integers and $\alpha_\alpha^* = \alpha([T_\alpha, e_{-\alpha}])$. We obtain $\alpha_\alpha^* \neq 0$ because $T_{-\alpha}$ must occur in $\sigma(a, b)$. It follows from Lemma 7, p. 80, that $-\alpha$ is a simple root and that $-2\alpha, -3\alpha, \dots$ are not roots. This holds for every root α . If we replace α by $-\alpha$ we see that α is a simple root, whereas $2\alpha, 3\alpha, \dots$ are not roots.

Now (41) may be written in the form

$$(42) \quad \sigma(a, b) = \sum_{i,j} \lambda^i \mu^j \operatorname{tr}(H_i H_j) + \sum_{\rho=\alpha, \beta, \dots} \sigma_\rho T_{-\rho} N_\rho$$

where $N_\alpha = \operatorname{tr}(E_\alpha E_{-\alpha}) \neq 0$. We obtain from (27) that N_α is a positive rational multiple of $\alpha_\alpha = \alpha[e_\alpha, e_{-\alpha}]$. Consequently we have $\alpha_\alpha \neq 0$.

We have for $a = b = \lambda^1 h_1 + \dots + \lambda^l h_l = h$

$$(43) \quad \sigma(h) = \sum \lambda^i \lambda^j \operatorname{tr}(H_i H_j)$$

This quadratic form is not degenerate since (42) is not degenerate. On the other hand, we get from (31) that

$$(44) \quad \sigma(h) = \sum_{\rho=\alpha, \beta, \dots} \rho(h)^2$$

We see that there are l linear ~~roots~~ independent roots among $\alpha, \beta, \dots$. The only element for which all the roots vanish is 0. But the roots vanish for the elements h of the commutator group f' of f according to Lemma 4, p. 75. Hence $f' = 0$; f is Abelian.

The bracket $[e_\alpha, e_\beta]$ belongs to $\mathcal{R}_{\alpha+\beta}$. Hence $[e_\alpha, e_\beta] = 0$ if $\alpha + \beta$ is neither 0 nor a root. If $\alpha + \beta$ is a root, we have

$[e_\alpha, e_\beta] = N_{\alpha, \beta} e_{\alpha+\beta}$ ($N_{\alpha, \beta}$ a scalar) since $\mathcal{R}_{\alpha+\beta}$ is of a dimension 1 and hence spanned by $e_{\alpha+\beta}$. Finally we have $[e_\alpha, e_{-\alpha}] = h_\alpha$.

Furthermore, if $\alpha \neq 0$ is a root, the results of Chapter 2, 2 hold, in particular Lemma 5, Lemma 6, (22) and (27).

Results: Let $\mathfrak{g}$ be semi-simple. The maximum nilpotent subgroup $\mathfrak{f}$ is Abelian. Let $h_1, \dots, h_\ell$ be a basis of $\mathfrak{f}$. The matrix $\mathcal{H}$ corresponding to $h = \lambda^1 h_1 + \dots + \lambda^\ell h_\ell$ in the adjoint representation of $\mathfrak{g}$ has 0 as a ℓ -fold characteristic root. The other characteristic roots $\alpha, \beta, \dots$ of $\mathcal{H}$ - the "roots" of $\mathfrak{g}$ - are all simple roots. They are linear forms in $\lambda^1, \dots, \lambda^\ell$

$$(45) \quad \alpha = \alpha_1 \lambda^1 + \dots + \alpha_\ell \lambda^\ell, \quad \beta = \beta_1 \lambda^1 + \dots + \beta_\ell \lambda^\ell, \dots$$

There occur ℓ linearly independent linear forms among $\alpha, \beta, \dots$.

If α is a fixed root, all the characteristic roots ρ of H , ($\rho = 0, \alpha, \beta, \dots$) are distributed in strings of the form

$$(46) \quad \rho - R\alpha, \dots, \rho - \alpha, \rho, \rho + \alpha, \dots, \rho + R\alpha$$

All the elements of this string are roots or 0, and it contains all roots whose difference from ρ is of the form $n \cdot \alpha$, n integer. The root α belongs to the string $-\alpha, 0, \alpha$. Therefore, $n\alpha$ is not a root for $n = \pm 2, \pm 3, \dots$

To every root α belongs an element e_α of $\mathfrak{g}$ such that $[h, e_\alpha] = \alpha e_\alpha$. Here e_α is determined uniquely but for a constant factor. The elements

$$(47) \quad h_1, h_2, \dots, h_\ell, e_\alpha, e_\beta, \dots$$

form a basis of $\mathfrak{g}$. The structure is given by

$$(48) \quad \left\{ \begin{array}{l} [h, h'] = 0 \text{ for } h, h' \text{ in } \mathfrak{f}; \\ [h, e_\alpha] = \alpha e_\alpha \\ [e_\alpha, e_\beta] = \begin{cases} 0, & \alpha + \beta \text{ neither a root nor } 0 \\ h_\alpha, & \alpha + \beta = 0 \\ N_{\alpha\beta} e_{\alpha+\beta}, & \alpha + \beta \text{ a root.} \end{cases} \end{array} \right. \quad (h = \lambda^1 h_1 + \dots + \lambda^\ell h_\ell)$$

Here $N_{\alpha\beta}$ is a scalar and h_α an element of $\mathfrak{g}$

$$(49) \quad h_\alpha = \alpha^1 h_1 + \dots + \alpha^l h_l$$

The number

$$(50) \quad \alpha_\alpha = \alpha^1 \alpha_1 + \dots + \alpha^l \alpha_l$$

is different from 0. The length $k + k' + 1$ of the string (46) is given by

$$(51) \quad g + 1 = 2 \frac{\Lambda_\alpha}{\alpha_\alpha} + 1$$

where $\Lambda = \rho + k\alpha$ is the "extreme" root in the string and Λ_α the value of Λ for $\lambda^1 = \alpha^1, \dots, \lambda^l = \alpha^l$. Let

$$(52) \quad a = \lambda^1 h_1 + \dots + \lambda^l h_l + \sigma_\alpha e_\alpha + \sigma_\beta e_\beta + \dots = h + \sigma_\alpha e_\alpha + \sigma_\beta e_\beta + \dots$$

be an arbitrary element of $\mathfrak{g}$. Let A be of the corresponding transformation $x \rightarrow [ax]$, (x ranges over $\mathfrak{g}$), in the adjoint representation. Then

$$(53) \quad \sigma(a) = \text{tr}(A^2) = \sigma(h) + \sum_{\rho=\alpha, \beta, \dots} \sigma_\rho \sigma_{-\rho} N_\rho = \sum_{\rho=\alpha, \beta, \dots} \rho(h)^2 + \sum_{\rho=\alpha, \beta, \dots} \sigma_\rho \sigma_{-\rho} N_\rho$$

is a non-degenerate quadratic form in $\lambda^1, \dots, \lambda^l, \sigma_\alpha, \sigma_\beta, \dots$.

Here

$$(54) \quad N_\alpha = \frac{1}{2} \alpha_\alpha \sum \frac{g(g+1)(g+2)}{1 \cdot 2 \cdot 3} = \frac{\alpha_\alpha}{2} \sum (g+2)$$

where $g+1$ is the length (51) of the string and where the sum extends over all the different strings. The number N_α is of the form $M_\alpha \cdot \alpha_\alpha$, M_α an integer ≥ 2 (as we are going to show). The bilinear form $\sigma(a, b)$ corresponding to $\sigma(a)$ satisfies

$$(55) \quad \sigma([c, a], b) + \sigma(a, [c, b]) = 0, \quad (a, b, c \text{ in } \mathfrak{g})$$

The form $\sigma(a, b)$ is invariant, $d\sigma = 0$ for the adjoint transformations $da = [ca]$, $db = [cb]$, c in $\mathfrak{g}$.

Remark: For the string $-\alpha, 0, \alpha$ we have $g = 2$, $\frac{g(g+1)(g+2)}{1 \cdot 2 \cdot 3} = 4$.

We consider another string, choosing β as the extreme root of the string

The negatives of these roots

$$-\rho, -\rho + \alpha, \dots, -\rho + g\alpha$$

are roots again. Obviously they form a string. We assume that both strings

are identical. Then $\rho = -\rho + g\alpha$, $\rho = \frac{g}{2}\alpha$. Here, g cannot be even because no integral multiple of α is a root and $\rho \neq \pm \alpha$. But g cannot be odd, because otherwise $\beta = \frac{\alpha}{2}$ would appear in our string

$\frac{g}{2}\alpha, (\frac{g}{2}-1)\alpha, \dots, -\frac{g}{2}\alpha$ and $\alpha = 2\beta$ could not be a root. Hence both strings are not identical. They have the same g . One number g in (54)

is 2 . The others appear in pairs of equal ones. Hence $N_\alpha = M_\alpha \cdot \alpha$

where M_α is an integer, not a half-integer, and $M \geq 2$.

An infinitesimal group $\mathfrak{g}$ is called the direct sum of two subgroups $\mathfrak{g}_1$ and $\mathfrak{g}_2$, if every element of $\mathfrak{g}$ can be represented as the sum of one element of $\mathfrak{g}_1$ and one element of $\mathfrak{g}_2$, and if furthermore $[ab] = 0$ for a in $\mathfrak{g}_1$ and b in $\mathfrak{g}_2$. One sees easily that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are invariant subgroups of $\mathfrak{g}$ in this case.

We shall prove as an application of our results

Theorem. Every semi-simple infinitesimal group can be represented as the direct sum of simple groups.

Proof. Let $\mathfrak{g}$ be semi-simple of order r , and $\mathfrak{g}_1$ be an invariant subgroup of order r_1 . We determine the set $\mathfrak{g}_2$ of elements a of $\mathfrak{g}$ for which

$$\sigma(a, x) = 0 \quad \text{for all } x \text{ in } \mathfrak{g}_1$$

This gives r_1 linear homogeneous equations for the parameters of a . Hence $\mathfrak{g}_2$ is a vector space of at least $r - r_1$ dimensions. If c is any element of $\mathfrak{g}$ we obtain from (55)

$$\sigma([ca], x) + \sigma(a, [cx]) = 0$$

If x and therefore $[cx]$ lies in $\mathfrak{g}_1$, and if a belongs to $\mathfrak{g}_2$, then the second summand is 0. Hence $\sigma([ca], x) = 0$ for all x in $\mathfrak{g}_1$. This proves that $[ca]$ belongs to $\mathfrak{g}_2$. The vector space $\mathfrak{g}_2$ is therefore an invariant subgroup of $\mathfrak{g}$.

Let a be an element of the intersection $\mathfrak{h}$ of $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Then $\sigma(a) = \sigma(a, a) = 0$ for all a . But $\mathfrak{h}$ is an invariant subgroup of $\mathfrak{g}$. If $\sigma^*(a)$ has the same significance for $\mathfrak{h}$ as $\sigma(a)$ has for $\mathfrak{g}$, we have $\sigma^*(a) = \sigma(a)$ for a in $\mathfrak{h}$ (compare the analogous remark concerning $\mathfrak{g}$ instead of $\mathfrak{h}$ on pages 81, ⁸² note. Criterion 1 yields that $\mathfrak{h}$ is solvable. Hence $\mathfrak{h} = 0$ because, by definition, the semi-simple group $\mathfrak{g}$ has no solvable invariant subgroup.

$\mathfrak{g}_1$ and $\mathfrak{g}_2$ are two vector spaces without common elements $\neq 0$. The sum of the dimensionality is at least r . Hence

$$\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{g}_2$$

in the sense of addition of vector spaces. But $[g_1, g_2]$ belongs to $\mathfrak{g}_1$ and $\mathfrak{g}_2$ if $g_1 \in \mathfrak{g}_1$ and $g_2 \in \mathfrak{g}_2$. Hence $[g_1, g_2] = 0$. Then $\mathfrak{g}$ is the direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Obviously, $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are semi-simple. If they are not simple we repeat the same argument and so on. We obtain finally a representation of $\mathfrak{g}$ as direct sum of simple groups.

4. The finite group (S) and the metric of the λ -space

We consider the linear forms in $\lambda^1, \dots, \lambda^r$

$$\xi = x_1 \lambda^1 + \dots + x_r \lambda^r$$

All these forms ξ form a vector space $\mathcal{L}$. The coefficients $x_1, \dots, x_r$ are the components of the "vector" ξ , $\xi = (x_1, \dots, x_r)$.

Let $\mathfrak{g}$ be a semi-simple group. Every root α determines a linear transformation S_α in $\mathcal{L}$

$$(56) \quad S_\alpha: \xi \longrightarrow \xi - \frac{2\xi_\alpha}{\alpha_\alpha} \alpha$$

$$\xi_\alpha = x_1 \alpha^1 + \dots + x_l \alpha^l$$

We call ξ an integral vector if $2\xi_\alpha/\alpha_\alpha$ is an integer for every root α

These integral vectors form a lattice. In particular the roots $\alpha, \beta, \dots$ are integral vectors (cf. (51)). Hence, the lattice is transformed into itself by (56). Let

$$\rho, \rho - \alpha, \dots, \rho - g\alpha$$

be a string of roots. Then $g = \frac{2\rho_\alpha}{\alpha_\alpha}$ according to (51). We have

$$\rho - \nu\alpha \longrightarrow \rho - \frac{2\rho_\alpha}{\alpha_\alpha} \alpha - \nu\left(\alpha - \frac{2\alpha_\alpha}{\alpha_\alpha} \alpha\right) = \rho - (g-\nu)\alpha$$

($\nu = 0, 1, 2, \dots, g$) under S_α . The transformation S_α transforms roots into roots.

Let (S) be the group generated by the transformations S_α . Every element S^* effects a permutation of the roots $\alpha, \beta, \dots$. Conversely, S^* is determined uniquely by this permutation because there are ℓ linearly independent roots. It follows that (S) is of finite order s . The number of different roots being $r-\ell$, we have $s \leq (r-\ell)!$

We replace e_α by $w_\alpha e_\alpha$, $w_\alpha \neq 0$ scalar, as we may. Then σ_α in (52) has to be replaced by $w_\alpha^{-1} \sigma_\alpha$ and hence N_α in (53) by $N_\alpha w_\alpha^{-1} w_{-\alpha}^{-1}$. This shows that we may assume $N_\alpha = 1$ in (53). We denote $\sigma(h)$ in (53) by $g(\lambda)$. This is a non-degenerate quadratic form in $\lambda^1, \dots, \lambda^r$, the sum of the squares of all roots $\alpha, \beta, \dots$. We put

$$(57) \quad g(\lambda) = \sum_{i,k} g_{i,k} \lambda^i \lambda^k, \quad g(\lambda, \mu) = \sum_{i,k} g_{i,k} \lambda^i \mu^k$$

We consider two elements a, b of $\mathfrak{g}$

$$(58) \quad \begin{cases} a = \sum \lambda^i h_i + \sum_{\substack{\rho=\alpha, \beta, \dots}} \sigma_\rho e_\rho = h + \sum \sigma_\rho e_\rho \\ b = \sum \mu^i h_i + \sum_{\substack{\rho=\alpha, \beta, \dots}} \tau_\rho e_\rho = h' + \sum \tau_\rho e_\rho \end{cases}$$

Then the bilinear form $\sigma(a, b)$ reads (cf. (53))

$$(59) \quad \sigma(a, b) = g(\lambda, \mu) + \sum_{\rho=\alpha, \beta, \dots} \sigma_{\rho} \tau_{-\rho}$$

We derive from (55) for $a = c$ that

$$(60) \quad \sigma(a, [ca]) = 0$$

We choose here $c = e_{\alpha}$ and have, according to (48)

$$[e_{\alpha} a] = -\alpha e_{\alpha} + \sigma_{-\alpha} h_{\alpha} + \sum \sigma_{\rho} N_{\alpha, \rho} e_{\alpha+\rho}$$

where the sum extends over all ρ for which $\alpha + \rho$ is a root again. We

take $b = [e_{\alpha}, a]$ in (58), and have (cf. (49))

$$h' = \sigma_{-\alpha} h_{\alpha}; \mu' = \sigma_{-\alpha} \alpha', \dots, \mu^l = \sigma_{-\alpha} \alpha^l$$

$$\tau_{\alpha+\rho} = \begin{cases} -\alpha & , \rho = 0 \\ \sigma_{\rho} N_{\alpha, \rho} & , \alpha + \rho \text{ root} \\ 0 & , \alpha + \rho \neq 0 \text{ and not a root.} \end{cases}$$

Putting this into (59) we obtain from (60) ($\alpha + \rho = -K$)

$$0 = \sigma_{-\alpha} g(\lambda, \alpha') + \sigma_{-\alpha} (-\alpha) + \sum \sigma_{\kappa} \sigma_{-\alpha-\kappa} N_{\alpha, -\alpha-\kappa}$$

In the sum, K runs over all roots for which $\rho = -\alpha - K$ is also a root.

This holds identically in λ and the σ 's. Hence

$$g(\lambda, \alpha') - \alpha = 0$$

$$(61) \quad N_{\alpha, \rho} + N_{\alpha, \kappa} = 0 \quad (\alpha, \kappa, \rho \text{ roots, } \alpha + \kappa + \rho = 0)$$

Using the notation (57), we write the first condition in the form

$$\sum \alpha_i \lambda^i = \sum g_{ik} \lambda^i \alpha^k$$

$$(62) \quad \alpha_i = \sum_k g_{ik} \alpha^k$$

The determinant $|g_{ik}|$ does not vanish $g(\lambda)$ being not degenerate. Let

$\|g^{ik}\|$ be the inverse matrix of g_{ik} :

$$(63) \quad \alpha^i = \sum_k g^{ik} \alpha_k$$

We define a metric in the vector space $\mathcal{L}$ by setting

$$(64) \quad (\xi, \eta) = \sum_{i, k} g^{ik} x_i y_k \quad \text{for } \xi = \sum x_i \lambda^i, \eta = \sum y_i \lambda^i$$

Equation (63) shows that $\alpha^1, \dots, \alpha^\lambda$ are the contravariant components of the vector α and the choice of notation α_i, α^i in (45) and (49) is justified. Furthermore, (63) enables us to find h_α if α and the metric are known. If α, β and $\alpha + \beta$ ^(are roots) $= \gamma$, (63) yields

$$(65) \quad h_{\alpha+\beta} = h_\alpha + h_\beta, \quad \alpha, \beta, \alpha+\beta \text{ roots}$$

Furthermore, we obtain from (63)

$$(66) \quad \xi_\alpha = \sum_i x_i \alpha^i = \sum g^{ik} x_i \alpha_k = (\xi, \alpha)$$

In particular, $\alpha_\alpha = (\alpha, \alpha)$

The transformation S_α , (56), can be written in the form

$$(67) \quad S_\alpha : \xi \rightarrow \xi' = \xi - 2 \frac{(\xi, \alpha)}{(\alpha, \alpha)} \alpha$$

It is a reflection in the plane orthogonal to α through the origin.

There are λ independent vectors among the roots of $\mathfrak{g}$. We can assume that they are chosen as the coordinate system in $\mathcal{L}$. Then $\lambda^1, \lambda^2, \dots, \lambda^\lambda$ are roots.

Let $\xi = x_1 \lambda^1 + \dots + x_\lambda \lambda^\lambda$ be an integral

vector. Then $\frac{2\xi_\alpha}{\alpha_\alpha}$ is an integer for $\alpha = \lambda^i$ ($i = 1, 2, \dots, \lambda$).

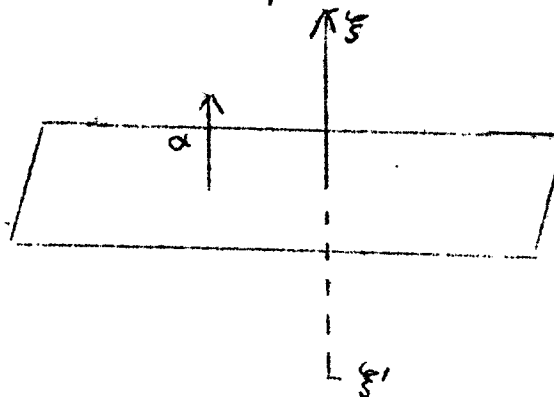
From (66) follows for $\alpha = \lambda^i$

$$(*) \quad 2 \frac{\xi_\alpha}{\alpha_\alpha} = 2 \frac{(\xi, \alpha)}{(\alpha, \alpha)} = 2 \sum_k g^{ik} x_k / g_{ii} = u_i, \quad (u_i \text{ integer}).$$

If, for the moment, $\xi = \lambda^k$, we see that $2g^{ik}/g_{ii}$ is an integer.

We can solve these linear equations for $x_1, \dots, x_\lambda$ since $|g^{ik}| \neq 0$.

We see that the x_k are rational numbers with a fixed denominator which is independent of the special integral vector ξ .



It follows then in a well known manner that there are λ integral vectors ξ such that every integral vector is a linear combination of them with integral coefficients. If we choose now these λ vectors ξ as the coordinate system, the integral vectors are just the vectors with integral components.

In particular, the roots ρ have all integral coefficients.

The quadratic form $Q(\lambda) = \sigma(h) = \sum_{\rho \in \alpha, \beta, \dots} \rho(h)^2$ has integral coefficients and is positive definite if we restrict λ to real values.

Results: It is possible to associate a "vector diagram" with every semi-simple group $\mathfrak{g}$ of order r . This consists of $r-\lambda$ vectors in an λ -dimensional vector space $\mathcal{L}$ in which a metric is defined by a non-degenerate quadratic form. λ is the order of the maximum Abelian subgroup $\mathfrak{f}$ of $\mathfrak{g}$. If α and β are two vectors of the diagram, the vector β' obtained from β by the reflection in the plane orthogonal to α belongs to the diagram. Furthermore, $\beta - \beta'$ is an integral multiple $g\alpha$ of α and all the vectors $\beta, \beta - \alpha, \dots, \beta - g\alpha$ belong to the diagram. All vectors of the diagram have integral coefficients when the coordinate system in $\mathcal{L}$ is properly chosen. The metric fundamental form is positive definite for this choice of the coordinate system if we restrict ourselves to vectors with real components.

There exists a finite group (S) of linear transformations in $\mathcal{L}$ which is generated by the reflections S_α in the plane orthogonal to α (α vector of the diagram). The whole diagram is transformed into itself under the transformations of (S).

The problem of determining all semi-simple groups was solved first by Cartan in his thesis. Another method was given by Artin and van der Waerden (Math. Zeitschr. 37, p.446) who constructed all possible vector dia-

grams. Coxeter constructed all finite groups generated by reflections. His results yield a third method (cf. the last section of Chapter 3.)

5. Determination of the infinitesimal group by the vector diagram

We want to prove now:

Theorem. The semi-simple infinitesimal group $\mathfrak{g}$ is uniquely determined by its vector diagram.

If the vector diagram is given, we know the roots α , hence the form

$g(\lambda) = \sum_{\alpha \in \Delta} \rho(\alpha)^2$, (57) and the h_α , (49) and (63). The whole constitution of $\mathfrak{g}$, (48), will be known if the $N_{\alpha\beta}$ are known.

We have the following arbitrariness in the construction of the diagram:

- 1) There are many possibilities to choose a maximum Abelian subgroup $\mathfrak{f}$. This is not of interest at this moment.
- 2) We can replace the basis $h_1, \dots, h_\ell$ of $\mathfrak{f}$ by another basis. This corresponds only to transition to another coordinate system in $\mathcal{L}$. If the coordinate system is chosen in the manner described in Chapter 2, 4, "Results", only unimodular integer coordinate transformations are permitted.
- 3) We can replace e_α by $\mu_\alpha e_\alpha$, μ_α scalar. The only necessary condition for the μ_α is that

$$(68) \quad \mu_\alpha \cdot \mu_{-\alpha} = 1$$

lest the normal form (58) of $\sigma(a)$ be changed. Then $N_{\alpha\beta}$ is to be replaced by

$$(69) \quad N'_{\alpha\beta} = N_{\alpha\beta} \frac{\mu_\alpha \mu_\beta}{\mu_{\alpha+\beta}}$$

(cf. (48)). Here α , β and $\alpha + \beta$ are roots ($N_{\alpha,\beta}$ is only defined for subscripts α , β satisfying this condition. We will call such systems N and N' , (69), equivalent systems.

We will derive now the conditions which are fulfilled by the $N_{\alpha\beta}$

We set $N_{\alpha\beta} = 0$ if $\alpha + \beta \neq 0$ and not a root. From (48) follows

$$(70) \quad N_{\alpha\beta} = -N_{\beta\alpha} \quad (\alpha + \beta \neq 0)$$

Three roots α, β, γ for which $\alpha + \beta + \gamma = 0$ will be called a triangle.

Equation (61) yields $N_{\alpha\beta} = -N_{\alpha\gamma} = N_{\gamma\alpha}$ and by permuting α, β, γ cyclically we have

$$(71) \quad N_{\alpha\beta} = N_{\beta\gamma} = N_{\gamma\alpha} \quad (\alpha, \beta, \gamma \text{ a triangle})$$

Another relation can be obtained from Jacobi's identity

$$[e_\alpha, [e_\beta, e_\gamma]] + [e_\beta, [e_\gamma, e_\alpha]] + [e_\gamma, [e_\alpha, e_\beta]] = 0$$

Let α, β, γ , and $\alpha + \beta + \gamma$ be roots and $\beta + \gamma \neq 0, \gamma + \alpha \neq 0,$

$\alpha + \beta \neq 0$. We obtain from (48)

$$(72) \quad \begin{aligned} & N_{\beta\gamma} [e_\alpha, e_{\beta+\gamma}] + N_{\gamma\alpha} [e_\beta, e_{\gamma+\alpha}] + N_{\alpha\beta} [e_\gamma, e_{\alpha+\beta}] = 0 \\ & N_{\beta\gamma} N_{\alpha, \beta+\gamma} + N_{\gamma\alpha} N_{\beta, \gamma+\alpha} + N_{\alpha\beta} N_{\gamma, \alpha+\beta} = 0 \end{aligned}$$

after suppressing $e_{\alpha+\beta+\gamma} \neq 0$. We denote the root $-\alpha$ by α' , $-\beta$ by

β' , and so on. If $\alpha + \beta + \gamma = -\delta$; $\alpha, \beta, \gamma, \delta$ form a quadrangle

of roots. These are four roots $\alpha, \beta, \gamma, \delta$ with vanishing sum

$\alpha + \beta + \gamma + \delta = 0$ for which $\beta + \gamma \neq 0, \gamma + \alpha \neq 0, \alpha + \beta \neq 0$. Then (71) gives us

$$N_{\alpha, \beta+\gamma} = N_{\delta\alpha} = -N_{\alpha\delta}, \quad N_{\beta, \gamma+\alpha} = -N_{\beta\delta}, \quad N_{\gamma, \alpha+\beta} = -N_{\gamma\delta}$$

and (72) reads

$$(72^*) \quad N_{\alpha\delta} N_{\beta\gamma} + N_{\beta\delta} N_{\gamma\alpha} + N_{\gamma\delta} N_{\alpha\beta} = 0 \quad (\alpha, \beta, \gamma, \delta \text{ quadrangle})$$

Let the root β belong to the α -string

$$\beta - R'\alpha, \dots, \beta - \alpha, \alpha, \beta + \alpha, \dots, \beta + R\alpha, \quad (\beta + R\alpha = 0)$$

If e_0 is a vector belonging to $\beta + R\alpha$, we constructed in (19) a vector

e_R belonging to β . Hence $e_R = c e_\beta$, $c \neq 0$ scalar, because there is only one independent vector belonging to β . We may assume $e_R = e_\beta$

If $e_\alpha \rightarrow E_\alpha, e_{-\alpha} \rightarrow E_{-\alpha}$ in the adjoint representation, we obtain from (1)

(21), (22), ($g = k+k'$)

$$E_{-\alpha} E_{\alpha} e_k = \mu_k E_{-\alpha} e_{k-1} = \mu_k e_k = (k\lambda_{\alpha} - \frac{k(k-1)}{2} \alpha_{\alpha}) e_k$$

$$E_{-\alpha} E_{\alpha} e_k = k(\frac{k+k'}{2} - \frac{k-1}{2}) \alpha_{\alpha} e_k = \frac{k(k'+1)}{2} \alpha_{\alpha} e_k$$

Now, $E_{\alpha} e_k = [e_{\alpha}, e_k]$, $E_{-\alpha} [e_{\alpha}, e_k] = [e_{-\alpha}, [e_{\alpha} e_k]]$

$$[e_{-\alpha}, [e_{\alpha} e_{\beta}]] = \frac{k(k'+1)}{2} \alpha_{\alpha} e_{\beta}$$

$$[e_{-\alpha}, [e_{\alpha}, e_{\beta}]] = [e_{-\alpha}, N_{\alpha\beta} e_{\alpha+\beta}] = N_{\alpha\beta} N_{-\alpha, \alpha+\beta} e_{\beta}$$

According to (71), we have $N_{-\alpha, \alpha+\beta} = N_{-\beta, -\alpha} = -N_{\alpha', \beta'}$. We denote the factor of e_{β} on the right side by $R_{\alpha\beta}$ and have

$$(73) \quad N_{\alpha\beta} N_{\alpha', \beta'} = -R_{\alpha\beta} = -\frac{k(k'+1)}{2} (\alpha, \alpha)$$

($\alpha, \beta, \alpha + \beta$ roots, $\beta + k\alpha$ and $\beta - k'\alpha$ extreme)

It is important that $R_{\alpha\beta}$ is known if the vector diagram is given. Obviously, $R_{\alpha\beta} > 0$.

We arrange the roots in the following way: We write $\alpha > 0$ if the first non-vanishing coefficient in α is positive. We write $\alpha > \beta$ if $\alpha - \beta > 0$.

After these preparations we will begin with the proof of the theorem stated above. We assume that there are two semi-simple groups $\mathfrak{g}$ and $\mathfrak{g}^*$ with the same vector diagram, the constants in (48) being $N_{\alpha\beta}$ and $N_{\alpha\beta}^*$ respectively. We have to show that both systems are equivalent.

Let ρ be a fixed root. We assume that the μ_{α} have been determined for $-\rho < \alpha < \rho$ such that (68) holds and

$$(74) \quad N_{\alpha\beta}^* = \frac{\mu_{\alpha} \mu_{\beta}}{\mu_{\alpha+\beta}} N_{\alpha\beta}$$

is true for $-\rho < \alpha, \beta, \omega < \rho$; $\alpha + \beta + \omega = 0$. We may then assume after replacing $N_{\alpha\beta}^*$ by an equivalent system that we have

$$(75) \quad N_{\alpha\beta}^* = N_{\alpha\beta}, \quad \mu_{\alpha} = 1 \text{ for } -\rho < \alpha, \beta, \omega < \rho; \alpha + \beta + \omega = 0.$$

We have to determine μ_p so that (74) holds for $-p \leq \alpha, \beta, \omega \leq p$ if one of α, β, ω has the value $\pm p$.

Let γ, δ be two roots with $-p \leq \gamma \leq p, -p \leq \delta \leq p$, $\gamma + \delta = p$. Then $0 < \gamma < p, 0 < \delta < p$. If α, β are two roots with $-p \leq \alpha \leq p, -p \leq \beta \leq p, \alpha + \beta = -p$, then $-p < \alpha < 0, -p < \beta < 0$ and $\alpha, \beta, \gamma, \delta$ form a quadrangle except in the case $\alpha = -\gamma$ or $-\delta$. From (72*) it follows that

$$N_{\alpha\beta} N_{\gamma\delta} = -N_{\alpha\delta} N_{\beta\gamma} - N_{\beta\delta} N_{\gamma\alpha}$$

The same relation holds for N^* . There are only N on the right side for which the indices and their sums lie between $-p$ and p . For instance

$-p < \alpha < 0, 0 < \delta < p, -p < \alpha + \delta < p$. The value of the right side is not changed by replacing N by N^* according to (75) (if, say, $\alpha + \delta$ is not a root, $N_{\alpha\delta} = 0 = N^*_{\alpha\delta}$). Hence

$$(76) \quad N_{\alpha\beta} N_{\gamma\delta} = N^*_{\alpha\beta} N^*_{\gamma\delta}$$

We determine now μ_{-p} so that for our fixed pair γ, δ holds

$$(77) \quad N^*_{\gamma\delta} = \mu_{-p} N_{\gamma\delta}$$

Then (76) implies (74) because of $\mu_\alpha = \mu_p = 1, -p = \alpha + \beta$. Hence

(74) holds for all α, β with $-p \leq \alpha, \beta \leq p, \alpha + \beta = -p$ except $\alpha = -\gamma, \beta = -\delta$ and $\alpha = -\delta, \beta = -\gamma$. But from (73) it can be seen that

$$(78) \quad N_{\alpha\beta} N_{\alpha'\beta'} = N^*_{\alpha\beta} N^*_{\alpha'\beta'}$$

because $R_{\alpha\beta}$ depends only on the vector diagram and is, consequently, the same for both groups $\mathcal{G}$ and $\mathcal{G}^*$. We apply this for $\alpha = \gamma, \beta = \delta$ and obtain from (77) that $N^*_{\gamma,\delta'} = N_{\gamma,\delta'} / \mu_{-p}$. Hence (74) is true for

$\alpha = \gamma' = -\gamma, \beta = -\delta$. Finally (74) is true in the other exceptional case $\alpha = -\delta, \beta = -\gamma$ because of (70).

We derive from (74), by using (71) and $\alpha + \beta + \omega = 0$, that

$$N_{\beta\omega}^* = \frac{\mu_\beta \mu_\omega}{\mu_{\beta+\omega}} N_{\beta\omega}, \quad N_{\omega\alpha}^* = \frac{\mu_\omega \mu_\alpha}{\mu_{\omega+\alpha}} N_{\omega\alpha}$$

if we set $\mu_\rho = 1/\mu_{-\rho}$ (The fractions are the same as in (74) since

$$\mu_\alpha = 1, \mu_\beta = 1, \mu_{\beta+\omega} = \mu_{-\alpha} = 1, \mu_{\alpha+\omega} = 1,$$

$$\mu_\omega = \mu_{-\alpha-\beta} = \mu_\rho = 1/\mu_{-\rho}.) \quad (74) \text{ holds therefore for all}$$

"triangles" α, β, ω with $-\rho \leq \alpha, \beta, \omega \leq \rho$ for which one of the roots at least is ρ . The same is true if one has the value $-\rho$ as is seen by using (78) and $\mu_\alpha = 1/\mu_{-\alpha}, \mu_\beta = 1/\mu_{-\beta}, \mu_\omega = 1/\mu_{-\omega}$.

We used the existence of roots γ, δ between $-\rho$ and ρ whose sum is ρ But if there are no such roots, we have no roots ^{either} between $-\rho$ and ρ whose sum is $-\rho$, since the negative of a root is always a root. In this case (74) is certainly true for all triangles α, β, ω and this finishes the proof of our theorem.

$$\text{with } -\rho \leq \alpha, \beta, \omega \leq \rho$$

We used only that $N_{\alpha\beta}$ and $N_{\alpha\beta}^*$ satisfy (70), (71) (72*) and that $R_{\alpha\beta}$ is the same number for both in (73). Then we could prove that $N_{\alpha\beta}^*$ and $N_{\alpha\beta}$ are equivalent sets.

If $N_{\alpha\beta}$ satisfies the conditions and if we set $N_{\alpha\beta}^* = -N_{\alpha'\beta'}$ then $N_{\alpha\beta}^*$ satisfies the conditions with the same $R_{\alpha\beta}$. Consequently there are μ_α such that

$$(79) \quad -N_{\alpha'\beta'} = \frac{\mu_\alpha \mu_\beta}{\mu_{\alpha+\beta}} N_{\alpha\beta} \quad ; (\alpha, \beta, -\alpha-\beta \text{ roots})$$

$$\mu_{\alpha'} = 1/\mu_\alpha$$

We put $\tilde{\mu}_\alpha = \sqrt{\mu_\alpha}$ where the sign in the radical is fixed so that $\tilde{\mu}_{\alpha'} = 1/\tilde{\mu}_\alpha$

From (79) follows

$$\begin{aligned} -N_{\alpha'\beta'} &= \frac{\tilde{\mu}_\alpha^2 \tilde{\mu}_\beta^2}{\tilde{\mu}_{\alpha+\beta}^2} N_{\alpha\beta} \\ -\frac{\tilde{\mu}_{\alpha'} \tilde{\mu}_{\beta'}}{\tilde{\mu}_{\alpha'+\beta'}} N_{\alpha'\beta'} &= \frac{\tilde{\mu}_\alpha \tilde{\mu}_\beta}{\tilde{\mu}_{\alpha+\beta}} N_{\alpha\beta} \end{aligned}$$

The equivalent system $\frac{\tilde{\mu}_{\alpha}\tilde{\mu}_{\beta}}{\tilde{\mu}_{\alpha\beta}} N_{\alpha\beta} = \tilde{N}_{\alpha\beta}$ satisfies the condition $-\tilde{N}_{\alpha'\beta'} = \tilde{N}_{\alpha\beta}$. We replace $N_{\alpha\beta}$ by the equivalent system $\tilde{N}_{\alpha\beta}$ as we may. We fix then the arbitrary scalar factor in e_{α} . We have

$$-N_{\alpha'\beta'} = N_{\alpha\beta}$$

From (73) follows $-N_{\alpha\beta}^2 = -R_{\alpha\beta}$. But $R_{\alpha\beta}$ is positive. We see that $N_{\alpha\beta}$ is real and $\neq 0$, ($\alpha, \beta, \alpha + \beta$ roots).

Results: The elements e_{α} in $\mathfrak{g}$ can be chosen such that $N_{\alpha\beta}$ is ^{isreal} in (48) and fulfills the condition $-N_{-\alpha, -\beta} = N_{\alpha\beta} \neq 0$ if α, β and $\alpha + \beta = -\gamma$ are roots. Under the same assumption $N_{\alpha\beta} = N_{\beta\gamma} = N_{\gamma\alpha}$ holds. We have $N_{\alpha\beta} = -N_{\beta\alpha}$. If α and β are roots but not $\alpha + \beta$, we set $N_{\alpha\beta} = 0$. Finally,

$$N_{\alpha\delta} N_{\beta\gamma} + N_{\beta\delta} N_{\gamma\alpha} + N_{\gamma\delta} N_{\alpha\beta} = 0$$

if $\alpha, \beta, \gamma, \delta$ are roots and $\alpha + \beta + \gamma + \delta = 0, \beta + \gamma \neq 0, \gamma + \alpha \neq 0, \alpha + \beta \neq 0$.

We assume now that it is possible to split the vector diagram into two parts, the first part lying in a subspace $\mathcal{L}_1$, and the second part in the subspace $\mathcal{L}_2$ orthogonal to $\mathcal{L}_1$ (i.e. every vector of $\mathcal{L}_2$ is orthogonal to every vector of $\mathcal{L}_1$). We choose the coordinates such that the first λ' components vanish for the elements of $\mathcal{L}_2$, and the last $\lambda - \lambda'$ components vanish for the vectors of $\mathcal{L}_1$. We have then $g_{ik} = 0$ for $i > \lambda', k \leq \lambda'$ in (64). It follows from (63) that the $\lambda - \lambda'$ last covariant components of α vanish also if α lies in $\mathcal{L}_1$.

The subspace $\mathfrak{g}_1$ of $\mathfrak{g}$ spanned by $\lambda' h_1 + \dots + \lambda' h_{\lambda'}$ and e_{α} for α in $\mathcal{L}_1$, forms a subgroup of $\mathfrak{g}$ as follows from (48). If, namely, α lies in $\mathcal{L}_1$, we have $\alpha^{(\ell+1)} = \dots = \alpha^{(\ell)} = 0$ and h_{α} belongs to $\mathfrak{g}_1$. If α and β be-

long to $\mathcal{L}_1$, so does $\alpha + \beta$.

Similarly, the elements

$$\lambda^{(e_1+1)} h_{e_1+1} + \dots + \lambda^{(e_l)} h_{e_l} \quad \text{and } e_\beta, \beta \text{ in } \mathcal{L}_2$$

define a subgroup $\mathcal{G}_2$. If α lies in $\mathcal{L}_1$ and β in $\mathcal{L}_2$, $\alpha + \beta$ is not a root, for $\alpha + \beta$ can lie neither in $\mathcal{L}_1$ (otherwise β would lie in $\mathcal{L}_1$) nor in $\mathcal{L}_2$. From (48) follows $[uv] = 0$ for u in $\mathcal{G}_1$, v in $\mathcal{G}_2$.

Hence $\mathcal{G}$ is the direct sum of $\mathcal{G}_1$ and $\mathcal{G}_2$

$$\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2$$

and $\mathcal{G}_1$ and $\mathcal{G}_2$ are invariant subgroups.

If, conversely, a group is the direct sum of two semi-simple groups, the vector diagram splits into two parts which lie in orthogonal spaces.

By splitting the vector diagram in mutually orthogonal parts which can not again be split themselves, we obtain again (cf. p. 88)

Theorem. Every semi-simple group is the direct sum of simple groups.

We have not proved yet that the vector diagram is uniquely determined by $\mathcal{G}$. There is an ambiguity with respect to the choice of $\mathcal{L}$. Nor have we proved that to every vector diagram belongs a group when all conditions are fulfilled.

6. The unitary restriction

Let $\mathcal{G}$ be a semi-simple group. We choose the basis (47) as in Chapter 2, secs. 4 and 5, such that all the roots have rational coefficients and that the $N_{\alpha\beta}$ are real. (Cf. the summaries on pp. 93, 99)

The quadratic form $\sigma(a)$ reads

$$\sigma(a) = \mathcal{G}(\lambda) + \sum \sigma_\alpha \sigma_{-\alpha} \quad \text{for } a = \sum_i \lambda^i h_i + \sum_\alpha \sigma_\alpha e_\alpha$$

We consider now only such elements a for which

a) $\lambda^j = 2\pi i \varphi^j$ is purely imaginary. We call $\varphi^1, \dots, \varphi^l$ the angles of a .

b) $\sigma_{-\alpha} = -\bar{\sigma}_{\alpha}$

These form obviously a vector space $\tilde{\mathfrak{N}}$ of r real dimensions. We can use as basis of $\tilde{\mathfrak{N}}$:

$$ih_1, \dots, ih_r, e_{\alpha} - e_{-\alpha}, i(e_{\alpha} + e_{-\alpha}), \dots$$

$\sigma(a)$ is negative definite for a in $\tilde{\mathfrak{N}}$ since $\mathcal{G}(\lambda)$ was positive definite for real λ .

We state now: $\tilde{\mathfrak{N}}$ is an infinitesimal group. We have to show that

the brackets of elements of the form h with purely imaginary λ and

$\sigma_{\alpha} e_{\alpha} + \sigma_{-\alpha} e_{-\alpha} = \sigma_{\alpha} e_{\alpha} - \bar{\sigma}_{\alpha} e_{-\alpha}$ belong to $\tilde{\mathfrak{N}}$. But we have $[h, h'] = 0$,

$$[h, \sigma_{\alpha} e_{\alpha} - \bar{\sigma}_{\alpha} e_{-\alpha}] = \sigma_{\alpha} \alpha e_{\alpha} - \bar{\sigma}_{\alpha} (-\alpha) e_{-\alpha}.$$

This belongs to $\tilde{\mathfrak{N}}$ because $-\alpha = \bar{\alpha}$ for purely imaginary λ .

$$(80) \left\{ \begin{aligned} & [\sigma_{\alpha} e_{\alpha} - \bar{\sigma}_{\alpha} e_{-\alpha}, \tau_{\beta} e_{\beta} - \bar{\tau}_{\beta} e_{-\beta}] \\ &= \sigma_{\alpha} \tau_{\beta} [e_{\alpha} e_{\beta}] + \bar{\sigma}_{\alpha} \bar{\tau}_{\beta} [e_{-\alpha} e_{-\beta}] - \sigma_{\alpha} \bar{\tau}_{\beta} [e_{\alpha} e_{-\beta}] - \bar{\sigma}_{\alpha} \tau_{\beta} [e_{-\alpha} e_{\beta}] \\ &= N_{\alpha\beta} \sigma_{\alpha} \tau_{\beta} e_{\alpha+\beta} + N_{-\alpha, -\beta} \bar{\sigma}_{\alpha} \bar{\tau}_{\beta} e_{-\alpha-\beta} - N_{\alpha, -\beta} \sigma_{\alpha} \bar{\tau}_{\beta} e_{\alpha-\beta} - N_{-\alpha, \beta} \bar{\sigma}_{\alpha} \tau_{\beta} e_{-\alpha+\beta} \end{aligned} \right.$$

if $\alpha \neq \pm \beta$. But we have $N_{-\alpha, -\beta} = -N_{\alpha\beta}$ = real, and correspondingly

$N_{\alpha, -\beta} = -N_{-\alpha, \beta}$ = real. Hence, (80) is an element of $\tilde{\mathfrak{N}}$.

Finally, if $\alpha = \pm \beta$, say $\alpha = \beta$, we have to replace the last two terms on the right side in (80) by

$$-\sigma_{\alpha} \bar{\tau}_{\beta} h_{\alpha} - \bar{\sigma}_{\alpha} \tau_{\beta} h_{-\alpha} = (-\sigma_{\alpha} \bar{\tau}_{\beta} + \bar{\sigma}_{\alpha} \tau_{\beta}) h_{\alpha}$$

This is an element of $\tilde{\mathfrak{N}}$ also.

We denote $\tilde{\mathfrak{N}}$ by the symbol $\mathfrak{Y}_u$ (unitary restricted group $\mathfrak{Y}$).

Then $\mathfrak{Y}_u$ can be obtained by imposing reality condition on an appropriately chosen system of parameters in $\mathfrak{Y}$. It follows that a linear function of the parameters of $\mathfrak{Y}$ vanishing for all elements of $\mathfrak{Y}_u$ vanishes identically.

7. Automorphisms of an infinitesimal group

Let $\mathcal{C}$ be a continuous group of transformations of a point field. An automorphism of $\mathcal{C}$ is a one-to-one mapping $c \rightarrow c^*$ of $\mathcal{C}$ upon itself such that from $c_1 \rightarrow c_1^*$, $c_2 \rightarrow c_2^*$ it follows that $c_1 c_2 \rightarrow c_1^* c_2^*$. The inner automorphisms are the automorphisms of the form $c \rightarrow c^* = s c s^{-1}$ (s fixed element of $\mathcal{C}$). Every automorphism transforms the unit element into itself. If we assume that c^* as function of c satisfies some conditions of differentiability, it follows that infinitesimal elements are transformed into infinitesimal elements. Every automorphism of $\mathcal{C}$ yields in this manner a mapping of the infinitesimal group $\mathfrak{g}$ of $\mathcal{C}$ upon itself. We see readily that this is a linear transformation of the vector space $\mathfrak{g}$ such that if $x \rightarrow x^*$, $y \rightarrow y^*$ it follows that $[x, y] \rightarrow [x^*, y^*]$. Such a mapping of an infinitesimal group $\mathfrak{g}$ is called an automorphism of $\mathfrak{g}$. * The inner automorphisms of $\mathfrak{g}$ are those which

* It should be noticed for the following considerations that the elements x of $\mathfrak{g}$ correspond to elements " $1 + \varepsilon x$ " of $\mathcal{C}$ in the neighborhood of the identity (ε infinitely small).

correspond to inner automorphisms $c \rightarrow s c s^{-1}$. This is the transformation corresponding to s in the adjoint representation of $\mathcal{C}$. If we take for s an element in the "neighborhood" of the identity, $s = 1 + \varepsilon a$ with a in $\mathfrak{g}$, we obtain an inner infinitesimal automorphism of $\mathfrak{g}$: $x \rightarrow [ax]$.

We consider an automorphism of $\mathfrak{g}$ in the "neighborhood" of the identical automorphism: $x \rightarrow \varphi(x) = x + \varepsilon \tau(x)$. Then $x \rightarrow \tau(x)$ is a linear transformation in $\mathfrak{g}$ and from $\varphi([x, y]) = [\varphi(x), \varphi(y)]$ it follows that

$$[x, y] + \varepsilon \tau([x, y]) = [x, y] + \varepsilon [\tau(x), y] + \varepsilon [x, \tau(y)] + \varepsilon^2(\dots),$$

$$(81) \quad \tau([x, y]) = [\tau(x), y] + [x, \tau(y)]$$

We call linear transformations $x \rightarrow \tau(x)$ of $\mathfrak{g}$ with the property (81) an infinitesimal automorphism of $\mathfrak{g}$.

Theorem. The infinitesimal automorphisms of a semi-simple infinitesimal group are all inner infinitesimal automorphisms $x \rightarrow [ax]$ of $\mathfrak{g}$ (E. Cartan)

Proof* We extend $\mathfrak{g}$ to a group $\mathcal{N}$ by adjoining one linearly inde-

* Cartan, Bulletin des Sciences math. (2) 49 (1925), p. 130.

pendent element τ to $\mathfrak{g}$ and defining the brackets by: $[\tau, a] = -[a, \tau] = \tau(a)$ for a in $\mathfrak{g}$, $[\tau, \tau] = 0$. We have to show that for three elements a, b, c of $\mathfrak{g}$ the Jacobi relation holds:

$$(82) \quad [a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$$

It is sufficient to consider merely the case a, b, c either $= \tau$ or a, b, c in $\mathfrak{g}$ because of the linear character of the bracket symbol. We can assume that at least one is $= \tau$ and we may take $a = \tau$. If b and c belong to $\mathfrak{g}$, (82) reads

$$\tau(b, c) = -[b, -\tau(c)] - [c, \tau(b)] = [b, \tau(c)] + [\tau(b), c]$$

and this follows from (82).

If $a = \tau$ and $b = \tau$, (82) is trivial.

$$[\tau, \tau(c)] + [\tau, -\tau(c)] + 0 = 0.$$

Let us assume that $\mathcal{N}$ is semi-simple. Then $\mathcal{N}$ is the direct sum of simple groups and -- as the proof on page 88/89 shows -- every invariant subgroup of $\mathcal{N}$ may be used as direct summand in a decomposition of $\mathcal{N}$. But $\mathfrak{g}$ is an invariant subgroup. Hence a formula

$$\mathcal{N} = \mathfrak{g} + \mathfrak{h}$$

holds where $\mathfrak{h}$ is an invariant subgroup. Its order is necessarily 1. Then, however, $\mathfrak{h}$ is solvable in contradiction to the assumption that $\mathcal{N}$ is semi-simple.

We see therefore that $\mathcal{N}$ is not semi-simple. Let A be a solvable invariant subgroup. The intersection of A and $\mathfrak{g}$ vanishes because it is a solv-

able invariant subgroup of the semi-simple group $\mathfrak{g}$. Then $[A, \mathfrak{g}]$ vanishes too as a subgroup of the intersection. Hence

$$N = \mathfrak{g} + A$$

where t has the order 1. We represent τ as sum of one element a of $\mathfrak{g}$ and one element t of A . Then, for x in $\mathfrak{g}$:

$$\tau(x) = [\tau, x] = [a+t, x] = [a, x] + [t, x] = [a, x], \text{ q.e.d.}$$

It follows from the Jacobi relation (82) that the transformations $x \rightarrow [ax]$ are infinitesimal ~~inner~~ automorphisms of $\mathfrak{g}$. Distinct elements a and b establish distinct automorphisms. If, namely, $[a, x] = [b, x]$ for all x in $\mathfrak{g}$, then $[a-b, x] = 0$ and $a-b$ would be the basis of a solvable invariant subgroup.

Going back to the continuous group C we know that the infinitesimal elements of the group of automorphisms of $\mathfrak{g}$ and of the adjoint group of $\mathcal{C}$ are identical. Both groups are groups of linear transformations of the vector space $\mathfrak{g}$. On effecting an infinite number of successive infinitesimal transformations we get:

Theorem. Let $\mathcal{C}$ be a semi-simple continuous group and $\mathfrak{g}$ be its infinitesimal group. The adjoint group of $\mathcal{C}$ is identical with a subgroup A_0 of the group A of automorphisms of $\mathfrak{g}$: A_0 consists of those elements of A which can be connected with the unit element (the identical automorphism).

We have therefore a topological characterization of the set A_0 of inner automorphisms of $\mathfrak{g}$ in the set A of all automorphisms.

We consider a fixed element of the adjoint group $x \rightarrow sxs^{-1}$ or

$1 + \varepsilon x \rightarrow s(1 + \varepsilon x)s^{-1}$. By multiplying an infinite number of such relations ~~for~~ ^{for} a fixed s (integration with respect to x) we obtain the mapping $c \rightarrow scs^{-1}$.

We see:

Theorem. Every automorphism of a semi-simple continuous group $\mathcal{C}$ which belongs to the class of the identity is an inner automorphism of $\mathcal{C}$.

It is easy to give examples of automorphisms of semi-simple groups which admit automorphisms not belonging to the class of the identity. Let $\mathcal{C}$ for instance, be the orthogonal group for an even number of dimensions. The mapping $c \rightarrow tct^{-1}$, c in $\mathcal{C}$, t a fixed orthogonal transformation of determinant -1 is an automorphism but not an inner automorphism of $\mathcal{C}$.

Let $\mathfrak{g}$ be a semi-simple infinitesimal group and $\mathfrak{g}_u$ its unitary restricted group. The determination of a such that

$$(83) \quad [ax] = \tau(x)$$

(where $\tau(x)$ is a given automorphism) depends on linear equations for the parameters of a . We choose the parameters in such a manner that the elements of $\mathfrak{g}_u$ are characterized by real values of the parameters. If $\tau(x)$ has the property that it transforms $\mathfrak{g}_u$ in itself ($\tau(x)$ in $\mathfrak{g}_u$ if x in $\mathfrak{g}_u$), then the equations (83) for the parameters of a have real coefficients. Since we have a solution, we have a real solution. We can find an a in $\mathfrak{g}_u$.

Every automorphism of $\mathfrak{g}_u$ can be extended to an automorphism of $\mathfrak{g}$. We see that the three preceding theorems hold for $\mathfrak{g}_u$ and $\mathcal{C}_u$ instead of $\mathfrak{g}$ and $\mathcal{C}$. Here $\mathcal{C}_u$ is the continuous group belonging to the "real" infinitesimal group $\mathfrak{g}_u$.

The automorphisms of $\mathfrak{g}$ are linear transformations $x \rightarrow \varphi(x)$ which have to satisfy

$$\varphi([e_i, e_k]) = [\varphi(e_i), \varphi(e_k)]$$

where $e_1, \dots, e_r$ form a basis of $\mathfrak{g}$. This is a finite set of algebraic conditions for the coefficients of φ . In the case of $\mathfrak{g}_u$ we have to add reality conditions.

The characteristic polynomial $f(\omega | a)$ of the matrix A corresponding to the element a of $\mathfrak{g}$ in the adjoint representation is independent of the choice of the basis of $\mathfrak{g}$. We take the basis $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_r)$ in order to find $f(\omega | \varphi(a))$. The matrix corresponding to $\varphi(a)$ then is the same matrix which corresponds to a if we take the basis $e_1, e_2, \dots, e_r$ of $\mathfrak{g}$ because from $[a, e_\kappa] = \sum a_{\kappa\lambda} e_\lambda$ follows $[\varphi(a), \varphi(e_\kappa)] = \sum a_{\kappa\lambda} \varphi(e_\lambda)$. Hence $f(\omega | a) = f(\omega | \varphi(a))$. The quadratic form $\sigma(a)$ is the sum of the squares of the roots of $f(\omega | a) = 0$. Consequently

$$\sigma(a) = \sigma(\varphi(a))$$

The quadratic form $\sigma(a)$ is invariant under the transformations $a \rightarrow \varphi(a)$.

If we restrict ourselves to the unitary group $\mathfrak{g}_u$ the form $\sigma(a)$ is definite. The automorphisms of $\mathfrak{g} : x \rightarrow \tau(x)$ leave a definite quadratic form invariant. They are defined by a finite number of algebraic equations for the coefficients of τ . Consequently they form a compact set A . The same is true for the part A_0 of A which is connected with the identity. We obtain:

Lemma: Let C be a semi-simple continuous group and $\mathfrak{g}$ its infinitesimal group. Let C_u and $\mathfrak{g}_u$ be the corresponding unitary restricted groups. The adjoint representation of C_u is compact.

There are, in general, many continuous groups belonging to the same infinitesimal group. For instance: The group of rotations of a circle and the group of translations of a straight line belong to the infinitesimal group of order 1. But there exists always a simply connected continuous group C to a given infinitesimal group $\mathfrak{g}$. Every other continuous group C^* belonging to $\mathfrak{g}$ is homomorphic to C and C is the simply connected covering manifold of C^* . *)

* Cf. the notes of last year's course, pp. 106-110.

We will call C the true group belonging to $\mathfrak{g}$. We are going to prove that the true group belonging to $\mathfrak{g}_u$ has only a finite number of sheets over the adjoint representation. Hence we will get the result that the true group belonging to $\mathfrak{g}_u$ is also compact.

8. Adjoint group and true group

Let $\mathfrak{g}$ be a semi-simple group. We call the elements of $\mathfrak{g}$ the main elements and the parameters $\lambda^1, \dots, \lambda^h$ in

$$a = h + \sum \sigma_\alpha e_\alpha, \quad h = h(\lambda^1, \dots, \lambda^h) = \sum \lambda^i h_i$$

the main parameters. The σ_α will be denoted as supplementary parameters.

We write $a = (\lambda^1, \dots, \lambda^h, \sigma_\alpha, \sigma_\beta, \dots)$

The adjoint group A of $\mathfrak{g}$ is a group of linear transformations in the vector space $\mathfrak{g}$. Its infinitesimal elements are given by $x \rightarrow [ax]$. The corresponding transformation of A in the "neighborhood" of the identity is

$x \rightarrow x' = x + [ax]dt$ (dt infinitely small) or $dx = [ax]dt$. Let κ^i be the main parameters of x , and τ_α the supplementary parameters. We have if $a = h$ is a main element

$$d\kappa^i = 0, \quad d\tau_\alpha = \alpha \tau_\alpha \cdot dt$$

On integrating we obtain the finite transformation

$$(84) \quad \kappa^{i'} = \kappa^i, \quad \tau_\alpha' = e^\alpha \tau_\alpha$$

These main elements (ε) form an χ -dimensional Abelian subgroup of A , (ε) being the diagonal matrix with χ elements 1 and $e^\alpha, e^\beta, \dots$ in the main diagonal.

We confine ourselves to the unitary restricted group A_u . Then

$$\lambda^j = 2\pi i \varphi^j \text{ is purely imaginary. We have } \alpha = 2\pi i(\alpha_1 \varphi^1 + \dots + \alpha_\ell \varphi^\ell)$$

We denote $\alpha_1 \varphi^1 + \dots + \alpha_\ell \varphi^\ell$ now by α . We must replace (84) by

$$(85) \quad \kappa^{i'} = \kappa^i, \quad \tau_\alpha' = e(\alpha) \tau_\alpha, \quad \alpha = \sum_j \alpha_j \varphi^j$$

This is the transformation $x \rightarrow x' = (\varepsilon)x(\varepsilon)^{-1}$ expressed in the parameters in accordance with the definition of the adjoint group.

Theorem. Every element t of A_u can be written in the form $u^{-1}(\varepsilon)u$ where u belongs to A_u and (ε) is a main element of A_u .

Proof. We assume that t has already been represented in the form of the theorem. If the equation $t + dt' = (u+du)^{-1}((\varepsilon)+d(\varepsilon))(u+du)$ is to hold, we must have *

* Cf. the analogous considerations in Chapter 1, 13, pp. 55-56.

$$(u+du)(t+dt) = (\varepsilon+d(\varepsilon))(u+du) \\ du \cdot t + u \cdot dt = (\varepsilon)du + d(\varepsilon)u$$

We multiply by $t^{-1}u^{-1} = u^{-1}(\varepsilon)^{-1}$ and get

$$du \cdot u^{-1} + u(dt \cdot t^{-1})u^{-1} = (\varepsilon)(du \cdot u^{-1})(\varepsilon)^{-1} + d(\varepsilon)(\varepsilon)^{-1}$$

We write δu for $du \cdot u^{-1}$, δt for $dt \cdot t^{-1}$, $\delta(\varepsilon)$ for $d(\varepsilon) \cdot (\varepsilon)^{-1}$.

These elements must be elements of $\mathfrak{g}_u$, in particular $\delta(\varepsilon) = 2\pi i h(d\varphi^1, \dots, d\varphi^l)$

Hence

$$(86) \quad u \cdot \delta t \cdot u^{-1} = 2\pi i h(d\varphi^1, \dots, d\varphi^l) + \{(\varepsilon)\delta u(\varepsilon)^{-1} - \delta u\}$$

On comparing the parameters on both sides we get by using (85) to obtain the parameters of $(\varepsilon)\delta u(\varepsilon)^{-1}$

$$(87a) \quad \lambda^j(u \cdot \delta t \cdot u^{-1}) = 2\pi i d\varphi^j$$

$$(87b) \quad \sigma_\alpha(u \cdot \delta t \cdot u^{-1}) = \rho(\alpha)\sigma_\alpha(\delta u) - \sigma_\alpha(\delta u) = \sigma_\alpha(\delta u)(\rho(\alpha) - 1)$$

If no α is an integer for (ε) we find $d\varphi^j$ from (87a) and

$\sigma_\alpha(\delta u)$ from (87b). We have no conditions for the main parameters of δu

They can be chosen arbitrarily as can be seen directly also.

We call an element (ε) regular if no α is an integer. We start from an element t of the form $u^{-1}(\varepsilon)u$ (for instance $t = (\varepsilon)$) with a regular (ε) . We obtain then an analogous representation for the neighboring

elements $t+dt$ and can continue as long as we do not come to a singular (ε) where one of the linear forms α has an integral value.

Let $\mathcal{T}$ be the set of elements of A_μ of the form $u^{-1}(\varepsilon)u$, u in A_μ , (ε) regular main element. $\mathcal{T}$ is open as follows from the above consideration. Let A be the boundary of $\mathcal{T}$. If t_n in $\mathcal{T}$ tends to a point t^* in A , we can represent $t_n = u_n^{-1}(\varepsilon)_n u_n$. Now, A_μ is compact. Consequently there exists a sequence $n = n_1, n_2, \dots$ of integers such that $u_n \rightarrow u^*$, $(\varepsilon)_n \rightarrow (\varepsilon)^*$ converge, where u^* and $(\varepsilon)^*$ belong to A_μ . The element (ε) must be a main element. We have by going to the limit:

$$t^* = u^{*-1}(\varepsilon)^* u^*$$

Naturally $(\varepsilon)^*$ here is singular.

The theorem will be proved when it is shown that A has three dimensions less than A_μ . Then A_μ will not be decomposed by A .

Let (ε) be a fixed singular element, say $e(p) = 1$ for (ε) where $2\pi i p$ is a root of η . The parameters of $(\varepsilon)x(\varepsilon)^{-1}$ for an x in $\mathfrak{g}_\mu$ are expressed in terms of the parameters of x by (85). We see that $x = x'$ for $x = e_p$ (i.e. for $\kappa^i = 0$, $\tau_\alpha = 0$ for $\alpha \neq p$, $\tau_p = 1$), for $x = e_{-p}$ and, of course, for $x = h^j$. These infinitesimal elements generate an $(\chi+2)$ -dimensional subgroup $\mathcal{V}$ of (finite) transformations v of A_μ with $v(\varepsilon)v^{-1} = (\varepsilon)$. Then $uv(\varepsilon)(uv)^{-1} = u(\varepsilon)u^{-1}$. The elements which can be represented in the form $u(\varepsilon)u^{-1}$, u in A_μ , form therefore an $(r-\chi-2)$ -dimensional manifold for a fixed singular (ε) . The singular (ε) however form an $(\chi-1)$ -dimensional manifold $(\alpha_1 \varphi^1 + \dots + \alpha_\chi \varphi^\chi = 0 \pmod{1} \text{ for at least one } \alpha)$. We obtain $(r-\chi-2) + (\chi-1) = r-3$ as the number of dimensions of A as stated above. This completes the proof.

Theorem. Let $\mathcal{Y}_u$ be a semi-simple infinitesimal group, unitary restricted. The universal covering manifold $\mathcal{C}_u$ of the adjoint representation A_u (i.e., the true group belonging to $\mathcal{Y}_u$) has only a finite number of sheets. Hence $\mathcal{C}_u$ is compact.

Proof. Let $t_0 = (\mathcal{E}_0)$ be a main element in A_u for which $\alpha, \beta, \dots$ are all different. Let t_0^* be an element of $\mathcal{C}_u$ which lies over t_0 . We consider a curve γ^* in $\mathcal{C}_u$ starting from t_0^* and ending in the point t_1^* lying over t_0 in a given sheet. The curve γ of A_u which corresponds to γ^* (the projection of γ^*) is closed. We can assume that the curve γ does not pass through points of A . This does not mean a material restriction since A has three dimensions less than A_u . We represent all t on γ in the form $t = u^{-1}(\mathcal{E})u$ such that u and $(\mathcal{E})$ depend continuously on t . Let u_1 and $(\mathcal{E}_1)$ be the values we obtain when t comes back to t_0 after running over γ . Then

$$(88) \quad (\mathcal{E}_0) = t_0 = u_1^{-1}(\mathcal{E}_1)u_1$$

Then the diagonal matrices $(\mathcal{E})$ and $(\mathcal{E}_1)$ have the same characteristic roots, perhaps in another order. Hence $e(\alpha), e(\beta), \dots$ undergo a permutation if t runs over γ from t_0 back to t_0 .

We assume now, that this is the identical permutation. Then $\alpha, \beta, \dots$ return also to their initial values because of $\alpha \neq 0, \beta \neq 0, \dots \pmod{1}$. We replace $\alpha, \beta, \dots$ in $(\mathcal{E})$ by $\gamma\alpha, \gamma\beta, \dots$. Let $(\mathcal{E}^{(\gamma)})$ and $t^{(\gamma)} = u^{-1}(\mathcal{E}^{(\gamma)})u$ be the matrices we obtain from $(\mathcal{E})$ and t in this manner. If t runs over γ , $t^{(\gamma)}$ runs over a curve $\gamma^{(\gamma)}$. We have $(\mathcal{E}_0) = (\mathcal{E}_1)$ in (88) under our present assumption. Therefore u_1 is commutative with $(\mathcal{E}_0) = (\mathcal{E}_1)$. By using the fact that $\alpha, \beta, \dots$ are all distinct for $(\mathcal{E}_0)$, one readily sees that u_1 is commutative with $\mathcal{E}_0^{(\gamma)} = \mathcal{E}_1^{(\gamma)}$. Hence $t_0^{(\gamma)} = t_1^{(\gamma)}$. Therefore $\gamma^{(\gamma)}$ is closed. If γ varies from 1 to 0, γ will be shrunk to a point. Hence $t_0^* = t_1^*$ on $\mathcal{C}_u$.

Now let us take the case of two closed curves γ_1, γ_2 starting in the same t_0 which give the same permutation of $\alpha, \beta, \dots$. If we run over γ_1 first and then over γ_2 in the opposite direction, we obtain a closed curve $\gamma_1 \gamma_2^{-1}$ belonging to the identity. It follows that the corresponding curves γ_1^* and γ_2^* end in the same point t_1^* in $\mathcal{C}_u$. The number of sheets of $\mathcal{C}_u$ is therefore at most equal to the number $(r-1)!$ of all possible permutations of $\alpha, \beta, \dots$.

The infinitesimal elements $h_1, \dots, h_\ell$ generate an ℓ -dimensional Abelian subgroup of $\mathcal{C}_u$ the elements $h^*(\varphi^1, \dots, \varphi^\ell)$ of which will be denoted as main elements again. The parameters can be chosen in such a manner that

$$h^*(\varphi^1, \dots, \varphi^\ell) h^*(\psi^1, \dots, \psi^\ell) = h^*(\varphi^1 + \psi^1, \dots, \varphi^\ell + \psi^\ell)$$

and that $h(\varphi^1, \dots, \varphi^\ell)$ is the element corresponding to $h^*(\varphi^1, \dots, \varphi^\ell)$ in the adjoint group.

It follows immediately from (85):

Theorem. The main elements $h^*(\varphi^1, \dots, \varphi^\ell)$ and $h^*(\psi^1, \dots, \psi^\ell)$ of $\mathcal{C}_u$ are represented by the same element in the adjoint group if and only if

$$\alpha_1 \varphi^1 + \dots + \alpha_\ell \varphi^\ell \equiv \alpha_1 \psi^1 + \dots + \alpha_\ell \psi^\ell \pmod{1}$$

for every root α .

* The roots in the original sense are $2\pi i \alpha$ according to the convention of p. 107.

We are now in a position to prove the first theorem of this paragraph for the true group.

Theorem. Every element t^* of $\mathcal{C}_u$ can be represented in the form $u^{*-1}(\varepsilon^*)u^*$ where u^* and the main element (ε^*) belong to $\mathcal{C}_u$. The $\varphi^1, \dots, \varphi^\ell$ belonging to (ε^*) can be chosen so that $k_\alpha \leq \alpha \leq k_\alpha + 1$ (k_α fixed integers) for every root α .

Proof. Let (ε_0) be a regular main element. Then $\alpha \not\equiv 0 \pmod{1}$ for every root α . Let (ε^*) be the corresponding main element in $\mathcal{C}_u$ with $k_\alpha < \alpha < k_\alpha + 1$. Then (ε^*) can be represented in the required form. If an element $t^* = u^{*-1}(\varepsilon^*)u^*$ has already been represented and every α lies between k_α and $k_\alpha + 1$ we consider the corresponding elements $t = u^{-1}(\varepsilon)u$ in A_u . There is a neighborhood of t whose element can be represented in the same form. If we choose this neighborhood sufficiently small and go back to $\mathcal{C}_u$, we obtain a neighborhood of t^* for the element of which the theorem holds.

Therefore we can prove the statement for all elements t^* with a projection t in A_u which can be connected with (ε_0) by a curve consisting of regular elements of A_u only (i.e. an element $u(\varepsilon)u^{-1}$ with a regular (ε)). Our above consideration shows that every element t^* of $\mathcal{C}_u$ is either of this form or at least a limit point of elements of this form. This proves the theorem.

If $t^* = u^{*-1}(\varepsilon^*)u^*$ and (ε^*) belongs to $\varphi^1, \dots, \varphi^l$, we say that $\varphi^1, \dots, \varphi^l$ are the angles of the element t^* . *

* These angles are not uniquely determined. We are going to show that every possible choice of $\varphi^1, \dots, \varphi^l$ can be obtained from a fixed one by an element of the finite group $(S) \pmod{1}$.

We consider once more the maximum Abelian subgroup $\mathfrak{f}$ of the semi-simple group $\mathfrak{g}$ which was the basis of all our consideration. We choose an element $\lambda^1 h_1 + \dots + \lambda^l h_l = h$, $\lambda^v = 2\pi i \varphi^v$ such that all roots α are different for this element. Then h is commutative only with the elements of $\mathfrak{f}$ as can be seen from (48). We call h a typical element of $\mathfrak{f}$.

Let $\tilde{\mathfrak{f}}$ be another maximum Abelian subgroup of $\mathfrak{g}_u$ and $\tilde{h}$ a typical element of $\tilde{\mathfrak{f}}$. Then according to the first theorem of this section there exist an element u in A_u and an element h_1 in $\mathfrak{f}$ such that $\tilde{h} = u^{-1}h_1u$. The elements of $u^{-1}\mathfrak{f}u$ are then commutative with $\tilde{h}$. Hence $u^{-1}\mathfrak{f}u \leq \tilde{\mathfrak{f}}$, $\mathfrak{f} \leq u\tilde{\mathfrak{f}}u^{-1}$; the max-

irreducibility property of f implies $f = u \tilde{f} u^{-1}$. It follows that if we start from $\tilde{f}$ instead of f , we have to replace e_α by $u^{-1} e_\alpha u$. We obtain the same vector diagram.

9. The characters of the irreducible representations

The main elements $h^*(\varphi^1, \varphi^2, \dots, \varphi^N)$ form a compact Abelian subgroup $\mathcal{H}^*$ of $\mathcal{C}_u$. Every irreducible representation of an Abelian group is of degree 1; every representation of a compact group is completely reducible.

Let $\mathcal{V}$ be representation of $\mathcal{C}_u$. We consider for the moment only the elements $h^*(\varphi^1, \dots, \varphi^N)$. The corresponding matrices form, of course, a representation of $\mathcal{H}^*$. The preceding remark shows that we may assume that all the matrices of this representation are diagonal matrices

$$(89) \quad h^*(\varphi^1, \dots, \varphi^N) \rightarrow \begin{pmatrix} E_1 & & \\ & E_2 & \\ & & \ddots \\ & & & E_N \end{pmatrix}$$

We consider next the infinitesimal elements η_n of $\mathcal{C}_u$. In this case, $E_1, E_2, \dots, E_N$ will be the weights $\Lambda_1, \Lambda_2, \dots, \Lambda_N$ of the representation and by iterating the infinitesimal transformations we see that for the finite element $h^*(\varphi^1, \dots, \varphi^N)$ we have

$$(89^*) \quad E_1 = e(\Lambda_1), \dots, E_N = e(\Lambda_N)$$

Lemma: Two main elements $h^*(\varphi^1, \dots, \varphi^N)$ and $h^*(\psi^1, \dots, \psi^N)$ of are identical if $\varphi^k = \psi^k \pmod{1}$.

Proof. If $\varphi^k = \psi^k \pmod{1}$, the quantities (89*) have the same value for both elements since the Λ_k are linear forms with integral coefficients. Both elements are represented by the same matrix in $\mathcal{V}$. It follows that the elements $h^*(\varphi^1, \dots, \varphi^N)$ and $h^*(\psi^1, \dots, \psi^N)$ are identical be-

cause the two corresponding matrices in any representation $\mathcal{G}$ are equal (of the notes of last year's course, p. 99).

The converse of this lemma is true also as we shall see later.

It follows that the angles $\varphi^1, \dots, \varphi^{\lambda}$ of an arbitrary element $t = u^{-1}h^*(\varphi^1, \dots, \varphi^{\lambda})u$ of $\mathcal{C}_u$ are determined only (mod 1). The value of the character χ of the representation $\mathcal{G}$ is the same for the element t as for $h^*(\varphi^1, \dots, \varphi^{\lambda})$: χ is a function of $\varphi^1, \dots, \varphi^{\lambda}$ and we have

$$(90) \quad \chi(t) = \chi(\varphi^1, \dots, \varphi^{\lambda}) = E_1 + \dots + E_N = e(\lambda_1) + \dots + e(\lambda_N)$$

We consider the parts $\mathcal{P}_1, \mathcal{P}_2, \dots$ in which the $\varphi^1, \dots, \varphi^{\lambda}$ space is divided by the hyperplanes $\alpha = 0, \alpha = \pm 1, \alpha = \pm 2, \dots, \beta = 0, \beta = \pm 1, \beta = \pm 2, \dots$. Every element t of $\mathcal{C}_u$ can be represented in the form $u^{-1}h^*(\varphi^1, \dots, \varphi^{\lambda})u$ where $(\varphi^1, \dots, \varphi^{\lambda})$ lies in a given $\mathcal{P}_k$ as follows from the theorem on p. 111. We consider the different points

$(\varphi^1, \dots, \varphi^{\lambda})$ in $\mathcal{P}_k$ which occur in the different possible representations of a given t . Let $(\varphi^1, \dots, \varphi^{\lambda})$ and $(\psi^1, \dots, \psi^{\lambda})$ be two such points. Then $h^*(\varphi^1, \dots, \varphi^{\lambda})$ and $h^*(\psi^1, \dots, \psi^{\lambda})$ are conjugate. So are the corresponding elements $h(\varphi^1, \dots, \varphi^{\lambda})$ and $h(\psi^1, \dots, \psi^{\lambda})$ in the adjoint group. The characteristic roots of such elements are

$$(90^*) \quad \underbrace{1, \dots, 1}_{\lambda\text{-times}}, e(\alpha), e(\beta), \dots$$

according to (85). We see that the values of $\alpha, \beta, \dots$ (mod 1) in $(\psi^1, \dots, \psi^{\lambda})$ are a permutation of the corresponding values in $(\varphi^1, \dots, \varphi^{\lambda})$. It follows that there are only a finite number of points $(\psi^1, \dots, \psi^{\lambda})$ in $\mathcal{P}_k$ such that $h^*(\psi^1, \dots, \psi^{\lambda})$ is conjugate to a given $h^*(\varphi^1, \dots, \varphi^{\lambda})$. Let Z be the number of such points $(\psi^1, \dots, \psi^{\lambda})$. Each of them gives rise to a representation $t = v^{-1}h^*(\psi^1, \dots, \psi^{\lambda})v$. We see in the same manner as in Chapter 2, 8, by continuation that every t in $\mathcal{C}_u$ has representations

$u^{-1}h^*(\varphi^1, \dots, \varphi^l)u$ where $(\varphi^1, \dots, \varphi^l)$ is one of $\mathbb{Z}$ possible points in $\mathcal{P}_k$. It is possible that for special elements t some of these points coincide but this can only be the case if the point $(\varphi^1, \dots, \varphi^l)$ lies on a hyperplane $\alpha - \beta = 0, \pm 1; \pm 2, \dots, \alpha - \gamma = 0, \pm 1, \pm 2, \dots, \beta - \gamma = 0, \pm 1, \pm 2, \dots$. Only a finite number of them has points in common with $\mathcal{P}_k$.

If we replace φ^j by $\varphi^j + k_j$, k_j a fixed integer, every $\mathcal{P}_k$ is transformed into another one because $\alpha, \beta, \dots$ have integral coefficients. We consider now two points $(\varphi^1, \dots, \varphi^l)$ and $(\psi^1, \dots, \psi^l)$ as identical if $\varphi^1 \equiv \psi^1, \dots, \varphi^l \equiv \psi^l \pmod{1}$. The whole space is decomposed in a finite number of $\mathcal{P}_k$. Let M be the sum of the numbers $\mathbb{Z}$ belonging to these single $\mathcal{P}_k$'s. We see that the angles $\varphi^1, \dots, \varphi^l$ of an element t of $\mathcal{C}_u$ can be chosen in M different ways such that

$$0 \leq \varphi^1 < 1, 0 \leq \varphi^2 < 1, \dots, 0 \leq \varphi^l < 1.$$

These M systems are different except when t lies in a subset of $\mathcal{C}_u$ with less than r dimensions.

We determine the volume dV of the part of $\mathcal{C}_u$ where the angles $\omega^1, \dots, \omega^l$ of t satisfy

$$\varphi^1 \leq \omega^1 \leq \varphi^1 + d\varphi^1, \dots, \varphi^l \leq \omega^l \leq \varphi^l + d\varphi^l$$

The consideration which we used in the case of the complex-group can be applied again. We obtain here

$$(91) \quad \begin{cases} dV = \gamma \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^l \\ \Delta = \prod_{\alpha > 0} s(\alpha/2) = \prod_{\alpha > 0} (e(\alpha/2) - e(-\alpha/2))^{*}), \Delta \bar{\Delta} = \prod_{\alpha > 0} (1 - e(\alpha)) \\ \gamma \text{ independent of } \varphi^1, \varphi^2, \dots, \varphi^l. \quad \text{or } \alpha < 0 \end{cases}$$

We denoted the linear transformation

$$\Lambda \rightarrow \Lambda - \frac{2\Lambda\alpha}{\alpha\alpha} \alpha$$

*) A linear form Λ with real coefficients is positive if the first non-vanishing coefficient is positive. We write $\Lambda > \Lambda^*$ if $\Lambda - \Lambda^*$ is positive. We have to multiply in (91) over all positive roots α .

by S_α . (S) was the group generated by S_α . An element S of (S) transforms weights into weights for every representation η . It follows from (90) that the character χ is transformed into itself. Furthermore, the set of positive roots $\alpha, \beta, \dots$ is transformed into a set of roots which contains either α or $-\alpha$, either β or $-\beta, \dots$. Hence $\Delta \rightarrow \pm \Delta$. We call S even or odd according to whether the + sign or the - sign holds.

We can write $\Delta = \sum \pm e(\psi)$ where every ψ is a linear form in $\varphi^1, \dots, \varphi^l$. We state that the coefficients are integers (not semi-integers). We have to show that $\frac{1}{2} \sum \rho = \Lambda$ has integral coefficients where the sum extends over all positive roots ρ . Such an integral form was characterized by the condition: $\frac{2\Lambda_\alpha}{\alpha_\alpha} = \frac{1}{\alpha_\alpha} \sum \rho_\alpha = \text{integer}$, for every root α . The transformation S_α transforms the root ρ into $\rho' = \rho - \frac{2\rho_\alpha}{\alpha_\alpha} \alpha$. Then $\rho'_\alpha = \rho_\alpha - 2\rho_\alpha = -\rho_\alpha$. If $\rho' > 0$ and $\rho' \neq \rho$, we take the terms $\rho_\alpha + \rho'_\alpha = 0$ together in our sum. If $\rho = \rho'$, we have $\rho_\alpha = 0$. If $\rho' < 0$ and $\rho \neq -\rho'$, we have $\frac{1}{\alpha_\alpha} (\rho_\alpha + (-\rho')_\alpha) = \frac{2\rho_\alpha}{\alpha_\alpha} = \text{integer}$. If $\rho' = -\rho$ then ρ and $-\rho$ belong to the same α -string of roots which is impossible for $\rho \neq \pm \alpha$ as has been shown on p. 88. If $\rho = \pm \alpha$, $\frac{\rho_\alpha}{\alpha_\alpha} = \pm 1$. Hence $\frac{1}{\alpha_\alpha} \sum \rho_\alpha = \text{integer}$.

The transformation S_α is odd: If by S_α a positive root ρ is transformed into a negative one ρ' , then the positive root $-\rho'$ is transformed into the negative one $-\rho$. Since $\rho = -\rho'$ only for one positive root $\rho = \pm \alpha$, the number of positive roots which are transformed into negative ones is odd, S_α is odd.

We multiply the character χ of a representation η by Δ :

$$\xi = \chi \cdot \Delta = (\sum e(\Lambda_\nu) \cdot \Delta) = \sum \pm e(\Phi)$$

is a finite Fourier series also; the Φ 's are integral linear forms. Furthermore ξ is alternating: It is invariant under even S and changes its sign under odd S.

Let Φ be the highest term occurring in such an alternating Fourier series. Then $\Phi \geq S\Phi$ for all S in (S) . Moreover, $\Phi > S\Phi$ for all odd S , because the terms $e(\Phi)$ and $-e(\Phi)$ would occur the same number of times if $S(\Phi) = \Phi$ for an odd Φ .

Here let $\Phi = k_1 \varphi^1 + \dots + k_\lambda \varphi^\lambda$. One proves easily $\Phi > 0$. *

* For the proof, cf. Weyl, Math. Zeitschrift 24, pp. 384-385.

We set

$$(92) \quad \xi(k_1, k_2, \dots, k_\lambda) = \sum_S \pm e(S\Phi)$$

If there exist $n(k_1, \dots, k_\lambda)$ even transformations S for which $S\Phi = \Phi$, the highest term in this elementary sum will be $n(k_1, \dots, k_\lambda) e(\Phi)$. Every finite alternating Fourier series ξ with integral coefficients is a linear combination of such $\frac{1}{n(k_1, \dots, k_\lambda)} \xi(k_1, \dots, k_\lambda)$. In particular

$$(93) \quad \xi = \chi \Delta = \sum \frac{g(k_1, \dots, k_\lambda)}{n(k_1, \dots, k_\lambda)} \xi(k_1, \dots, k_\lambda),$$

$g(k_1, \dots, k_\lambda)$ integer

$$\text{Let } \xi(k'_1, \dots, k'_\lambda) = \sum_{S'} \pm e(S'\Phi'), \quad \Phi' = k'_1 \varphi^1 + \dots + k'_\lambda \varphi^\lambda.$$

Then $\xi(k_1, \dots, k_\lambda) \cdot \overline{\xi(k'_1, \dots, k'_\lambda)} = \sum_{S, S'} \pm e(S\Phi - S'\Phi')$. Hence

$$(94) \quad \int_0^1 \dots \int_0^1 \xi(k_1, \dots, k_\lambda) \overline{\xi(k'_1, \dots, k'_\lambda)} d\varphi^1 \dots d\varphi^\lambda = \begin{cases} 0 & \text{for } \Phi \neq \Phi' \\ sn(k_1, \dots, k_\lambda) & \text{for } \Phi = \Phi' \end{cases}$$

where s is the order of (S) . This integral is, namely, equal to the number of vanishing $S\Phi - S'\Phi'$.

One proves without difficulty that for an alternating finite Fourier series $\sum$ the quotient ξ/Δ is a finite Fourier series itself. * If one ap-

* See note above.

plies the representation (93) on Δ itself, $\chi = 1$, every term on the right-hand side is divisible by Δ . It follows readily that Δ itself is an elementary sum. The leading term $e(\Phi_0) = e(r_1 \varphi^1 + \dots + r_\chi \varphi^\chi)$ occurs in

$\Delta = \sum_{p \geq 0} (e(\frac{p}{2}) - e(-\frac{p}{2}))$ exactly once, we have

$$(95) \quad \Delta = \frac{1}{n(r_1, \dots, r_\chi)} \xi(r_1, \dots, r_\chi)$$

We proceed now in the following manner which is quite analogous to the method we used in the case of the complex group: If χ is a character of an irreducible representation, we have

$$\int \chi \bar{\chi} dV = V$$

where the integral extends over $\mathcal{C}_n$ and V is the volume of $\mathcal{C}_n$. We use

(91) and obtain -- since χ depends only on $\varphi^1, \dots, \varphi^\chi$ --

$$(96) \quad \int_0^1 \int_0^1 \dots \int_0^1 \chi \bar{\chi} \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^\chi = W$$

where W has the same value for every irreducible character χ . ($W = \frac{1}{\chi} MV$).

We put here $\chi = 1$. From (94) and (95) follows

$$W = \frac{S}{n(r_1, \dots, r_\chi)}$$

We apply now (96) for any irreducible character χ and use (93) and (94). We get

$$(96) \quad \sum \frac{g(\kappa_1, \dots, \kappa_\chi)^2}{n(\kappa_1, \dots, \kappa_\chi)} S = W = \frac{S}{n(r_1, \dots, r_\chi)}$$

Let $e(\Psi)$ be the highest term in χ . The highest terms in Δ and

$\xi = \chi \Delta$ being $e(\Phi_0)$ and $e(\bar{\Phi})$ respectively, we have $\bar{\Phi} = \Psi + \bar{\Phi}_0$.

Now $\Psi \geq S \Psi$, $\bar{\Phi}_0 \geq S \bar{\Phi}_0$. If therefore $\bar{\Phi}_0 > S \bar{\Phi}_0$ then we have $\bar{\Phi} > S \bar{\Phi}$

and, consequently, $n(\kappa_1, \dots, \kappa_\chi) \leq n(r_1, \dots, r_\chi)$. Hence

$$\sum g(\kappa_1, \dots, \kappa_\chi)^2 \leq 1$$

But the $g(\kappa_1, \dots, \kappa_\chi)$ are integers and $g(k_1, \dots, k_\chi) > 0$. We have then

$g(k_1, \dots, k_\chi) = 1$, all other $g(\kappa_1, \dots, \kappa_\chi) = 0$, and $n(k_1, \dots, k_\chi) =$

$= n(r_1, \dots, r_\chi)$ in (96). We derive from (93) and (95):

$$(97) \quad \chi = \frac{\xi}{\Delta} = \frac{\xi(k_1, \dots, k_\ell)}{n(k_1, \dots, k_\ell)}; \quad \frac{\xi(r_1, \dots, r_\ell)}{n(r_1, \dots, r_\ell)} = \frac{\xi(k_1, \dots, k_\ell)}{\xi(r_1, \dots, r_\ell)}$$

$$\chi = \frac{\xi(k_1, \dots, k_\ell)}{\xi(r_1, \dots, r_\ell)}$$

Now, one can prove that $n(k_1, \dots, k_\ell) = n(r_1, \dots, r_\ell) = 1$, cf. Weyl, Mathematische Zeitschrift 24, p. 791. *

* A direct proof will be given in Chapter 3, sec. 4, theorem on p. 164.

Cartan proved by explicit construction that every ^(integral) linear form ψ is the highest weight of a representation if $\psi \geq S\psi$ for all S in (S) . Hence all expressions (97) are characters.

Another method to obtain this last result is the following one: The linear transformation $\Phi \rightarrow S\Phi$ permutes the weights of a representation only. If $(\varphi^1, \dots, \varphi^\ell)$ and $(\psi^1, \dots, \psi^\ell)$ are two points which are connected by a transformation of (S) , all the characters χ of $\mathfrak{G}_u$ have the same value for both systems, i.e. for $h^*(\varphi^1, \dots, \varphi^\ell)$ and $h^*(\psi^1, \dots, \psi^\ell)$. Consequently both elements are conjugate in $\mathfrak{G}_u$ (cf. the notes of last year's course, p. 99). If one can prove* the converse, that $h^*(\varphi^1, \dots, \varphi^\ell)$ and

* Cf. Chapter 3, sec. 5, theorem 6, p. 177.

$h^*(\psi^1, \dots, \psi^\ell)$ are only conjugate if this condition is fulfilled, it follows that the expression (97) is a class function (i.e. has a constant value for every class of conjugate elements). If it would not be a character it would be orthogonal to all the characters according to (94) because all the characters are of the form (97) as we have seen. But the characters form a complete orthogonal system. One obtains a contradiction.

Theorem. If $\psi = m_1 \varphi^1 + \dots + m_\ell \varphi^\ell$ is an integral linear form such that $\psi \geq S\psi$ for all S in (S) , then there is an irreducible representation with the highest weight ψ . Its character is given by (97) where

$k_1 = m_1 + r_1, \dots, k_\lambda = m_\lambda + r_\lambda$. Here $\mathcal{C}(r_1 \varphi^1 + \dots + r_\lambda \varphi^\lambda)$ is the highest term in $\sum_{\alpha > 0} S(\frac{\alpha}{2})$.

The transition from the unitary restricted group to the unrestricted group can be effected without difficulty.

We can use the same device as in the case of the complex-group in order to compute the degree of the representation. We obtain the result:

$$N = \frac{\prod_{\alpha > 0} (\alpha^1 A_1 + \dots + \alpha^\lambda A_\lambda)}{\prod_{\alpha > 0} (\alpha^1 \tau_1 + \dots + \alpha^\lambda \tau_\lambda)} = \frac{\prod_{\alpha > 0} (\alpha, K)}{\prod_{\alpha > 0} (\alpha, \tau)}$$

where $A_1 = m_1 + r_1, \dots, A_\lambda = m_\lambda + r_\lambda$.

Every linear integral form Λ is a weight because the highest form $\underline{\psi}$ among the $S\Phi$ has the property $S\underline{\psi} \geq \underline{\psi}$. If two elements $h^*(\varphi^1, \dots, \varphi^\lambda)$ and $h^*(\psi^1, \dots, \psi^\lambda)$ are equal, then necessarily $\mathcal{C}(\Lambda)$ has the same value for $(\varphi^1, \dots, \varphi^\lambda)$ and $(\psi^1, \dots, \psi^\lambda)$ (cf. (89) and (89*)). Hence $\Lambda(\varphi^1, \dots, \varphi^\lambda) \equiv \Lambda(\psi^1, \dots, \psi^\lambda) \pmod{1}$ for every integral Λ . Then $\varphi^1 \equiv \psi^1, \dots, \varphi^\lambda \equiv \psi^\lambda \pmod{1}$.

This shows that the converse of the lemma on p. 113 holds.

Chapter 3.

Seminar on various topics in group theory

1. Connection between the Betti numbers of the manifold of a semi-simple group and the number of integral-invariants *

* E. Cartan, Ann. Soc. Polonaise de Math. 8 (1929), p. 181.

Given a compact manifold M whose points are given by coordinates $x = (x^1, x^2, \dots, x^r)$. We consider differentials of p^{th} order on M

$$(1) \quad \omega = \sum_{i_1 < i_2 < \dots < i_p} f_{i_1 \dots i_p}(x) dx^{i_1} dx^{i_2} \dots dx^{i_p}$$

These are the terms under the integral sign in an p -dimensional integral on M

Here $dx^{i_1}, \dots, dx^{i_p}$ is to be considered as the component $(i_1, \dots, i_p)$ of a skew-symmetric tensor of rank p : If $y = (y^1, y^2, \dots, y^r)$ is another coordinate system on M

$$(2) \quad y^i = \varphi^i(x^1, \dots, x^r) = \varphi^i(x)$$

$$(3) \quad dy^i = \sum \frac{\partial \varphi^i}{\partial x^k} dx^k$$

We set

$$(4) \quad dy^{i_1} \dots dy^{i_p} = \sum_{k_1 < k_2 < \dots < k_p} \frac{\partial(\varphi^{i_1}, \dots, \varphi^{i_p})}{\partial(x^{k_1}, \dots, x^{k_p})} dx^{k_1} \dots dx^{k_p}$$

$$(5) \quad \sum f_{i_1 \dots i_p} dy^{i_1} \dots dy^{i_p} = \sum dx^{k_1} \dots dx^{k_p} \sum f_{i_1 \dots i_p} \frac{\partial(\varphi^{i_1}, \dots, \varphi^{i_p})}{\partial(x^{k_1}, \dots, x^{k_p})}$$

The transformation (4) is exactly the transformation of the components of a skew-symmetric tensor of rank p which is induced by the transformation (2) in the vector space. The formula (5) is the formula we have to use for the transformation of p -dimensional integrals. In accordance with our agreement, we set

$$(6) \quad dx^i dx^k = -dx^k dx^i, \quad dx^i dx^i = 0$$

If we define $f_{i_1, \dots, i_p}$ for all the values of $i_1, \dots, i_p$,

($i_p = 1, 2, \dots, r$) so that it is skew-symmetric with respect to $i_1, \dots, i_p$, we have

$$(7) \quad \omega = \frac{1}{p!} \sum_{i_1, \dots, i_p=1}^n f_{i_1 \dots i_p} dx^{i_1} \dots dx^{i_p}$$

The differentials (for a fixed x) are the elements of the dual space of the tensor space.

Suppose that p directions $d_1, d_2, \dots, d_p$ in the point x are given, $d_p = (d_p^1, \dots, d_p^p)$. We can derive a skew-symmetric tensor $[d_1, d_2, \dots, d_p]$ from these p vectors (cf. p. 21). The expression obtained from ω by replacing $dx^{i_1}, \dots, dx^{i_p}$ by the corresponding component of $[d_1, d_2, \dots, d_p]$ will be denoted by $\omega(d_1, d_2, \dots, d_p)$.

We consider more general sums of differentials of different order, differentials of order 0 being functions $f(x^1, \dots, x^r)$. These sums form the elements of an associative algebra if multiplication is defined by means of (6).

The derivative of a differential ω , (7), is defined by

$$(8) \quad \omega' = \frac{1}{p!} \sum_{i_1, \dots, i_p=1}^r d f_{i_1 \dots i_p} dx^{i_1} \dots dx^{i_p} = \frac{1}{p!} \sum_{i_1, \dots, i_p=1}^r \sum_j \frac{\partial f_{i_1 \dots i_p}}{\partial x^j} dx^j dx^{i_1} \dots dx^{i_p}$$

it is a differential of order $p+1$. The following rules can be proved without great difficulty*:

* For proofs cf. Cartan's book: *Lecons sur les invariants integraux*, Paris, 1922, pp. 65-73
In particular, pp. 71-73 for (13), p. 67 for (12), p. 67 for (11)

$$(9) \quad (\omega_1 + \omega_2)' = \omega_1' + \omega_2', \quad (\omega_1 \omega_2)' = \omega_1' \omega_2 + (-1)^s \omega_1 \omega_2'$$

where ω_1 and ω_2 are two differentials, s being the order of ω_2 . The derivative of the derivative is always 0.

$$(10) \quad (\omega')' = 0$$

Let $d_1, d_2, \dots, d_{p+1}$ be $p+1$ directions in x . We have

$$(11) \quad \omega'(d_1, \dots, d_{p+1}) = \sum_{v=1}^{p+1} (-1)^{v+1} d_v \omega(d_1, \dots, d_{v-1}, d_{v+1}, \dots, d_{p+1})$$

Here $d_V \omega = \sum_p \frac{\partial \omega}{\partial x^p} d_V^p$ is obtained from ω by differentiation in the direction d_V , the $d_\mu^1, \dots, d_\mu^r$ being considered as constants for this differentiation. From (11) can be easily derived that the formation of the derivative is invariant with respect to transformation (2).

Let R be a $(p+1)$ -dimensional subset of M satisfying certain conditions with regard to differentiability etc. Let R be bounded by the p -dimensional set B . Then

$$(12) \quad \int_R \omega' = \int_B \omega \quad (\text{generalized theorem of Stokes})$$

A differential ω is exact if the derivative ω' vanishes. In particular, every derivative is exact according to (10). Conversely, every exact differential can be represented as the derivative of a differential Ω of $(p-1)^{\text{th}}$ order:

$$(13) \quad \omega = \Omega' \quad \text{in a neighborhood of a given point.}$$

In general, the same is not true "im Grossen". If ω is a derivative on the whole manifold M we say that ω is homologous 0, $\omega \sim 0$.

If ω is exact, $\omega' = 0$, and $\mathcal{C}$ is a p -cycle homologous 0 on M , we can derive from (12) that $\int_{\mathcal{C}} \omega = 0$. If $\omega \sim 0$ it follows from (13) and (12) applied on $\mathcal{C}$ that $\int_{\mathcal{C}} \omega = 0$ for all p -cycles $\mathcal{C}$ on M .

G. de Rham (Journal de mathématiques pures et appliquées (9) 10 (1931), pp. 165-190) proved: 1) if, conversely, $\int_{\mathcal{C}} \omega = 0$ for all p -cycles $\mathcal{C}$, then $\omega \sim 0$; 2) To every p -cycles $\mathcal{C}_k$ belongs an exact differential ω such that

$$\int_{\mathcal{C}_1} \omega = 1 \text{ but } \int_{\mathcal{C}_k} \omega = 0 \text{ for all } p\text{-cycles } \mathcal{C}_k \text{ which are not homologous } *.$$

A system $\omega^1, \dots, \omega^h$ of exact differentials of the same order p is linearly independent if no relation

$$c_1 \omega^1 + \dots + c_h \omega^h \sim 0$$

with constant coefficients $c_1, \dots, c_h \neq (0, \dots, 0)$ holds. The maximum

number of linearly independent exact differentials of order p equals the p^{th} Betti number B_p of M according to de Rham's results.

We assume now that M is a compact continuous group, generated by infinitesimal elements. If x is a point of M and $x+dx$ is a neighboring element, then $(x+dx)x^{-1} = 1 + \delta x$ is an infinitesimal transformation of the group. The components $(\delta x^1, \delta x^2, \dots, \delta x^r)$ of this element are obtained from the components $(dx^1, \dots, dx^r)$ of dx by a linear transformation $\delta x^i = \sum \lambda_{ik}^i dx^k$ of non-vanishing determinant [p. 34]. Every differential

$$\omega = \sum_{i_1, \dots, i_p} f_{i_1, \dots, i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p} \quad (14)$$

$$\omega = \sum_{i_1, \dots, i_p} g_{i_1, \dots, i_p} \delta x^{i_1} \wedge \dots \wedge \delta x^{i_p}$$

where the $g_{i_1, \dots, i_p}$ also are skew-symmetric.

A differential ω is called invariant if it is invariant with respect to right and left hand translation on M : $x \rightarrow x' = xa$ and $x \rightarrow x'' = ax$ respectively where a is an arbitrary group element. We have for right-hand translation $(x+dx) \cdot x^{-1} \rightarrow (x+dx)a \cdot a^{-1}x^{-1} = (x+dx)x^{-1}$, $\delta x \rightarrow \delta x' = \delta x$.

The invariance of ω under $x \rightarrow x'$ gives hence $g_{i_1, \dots, i_p}(x) = g_{i_1, \dots, i_p}(x')$.

The coefficients $g_{i_1, \dots, i_p}$ in (14) must be constant.

In the case of left-hand translation, we have

$$(x + dx) \cdot x^{-1} \rightarrow a(x+dx) \cdot x^{-1} \cdot a^{-1}, \quad \delta x \rightarrow \delta x'' = a \delta x a^{-1}$$

The δx undergo, therefore, the transformation A of the adjoint group corresponding to a . Thus we have

Lemma 1: A differential is invariant if, when written in the form (14), the coefficients are constant and the expression is invariant under the transformations of the adjoint group: $\delta x \rightarrow a \cdot \delta x \cdot a^{-1}$

We shall prove now

Lemma 2: Every invariant differential is exact.

Proof. Let (14) be an invariant differential. If $e_1, \dots, e_r$

form the basis of the infinitesimal group, and

$$[e_\sigma e_\lambda] = \sum c_{\sigma\lambda}^{(K)} e_K$$

then the matrix E_σ associated with e_σ in the adjoint group is given by

$\|c_{\sigma\lambda}^{(K)}\|$ We apply to δx^i the corresponding transformation in the neighborhood of the identity (h infinitesimal)

$$\delta x^K \rightarrow \delta x^K + h \sum c_{\sigma\sigma}^{(K)} \delta x^\sigma$$

Now (14) is invariant under this transformation since it is invariant for the adjoint group. Hence

$$(15) \quad 0 = \sum g_{i_1 \dots i_p} \left(c_{\sigma\sigma}^{i_1} \delta x^\sigma \delta x^{i_2} \dots \delta x^{i_p} + c_{\sigma\sigma}^{i_2} \delta x^{i_1} \delta x^\sigma \delta x^{i_3} \dots \delta x^{i_p} \right. \\ \left. + \dots + c_{\sigma\sigma}^{i_p} \delta x^{i_1} \dots \delta x^{i_{p-1}} \delta x^\sigma \right)$$

holds for every σ .

We express in (14) the δx in terms of the dx , $\delta x^i = \sum \lambda_{K_i}^i dx^K$

and obtain

$$\omega = \sum_{i_1, \dots, i_p, K_1, \dots, K_p} g_{i_1 \dots i_p} \lambda_{K_1}^{i_1} \dots \lambda_{K_p}^{i_p} dx^{K_1} \dots dx^{K_p}$$

The derivative ω' of ω is, therefore, given by

$$\omega' = \sum \frac{\partial (g_{i_1 \dots i_p} \lambda_{K_1}^{i_1} \dots \lambda_{K_p}^{i_p})}{\partial x^j} dx^j dx^{K_1} \dots dx^{K_p} \\ = \sum_{\substack{i_1, \dots, i_p \\ K_1, \dots, K_p}} g_{i_1 \dots i_p} \left(\frac{\partial \lambda_{K_1}^{i_1}}{\partial x^j} \lambda_{K_2}^{i_2} \dots \lambda_{K_p}^{i_p} + \lambda_{K_1}^{i_1} \frac{\partial \lambda_{K_2}^{i_2}}{\partial x^j} \lambda_{K_3}^{i_3} \dots \lambda_{K_p}^{i_p} + \dots \right) dx^j dx^{K_1} \dots dx^{K_p}$$

the g 's being constant. We consider the term

$$S = \sum g_{i_1 \dots i_p} \frac{\partial \lambda_{K_1}^{i_1}}{\partial x^j} \lambda_{K_2}^{i_2} \dots \lambda_{K_p}^{i_p} dx^j dx^{K_1} \dots dx^{K_p}$$

If we interchange the summation indices j and K_1 , we have

$$S = \sum g_{i_1 \dots i_p} \frac{\partial \lambda_{K_1}^{i_1}}{\partial x^{K_1}} \lambda_{K_2}^{i_2} \dots \lambda_{K_p}^{i_p} (-dx^j dx^{K_1}) \dots dx^{K_p}$$

because $dx^{K_1} dx^j = -dx^j dx^{K_1}$, and hence

$$2S = \sum g_{i_1 \dots i_p} \left(\frac{\partial \lambda_{i_1}^{j_1}}{\partial x^{j_1}} - \frac{\partial \lambda_{j_1}^{i_1}}{\partial x^{i_1}} \right) \lambda_{i_2}^{j_2} \dots \lambda_{i_p}^{j_p} dx^{j_1} dx^{j_2} \dots dx^{j_p}$$

Now the following fundamental formula holds in Lie's theory, [p. 36]

$$\frac{\partial \lambda_{\alpha}^{\beta}}{\partial x^{\beta}} - \frac{\lambda_{\beta}^{\alpha}}{\partial x^{\alpha}} + \sum c_{\sigma\tau}^{\rho} \lambda_{\alpha}^{\sigma} \lambda_{\beta}^{\tau} = 0$$

Therefore, we have

$$2S = \sum g_{i_1 \dots i_p} c_{\sigma\tau}^{i_1} \lambda_{i_1}^{\sigma} \lambda_{j_1}^{\tau} \lambda_{i_2}^{j_2} \dots \lambda_{i_p}^{j_p} dx^{j_1} dx^{j_2} \dots dx^{j_p}$$

and if we introduce the $\delta x^i = \sum \lambda_{i_1}^{i_1} dx^{i_1}$ again

$$2S = \sum g_{i_1 \dots i_p} c_{\sigma\tau}^{i_1} \delta x^{\sigma} \delta x^{\tau} \delta x^{i_2} \dots \delta x^{i_p}$$

The same method can be applied in the other terms of ω' . We obtain so

$$2\omega' = \sum g_{i_1 \dots i_p} \delta x^{\tau} \left(c_{\sigma\tau}^{i_1} \delta x^{\sigma} \delta x^{i_2} \dots \delta x^{i_p} + c_{\tau\sigma}^{i_2} \delta x^{i_1} \delta x^{\sigma} \delta x^{i_3} \dots \delta x^{i_p} + \dots \right)$$

We multiply (15) by δx^{τ} and sum over τ . We have then exactly the same expression on the right-hand side. Hence $\omega' = 0$ as required by the lemma.

Lemma 3: Every exact differential ω is homologous to an invariant differential η .

Proof. We perform a left-hand translation $x \rightarrow y = s \cdot x$ in M , $y = \varphi(x)$. We replace x by y , dx^i by dy^i in ω , and introduce x again by means of (2), (3) and (5). We obtain a new differential which may be denoted by $s\omega$. Let $\mathcal{C}$ be a p -cycle on M , and $s\mathcal{C}$ its image by the translation.

Then

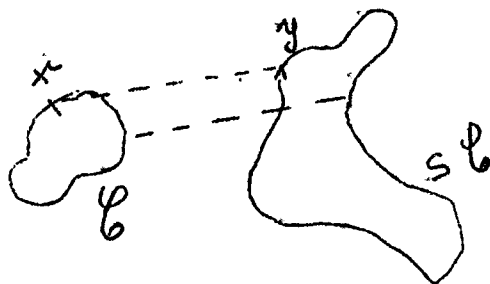
$$\int_{s\mathcal{C}} \omega = \int_{\mathcal{C}} s\omega$$

according to (5).

But $s\mathcal{C}$ can be obtained from $\mathcal{C}$ by a continuous deformation,

the group being connected. As we have seen, we then have $\int_{\mathcal{C}} \omega = \int_{s\mathcal{C}} \omega$ because ω is exact and $\mathcal{C}$ and $s\mathcal{C}$ are homologous. Hence

$$\int_{\mathcal{C}} \omega = \int_{s\mathcal{C}} \omega$$



We integrate with respect to s over $\mathcal{M}$ and obtain

$$\int_{\mathcal{C}} \omega = \int_{\mathcal{C}} \int_{\sin \mathcal{M}} s \omega$$

the invariant volume measure on $\mathcal{M}$ being chosen so that $\mathcal{M}$ itself has the volume 1. Here $\int_{\sin \mathcal{M}} s \omega = \mathcal{N}$ is an exact differential (integration with respect to s , not with respect to ω !). Obviously, $\mathcal{N}$ is invariant under left translation.

By using the same argument with regard to right-hand translation, we obtain a differential $\int_{\mathcal{C}} \mathcal{N} s = \mathcal{Y}$ which is invariant under left and right-hand translation. For every p -cycle we have

$$\int_{\mathcal{C}} \omega = \int_{\mathcal{C}} \mathcal{Y}.$$

It follows from de Rham's results (cf. above) that if the integral of an exact differential vanishes for every p -cycle, the differential is homologous 0.

Hence $\omega - \mathcal{Y} \sim 0$ and this proves the lemma. Cartan gave a different proof:

He shows that $\omega \sim s \omega$. Then by integration with respect to s he obtains (the same notations as above) $\omega \sim \mathcal{N}$ and hence similarly $\omega \sim \mathcal{Y}$. In order to prove $s \omega \sim \omega$ it is sufficient to consider only infinitesimal transformations s of $\mathcal{M}$ since the result for a finite s then can be obtained by repeating infinitesimal transformations (integration with respect to a parameter).

Let δs be a fixed infinitesimal element of the group. The element $1 + \delta s$ transforms every element x of $\mathcal{M}$ into a neighboring element

$x + dx = (1 + \delta s) \cdot x$. The components of δs may be written in the form

$$\delta s^i = \varepsilon \xi^i, \quad \varepsilon \text{ infinitesimal and the } \xi^i \text{ being constants. The } \delta s^i$$

and the components dx^i of dx are connected by formulas

$$dx^i = \sum \mu_{\mathcal{R}}^i(x) \delta s^{\mathcal{R}} = \varepsilon \sum \mu_{\mathcal{R}}^i \xi^{\mathcal{R}}$$

We apply (11). Because of $\omega' = 0$, we have

$$(16) \quad \sum_{v=1}^{p+1} (-1)^{v+1} d_v \omega(d_1, \dots, d_{v-1}, d_{v+1}, \dots, d_{p+1}) = 0$$

$$d_{p+1} \omega(d_1, \dots, d_p) = \sum_{v=1}^p (-1)^{v+1} d_v \omega(d_1, \dots, d_{v-1}, d_{v+1}, \dots, d_{p+1}) \cdot (-1)^p$$

We replace now $d_{p+1} = (d_{p+1}^1, \dots, d_{p+1}^r)$ by the vector $(\sum \mu_k^i \xi^k, \dots, \sum \mu_k^r \xi^k)$.

To be true, the components are not constants. But (11) still holds if we agree to set

$$d_{p+1}(d_p^i) = d_p d_{p+1}^i = d_p \left(\sum_k \mu_k^i \xi^k \right) = \sum_{k, \sigma} \frac{\partial \mu_k^i}{\partial x^\sigma} d_p^\sigma \xi^k, \quad (p \leq$$

and to take this into account when effecting the differentiations

$d_{p+1} \omega(d_1, \dots, d_p)$ and $d_v \omega(\dots)$ in (16).

Now $\omega(d_1, \dots, d_{v-1}, d_{v+1}, \dots, d_p, d_{p+1})(-1)^p$ may be written in the form $\mathcal{Q}(d_1, \dots, d_{v-1}, d_{v+1}, \dots, d_p)$ where $\mathcal{Q}$ is a differential of order $p-1$. If we use (11) for $\mathcal{Q}$ instead of ω we obtain from (16)

$$d_{p+1} \omega(d_1, \dots, d_p) = \mathcal{Q}'(d_1, \dots, d_p)$$

But the left-hand side is the limit of $\frac{\omega^* - \omega}{\varepsilon}$ for $\varepsilon \rightarrow 0$ where ω^* is obtained from ω by replacing x by $x + dx = x + \varepsilon d_{p+1}$ and d_p^i by $d_p^i + \varepsilon d_{p+1}(d_p^i)$. We have therefore $\omega^* = (1 + \mathcal{I}_s) \omega$ and hence

$$(1 + \mathcal{I}_s) \omega(d_1, \dots, d_p) - \omega(d_1, \dots, d_p) = \varepsilon \mathcal{Q}'(d_1, \dots, d_p)$$

This holds for any p vectors $d_1, \dots, d_p$. Hence

$$(1 + \mathcal{I}_s) \omega - \omega = \varepsilon \mathcal{Q}'$$

The statement $s \omega - \omega \sim 0$ is, therefore, true for any $s = 1 + \mathcal{I}_s$ in the neighborhood of the identity. We obtain the statement for an arbitrary s by integrating with respect to $\mathcal{I}_s, \varepsilon$.

Lemma 4: If ω is invariant and $\omega \sim 0$, we have $\omega = 0$.

Proof. Suppose that $\omega = \mathcal{Q}'$. We have for s in $\mathcal{M}$

$$\omega = s \omega = s \mathcal{Q}' = (s \mathcal{Q})'$$

and on integrating with respect to s

$$\omega = \int_{\text{sim } \mathcal{M}} (s \mathcal{Q})' = \left(\int_s s \mathcal{Q} \right)' = \textcircled{4}'$$

Here $\textcircled{4}$ is invariant with respect to left-side translation on $\mathcal{M}$. Now we perform the same argument with regard to right-hand translation

$$\omega = \omega S = (\omega S)' = (\omega S)'$$

$$\omega = \int (\omega S)' = \left(\int \omega S \right)' = \Phi'$$

where Φ is invariant. From Lemma 2 follows then that Φ is exact, $\Phi' = 0$, i.e. $\omega = 0$, q.e.d.

We are now in a position to prove:

Theorem. The p^{th} Betti number B_p of a compact connected continuous group is equal to the number of linearly independent invariant differentials of order p on M , linear independence being understood in the ordinary sense.

Proof. The number B_p equals the maximum number of linearly independent exact differentials of order p on M . We may replace here the exact differentials by invariant differentials according to Lemma 3. But invariant differentials are linearly independent in the sense of homology, if and only if they are linearly independent in the ordinary sense according to Lemma 4. This proves the theorem.

The integral $J(R) = \int_R \omega$, defined for any r -dimensional R on M is called invariant, if ω is invariant. Then $J(R) = J(R')$ if R is carried over into R' by a (left or right-hand) translation. The Betti number B_r is the number of linearly independent r -fold invariant integrals.

If we make use of Lemma 1 we may formulate the last theorem in the following form.

Theorem. Let $A^{(p)}$ be the transformation which the skew-symmetric tensors of rank p undergo if the transformation A of the adjoint group $\mathcal{A}$ of M is performed in the vector space. Then the Betti number B_p equals the number of linearly independent linear forms

$$\sum_{i=1}^N a_i z_i \quad (a_i \text{ constant, } N = \binom{n}{p})$$

which are invariant if the z_i undergo the transformation $A^{(p)}$ for every A in $\mathcal{A}$.

Let $A = A(s) = \| a_{ik}(s) \|$ be the element corresponding to s in a unitary restricted semi-simple group M .

$$(17) \quad |t \cdot 1 + A(s)| = t^r + g_1(s)t^{r-1} + \dots + g_r(s)$$

Here $g_p(s)$ is the p^{th} elementary symmetric function of the characteristic roots $\lambda_1, \lambda_2, \dots, \lambda_r$ of A . As is easily proved, the characteristic roots of $A^{(p)}$ are the products $\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_p}$, $i_1 < i_2 < \dots < i_p$. Hence

$$(18) \quad \text{tr}(A^{(p)}) = g_p(s)$$

We use now the results and notations of Chapter 2. We denote by $\{1\}$ with the representation of degree 1 which associates the number 1 with every s . If $T(s)$ is an irreducible representation and $\chi(s) = \text{tr}(T(s))$ its character, we have (Chapter 2, formulas (94), (97) and (95), and (91))

$$(19) \quad \int_0^1 \dots \int_0^1 \chi \cdot 1 \cdot \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^l = \begin{cases} 0 & T(s) \neq \{1\} \\ c \neq 0 & T(s) = \{1\} \end{cases}$$

$$\Delta = \prod_{\alpha > 0} (e(\frac{\alpha}{2}) - e(-\frac{\alpha}{2})), \quad \Delta \bar{\Delta} = \prod_{\alpha} (1 - e(\alpha))$$

where α ranges over the "roots" of the group.

It follows in the case of a reducible $T(s)$ that $\frac{1}{c} \int_0^1 \dots \int_0^1 \chi \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^l$ equals the number of times $T(s)$ contains $\{1\}$ as an irreducible constituent.

Then in an appropriate coordinate system all matrices of $T(s)$ have the form

$$T(s) = \left\| \begin{array}{c|c} 1 & 0 \\ \hline 0 & * \end{array} \right\| \left. \vphantom{\begin{array}{c|c} 1 & 0 \\ \hline 0 & * \end{array}} \right\} q \text{ rows}$$

since the representation $T(s)$ is completely reducible.

The first q coordinates are invariants under the transformations of $T(s)$. If there should exist a $(q+1)^{\text{st}}$ linearly independent linear invariant, we could choose this linear form as $(q+1)^{\text{st}}$ coordinate and we see that $T(s)$ would contain one more constituent $\{1\}$.

Thus q is the number of linearly independent invariants for the representation $T(s)$.

The last theorem in connection with (18) gives

$$\frac{1}{c} \int_0^1 \dots \int_0^1 g_p(s) \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^l = B_p$$

Hence we obtain from (17) that

$$(20) \quad \frac{1}{c} \int_0^1 \dots \int_0^1 |t \cdot 1 + A| \Delta \bar{\Delta} d\varphi^1 \dots d\varphi^l = t^r + B_1 t^{r-1} + \dots + B_r = P(t)$$

We call this polynomial the Poincaré-polynomial of M ; the coefficients are the Betti numbers.

The characteristic roots of $|t \cdot 1 + A|$ are the values $t + \lambda_1, \dots,$

$t + \lambda_r$. The determinant is the product of the characteristic roots,

$$|t \cdot 1 + A| = \prod_i (t + \lambda_i)$$

According to Chapter 2, (90*), λ of the characteristic roots

$\lambda_1, \dots, \lambda_r$ of A have the value 1, the others are $e(\alpha)$ where the α are the roots of the group again. Hence $|t \cdot 1 + A| = (t+1)^l \prod_{\alpha} (t + e(\alpha))$.

Hence it follows from (20) and (19) that

$$P(t) = \frac{1}{c} \int_0^1 \dots \int_0^1 (t+1)^l \prod_{\alpha} (t + e(\alpha)) \prod_{\alpha} (1 - e(\alpha)) d\varphi^1 \dots d\varphi^l$$

With respect to $\varphi^1, \dots, \varphi^l$ every term under the integral sign is a finite

Fourier series $\sum q_{m_1, \dots, m_l} e^{(m_1 \varphi^1 + \dots + m_l \varphi^l)}$ with integral

$m_1, \dots, m_l$. Hence the integrand has the same form and the value of the integral is the term independent of $\varphi^1, \dots, \varphi^l$ in the Fourier series of

$$\frac{1}{c} (1+t)^l \prod_{\alpha} (t + e(\alpha)) \prod_{\alpha} (1 - e(\alpha))$$

In particular, the highest coefficient 1 of $P(t)$ is the constant term in the Fourier series of

$$(21) \quad \frac{1}{c} \prod_{\alpha} (1 - e(\alpha))$$

Furthermore, for $t = 1$, we see that $P(1)$ is the constant term in the

Fourier series of

$$\frac{2^l}{c} \prod_{\alpha} (1 + e(\alpha)) \prod_{\alpha} (1 - e(\alpha)) = \frac{2^l}{c} \prod_{\alpha} (1 - e(2\alpha)).$$

But this Fourier series is obtained from (21) by replacing φ^i by $2\varphi^i$ and multiplying by 2^{λ} . Hence the constant term is 2^{λ} . This gives

$$(22) \quad P(1) = 1 + B_1 + B_2 + \dots + B_r = 2^{\lambda}.$$

In the case of the four "great" classes of simple groups Pontrjagin determined the B_r by direct topological methods (see below, sec. 3). In the following sec. 2 the invariant integrals of these groups will be constructed and Pontrjagin's result will be derived from the preceding theorem.

2. Construction of the invariant integrals for the manifold of a compact simple group*

* Cf. a note by R. Brauer which will be published soon in the Comptes rendus. (Volume 201, (1935), p. 419)

There exist four "great" classes of simple groups. The unimodular group $\mathcal{U}_n$ in n dimensions, the orthogonal group $\mathcal{O}_n$ for an odd number $n = 2\nu + 1$ of dimensions, the complex-group $\mathcal{L}_n$ in n dimensions, and the orthogonal group $\mathcal{A}_n$ for an even number $n = 2\nu$ of dimensions. Besides these groups there exist only five "exceptional" simple groups (cf. the last section of this chapter).

We shall construct a complete system of linearly independent invariant differentials of order p . The method is almost the same for the four classes of simple groups. So we deal mainly with the first case and give afterwards the changes which are required in the other cases. In the following consideration it is of no influence whether we consider the unitary restricted groups or the original groups.

We take the full linear group $\mathcal{G}$ of all linear transformations of an n -dimensional vector space (R) instead of the unimodular group $\mathcal{U}_n$.

The elements x of $\mathcal{G}$ are matrices $x = \|x_k^i\|$ and so are the infinitesimal transformations $\delta x = \|\delta x_k^i\|$. The infinitesimal elements form a vector space $\mathcal{V}$ of dimension $r = n^2$. The element h_x which corresponds to x in the adjoint representation $\mathcal{L}_x$ of $\mathcal{G}$ is given by $\|\delta x_k^i\| \rightarrow \|\tilde{\delta} x_k^i\| = x \cdot \|\delta x_k^i\| \cdot x^{-1}$. If $x'^{-1} = \|\xi_i^k\|$ is the contragredient matrix of x , we have

$$(23) \quad \delta x_k^i \rightarrow \tilde{\delta} x_k^i = \sum_p x_p^i \xi_k^p \delta x_p^q$$

Let $y = (y^1, y^2, \dots, y^n)$ be a vector of $\mathcal{R}$ and

$u = (u_1, u_2, \dots, u_n)$ a vector of the dual space $\mathcal{R}^*$. If x is transformed in $\mathcal{R}$ and x'^{-1} in $\mathcal{R}^*$, the components $y^i u_k$ of $y \times u$ undergo the same transformation $h(x) = x \times x'^{-1}$ as the δx_k^i in (23), we may consider

$y \times u$ as elements of $\mathcal{V}$.

We choose p vectors $Y_1, Y_2, \dots, Y_p$ of $\mathcal{R}$ and p vectors

$U_1, U_2, \dots, U_p$ of $\mathcal{R}^*$ whose coordinates are independent variables, and put

$$(24) \quad W_v = Y_v \times U_v^*$$

$$(25) \quad Z = [W_1, W_2, \dots, W_p]^*$$

* Cf. the definitions on p. 21.

Then W_v may be considered as vectors of $\mathcal{V}$. We distinguish the r coordinates in $\mathcal{V}$ by a pair i, k of indices; $i, k = 1, 2, \dots, n$, as we did in δx_k^i . Thus the components of W_v will be denoted by $W_v^{(i)}(k)$ and the components of the skew-symmetric tensor Z by

$$(26) \quad Z_{(k_1, \dots, k_p)}^{(i_1, \dots, i_p)}$$

The quantities (26) are linear homogeneous in the components of every Y_v and U_v , they are linearly independent.

Let ω be an invariant differential. We write it in the form (14)

$$(27) \quad \omega = \sum \varepsilon_{(k_1, \dots, k_p)}^{(i_1, \dots, i_p)} \delta x_{k_1}^{i_1}, \dots, \delta x_{k_p}^{i_p}$$

The coefficients g are constant according to Lemma 1 of the last section. Furthermore, if $\int x_k^i$ undergoes a transformation (23) of the adjoint group, then (27) is invariant. Here the product $\int x_{k_1}^{i_1}, \dots, \int x_{k_p}^{i_p}$ has to be considered as $(\begin{smallmatrix} i_1, \dots, i_p \\ k_1, \dots, k_p \end{smallmatrix})$ -component of a skew-symmetric tensor of rank p associated with the space $\mathcal{H}$. Conversely, if the left side in (27) is invariant under these transformations then ω is an invariant differential.

We replace now $\int x_{k_1}^{i_1}, \dots, \int x_{k_p}^{i_p}$ in (27) by the component (26) of Z .

We obtain a function

$$(28) \quad f = f(Y_1, \dots, Y_p, U_1, \dots, U_p)$$

with the following properties:

- a) f is linear and homogeneous in every Y_ν and U_ν
- b) f is invariant if the Y_ν are transformed cogrediently by a g of $\mathcal{G}$ and the U_ν by the contragredient transformation g^{-1} .
- c) If we apply a permutation π of $1, 2, \dots, p$ to the indices of $Y_1, Y_2, \dots, Y_p$ and of $U_1, U_2, \dots, U_p$, then f is transformed into $\pi(f) = \delta_\pi \cdot f$, $\delta_\pi = 1$ for even π and $\delta_\pi = -1$ for odd π .

Conversely, let f be a function with these properties. From a) it follows that we may express f linearly by the components $W_{\mathcal{S}}^{(i)}(k)$ of the vectors

(24) only. We put

$$(29) \quad f = \sum b(\begin{smallmatrix} i_1, \dots, i_p \\ k_1, \dots, k_p \end{smallmatrix}) W_1^{(i_1)}(k_1) W_2^{(i_2)}(k_2) \dots W_p^{(i_p)}(k_p)$$

The property c) yields

$$(30) \quad f = \frac{1}{p!} \sum_{\pi} \delta_\pi \pi(f)$$

where the sum extends over all $p!$ permutations π . We express the ~~left~~ ^{right} hand side by means of (29). The factor of $b(\begin{smallmatrix} i_1, \dots, i_p \\ k_1, \dots, k_p \end{smallmatrix})$ is the determinant

$$(31) \quad \begin{vmatrix} W_1^{(i_1)}(k_1) & W_1^{(i_2)}(k_2) & \dots & W_1^{(i_p)}(k_p) \\ \dots & \dots & \dots & \dots \\ W_p^{(i_1)}(k_1) & W_p^{(i_2)}(k_2) & \dots & W_p^{(i_p)}(k_p) \end{vmatrix}$$

which is equal to the expression (26) according to the definition of (25). Hence f is a linear function of the expressions (26). Let Z^* be any skew-symmetric tensor of rank p in the space $\mathcal{X}$. We replace (26) in f by the corresponding component of Z^* . It follows from the linear independence of the expressions (26) that f still is invariant if a transformation of the adjoint group is performed in S . In particular, if we replace (26) by

$$\delta x_{k_1}^{i_1} \delta x_{k_2}^{i_2} \dots \delta x_{k_p}^{i_p}, \text{ we obtain an invariant differential,}$$

Hence our task is to construct a complete system of linearly independent functions f which possess the properties a), b), c).

The "fundamental theorem" of the theory of invariants states *: Every

* For the proof of.: Study, *Invariantentheorie*, Braunschweig 1923; Weyl, *Math. Zeitschr.* 20, 10. 131.

invariant of the full linear group GL which is integral rational in a finite number of contravariant vectors $Y_1, Y_2, \dots$ and covariant vectors $U_1, U_2, \dots$ can be expressed as an integral rational function of expressions Q of the following types: 1) The determinants formed of n of the vectors Y_p ; 2) The determinants of n of the vectors U_p ; 3) The inner products $(Y_p U_\sigma)$.

In our case, f is linear in every Y_p and every U_σ , $p = 1, 2, \dots, p$ and $\sigma = 1, 2, \dots, p$. Hence every term Q of the first kind must appear multiplied by a term Q of the second kind. The multiplication theorem for determinants allows expressing a product of one factor of the first kind by a factor of the second in terms of expressions Q of the third kind. Hence we need only such inner products $(Y_p U_\sigma)$ in the representation of f . Because of the property a) we have a formula

$$(32) \quad f = \sum_p a_p (y_1 u_{p_1}) (y_2 u_{p_2}) \dots (y_p u_{p_p})$$

where $\rho : \nu \rightarrow \rho_\nu$ is a permutation of $1, 2, \dots, p$ and the a_ρ 's are constant. If g is any function of $Y_1, \dots, Y_p, U_1, \dots, U_p$, we set

$$(33) \quad T(g) = \frac{1}{p!} \sum_{\pi} \delta_{\pi} \cdot \pi(g)$$

where the sum extends over all $p!$ permutations of $1, 2, \dots, p$. Obviously

$$T(g_1 + g_2) = T(g_1) + T(g_2), \quad T(cg) = cT(g) \text{ for constant } c.$$

Furthermore, we set for any permutation $\rho : \nu \rightarrow \rho_\nu$ of

$1, 2, \dots, p$:

$$(34) \quad \mathcal{M}_\rho = (Y_1 U_{\rho_1}) (Y_2 U_{\rho_2}) \dots (Y_p U_{\rho_p}).$$

We apply the operation T on (32) and make use of (34) and (30). We have then

$$f = \sum_{\rho} a_{\rho} T(\mathcal{M}_\rho)$$

Every invariant f with the properties a), b), c) is a linear combination of $T(\mathcal{M}_\rho)$ with constant coefficients.

The $T(\mathcal{M}_\rho)$ themselves possess the properties a), b), c). For a) is evident, b) follows from the fact that (34) is an invariant, and finally, we have for every function g

$$\rho(T(g)) = \frac{1}{p!} \sum_{\pi} \delta_{\pi} \rho \pi(g) = \frac{\delta_{\rho\pi}}{p!} \sum_{\pi} \delta_{\pi} \rho \pi(g); \quad T(\rho g) = \frac{\delta_{\rho\pi}}{p!} \sum_{\pi} \delta_{\pi} \rho \pi(g)$$

But $\rho\pi$ and $\pi\rho$ range over all the permutations of $1, 2, \dots, p$. Hence

$$(35) \quad \rho(T(g)) = \delta_{\rho} T(g) = T(\rho(g))$$

In particular, $T(\mathcal{M}_\rho)$ has the property c).

Let $\pi : \nu \rightarrow \pi_\nu$ and $\rho : \nu \rightarrow \rho_\nu$ be two permutations. We compute $\pi T(\mathcal{M}_\rho) = T(\pi(\mathcal{M}_\rho))$. From (34) follows

$$\pi(\mathcal{M}_\rho) = (Y_{\pi_1} U_{\pi_{\rho_1}}) (Y_{\pi_2} U_{\pi_{\rho_2}}) \dots (Y_{\pi_p} U_{\pi_{\rho_p}}).$$

This expression has the form (34) itself. If $\pi_\lambda = \nu$, Y_ν is multiplied by

$U_{\pi_{\rho_\lambda}}$. The permutation $\pi\rho\pi^{-1}$ transforms ν into π_{ρ_λ} . Hence

* In the theory of permutation groups it is customary to write the factors in the opposite order as we do here in accordance with our general convention concerning transformations. If ρ transforms $\lambda \rightarrow \rho_\lambda$ and π transforms $\rho_\lambda \rightarrow \pi_{\rho_\lambda}$, we say here that $\pi\rho$ transforms $\lambda \rightarrow \pi_{\rho_\lambda}$.

$\pi(M_\rho) = M_{\pi\rho\pi^{-1}}$. Then (35) yields

$$(36) \quad T(M_{\pi\rho\pi^{-1}}) = \delta_\pi T(M_\rho).$$

We represent the permutation ρ as product of cyclic permutations

$$(37) \quad \rho = (\alpha_1, \alpha_2, \dots, \alpha_{h_1})(\beta_1, \beta_2, \dots, \beta_{h_2}) \dots$$

where every index $1, 2, \dots, p$ appears in exactly one parenthesis

$$p = h_1 + h_2 + \dots, \quad 1 \leq h_1 \leq h_2 \leq \dots$$

Then we have evidently

$$(38) \quad \pi\rho\pi^{-1} = (\pi\alpha_1, \pi\alpha_2, \dots, \pi\alpha_{h_1})(\pi\beta_1, \pi\beta_2, \dots, \pi\beta_{h_2}) \dots$$

We assume that there is an odd permutation π which is commutative with ρ ,

$\pi\rho\pi^{-1} = \rho$. From (36) follows $T(M_\rho) = -T(M_\rho)$, $T(M_\rho) = 0$. There are two cases where the assumption holds: 1) Let one of the h_ν be even, for instance h_1 even.

Then we take $\pi = (\alpha_1, \alpha_2, \dots, \alpha_{h_1})$ and π is odd.

From (37) and (38) follows $\rho = \pi\rho\pi^{-1}$ 2) Let two of the h_ν be equal,

for instance $h_1 = h_2$. Then we take $\pi = (\alpha_1, \beta_1)(\alpha_2, \beta_2) \dots (\alpha_{h_1}, \beta_{h_1})$.

We need only consider the case that h_1 is odd. Here π is odd because it is

the product of h_1 transpositions. According to (38) we have $\pi\rho\pi^{-1} = \rho$,

only the first two factors of ρ being interchanged. Thus in these two cases

we have $M_\rho = 0$.

In general, we can determine π in such a manner that on the right-hand side in (37) the indices $1, 2, \dots, p$ appear in natural order. Then

$$TM_\rho = \pm TM_{\pi\rho\pi^{-1}}$$

We have proved that every f with the properties a), b), c) is a linear combination of the M_ρ . Now we see that we need only the following ρ

$$\rho = (1, 2, \dots, h_1)(h_1 + 1, h_1 + 2, \dots, h_1 + h_2) \dots$$

$$(39) \quad p = h_1 + h_2 + \dots, \quad 0 < h_1 < h_2 < \dots, \quad h_\nu \equiv 1 \pmod{2}$$

$$M_\rho = \{(y_1, u_2)(y_2, u_3) \dots (y_{h_1-1}, u_{h_1})(y_{h_1}, u_1)\} \{(y_{h_1+1}, u_{h_1+2}) \dots\} \dots$$

where the second parenthesis refers to the second cycle in ρ etc.

We denote the component of all the vectors by upper indices. Then we have for the first parenthesis in M_ρ ($h = h_1$)

$$\sum_{i_1, \dots, i_p} y_{i_1}^{i_1} u_{i_2}^{i_2} y_{i_2}^{i_2} u_{i_3}^{i_3} \dots y_{i_{h-1}}^{i_{h-1}} u_{i_h}^{i_h} y_{i_h}^{i_h} u_{i_1}^{i_1}$$

Now $Y_\rho^\alpha U_\rho^\beta$ is the component $w_\rho(\alpha_\beta)$ of W_ρ ; (24) and similar formulas hold for the other parentheses. Hence

$$(40) \quad M_\rho = \sum W_1(i_1) W_2(i_2) \dots W_{h_1-1}(i_{h_1-1}) W_{h_1}(i_{h_1}) \sum W_{h_1+1}(i_{h_1+1}) \dots \sum \dots$$

We form now the differential

$$(41) \quad \omega = \sum_{i_1, i_2} \delta x_{i_1}^{i_1} \delta x_{i_2}^{i_2} \dots \delta x_{i_{h_1-1}}^{i_{h_1-1}} \sum \delta x_{i_{h_1+1}}^{i_{h_1+1}} \dots \sum \dots$$

We replace here $\delta x_{k_1}^{i_1} \delta x_{k_2}^{i_2} \dots$ by the component $z_{k_1, \dots, k_p}^{i_1, \dots, i_p}$ which is given by the determinant (31)

$$z_{k_1, \dots, k_p}^{i_1, \dots, i_p} = \sum_{\pi} \delta_{\pi} W_{\pi_1}(i_1) W_{\pi_2}(i_2) \dots W_{\pi_p}(i_p)$$

where π ranges over all permutations, $\pi: \gamma \rightarrow \pi_\gamma$. But the right-hand side here can be written in the form

$$z_{k_1, \dots, k_p}^{i_1, \dots, i_p} = \sum_{\pi} \delta_{\pi} \pi(W_1(i_1) \dots W_p(i_p)) = T(W_1(i_1) \dots W_p(i_p))$$

Hence we obtain when we perform the replacement

$$\sum_{i_1, i_2, \dots} T(W_1(i_1) W_2(i_2) \dots W_{h_1}(i_{h_1}) \cdot W_{h_1+1}(i_{h_1+1}) \dots) = T(\sum \dots) =$$

according to (40). Consequently, ω is exactly the invariant differential belonging to our invariant $f = T(M_\rho)$. We define

$$(42) \quad \Omega_h = \sum_{i_1, \dots, i_h} \delta x_{i_1}^{i_1} \delta x_{i_2}^{i_2} \dots \delta x_{i_h}^{i_h} = \text{tr}(\|\delta x_i^k\|^h)$$

Then (41) has the form

$$(43) \quad \Omega = \Omega_{h_1} \Omega_{h_2} \dots$$

The multiplication of differentials has been defined in sec. 1. We have the result: Every invariant differential can be expressed as a linear combination of the products (43) where the h_ν satisfy the conditions (39). Now the lemma holds:

Lemma: If $p \geq 2n+1$, $\mathcal{Q}_p$ itself can be represented as a linear combination of products (42) where all the indices are smaller than p .

Proof. We may assume that $p = 2q+1$ odd, $q > n$. We multiply the determinants of degree

$$\begin{vmatrix} Y_{s_1} & Y_{s_2} & \dots & Y_{s_q} \end{vmatrix} \cdot \begin{vmatrix} U_{t_1} & U_{t_2} & \dots & U_{t_q} \end{vmatrix} = P.$$

Here the elements in the ν -th column are the components of Y_{s_ν} (U_{t_ν} respectively) in the first n rows, whereas the last $q-n$ rows contain merely 0's.

Then, of course, $P = 0$. We multiply columns by columns and obtain:

$\det \begin{vmatrix} (Y_{s_1} \cdot U_{t_k}) \end{vmatrix} = 0$. We choose now $(s_1, s_2, \dots, s_q) = (2, 4, 6, \dots, p-1, p)$

and $(t_1, \dots, t_q) = (3, 5, 7, \dots, p, 1)$. If we expand the determinant and

multiply by $(Y_1 U_2)(Y_3 U_4) \dots (Y_{p-2} U_{p-1})$, we obtain

$$(44) \quad 0 = (y_1 u_2)(y_3 u_4) \dots (y_{p-2} u_{p-1}) \sum_{\gamma} \delta_\gamma (y_2 u_{\gamma_1})(y_4 u_{\gamma_2}) \dots (y_{p-1} u_{\gamma_{q-1}})(y_p u_{\gamma_q})$$

where $(\gamma_1, \gamma_2, \dots, \gamma_q)$ is obtained from $(3, 5, \dots, 7, p, 1)$ by a permutation γ of these q numbers and $\delta_\gamma = 1$ if γ is even, $\delta_\gamma = -1$ if γ is odd. The left-hand side is a sum of expressions $\pm M_\rho$, (34). We are inter-

ested only in such ρ as are cyclic of order p . The permutation ρ belonging to the term written down in (44) is given by

$$\rho = \begin{pmatrix} 1 & 2 & \dots & p \\ \rho_1 & \rho_2 & \dots & \rho_p \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & p-2 & p-1 & p \\ 2 & \gamma_1 & 4 & \gamma_2 & \dots & p-1 & \gamma_{q-1} & \gamma_q \end{pmatrix}$$

We have then $\gamma^{-1}\rho = (1, 2, 3, \dots, p-2, p-1, p)$. If ρ is a cycle of length p as is the product $\gamma^{-1}\rho$, then ρ and $\gamma^{-1}\rho$ both are even since p is odd. Then γ^{-1} and γ are even, $\delta_\gamma = 1$. The terms in question appear with the + sign in (44).

If ρ is a cycle, it has the form

$$\rho = (p, \overset{\text{odd}}{j_1}, \overset{\text{even}}{j_2}, \dots, \overset{\text{odd}}{j_{\mu-1}}, \overset{\text{even}}{j_{\mu}}, \dots)$$

Let π denote the permutation which transforms the indices of the parenthesis in $p, 1, 2, \dots, p-1$ in this order. Then

$$(45) \quad \pi \rho \pi^{-1} = (p, 1, 2, \dots, p-1) = (1, 2, \dots, p)$$

according to (37), (38). π does not alter the symbol p . It transforms consecutive pairs $2k-1, 2k$ into a pair of two consecutive indices, the first being odd, the second being even. It follows very easily that π is even. If

$\rho_0 = (1, 2, \dots, p)$, we have $T(M_\rho) = T(M_{\rho_0})$ according to (45) and (36).

Now (44) has the form

$$(46) \quad 0 = M_{\rho_0} + M_{\rho_1} + \dots + M_{\rho_k} + \sum M_\sigma$$

where $\rho_0, \rho_1, \dots, \rho_k$ are the occurring cycles of lengths p , and the σ are products of shorter cycles. It is important that $\rho_0 = (1, 2, \dots, p)$ occur in fact (for $j_1 = 3, j_2 = 5, \dots, j_{q-1} = p, j_q = 1$). We apply the operation T . We have

$$(k+1)TM_\rho + \sum TM_\sigma = 0$$

The same relation must hold if we replace (26) by the differential

$\delta_{x_{k_1}^{i_1}} \dots \delta_{x_{k_p}^{i_p}}$. But then TM_{ρ_0} gives Ω_p because of $\rho_0 = (1, 2, \dots, p)$. Moreover TM_σ is a product of several Ω_{h_v} with $h_v < p$ since σ is a product of cycles of length $h_v < p$. (Cf. (40), (41), (43)) This proves the lemma.

The lemma shows in connection with our former results that every invariant differential is a linear combination of invariants of the type

$$(47) \quad \omega = \Omega_{h_1} \Omega_{h_2} \Omega_{h_3} \dots$$

with

$$(48) \quad p = h_1 + h_2 + \dots; \quad h_1 < h_2 < h_3 < \dots < 2n+1; \quad h_v \equiv 1 \pmod{2}.$$

The number N of these ω is, therefore, equal to the number of decompositions

of p into a sum of different odd integers $\leq 2n-1$. Consequently, N is the coefficient of t^p in

$$(49) \quad \mathcal{F}(t) = (1+t)(1+t^3)\dots(1+t^{2n-1})$$

The number N^* of linearly independent invariants of order p cannot be larger than the number N of ω 's, $N^* \leq N$. We have $N^* = B_p$ according to the theorem of the last section where B_p is the p^{th} Betti number of the group space of the unitary restricted full linear group, i.e. the unitary group $\mathcal{U}$. Hence the polynomial

$$(50) \quad F^*(t) = 1 + B_1 t + B_2 t^2 + \dots + B_{r-1} t^{r-1} + t^r$$

has the property that every coefficient is not larger than the corresponding coefficient in (49). For $t = 1$, (49) has the value 2^n . According to (22)

$$\mathcal{F}^*(1) = 2^n \text{ also } *) \quad \mathcal{F}(1) = \mathcal{F}^*(1). \quad \text{This is possible only if no coefficient}$$

* The group $\mathcal{U}$ is, to be true, not semi-simple. But it can be easily seen that the proof in sec. 1 still holds in this case. The quantity χ has the value n here. Another way of proving $F^*(1) = 2^n$ is the following: The unitary group $\mathcal{U}$ is the direct product of the unitary unimodular group $\mathcal{U}_0$ and a 1-sphere. Hence the Poincare polynomial (50) of $\mathcal{U}_0$ is the product of the Poincare polynomial $\mathcal{F}^*(t)$ of $\mathcal{U}_0$ by $1+t$. From (22) applied to $\mathcal{U}_0$ follows $F_0(1) = 2 \cdot \mathcal{F}_0^*(1) = 2 \cdot 2^{n-1} = 2^n$.

in $F(t)$ is larger than the corresponding coefficient in $F^*(t)$. Therefore

$$F(t) = \mathcal{F}^*(t). \quad \text{Hence}$$

$$(\mathcal{U}) : 1 + B_1 t + B_2 t^2 + \dots + B_{r-1} t^{r-1} + t^r = (1+t)(1+t^3)\dots(1+t^{2n-1})$$

Our proof shows that the number of linearly independent invariant differentials of order p is equal to the number of ω , (47). This is only if the ω 's are linear themselves since every invariant differential of order p is a linear combination of such ω 's.

We express in Ω_p , (42), the δx_k^i by the dx_k^i . We have

$$x + dx = (1 + \delta x) \cdot x, \quad dx_k^i = \|\delta x_k^i\| \cdot \|x_k^i\|. \quad \text{If } x^{-1} = \|\xi_k^i\| \text{ we have}$$

$$\delta x_k^i = \|\delta x_k^i\| \cdot \|\xi_k^i\|$$

$$\delta_{x_k}^i = \sum_j dx_j^i \xi_k^j$$

We have proved now

Theorem 1. A complete system S of linearly independent invariant differentials of order p in the group space of the full linear group is obtained in the following way: set

$$\Omega_n = \sum \xi_{i_1}^{j_1} \xi_{i_2}^{j_2} \dots \xi_{i_n}^{j_n} dx_{j_1}^{i_1} dx_{j_2}^{i_2} \dots dx_{j_n}^{i_n} \quad dx_{j_n}^{i_n} = -\text{tr}((dx \cdot x^{-1})^n)$$

(x an arbitrary element of the group, $x^{-1} = \|\xi_k^i\|^{-1}$). The system S consists of all products $\Omega_{h_1} \Omega_{h_2} \Omega_{h_3} \dots$ where

$$h_1 + h_2 + \dots = p, \quad h_1 < h_2 < h_3 < \dots < 2n+1, \quad h_i \equiv 1 \pmod{2}$$

In the expression $-\text{tr}((dx \cdot x^{-1})^h)$, for Ω_h it is to be noticed that the dx_k^i are not commutative. The $-$ sign is necessary because of this fact.

Let x be an element of $\mathcal{G}$. Then x has the form $x = x_0 \cdot \gamma$ where γ is scalar and x_0 belongs to the unimodular group $\mathcal{U}_n$. In the adjoint representation $\|\delta x_k^i\| \rightarrow x \cdot \|\delta x_k^i\| \cdot x^{-1}$, x and x_0 are represented by the same transformation $h(x)$. We denoted the adjoint representation by $\mathcal{H}$, the space δx_k^i by $\mathcal{V}$. Let χ be the trace of $\|\delta x_k^i\|$, $\chi = \sum \delta x_k^i$. Then χ is invariant for all $h(g)$.

The infinitesimal elements of $\mathcal{U}_n$ are the matrices $\|\delta x_k^i\|$ with vanishing trace, i.e. the elements of $\mathcal{V}$ with $\chi = 0$. The adjoint group $\mathcal{H}^*$ of $\mathcal{U}_n$ takes place in this subspace $\mathcal{V}^*$ of $\mathcal{V}$. If x_0 in $\mathcal{U}_n$, the corresponding transformation $h^*(x_0)$ in $\mathcal{H}^*$ is given by $\|\delta x_k^i\| \rightarrow x_0 \cdot \|\delta x_k^i\| \cdot x_0^{-1}$ where $\sum \delta x_k^i = 0$. Hence $h^*(x_0)$ is the transformation which is induced by $h(x_0)$ in the hyperplane $\chi = 0$. If we choose a new coordinate system such that all the basis vectors except the first lie in $\chi = 0$, and the first vector is

$$\|\delta x_k^i\| = 1, \text{ we have}$$

$$(52) \quad h(x) = h(x_0) = \begin{pmatrix} 1 & 0 \\ 0 & h^*(x_0) \end{pmatrix}$$

Let $\delta y^0, \delta y^1, \dots, \delta y^{n^2-1}$ be the new coordinates, $\delta y^0 = \frac{1}{n} \chi$.

An invariant differential ω of order p in the group space $\mathcal{O}_n$ can be expressed by $\delta y^1, \dots, \delta y^{n^2-1}$ only. But because of the form (52) of the transformations $h(g)$ it is invariant if we subject the $\delta y^0, \delta y^1, \dots, \delta y^{n^2-1}$ to the transformation $h(x)$. Hence it is an invariant differential in $\mathcal{G}$. According to our last theorem we may represent ω as a linear combination of products $\mathcal{O}_{h_1} \mathcal{O}_{h_2} \dots$ where the indices satisfy (48). We set $\delta y^0 = 0, \omega$; is not changed. But $\mathcal{O}_1 = \sum \delta x_1^i$ according to (42), hence $\mathcal{O}_1 = \chi = n \delta y^0 = 0$. If $\mathcal{O}_h^*$ is the value of $\mathcal{O}_h$ for $\delta y^0 = 0$, we get that ω is a linear combination of products $\mathcal{O}_{h_1}^* \mathcal{O}_{h_2}^* \dots$ where

$$(53) \quad h_1 + h_2 + \dots = p, \quad 1 < h_1 < h_2 < \dots < 2n-1, \quad h_v \equiv 1 \pmod{2}$$

The number N of linearly independent invariant differentials ω of order p in the space $\mathcal{O}_n$ is, therefore, not larger than the coefficient of t^p in

$$(1+t^3)(1+t^5)\dots(1+t^{2n-1})$$

We denote by B_p now the Betti-numbers of $\mathcal{O}_n$ and set $r = n^2-1$. The same argument which we used in connection with (50) and (49) yields

$$(\mathcal{O}_n) : 1 + B_1 t + B_2 t^2 + \dots + B_{r-1} t^{r-1} + t^r = (1+t^3)(1+t^5)\dots(1+t^{2n-1})$$

Furthermore, the products $\mathcal{O}_{h_1}^* \mathcal{O}_{h_2}^* \dots$ where (53) holds are linearly independent. So are the products $\mathcal{O}_{h_1} \mathcal{O}_{h_2} \dots$. Now $\mathcal{O}_h$ is invariant in $\mathcal{O}_n$ also. Consequently $\mathcal{O}_{h_1} \mathcal{O}_{h_2} \dots$ form a basis for the invariant differentials of order p in $\mathcal{O}_n$.

Theorem 2. Theorem 1 holds for the unimodular group instead of the full linear group if the inequality $h_1 > 1$ is added to the conditions for the h_v .

We consider now the orthogonal group. The considerations here are very similar to those we made in the case of the full linear group and so we give only the changes which have to be made. The infinitesimal elements are given here by

$\|\delta x_k^i\|$ with $\delta x_k^i = -\delta x_i^k$. Therefore we have to replace (24) by
 (24') $W_\nu = [Y_\nu, U_\nu]$

Then the components of the W_ν undergo the same transformation (23) as the δx_k^i . There is no difference between covariant vectors U_ν and contravariant vectors Y_ν in this case. According to (24') the function f , (28), has the following property d) besides a), b), c) on p. 134:

d) f changes its sign if Y_ν and U_ν are interchanged.

On the other hand, if f has this property d) it contains the components Y_ν^i of Y and U_ν^i of U_ν only in the combinations

$$W_\nu \binom{i}{k} = Y_\nu^i U_\nu^k - Y_\nu^k U_\nu^i$$

It follows that every f with the properties a), b), c), d) possesses a representation (29). We see that our task is the construction of a complete system of linearly independent f with these properties.

The fundamental theorem of the theorem of invariants states in the case of the orthogonal group that every invariant may be expressed in terms of inner products of two vectors $(Y_\mu Y_\nu)$, $(Y_\mu U_\nu)$, $(U_\mu U_\nu)$ and of determinants formed of n of the vectors Y_μ or U_ν . (There is no difference between the Y_μ and U_ν now.) The product of two determinants can be expressed by means of inner products. If n is odd it follows from the property a) that the case of a term which contains one determinant cannot occur. Such a term would contain an odd number of vectors Y_ν, U_ν .

In the case of an even n we consider now only f which are invariant if the Y_ν, U_ν undergo an improper orthogonal transformation of determinant -1 . We call these f the even invariants f . The other f will be taken into account only later. Under this restriction no determinant will appear in the representation of f . We will have a formula (32) with the difference that factors $(Y_\mu Y_\nu)$ and $(U_\mu U_\nu)$ are allowed.

By Π we understand now a permutation of $Y_1, \dots, Y_p, U_1, \dots, U_p$

which is composed by permutations of the following kind

1) Application of a permutation of $1, 2, \dots, p$ to the indices of $Y_1, \dots, Y_p$ and of $U_1, \dots, U_p$.

2) Interchanging of Y_μ and U_μ .

A permutation of the second kind is to be considered as an odd permutation whereas the character of permutations of the first kind is the same as the character of the permutation of $1, 2, \dots, p$. A product of permutations 1), 2) is odd if an odd number of factors is odd, and even in the other case. We set $\delta_\Pi = -1$ and $\delta_\Pi = 1$ accordingly. Then (33) is to be replaced by

$$(33') \quad T(g) = \frac{1}{p! 2^p} \sum_{\Pi} \delta_\Pi \Pi(g)$$

the factor $p! 2^p$ being the order of the group of the Π 's. If g has the properties c) and d), we have $T(g) = g$. Furthermore (35) holds again. We see that every f is a linear combination of expressions $T(M)$ where M is a product of p factors $(Y_\mu U_\nu)$, $(Y_\mu Y_\nu)$, $(U_\mu U_\nu)$ every Y_μ and every U_ν appearing just in one factor. By a permutation Π such an M can be transformed into an expression M_ρ , (34), and because of (35) only such TM_ρ are necessary, ρ being a permutation of $1, 2, \dots, p$.

Let one of the h_ν in (37) be of the form $4x+1$, for instance $h = h_1$.

We apply the permutation $(\alpha_2, \alpha_h)(\alpha_3, \alpha_{h-1}) \dots (\alpha_{\frac{h+1}{2}}, \alpha_{\frac{h+3}{2}})$ and interchange every Y_{α_ν} with U_{α_ν} . This permutation Π is odd because of $h \equiv 1 \pmod{4}$. It transforms $(h = h_1)$

$$(Y_{\alpha_1} U_{\alpha_2})(Y_{\alpha_2} U_{\alpha_3}) \dots (Y_{\alpha_{h-1}} U_{\alpha_h})(Y_{\alpha_h} U_{\alpha_1})$$

and hence M_ρ into itself, the other factors of M_ρ being unaltered. From (35) follows $T(\Pi(M_\rho)) = -T(M_\rho)$ and because of $\Pi(M_\rho) = M_\rho$, we have

$T(M_p) = -T(M_p)$, $T(M_p) = 0$. Consequently we need only such M_p where all the numbers h_ν are of the form $4x+1$. If π is a permutation which does not interchange any Y_μ with U_μ , (36) holds again. Therefore it is sufficient to consider the M of the form (39). In (40) $W_\nu^{(i)}{}_k$ is $= Y_\nu^i U_\nu^k - Y_\nu^k U_\nu^i$ now according to (24'). We see that the differential ω , (41), gives exactly $T(M_p)$ again if we replace $\delta x_{k_1}^{i_1} \dots \delta x_{k_p}^{i_p}$ by the corresponding component of Z , (25). The terms in (33') which did not appear in (33) are accounted for by the fact that we have $W_\nu^{(i)}{}_k = Y_\nu^i U_\nu^k - Y_\nu^k U_\nu^i$ instead of $W_\nu^{(i)}{}_k = Y_\nu^i U_\nu^k$ in the case of the full linear group.

We obtain in this way

Theorem 3. A complete system of linearly independent invariant differentials of order p in the group space of the orthogonal group $\mathcal{L}_n$, n odd, is formed by the products $\mathcal{Q}_{h_1} \mathcal{Q}_{h_2} \dots$ where

$$h_1 + h_2 + \dots = p, \quad h_1 < h_2 < \dots < 2n+1, \quad h_\nu \equiv -1 \pmod{4}$$

Here $\mathcal{Q}_h$ is the same as in Theorem 1.

$$\text{We have } r = \frac{n(n-1)}{2}$$

$$(\mathcal{L}_n) \quad 1 + B_1 t + \dots + B_{r-1} t^{r-1} + t^r = (1+t^3)(1+t^7) \dots (1+t^{2n+1}) \quad 2n-3$$

In the case of an even n , we constructed only the even invariants, for which (27) is invariant also if δx_k^i undergo a transformation of the adjoint group of the improper orthogonal group. An odd invariant is one which changes its sign for this class of transformations. Then every ω is a sum of an even and an odd invariant.

The adjoint group $\mathcal{G}$ of the orthogonal group is orthogonal itself. In the space $\mathcal{R}$ in which $\mathcal{G}$ takes place, there is a one-to-one correspondence between skew-symmetrical tensors of rank p and $r-p$ which is invariant for all proper orthogonal transformations: If $a(i_1, i_2, \dots, i_p)$ is a skew-symmetric tensor of rank p , the corresponding tensor is given by

$$a^*(j_1, j_2, \dots, j_{r-p}) = a(i_1, i_2, \dots, i_p)$$

where $i_1, i_2, \dots, i_p, j_1, j_2, \dots, j_{r-p}$ is an even permutation of $1, 2, \dots, r$.

In particular, this construction coordinates a tensor of rank $r-p$ to

$$\int x_{k_1}^{i_1} \dots \int x_{k_p}^{i_p} \text{ in (27). We replace } \int x_{k_1}^{i_1} \dots \int x_{k_p}^{i_p} \text{ by } \int x_{\lambda_1}^{j_1} \int x_{\lambda_2}^{j_2} \dots \int x_{\lambda_{r-p}}^{j_{r-p}} \text{ where}$$

$$(54). \quad (i_1 k_1)(i_2 k_2) \dots (i_p k_p)(j_1 \lambda_1)(j_2 \lambda_2) \dots (j_{r-p} \lambda_{r-p})$$

are all $r = \frac{n(n-1)}{2}$ pairs (i, k) , $i < k$, in such an order that (54) can be transformed into

$$(1, 2)(1, 3), \dots, (1n)(23)(24), \dots, (2n), \dots, (n-1, n)$$

by an even permutation of the parentheses (i, k) . If we replace

$$\int x_{k_1}^{i_1} \dots \int x_{k_p}^{i_p} \text{ by } \int x_{\lambda_1}^{j_1} \dots \int x_{\lambda_{r-p}}^{j_{r-p}} \text{ we obtain a new invariant } \omega^* \text{ from}$$

(27). If ω is even, ω^* is odd as is easily seen. Every odd invariant is obtained in this way.

By a device similar to the proof of the lemma on p. 139 one sees that

$\mathcal{O}_{2n-1}$ can be expressed by means of $\mathcal{O}^*$ with smaller indices.

Theorem 4. In the case of the orthogonal group $\mathcal{O}_n$ for an even number n of dimensions a complete system of linearly independent even invariants is formed by the products $\mathcal{O}_{h_1} \mathcal{O}_{h_2} \dots = \omega$ with

$$h_1 + h_2 + \dots = p, \quad h_1 < h_2 < \dots < 2n-1, \quad h_v \equiv -1 \pmod{4}$$

The ω^* then form a basis for the odd invariant differentials of order $r-p$.

We set

$$(1+t^3)(1+t^7) \dots (1+t^{2n-5}) = g(t) = z_0 + z_1 t + z_2 t^2 + \dots$$

Then the maximum number of linearly independent even invariant differentials of order p is equal to z_p , the maximum number of linearly independent odd invariant differentials of order p is equal to z_{r-p} . Hence $B_p = z_p + z_{r-p}$, $r = \frac{n(n-1)}{2}$

$$1 + B_1 t + B_2 t^2 + \dots + t^r = g(t) + t^r g\left(\frac{1}{t}\right)$$

The degree of $g(t)$ is equal to $\frac{1}{2}(2n-2)(\frac{n}{2}-1) = \frac{(n-1)(n-2)}{2}$. Then

$$t^{\frac{(n-1)(n-2)}{2}} g\left(\frac{1}{t}\right) = g(t), \quad t^r g\left(\frac{1}{t}\right) = t^{r - \frac{(n-1)(n-2)}{2}} g(t) = t^{n-1} g(t)$$

$$(\mathcal{A}_n) \quad 1 + B_1 t + B_2 t^2 + \dots + B_{r-1} t^{r-1} + t^r = (1+t^3)(1+t^7) \dots (1+t^{2n-5})(1+t^{n-1})$$

In the case of the complex-group the final result is the same as in the case of B_n . There are some slight differences in the proof which come from the different form of the condition for the infinitesimal elements and from a difference in the form of the fundamental theorem of the theory of invariants for both groups.

Theorem 5. In Theorem 3, the group $\mathcal{L}_n$ can be replaced by the complex-group $\mathcal{K}_n$.

Since n is even here, we have $r = \frac{n(n+1)}{2}$

$$(\mathcal{K}_n) \quad 1 + B_1 t + \dots + B_{r-1} t^{r-1} + t^r = (1+t^3)(1+t^7) \dots (1+t^{2n-1})$$

3. Direct determination of the Betti numbers of the group manifolds of the compact simple groups *

* L. Pontrjagin, Comptes Rendus 200 (1935), p. 1277.

The Betti numbers of the manifolds of the compact simple groups $\mathcal{O}_n$, $\mathcal{L}_n$, $\mathcal{K}_n$, and $\mathcal{A}_n$ are given by the formulas $(\mathcal{O}_n)$, $(\mathcal{L}_n)$, $(\mathcal{K}_n)$ and $(\mathcal{A}_n)$ in sec. 2. These formulas have been proved by L. Pontrjagin by direct topological methods. Pontrjagin published the proof only for the case $\mathcal{O}_n$ and remarks that the proof is similar in the other cases. We give now Pontrjagin's proof for $\mathcal{O}_n$.

Let $\mathcal{O}_n$ be the group of all unitary unimodular matrices $\|a_{ik}\|$. If $a_{11} = 1$, we have

$$a_{12} = a_{13} = \dots = a_{1n} = 0, \quad a_{21} = a_{31} = \dots = a_{n1} = 0$$

These elements form a subgroup isomorphic with $\mathcal{O}_{n-1}$ which may be identified

with $\mathcal{O}_{n-1}$

Let $f = (f_1, f_2, \dots, f_n)$ be a matrix with one row, n columns and

$$(55) \quad f \cdot \bar{f}' = f_1^2 + f_2^2 + \dots + f_n^2 = 1$$

For any number α of absolute value 1, the matrix

$$F = 1 + (\alpha - 1)\bar{f}' \cdot f$$

is unitary. We have, namely, if we use (55)

$$\begin{aligned} \bar{f}' \cdot F &= (1 + (\alpha - 1)\bar{f}' \cdot f)(1 + (\alpha - 1)\bar{f}' \cdot f) \\ &= 1 + (\alpha - 1)\bar{f}' \cdot f + (\alpha - 1)\bar{f}' \cdot f + (\alpha - 1)\bar{f}' \cdot f \cdot \bar{f}' \cdot f (\alpha - 1) \\ &= 1 + (\alpha - 1 + \alpha - 1 + (\alpha - 1)(\alpha - 1)\bar{f}' \cdot f) = 1 \end{aligned}$$

If the coordinate system is chosen in such a manner that $f = (1, 0, \dots, 0)$, we see that $\bar{f}' \cdot f$ has the characteristic roots 1, 0, ..., 0. The characteristic roots of F are therefore $\alpha, 1, 1, \dots, 1$. Hence $\det(F) = \alpha$. The matrix

$$(56) \quad U = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & & -\alpha^{-1} \end{pmatrix} F = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & & -\alpha^{-1} \end{pmatrix} (1 + (\alpha - 1)\bar{f}' \cdot f)$$

is unitary and unimodular. We have in U

$$(57) \quad a_{11} = -1 + (1 - \alpha)\bar{f}_1' f_1, \quad a_{ij} = (1 - \alpha)\bar{f}_1' f_j, \quad (j \geq 2)$$

If we multiply f by a number of absolute value 1, the matrix U is not changed.

We denote the set of all such U by c .

We determine the intersection of c and $\mathcal{O}_{n-1}$. For the elements of it we have $1 = a_{11} = -1 + (1 - \alpha)|f_1|^2$. From $|f_1| \leq 1$, $|\alpha| = 1$ it follows that $\alpha = -1$, $|f_1| = 1$. Then, however, $f_2 = f_3 = \dots = f_n = 0$ and $U = 1$. The intersection of $\mathcal{O}_{n-1}$ and c consists of 1 only.

We set

$$a_{jk} = a'_{jk} + ia''_{jk}, \quad f = f'_v + if''_v, \quad \alpha = -e^{i\varphi}$$

In a neighborhood of the identity, we choose $f_2', f_2'', \dots, f_n', f_n''$ and φ as parameters. Then f_1 is nearly of the absolute value 1, and we may choose it as a real number (after multiplying f by an appropriate number of absolute value 1). We have, according to (55):

$$f_1' = \sqrt{1 - \sum_{j=2}^n (f_j'^2 + f_j''^2)} = 1 + \dots; \quad f_1'' = 0$$

The terms omitted in f_1' are of second orders in the parameters. In the same sense, $\alpha = -1 - i\varphi + \dots$ and

$$(58) \quad a_{11}'' = \varphi + \dots, \quad a_{1j}' = 2f_j' - \varphi f_j'', \quad a_{1j}'' = \varphi f_j' + 2f_j''$$

We denote the parameters $f_2', f_2'', \dots, f_n', f_n'', \varphi$ by

$$\lambda_1, \lambda_2, \dots, \lambda_{2n-1}. \quad \text{In the identity, we have } \lambda_1 = \lambda_2 = \dots = \lambda_{2n-1} = 0$$

and as we readily see from (58)

$$(59) \quad \Delta = \frac{\partial(a_{11}'', a_{12}', a_{12}'', \dots, a_{1n}', a_{1n}'')}{\partial(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_{2n-1})} \neq 0.$$

One sees that c has $2n-1$ dimensions. It is, of course, a cycle on $\mathcal{A}_n$.

Let V be any element in the neighborhood of a fixed element P_0 of $\mathcal{A}_{n-1}$. The first row in V is nearly 1, 0, ..., 0, and we can determine a U in c such that $a_{11}'', a_{12}', a_{12}'', \dots, a_{1n}', a_{1n}''$ have the same values in U as in V . The same is true for a_{11}' by itself. The first rows in U and V are identical. So are the first rows in $U \cdot U^{-1}$ and $V \cdot U^{-1}$. Hence the first row in VU^{-1} is 1, 0, ..., 0. Therefore $VU^{-1} = P$ belongs to $\mathcal{A}_{n-1}$. We obtain

$$(60) \quad V = PU \quad P \text{ in } \mathcal{A}_{n-1}, U \text{ in } c \text{ in the neighborhood of } 1$$

Let $\mu_1, \mu_2, \dots, \mu_{n^2-2n}$ be regular parameters in $\mathcal{A}_{n-1}$ in the neighborhood of P_0 . Then the elements of a neighborhood of P_0 in $\mathcal{A}_n$ can be expressed in terms of the parameters $\lambda_1, \dots, \lambda_{2n-1}$ of U and of the parameters

$\mu_1, \dots, \mu_{n^2-2n}$ of P in (60). We form the Jacobian matrix

$$\mathcal{M} \left| \frac{\partial (a'_{11}, a'_{12}, a''_{12}, \dots, a'_{1n}, a''_{1n}, \dots, a'_{k\lambda}, a''_{k\lambda}, \dots)}{\partial (\lambda_1, \dots, \lambda_{2n-1}, \mu_1, \dots, \mu_{n^2-2n})} \right|_{\lambda_k=0, \mu_k=\mu}$$

where $\mu_k^{(0)}$ are the parameters belonging to $\mathcal{P}_0$. For the differentiation with regard to λ_k we may put $\mu_k = \mu_k^{(0)}$, $P = P_0 < \mathcal{O}_{n-1}$. The first row in V is the same then as in U . We obtain that in the first $2n-1$ rows and columns in $\mathcal{M}$ we have a matrix of determinant Δ , (59). During the differentiation with regard to μ_k we may put $\lambda_k = 0$, $U = 1$, $V = P$. We have zeros in the first $2n-1$ rows and the last n^2-2n columns since a_{1k} is constant in $\mathcal{P}$. There is, however, a minor different from 0 in these n^2-2n columns because the μ 's are regular parameters of $\mathcal{P}$. Hence $\mathcal{M}$ is of rank $2n-1 + n^2-2n = n^2-1$. The parameters $\lambda_1, \dots, \lambda_{2n-1}, \mu_1, \mu_2, \dots$ are, consequently, regular parameters of $\mathcal{O}_n$ in a neighborhood of $\mathcal{P}_0$.

Let ρ and σ be two cells on $\mathcal{O}_{n-1}$ of complementary dimensionalities which have an intersection of index +1 in $\mathcal{P}_0$. We may assume that the parameters μ_k are chosen in such a manner that $\mu_1, \mu_2, \dots, \mu_\chi$ are positively oriented coordinates in ρ , whereas $\mu_{\chi+1} = \text{const.}$, $\mu_{\chi+2} = \text{const.}, \dots, \mu_{n^2-2n} = \text{const.}$ in σ . In σ , however, we have $\mu_1 = \text{const.}$, $\mu_2 = \text{const.}, \dots, \mu_\chi = \text{const.}$. Here $\mu_{\chi+1}, \mu_{\chi+2}, \dots$ are positively oriented coordinates. We denote by $\rho \cdot \tilde{c}$ the product of all elements of ρ by elements of a neighborhood $\tilde{c}$ of 1 on c . In $\mathcal{O}_n$, $\rho \cdot \tilde{c}$ is given by $\mu_{\chi+1} = \text{const.}, \dots, \mu_{n^2-2n} = \text{const.}$ and σ is given by $\lambda_1 = \text{const.}, \lambda_2 = \text{const.}, \dots, \lambda_{2n-1} = \text{const.}$, $\mu_1 = \text{const.}, \dots, \mu_\chi = \text{const.}$. In $\rho \cdot \tilde{c}$ we may consider $\lambda_1, \dots, \lambda_{2n-1}, \mu_1, \dots, \mu_\chi$ as coordinates. Thereby, an orientation of $\rho \cdot \tilde{c}$ is fixed. We see that $\rho \cdot \tilde{c}$ and σ have 1 as index of intersection.

Let β and γ be two cycles on $\mathcal{O}_{n-1}$ of complementary dimensionalities. We denote by $\beta \cdot c$ the cycle of $\mathcal{O}_n$ which is obtained by multiplying

every element of β by every element of c . If P_0 is a point of intersection of $\beta \cdot c$ and γ , we have $P_0 U = P$, where P and P_0 belong to $\mathcal{U}_{n-1}$ and U belongs to c . Then U must belong to the intersection of $\mathcal{U}_{n-1}$ and c . Hence $U = 1$, $P = P_0$. We have to consider only points in a neighborhood $\tilde{c}$ of the identity on c in order to investigate the intersection of βc and γ . The group manifolds are orientable. The orientation can be defined by means of left-hand translation of an indicatrix arbitrarily defined in the identity. Thus $\mathcal{U}_{n-1}$ and $\mathcal{U}_n$ are orientable. An indicatrix in $\mathcal{U}_{n-1}$ and an indicatrix in c together define an indicatrix in $\mathcal{U}_n$ in the manner described. If it should be possible to invert the indicatrix in c by shifting it in c , the same would be true for the indicatrix in $\mathcal{U}_n$. Hence c is orientable.

Let ν be the index of intersections of the two cycles β and γ in $\mathcal{U}_{n-1}$. The above consideration shows that βc and γ have also the index of intersection ν . In particular we see in the case that β is a point and $\gamma = \mathcal{U}_{n-1}$ that c and $\mathcal{U}_{n-1}$ have the index of intersection 1.

Suppose that

$$(61) \quad a_1, a_2, \dots, a_m$$

form a basis of the cycles of all possible dimensionalities in $\mathcal{U}_{n-1}$ in the sense of homology with division. We are going to prove that

$$(62) \quad a_1, a_2, \dots, a_m, a_1^c, a_2^c, \dots, a_m^c$$

form a corresponding basis in $\mathcal{U}_n$.

If the cycles (62) should not be linearly independent, we should have a homology

$$(63) \quad a^{\rho+2n-1} + b^{\rho} c \sim 0 \quad \text{on } \mathcal{U}_n$$

where $a^{\rho+2n-1}$ and b^{ρ} are cycles on $\mathcal{U}_{n-1}$, the number of dimensions being given by the upper indices. We assume first that $b^{\rho} \neq 0$ on $\mathcal{U}_{n-1}$. In this case there exist cycles b^* on $\mathcal{U}_{n-1}$ of $(n-1)^2 - 1 - \rho$ dimensions such that the

index of intersection ν of b^p and b^* is different from 0 on $\mathcal{U}_{n-1}$. The index of intersection of b^p and b^* on $\mathcal{U}_n$ has the same value ν . It follows from (63) that a^{p+2n-1} and b^* have an index of intersection different from 0. Both a^{p+2n-1} and b^* are on $\mathcal{U}_{n-1}$. We form $\mathcal{U}_{n-1}Q$ and change here Q continuously from 1 to an element Q_0 not in $\mathcal{U}_{n-1}$. The cosets $\mathcal{U}_{n-1}$ and $\mathcal{U}_{n-1}Q_0$ have no common element. This deformation of $\mathcal{U}_{n-1}$ transforms a^{p+2n-1} in a cycle $\tilde{a}$ which is on $\mathcal{U}_{n-1}Q_0$. Then $\tilde{a}$ has no common element with $\mathcal{U}_{n-1}$ or b^* . The index of intersection which is invariant by deformation does not vanish. This gives a contradiction. Hence $b^p \sim 0$ on $\mathcal{U}_{n-1}$. From (63) follows $a^{p+2n-1} \sim 0$ on $\mathcal{U}_n$. If a^{p+2n-1} should be $\neq 0$ on $\mathcal{U}_{n-1}$, we could determine an a^* on $\mathcal{U}_{n-1}$ such that the index of intersection of a^{p+2n-1} and a^* on $\mathcal{U}_{n-1}$ would be different from 0. The same would be true for a and a^*c on $\mathcal{U}_n$. Hence $a^{p+2n-1} \neq 0$ on $\mathcal{U}_n$.

The relation (63) implies therefore $a^{p+2n-1} \sim 0$, $b^p \sim 0$ on $\mathcal{U}_{n-1}$. The cycles (62) are linearly independent. The sum of all Betti numbers of $\mathcal{U}_n$ has at least the value $2m$ where m is the sum of all Betti numbers of $\mathcal{U}_{n-1}$. If we use Cartan's result (sec. 1) that the sum of all Betti numbers of $\mathcal{U}_N$ has the value 2^{N-1} , we have $m = 2^{n-2}$ and the number of cycles in (62) is equal to the sum $2m = 2^{n-1}$ of all Betti numbers of $\mathcal{U}_n$. The cycles (62) form a basis.

It is possible to prove this without using Cartan's result. Let $a_1^p, a_2^p, \dots, a_k^p$ be the cycles of dimension p among (62). If b is a cycle of dimension $\sigma = (n-1)^2 + 1 - p$ on $\mathcal{U}_{n-1}$ such that the index of intersection of b with every a_k^p vanishes, we have $b \sim 0$. Let

$$(64) \quad b = x_1 a_1^\sigma + x_2 a_2^\sigma + \dots + x_k a_k^\sigma$$

where $a_1^\sigma, \dots, a_k^\sigma$ are the cycles of dimension σ among the cycles (62). The number k is the same as for the dimension p . The indices of intersection $i_1, i_2, \dots, i_k$ of b with $a_1^p, a_2^p, \dots, a_k^p$ give k linear homogeneous functions

of $x_1, x_2, \dots, x_k$. We have seen that they vanish simultaneously only if $b \sim 0$, i.e. $x_1 = x_2 = \dots = x_k$. Hence the determinant of the linear functions is different from 0.

Let z be a cycle of $\sigma + 2n-1$ dimensions. Let $i_1^*, i_2^*, \dots, i_k^*$ be the indices of intersection of z with $a_1^{\mathcal{P}}, a_2^{\mathcal{P}}, \dots, a_k^{\mathcal{P}}$ on $\mathcal{Q}_n$. We can determine x_K in (64) such that $i_V = i_V^*$ ($V = 1, 2, \dots, k$) the determinant of these linear equations being $\neq 0$. The x_K then are rational. If we multiply them by their ~~greatest~~ ^{smallest} common denominator we obtain a cycle b with integral coefficients x_K such that $i_K = ti_K^*$. The index of intersection of bc with i_K on $\mathcal{Q}_n$ has the same value i_K as we have seen before. It follows that the index of intersection of $z' = tz - bc$ with $a_K^{\mathcal{P}}$ vanishes for $K = 1, 2, \dots, k$.

Pontrjagin's duality theorem * shows then that z' is homologous to a

* Math. Ann. 105, p. 190, equation (3).

cycle z'' on $\mathcal{Q}_n - \mathcal{Q}_{n-1}$.

We consider the function $f = a'_{11}$ on $\mathcal{Q}_n$. A simple consideration shows that one may use a'_{11} as one of the parameters of $\mathcal{Q}_n$ except if $a_{11} = 1$ or $a_{11} = -1$, i.e. except in the points of $\mathcal{Q}_{n-1}$ and of a certain coset $\mathcal{Q}_{n-1}Q = \mathcal{Q}_{n-1}^*$. In the other points of $\mathcal{Q}_n$, the total differential df is therefore different from 0.

Let $y_1, y_2, \dots$ be the parameters for the moment. We construct the trajectories * with respect to f . Those are the solutions of the equations

* Cf. Morse, Trans. Am. Math. Soc., 27, p. 345, for the properties of trajectories.

$$\frac{dy_1}{d\tau} = \frac{\partial f}{\partial y_1}, \frac{dy_2}{d\tau} = \frac{\partial f}{\partial y_2}, \dots$$

The parameter τ can be chosen such that we have $\tau = f(y_1, y_2, \dots)$ in every point. There is one and only one trajectory through every point in which df does not vanish. In our case these are the points with $a_{11} \neq \pm 1$ and we may use a_{11} as parameter. If we continue the trajectory in the direction of decreasing a_{11} we shall come into an arbitrarily small neighborhood N of $\mathcal{O}_{n-1}^*$.

We deform z'' along these trajectories. Since z'' has no points in common with $\mathcal{O}_{n-1}$ we can deform z'' so that it will belong to N . Then we can deform z'' also in such a way that it belongs to $\mathcal{O}_{n-1}^* = \mathcal{O}_{n-1}Q$. Now we change Q continuously from Q to 1. Thereby $\mathcal{O}_{n-1}^*$ is deformed into $\mathcal{O}_{n-1}$.

We obtain finally a cycle $\tilde{z}$ on $\mathcal{O}_{n-1}$,

$$\tilde{z} \sim z'' \sim z' = tz - bc \quad \text{on } \mathcal{O}_n$$

We can represent z here by the basis (61). According to (64) we have

then a formula

$$tz \sim \sum u_v a_v^p + \sum x_v a_v^c c \quad \text{on } \mathcal{O}_n$$

where $t \neq 0$ is an integer. This proves that (63) is a basis in the sense of homologies with division. It follows immediately from (63) that the Poincaré' polynomial of $\mathcal{O}_n$ is obtained from the Poincaré' polynomial of $\mathcal{O}_{n-1}$ by multiplication by $1 + t^{2n-1}$, $2n-1$ being the dimensionality of c . Hence

$$(\mathcal{O}_n) \quad 1 + B_1 t + \dots + B_{n-2} t^{n^2-2} + t^{n^2-1} = (1+t^3)(1+t^5) \dots (1+t^{2n-1})$$

The method yields an actual construction of a basis for the cycles.

4. The fundamental region of finite groups generated by reflections *

* E. Cartan, Annali di Mat. (4) 4, p. 209, Chap. I; and 5, p. 253 (1928).
Cf. also H.S.M. Coxeter, Annals of Math 35 (1934), p. 588.

Every semi-simple infinitesimal group is associated with a finite group (S) of linear transformations of a Euclidean vector space where (S) can be generated by reflections. We shall investigate the properties of such groups (S).

In sec. 6, the actual construction of all the groups (S) will be given. This yields the determination of all semi-simple continuous groups.

Let $\varphi^1, \varphi^2, \dots, \varphi^\lambda$ be the coordinates in the λ -dimensional vector space. Let $\sum g_{ik} \varphi^i \varphi^k$ be the metric fundamental form. The covariant components φ_i of a vector are connected with the contravariant components φ^i by the equations

$$\varphi_i = \sum_k g_{ik} \varphi^k, \quad g_{ik} = g_{ki}$$

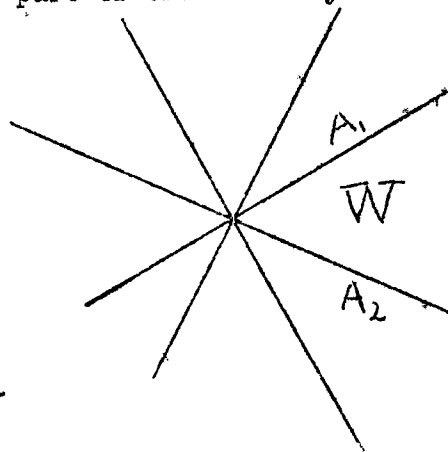
We set $\varphi^i = \sum_k g^{ik} \varphi_k$ and have (as usual)

$$(\varphi, \psi) = \sum_{ik} g_{ik} \varphi^i \psi^k = \sum \varphi^i \psi_i = \sum \varphi_i \psi^i = \sum g^{ik} \varphi_i \psi_k$$

Let (S) be a finite group of linear transformation which can be generated by reflections. If t is a reflection in a hyperplane E , and s an element of (S), then sts^{-1} is the reflection in sE .

By assumption, (S) is generated by the reflections $s_1, s_2, \dots, s_k$ in certain hyperplanes $E_1, E_2, \dots, E_k$. The preceding remark shows that (S) contains the reflection in sE_k too. We denote the set of these hyperplanes by $E_1, E_2, \dots, E_n$ and the corresponding reflections by $s_1, s_2, \dots, s_n$, ($n \geq k$).

The whole space is divided by the hyperplanes E_ν , ($\nu = 1, 2, \dots, n$), in a finite number of "chambers". Let W be such a chamber and $A_1, A_2, \dots, A_m$ be the hyperplanes E_ν which contain an $(\lambda-1)$ -dimensional part of the boundary of W . We call the A_μ the "walls" of W . Two chambers W_1 and W_2 are "neighboring" if they have an $(\lambda-1)$ -dimensional part of their boundary in common. If S belongs to (S), SW is a chamber and its walls are $SA_1, SA_2, \dots, SA_m$. Let $r_1, r_2, \dots, r_m$ be the reflections in $A_1, A_2, \dots, A_m$. The reflection in SA_μ is $Sr_\mu S^{-1}$.



Every chamber W^* can be joined to W by a chain of chambers

$$W_0 = W, W_1, W_2, \dots, W_k = W^*$$

such that $W_{\nu-1}$ and W_ν are neighboring. Suppose that we know that

$W_{\nu-1} = S_{\nu-1} W$, $S_{\nu-1}$ in (S) . If $W_{\nu-1}$ and W_ν have an $(\lambda-1)$ -dimensional part of the wall $S_{\nu-1} A_{\sigma_\nu}$ as common border, the reflection $S_{\nu-1} r_{\sigma_\nu} S_{\nu-1}^{-1}$ transforms $W_{\nu-1}$ in W_ν . We have

$$(1) \quad W_\nu = (S_{\nu-1} r_{\sigma_\nu} S_{\nu-1}^{-1}) W_{\nu-1} = S_{\nu-1} r_{\sigma_\nu} S_{\nu-1}^{-1} S_{\nu-1} W = S_{\nu-1} r_{\sigma_\nu} W.$$

Hence we may set $S_\nu = S_{\nu-1} r_{\sigma_\nu}$. We have

$$(2) \quad W_0 = W, W_1 = r_{\sigma_1} W, W_2 = r_{\sigma_1} r_{\sigma_2} W, \dots, W^* = W_k = r_{\sigma_1} r_{\sigma_2} \dots r_{\sigma_k} W$$

Conversely, if

$$(3) \quad R = r_{\sigma_1} r_{\sigma_2} \dots r_{\sigma_k}$$

is any product of r_{σ} , then $W_{\nu-1}$ and W_ν in (2) are neighboring and an $(\lambda-1)$ -dimensional part of $r_{\sigma_1} r_{\sigma_2} \dots r_{\sigma_{\nu-1}} A_{\sigma_\nu}$ belongs to the boundary of both.

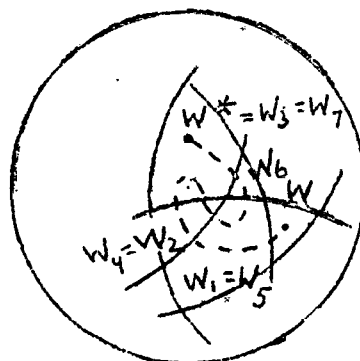
Every E_μ is the wall of a chamber, say of W^* . But according to (2)

and (3) the walls of W^* are $RA_1, RA_2, \dots, RA_m$. If $E_\mu = RA_\nu$, the reflection s_μ in E_μ is given by $R r_\nu R^{-1}$. Consequently, s_μ can be written as the product of reflections $r_1, r_2, \dots, r_m$. The s_μ generate the whole group (S) . Hence

Theorem. Every element of (S) can be written as a product (3) of factors $r_1, r_2, \dots, r_m$.

We construct the unit sphere about the origin. The division of space into chambers yields a division of the sphere into $(\lambda-1)$ -cells. We denote these cells by the same symbols as the corresponding chambers. We associate a path on the sphere with the chain of chambers (2). We start from a point of W_0 , cross the common border to W_1 and pass from W_1 to W_2 and so on until we come to W_k .

Conversely every such path Π which avoids the $(\lambda-2)$ -cells on the sphere belongs to a chain (2) of chambers and to a product (3).



We assume now that $W^* = W$, $RW = W$. Then Π may be chosen as a closed path. We can shrink Π on the sphere to its initial point in W . We may effect this in such a manner that Π always avoids the $(\chi-4)$ -cells on our $(\chi-1)$ -sphere. We ask how the associated element R will change during the deformation.

There are two cases which we have to consider:

1) Π passes from a cell W_{V-1} to a neighboring cell W_V and hence back to W_{V-1} . It is deformed in such a manner that it does not leave

W_V . We have to replace $\{W_{V-1}, W_V, W_{V-1}\}$ in the chain by $\{W_{V-1}\}$. If $W_{V-1} = S_{V-1}W$, and

$S_V = S_{V-1} r_{\sigma_V}$ it follows from (1) and $r_{\sigma}^{-1} = R_{\sigma}$ that $W_{V+1} = W_{V-1} = S_V = S_{V-1} r_{\sigma_V} S_{V-1}^{-1} \cdot W_V = (S_V r_{\sigma_V} S_V^{-1}) W_V$. Hence $S_{V+1} = S_{V-1} r_{\sigma_V} r_{\sigma_V} = S_{V-1}$

Before the deformation, R has the form

$$R = \tau_{\sigma_1} \tau_{\sigma_2} \dots \tau_{\sigma_{V-1}} (\tau_{\sigma_V} r_{\sigma_V} \tau_{\sigma_{V+2}} \dots \tau_{\sigma_R})$$

After the deformation the two factors r_{σ_V} in the parenthesis will not occur.

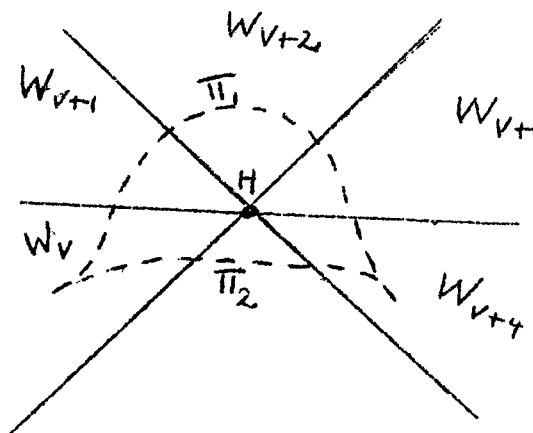
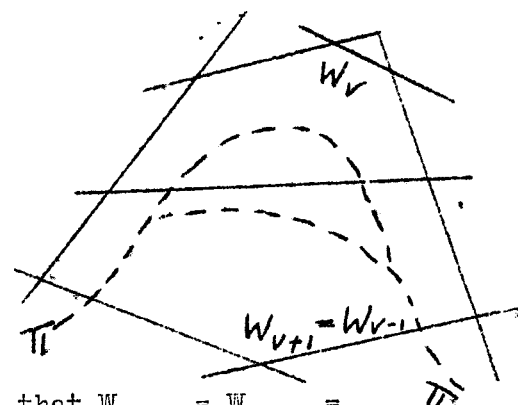
$$R = \tau_{\sigma_1} \tau_{\sigma_2} \dots \tau_{\sigma_{V-1}} \tau_{\sigma_{V+2}} \dots \tau_{\sigma_R}$$

We get that R does not change its value since $r_{\sigma_V}^2 = 1$.

The inverse deformation has the effect that two equal factors, $r_{\sigma_V} r_{\sigma_V}$ are inserted at some place in R .

2) Π during the deformation passes an $(\chi-3)$ -cell H on the sphere. The planes

E_{μ} which pass through H have an $(\chi-2)$ -dimensional subspace $\tilde{H}$ of the original space in common. Let δ be the 2-dimensional plane orthogonal to $\tilde{H}$ and P its point of intersection with $\tilde{H}$. The E_{μ} through H intersect δ in straight lines and the chambers which contain $\tilde{H}$ on their boundary are represented by sectors



in the figure. We may assume without restriction that π after surrounding H returns to the same chamber (cf. the second figure. We can effect this by additional deformations of type 1).

Let $W_v, W_{v+1}, W_{v+2}, \dots, W_{v+2\tau-1}$ be the chambers which have H on their boundary; their number is even. Let u and v be the two reflections which turn W_v into W_{v+1} and $W_{v+2\tau-1}$ respectively. The situation is completely described by the situation in the 2-dimensional cut with δ . The product uv is a rotation about P through the angle $2\pi/\tau$. We have

$$(uv)^{\tau} = 1$$

The reflection which transforms $W_{v+\mu}$ into $W_{v+\mu+1}$ is given by $(uv)^{\mu} u$. We

have to determine $r_{\sigma_{v+1}}, r_{\sigma_{v+2}}, \dots$ successively, so that if we set

$S_{v+\mu} = S_v r_{\sigma_{v+1}} r_{\sigma_{v+2}} \dots r_{\sigma_{v+\mu}}$ the reflection transforming $W_{v+\mu}$ into $W_{v+\mu+1}$ is given by $S_{v+\mu}^{-1} r_{\sigma_{v+\mu+1}} S_{v+\mu}$. Hence

$$S_v r_{\sigma_{v+1}} \dots r_{\sigma_{v+\mu}} r_{\sigma_{v+\mu+1}} r_{\sigma_{v+\mu}} \dots r_{\sigma_{v+1}} S_v^{-1} = (uv)^{\mu} u$$

It follows easily that

$$r_{\sigma_{v+1}} = S_v^{-1} u S_v, r_{\sigma_{v+2}} = S_v^{-1} v S_v, r_{\sigma_{v+3}} = S_v^{-1} u S_v,$$

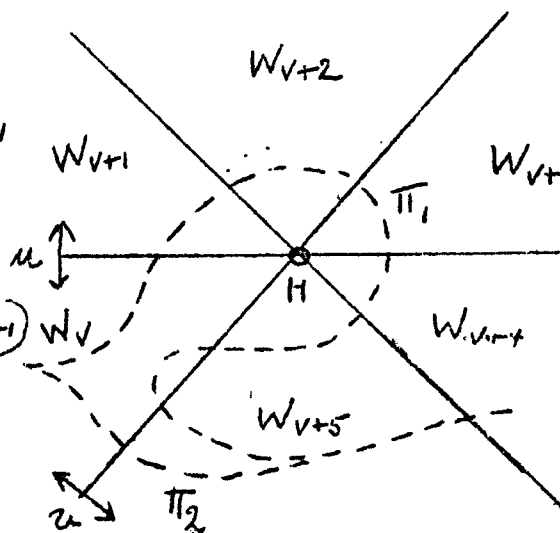
$$r_{\sigma_{v+4}} = S_v^{-1} v S_v, \dots, r_{\sigma_{v+2\tau}} = S_v^{-1} v S_v$$

After the deformation, the part of π which we consider will not leave W_v .

Hence the factors $r_{\sigma_{v+1}}, r_{\sigma_{v+2}}, \dots, r_{\sigma_{v+2\tau}}$ have to be dropped in R . Their product has, however, the value

$$(4) \quad r_{\sigma_{v+1}} r_{\sigma_{v+2}} \dots r_{\sigma_{v+2\tau}} = S_v^{-1} (uv)^{\tau} S_v = 1.$$

Thus R does not change its value. In the case of the inverse deformation, factors whose product is 1 have to be inserted in R .



We can shrink π to a point by means of deformations of these types.

R does not change its value. If π consists of one point only we have $R = 1$.

Consequently we had had $R = 1$ before. The only element of (S) which transforms W into itself is $R = 1$. Hence

Theorem. Each chamber W is a fundamental region for the group (S) .

If a product R , (3), has the value 1 we can transform R into 1 by inserting or dropping factors $r_i^2 = 1$ or (4). (S) can therefore be defined by these relations. Now $S_v^{-1} u S_v$ and $S_v^{-1} v S_v$ are reflections r_i and r_k in two walls which form an angle of π/τ in an $(\lambda-2)$ -dimensional edge of W . Therefore (4) can be written in the form $(r_i r_k)^{\tau} = 1$. We have

Theorem. The group (S) can be defined by the relations

$$r_i^2 = 1, (r_i r_k)^{\tau_{ik}} = 1$$

Here the r_i correspond to the walls of the fundamental region. We have

$\tau_{ik} = 0$ if the associated walls have no $(\lambda-2)$ -dimensional edge on the boundary of W in common. In the other case, π/τ_{ik} is the angle between these walls.

We assume now that there are λ linearly independent hyperplanes among the E_μ . In the other case the E_μ have a linear manifold K in common. By cutting the whole configuration by the linear space $\mathcal{L}$ orthogonal to K , we can reduce the whole problem to the investigation of the group (S') induced by (S) in $\mathcal{L}$ and $\mathcal{L}$ has a smaller number of dimensions than $\mathcal{L}$. In the case of a semi-simple group $\mathfrak{g}$ the assumption concerning (S) is fulfilled.

Under our assumption, there are λ independent ones among the walls $A_1, A_2, \dots, A_m$ of the chamber W . Hence $m \geq \lambda$. We are going to prove:

Theorem. The number of walls of the chamber W is equal to the number of dimensions λ (under our assumption).

Proof. Let $A^{(i)} = 0$ be the equation of the wall A_i written so that W lies in $A^{(i)} > 0$.

$$A^{(i)} = a_1^{(i)} \varphi^1 + \dots + a_\gamma^{(i)} \varphi^\gamma$$

Then $a_1^{(i)}, a_2^{(i)}, \dots, a_\gamma^{(i)}$ are the covariant components of the vector n_i orthogonal to $A^{(i)} = 0$. The reflection in A_i is given by

$$(5) \quad r_i \quad p \rightarrow p' - 2 \frac{(n_i, p)}{(n_i, n_i)} n_i$$

In particular: $n_k \rightarrow n_k - \frac{2(n_i, n_k)}{(n_i, n_i)} n_i$

$$(6) \quad A^{(k)} = 0 \rightarrow A^{(k)} - \mu_{ik} A^{(i)} = 0, \quad \mu_{ik} = \frac{2(n_i, n_k)}{(n_i, n_i)}$$

We show first that $\mu_{ik} \leq 0$ for $i \neq k$. Suppose that one of the μ_{ik}

is positive, say $\mu_{12} > 0$. We consider the region I defined by

$$(I) \quad A^{(2)} > 0, A^{(3)} > 0, \dots, A^{(m)} > 0.$$

Then in (I), $A^{(1)} - \mu_{12} A^{(2)} \neq 0$. In the other case we should have vectors in I where $A^{(1)} - \mu_{12} A^{(2)} = 0$ and hence $A^{(1)} > 0$. Such a point would belong to

W. Then the hyperplane $A^{(1)} - \mu_{12} A^{(2)} = 0$ which, according to (6), is one of the planes E_i , would have points in common with the chamber W. That is impossible.

Secondly, let II be the region defined by

$$(II) \quad A^{(1)} > 0, A^{(3)} > 0, \dots, A^{(m)} > 0$$

Again $A^{(1)} - \mu_{12} A^{(2)}$ cannot vanish in II because otherwise we should have $A^{(2)} > 0$ and the point would belong to W.

Therefore $A^{(1)} - \mu_{12} A^{(2)}$ is of constant sign in I and in II. In

both cases the sign is the same since I and II have W in common. But if

$A^{(1)} - \mu_{12} A^{(2)} > 0$ in I, then $A^{(1)} > 0$ in I, and I is identical with W. Then, however, $A^{(1)} = 0$ would not be a wall of W. Similarly, if $A^{(1)} - \mu_{12} A^{(2)} < 0$,

we should have $A^{(2)} > \frac{1}{\mu_{12}} A^{(1)} > 0$ in II, II would be identical with W, and

$A^{(2)} = 0$ would not be a wall. The assumption $\mu_{12} > 0$ leads to a contradiction;

we have $\mu_{12} \leq 0, \mu_{ik} \leq 0$ for $i \neq k$.

If we choose the coordinates properly we may assume that

$$(7) \quad A^{(1)} = \varphi^1, \quad A^{(2)} = \varphi^2, \dots, \quad A^{(\lambda)} = \varphi^\lambda$$

Suppose that $m > \lambda$ and let $A = A^{(m)} = 0$ be another wall. We write μ for μ_m , a_i for $a_i^{(m)}$, a^i for $a^{(m)i}$.

The covariant components of $\mu^{(i)}$ vanish except the i^{th} one which has the value 1, ($i = 1, 2, \dots, \lambda$). Hence

$$0 \geq \mu_{ik} = \frac{2(\mu_i \mu_k)}{(\mu_k \mu_k)} = \frac{2g^{ik}}{g^{kk}}; \quad i, k = 1, 2, \dots, \lambda; \quad i \neq k.$$

Since $\sum g^{ik} \varphi_i \varphi_k$ is positive definite, we have $g^{kk} > 0$

$$(8) \quad 0 \geq g^{ik} \quad (i \neq k)$$

Similarly we have for $i = 1, 2, \dots, \lambda$

$$0 \geq \mu_{im} = 2 \frac{(\mu_i \mu)}{(\mu_i \mu_i)} = 2 \frac{a^i}{g^{ii}}$$

$$0 \geq a^i$$

The coordinates a_i , however, are not all ≤ 0 because $\varphi^1 > 0$,

$\varphi^2 > 0, \dots, \varphi^\lambda > 0$; $a_1 \varphi^1 + \dots + a_\lambda \varphi^\lambda > 0$ hold for the points of W . We

choose the notation such that $a_1 > 0, \dots, a_p > 0, a_{p+1} \leq 0, \dots, a_\lambda \leq 0$. We

have

$$0 \geq a^i = \sum_{j=1}^p g^{ij} a_j + \sum_{j=p+1}^{\lambda} g^{ij} a_j$$

In the second sum $g^{ij} \leq 0$ for $i = 1, 2, \dots, p$ according to (8) and $a_j \leq 0$.

Hence

$$0 \geq \sum_{j=1}^p g^{ij} a_j \quad (i = 1, 2, \dots, p)$$

$$0 \geq \sum_{j=1}^p g^{ij} a_i a_j$$

This can hold only for $a_i = \dots = a_p = 0$ since $\sum g^{ij} \varphi_i \varphi_j$ is positive definite. But we had $a_1 > 0, \dots, a_\lambda > 0$. We have a contradiction and the theorem is proved.

Let $\Lambda = x_1 \lambda^1 + \dots + x_\chi \lambda^\chi$ be a linear form. We denote by $s\Lambda$ the linear form which is obtained by applying the element s in (S) to the vector ξ with the covariant components $(x_1, \dots, x_\chi)$. There is a "dominant" form among the $s\Lambda$ which is not lower than any other of the $s\Lambda$ *. Let us assume

* We write $\Lambda > 0$ if the first non-vanishing coefficient in Λ is positive. Furthermore $\Lambda_1 > \Lambda_2$ if $\Lambda_1 - \Lambda_2 > 0$ (cf. Chap. 2.)

that Λ itself is dominant.

We take $\Lambda^{(i)}$ in the form (7) again. According to (5) we have for r_i

$$\xi \rightarrow \xi - \frac{2(\alpha_i \xi)}{(\alpha_i \alpha_i)} \alpha_i = \xi - \frac{2x_i^2}{g^{ii}} \alpha_i$$

$$\Lambda \rightarrow \Lambda - \frac{2x_i^2}{g^{ii}} \varphi^i = r_i(\Lambda)$$

Since $r_i \Lambda \leq \Lambda$, we have $x_i^2 \geq 0$.

The dominant linear forms have therefore the property that the vector ξ (whose covariant components are the coefficients of Λ) lie in W or on its boundary *. Conversely, if ξ lies in W , the vectors $s\xi$, $s \neq 1$, do not lie

* The vectors of W are those whose contravariant components are positive, $\varphi^1 > 0$, $\varphi^2 > 0$, ..., $\varphi^\chi > 0$.

in W . One of the $s\Lambda$ is dominant. The only possibility is that Λ itself is dominant. If ξ lies on the boundary of W none of the vectors $s\xi$ lies in the interior of W . If $s\xi$ lies on the boundary of W , say on a k -dimensional linear part E of the boundary, then sW contains E as a k -dimensional linear part of its boundary. One proves without difficulty that it is possible to carry sW over into W by a succession of reflections in hyperplanes which contain E . Let s_0 be this product of reflections. Then $s_0 \cdot sW = W$. We have, therefore, $s_0 s = 1$. The element $s\xi$, however, is not altered by s_0 since it lies on E .

Hence $s_0 \cdot s \xi = s \xi$. From $s_0 s = 1$ follows $\xi = s \xi$. We prove now as above that Λ is dominant. If ξ lies on the boundary of W , we have reflections s in (S) such that $s \Lambda = \Lambda$.

Theorem. If Λ is dominant, i.e. $\Lambda \geq s \Lambda$ for all s in (S) , the vector ξ whose covariant components are the coefficients of Λ lies in W or on the boundary of W . In the first case we have $\Lambda > s \Lambda$ for $s \neq 1$. In the latter case there exist elements s of odd type in (S) such that $s \Lambda = \Lambda$.

It is interesting that the order of $\varphi^1, \varphi^2, \dots, \varphi^l$ is of no influence in our criterion. A part of this theorem has been used in Chap. 2, 9. We obtained ⁽²⁵⁾ there in a more complicated manner.

The determination of the characters can be effected in a simpler way, if this theorem is known.

5. Direct characterization of the finite group (S) associated with a semi-simple continuous group*

* E. Cartan, Annali di Mat. (4), 4 and 5.

Let $\mathfrak{g}$ be a semi-simple infinitesimal group * of order r . There exist

* Cf. the results of Chap. 2, secs. 3 and 4 for the following considerations.

$(r-1)$ roots of $\mathfrak{g}$ which are linear forms in l parameters $\lambda^1, \lambda^2, \dots, \lambda^l$. If the parameters in $\mathfrak{g}$ are properly chosen the coefficients of these roots are real (even integers). If α is a root, so is $-\alpha$. No other multiple $k\alpha$, $k = \pm 2, \pm 3, \dots$ is a root.

Let α be $\alpha_1 \lambda^1 + \alpha_2 \lambda^2 + \dots + \alpha_l \lambda^l$. We consider α as a vector with the covariant components $(\alpha_1, \alpha_2, \dots, \alpha_l)$. The vector diagram consisting of these roots has the property that the reflection S_α in the hyperplane orthogonal to α permutes all the roots (for every fixed α). If β

is another root, we have

$$(1) \quad S_\alpha : \beta \rightarrow \beta - \mu_{\alpha\beta} \alpha, \quad \mu_{\alpha\beta} = 2 \frac{(\alpha, \beta)}{(\alpha, \alpha)}$$

$$(2) \quad \mu_{\alpha\beta} = \text{integer}$$

$$(3) \quad \beta - \tau \alpha \quad \text{is a root, } (\tau \text{ integer between 0 and } \mu_{\alpha\beta})$$

The inner products are defined by means of a definite quadratic form

$$\sum g_{ik} \lambda^i \lambda^k$$

$$(\alpha, \beta) = \sum g_{ik} \alpha^i \beta^k = \sum g^{ik} \alpha_i \beta_k = \sum \alpha^i \beta_i = \sum \alpha_i \beta^i$$

where, of course, $\alpha_i = \sum g_{ik} \alpha^k$, $\alpha^i = \sum g^{ik} \alpha_k$.

We shall deal with the unitary restricted group U_n (Chap. 2, 6).

Here the λ^j are purely imaginary, $\lambda^j = 2\pi i \varphi^j$. We drop the factor $2\pi i$ in the roots (as we did in Chap. 2). We have

$$(4) \quad \alpha = \alpha_1 \varphi^1 + \dots + \alpha_\ell \varphi^\ell$$

The reflections S_α generate a finite group (S). The reflecting hyperplanes E are given by $\alpha = 0$. This system of hyperplanes contains with every E all the sE, s in (S). Therefore the results of the preceding section can be applied. There are ℓ independent roots.

Let W be a chamber. Its walls are given by equations $\alpha^{(1)} = 0$,

(2) $\alpha^{(\ell)} = 0$ where the $\alpha^{(r)}$ are roots. We may choose the coordinate system such that

$$(5) \quad \alpha^{(1)} = \varphi^1, \quad \alpha^{(2)} = \varphi^2, \dots, \alpha^{(\ell)} = \varphi^\ell.$$

The reflection $S_{\alpha^{(r)}} = S_r$ has the effect on a root β that an integral multiple of φ^r is subtracted from β . We can obtain each of the hyperplanes $\rho = 0$ by a succession of reflections S_r starting from a wall $\alpha^{(r)} = 0$. By this process we always get hyperplanes of the form

$$\sigma = c_1 \varphi^1 + c_2 \varphi^2 + \dots + c_\ell \varphi^\ell = 0, \quad c_k \text{ integers}$$

Hence ρ must be equal to such a σ but for a constant factor $\sigma = \gamma \cdot \rho$,

$\gamma = \text{constant}$. Then we have, according to (2) and (3)

$$\mu_{\rho\sigma} = \frac{2(\rho, \sigma)}{(\rho, \rho)} = 2\gamma = \text{integer } N$$

If N is even, the root σ is an integral multiple of the root ρ , and hence $\sigma = \pm \rho$. If $N = 2n+1$ is odd, we have $\gamma = \frac{1}{2}N = n + \frac{1}{2}$. It follows from (3) that $\sigma - \tau\rho = (n + \frac{1}{2})\rho - \rho$ is a root if τ is any integer between 0 and $2n+1$. Then, however, $\frac{1}{2}\rho$ is a root and this is impossible since $2 \cdot (\frac{1}{2}\rho) = \rho$ is a root. Hence $\sigma = \pm \rho$. We see that every root ρ has integral coefficients.

The coefficients of

$$\rho = \rho_1 \varphi^1 + \dots + \rho_\lambda \varphi^\lambda$$

are either all ≥ 0 or all ≤ 0 . In the other case we could find values such that $\varphi^1 > 0, \dots, \varphi^\lambda > 0$ and $\rho = 0$. Then W would have points in common with $\rho = 0$ and that is impossible according to the definition of W .

To every system $\varphi^1, \varphi^2, \dots, \varphi^\lambda$ belongs a "main element" $h(\varphi^1, \varphi^2, \dots, \varphi^\lambda)$ of the adjoint continuous group A_n and a main element $h^*(\varphi^1, \dots, \varphi^\lambda)$ of the true group $\mathcal{C}_n$. (cf. Chap. 2, sec. 8). We have

$$h^*(\varphi^1, \dots, \varphi^\lambda) h^*(\psi^1, \dots, \psi^\lambda) = h^*(\varphi^1 + \psi^1, \dots, \varphi^\lambda + \psi^\lambda)$$

and $h(\varphi^1, \dots, \varphi^\lambda)$ is the element corresponding to $h^*(\varphi^1, \dots, \varphi^\lambda)$ in the adjoint representation. The questions which are to be answered are

1. When is $h(\varphi^1, \dots, \varphi^\lambda) = h(\psi^1, \dots, \psi^\lambda)$? When is

$$h^*(\varphi^1, \dots, \varphi^\lambda) = h^*(\psi^1, \dots, \psi^\lambda)?$$

2. When are $h(\varphi^1, \dots, \varphi^\lambda)$ and $h(\psi^1, \dots, \psi^\lambda)$ conjugate in A_n ?

When are $h^*(\varphi^1, \dots, \varphi^\lambda)$ and $h^*(\psi^1, \dots, \psi^\lambda)$ conjugate in $\mathcal{C}_n$?

The importance of these problems can be seen from Chap. 2, secs. 8 and 9. (part of the first)

9. The first question has been answered in Chap. 2, §, by an elementary device.

We saw (Chap. 2, equation (85)) that $h(\varphi^1, \dots, \varphi^\lambda)$ is a diagonal matrix with the diagonal elements

(6) 1 (λ -times), $e(\alpha)$ for all roots α .

We could easily infer that $h(\varphi^1, \dots, \varphi^\lambda) = h(\psi^1, \dots, \psi^\lambda)$ if and only if

(7) $\alpha(\varphi^1, \dots, \varphi^\lambda) \equiv \alpha(\psi^1, \dots, \psi^\lambda) \pmod{1}$ for every root α .

Sufficient conditions for 1 and 2 have been given in Chap. 2, 9, by using the orthogonality properties. We shall use here more elementary methods. We prove first:

Theorem 1. If an element s of S transforms * the system $(\varphi^1, \dots, \varphi^\lambda)$

* There are two possibilities of interpreting this expression. 1) We consider $(\varphi^1, \dots, \varphi^\lambda)$ and $(\psi^1, \dots, \psi^\lambda)$ as contravariant components of vectors in the vector space in which S is performed. 2) Let Φ be any linear form $m_1 \varphi^1 + \dots + m_\lambda \varphi^\lambda$. We consider Φ as a covariant vector and define $S\Phi$ by applying the transformation on the m_i whereas the φ^i are constants. In particular the linear forms $\Phi = \varphi^i$ undergo a linear transformation. Both possibilities of defining $S\varphi^i$ are not identical, but one proves easily that if $S\varphi^i = \psi^i$ in the first sense, we have $S^{-1}\varphi^i = \psi^i$ with respect to the second definition. Thus it does not matter which definition is used if we are concerned with the question whether or not $(\varphi^1, \dots, \varphi^\lambda)$ and $(\psi^1, \dots, \psi^\lambda)$ are equivalent with respect to the group (S) .

into $(\psi^1, \dots, \psi^\lambda)$, then $h^*(\varphi^1, \dots, \varphi^\lambda)$ and $h^*(\psi^1, \dots, \psi^\lambda)$ are conjugate elements of $\mathcal{C}_u$

We consider first infinitesimal elements

$$(8) \quad h = \varphi^1 h_1 + \dots + \varphi^\lambda h_\lambda, \quad h' = \psi^1 h_1 + \dots + \psi^\lambda h_\lambda$$

instead of $h^*(\varphi_1, \dots, \varphi_\lambda)$, $h^*(\psi_1, \dots, \psi_\lambda)$.

We assume that a transformation S_α transforms $(\varphi^1, \dots, \varphi^\lambda)$ into $(\psi^1, \dots, \psi^\lambda)$.

To the root α corresponds an element e_α of $\mathfrak{g}$ such that

$$(9) \quad [he_\alpha] = \alpha e_\alpha, \quad [e_\alpha, e_{-\alpha}] = h_\alpha$$

(Chap. 2, (48)). The following elements belong to $\mathfrak{g}_u$:

$$u = \frac{e_\alpha - e_{-\alpha}}{\sqrt{2} \sqrt{(\alpha, \alpha)}}, \quad v = \frac{e_\alpha + e_{-\alpha}}{i\sqrt{2} \sqrt{(\alpha, \alpha)}}, \quad w = \frac{h_\alpha}{i\sqrt{(\alpha, \alpha)}}$$

From (9) follows (because of $[h_\alpha, e_\alpha] = \alpha_\alpha e_\alpha = (\alpha, \alpha) e_\alpha$)

$$[uv] = w, \quad [v, w] = u, \quad [wu] = v$$

Consequently, u, v and w are the basis of a subgroup of order 3 isomorphic to the 3-dimensional infinitesimal rotation group. We form the element $1 + \varepsilon u$, ε infinitesimal, in the neighborhood of the identity in $\mathcal{G}_\mu$. The corresponding transformation of the adjoint representation is given by

$$x \rightarrow x' = x + \varepsilon [ux]$$

$$dx = x' - x = \varepsilon [ux]$$

In particular $dv = \varepsilon w$, $dw = -\varepsilon v$. We put $\varepsilon = dt$ and integrate with regard to t (repeat $1 + \varepsilon u$ again and again). The corresponding finite transformation $\mathcal{U}$ is represented in A_μ by

$$w \rightarrow w' = w \cos t - v \sin t$$

$$v \rightarrow v' = w \sin t + v \cos t$$

We have namely, $\frac{dw'}{dt} = -v'$, $\frac{dv'}{dt} = w'$ and $w' = w$, $v' = v$ for $t = 0$.

We put $t = \pi$. Then $w \rightarrow w' = -w$. Consequently the adjoint transformation of U carries h_α over into $-h_\alpha$ or, in other words,

$$(10) \quad U h_\alpha U^{-1} = -h_\alpha$$

On the other hand, it follows from (9) that

$$[e_\alpha, h - \frac{\alpha}{(\alpha, \alpha)} h_\alpha] = 0, \quad [e_{-\alpha}, h - \frac{\alpha}{(\alpha, \alpha)} h_\alpha] = 0$$

hold. Hence $[u, h - \frac{\alpha}{(\alpha, \alpha)} h_\alpha] = 0$. The corresponding integration as above shows that

$$\mathcal{U}(h - \frac{\alpha}{(\alpha, \alpha)} h_\alpha) \mathcal{U}^{-1} = h - \frac{\alpha}{(\alpha, \alpha)} h_\alpha$$

This gives in connection with (10) the relation

$$\mathcal{U} h \mathcal{U}^{-1} = h - \frac{2\alpha}{(\alpha, \alpha)} h_\alpha = \sum \varphi^i h_i - \frac{2 \sum \alpha_i \varphi^i}{(\alpha, \alpha)} \sum \alpha^i h_i$$

$$\mathcal{U}(\varphi^1 h_1 + \dots + \varphi^l h_l) \mathcal{U}^{-1} = \varphi^1 h_1 + \dots + \varphi^l h_l$$

where $(\varphi^1, \dots, \varphi^l)$ is the image of $(\varphi^1, \dots, \varphi^l)$ under S_α Hence

$Uh \cdot U^{-1} = h''$ where U is independent of $(\varphi^1, \dots, \varphi^{\lambda})$.

Every element S of (S) is a product of elements S_{α} . We see that there exists an element T of $\mathcal{C}_u$ for every S in (S) with the following property: We have $T(\varphi^1 h_1 + \dots + \varphi^{\lambda} h_{\lambda}) T^{-1} = \psi^1 h_1 + \dots + \psi^{\lambda} h_{\lambda}$ if S transforms $(\varphi^1, \dots, \varphi^{\lambda})$ into $(\psi^1, \dots, \psi^{\lambda})$. Here T is the product of the elements U corresponding to the single factors S_{α} of S . On integrating with respect to $\varphi^1, \dots, \varphi^{\lambda}$ (repeating $1 + \varepsilon h(\varphi^1, \dots, \varphi^{\lambda})$ again and again) we obtain

$$Th^*(\varphi^1, \dots, \varphi^{\lambda})^{-1} = h^*(\psi^1, \dots, \psi^{\lambda}), \quad \text{in } \mathcal{C}_u$$

for finite main elements $h^*(\varphi^1, \dots, \varphi^{\lambda})$ and $h^*(\psi^1, \dots, \psi^{\lambda})$ which are connected by a transformation S of (S) . This proves the theorem. Of course the same theorem holds for main elements $h(\varphi^1, \dots, \varphi^{\lambda})$ and $h(\psi^1, \dots, \psi^{\lambda})$ of the adjoint representation A_u .

We construct all the hyperplanes

$$\alpha = 0, \pm 1, \pm 2, \dots \quad \text{for every root } \alpha$$

in the $\varphi^1 - \varphi^2 - \dots - \varphi^{\lambda}$ space. Let W^* be the piece of W which is cut off from W by these hyperplanes and contains the origin on its boundary. The boundary

of W^* is obtained by putting $\alpha^{(1)} = 0$ or $\alpha^{(2)} = 0, \dots,$

or $\alpha^{(2)} = 0$, or by putting some positive roots

$\beta^{(1)}, \beta^{(2)}, \dots, \beta^{(m)} = 1$. If, namely, a point with $\rho = k$, ρ a root, $k = 2$ or 3

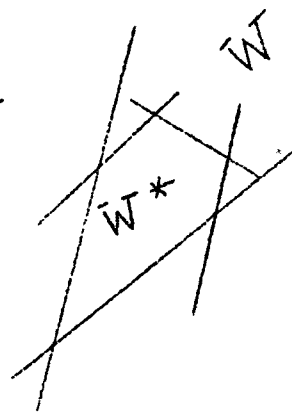
or $4, \dots$, should be on the boundary of W^* , then W^* would be split by $\rho = 1$.

If we had $\rho = -k$, ($k = 1, 2, \dots$), on the boundary of W^* , then ρ must be a negative root because $\varphi^1, \dots, \varphi^{\lambda}$ are ≥ 0 on the boundary of W . In this

case we replace the root ρ by the positive root $-\rho$.

It follows from ^(the last) Theorem on p. III that every element of $\mathcal{C}_u$ can

be represented in the form



$$U^* h^*(\varphi^1, \dots, \varphi^{\lambda}) U^{*-1}$$

where U^* belongs to $\mathcal{C}_m$ and $(\varphi^1, \dots, \varphi^{\lambda})$ is a point of W^* or of its boundary. In the latter case we called the element a singular element. The corresponding theorem is true for the adjoint representation. Every element of A_m can be represented in the form

$$U h(\varphi^1, \dots, \varphi^{\lambda}) U^{-1}$$

U in A_m , $(\varphi^1, \dots, \varphi^{\lambda})$ in W^* or on its boundary. The points of W^* represent, so to speak, the classes of conjugate elements in $\mathcal{C}_m$ and A_m . This is, however, not quite true, because it is conceivable that two elements $h^*(\varphi^1, \dots, \varphi^{\lambda})$ and $h^*(\psi^1, \dots, \psi^{\lambda})$, $((\varphi^1, \dots, \varphi^{\lambda})$ and $(\psi^1, \dots, \psi^{\lambda})$ in W^* are conjugate themselves.

We investigate this first for the group A_m . Let $(\varphi_0^1, \dots, \varphi_0^{\lambda})$ be a point of W^* such that the roots α have values distinct (mod 1) in this point. Suppose that

$$(11) \quad U h(\varphi_0^1, \dots, \varphi_0^{\lambda}) U^{-1} = h(\psi_0^1, \dots, \psi_0^{\lambda}),$$

U in A_m , $(\psi_0^1, \dots, \psi_0^{\lambda})$ in W^* or on its boundary.

Let $\rho_1, \rho_2, \dots, \rho_{r-\lambda}$ be the $r-\lambda$ roots α . $h(\varphi_0^1, \dots, \varphi_0^{\lambda})$ is a diagonal matrix. The main diagonal consists of

$$\underbrace{1, 1, \dots, 1}_{\ell \text{ times}}, e(\rho_1), e(\rho_2), \dots, e(\rho_{r-\lambda})$$

according to (6) where $\rho_1, \rho_2, \dots, \rho_{r-\lambda}$ are to be taken in $(\varphi_0^1, \dots, \varphi_0^{\lambda})$.

The corresponding facts hold for $h(\psi_0^1, \dots, \psi_0^{\lambda})$. The matrix equation (11) shows that the $e(\rho_v(\psi_0^1, \dots, \psi_0^{\lambda}))$ form a permutation of the

$e(\rho_v(\varphi_0^1, \dots, \varphi_0^{\lambda}))$, say

$$(12) \quad e(\rho_v(\psi_0^1, \dots, \psi_0^{\lambda})) = e(\rho_v(\varphi_0^1, \dots, \varphi_0^{\lambda}))$$

Moreover U has the form

$$(13) \quad U = \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix}$$

where U_1 is of degree λ and there is only one element different from 0 in the ν^{th} row of U_2 , namely $\binom{m}{\nu}$ the ν' -th column.

The elements on the left side of (12) are different from 1 since the elements of the right-hand side are different from 1. Consequently $(\psi^1_0, \dots, \psi^\lambda_0)$ lies in W^* , not on the boundary.

We now form

$$Uh(\varphi^1, \dots, \varphi^\lambda)U^{-1} = h'$$

for any $(\varphi^1, \dots, \varphi^\lambda)$ in W^* . It follows from the form of U and of $h(\varphi^1, \dots, \varphi^\lambda)$ that h' is a diagonal matrix. The main diagonal consists of

$$(14) \quad 1, \dots, 1, e(\rho_1(\varphi^1, \dots, \varphi^\lambda)), e(\rho_2(\varphi^1, \dots, \varphi^\lambda)), \dots$$

The element h' is commutative with all the elements of the maximum Abelian subgroup $\mathcal{f}$. So is the variation dh' of h' corresponding to an infinitesimal variation $d\varphi^1, \dots, d\varphi^\lambda$ of $\varphi^1, \dots, \varphi^\lambda$. Then the infinitesimal element $\delta h' = dh' \cdot h'^{-1}$ of $\mathcal{g}$ is commutative with the elements of $\mathcal{f}$ and, consequently, $\delta h'$ belongs to $\mathcal{f}$.

We start now from $\varphi^1 = \varphi^1_0, \dots, \varphi^\lambda = \varphi^\lambda_0$, and vary

$\varphi^1, \dots, \varphi^\lambda$ continuously. We see that h' always is a main element $h(\psi^1, \dots, \psi^\lambda)$ of A_n

$$(15) \quad Uh(\varphi^1, \dots, \varphi^\lambda)U^{-1} = h(\psi^1, \dots, \psi^\lambda)$$

$$(16) \quad e(\rho_\nu(\psi^1, \dots, \psi^\lambda)) = e(\rho_{\nu'}(\varphi^1, \dots, \varphi^\lambda))$$

As long as $(\varphi^1, \dots, \varphi^\lambda)$ stays inside of W^* , the right-hand side of (16) is different from 1 and $\psi^1, \dots, \psi^\lambda$ does not leave W^* . Let $\rho_{\nu'}$ be a positive root. The value of it in W^* lies between 0 and 1. If ρ_ν is positive also, it follows from (16) that

$$(17a) \quad \rho_\nu(\psi^1, \dots, \psi^\lambda) = \rho_{\nu'}(\varphi^1, \dots, \varphi^\lambda), \quad (\rho_\nu > 0, \rho_{\nu'} > 0)$$

If ρ_ν is negative, $-\rho_\nu$ lies between 0 and 1 in W^* and from (16) it follows that

$$(17b) \quad \rho_v(\psi^1, \dots, \psi^\ell) = \rho_v(\varphi^1, \dots, \varphi^\ell) - 1, \quad (\rho_v < 0, \rho_v > 0)$$

The first ℓ of the roots are $\varphi^1, \varphi^2, \dots, \varphi^\ell$ according to (5). We see that $\psi^1, \psi^2, \dots, \psi^\ell$ are linear functions of $\varphi^1, \dots, \varphi^\ell$ and vice versa. If $(\varphi^1, \dots, \varphi^\ell)$ moves in W^* , so does $(\psi^1, \dots, \psi^\ell)$. The point $(\varphi^1, \dots, \varphi^\ell)$ lies on the boundary of W^* when at least one ρ_v has an integral value. Then $(\psi^1, \dots, \psi^\ell)$ lies on the boundary too. The equations (17a) and (17b) give a one-to-one linear mapping of W^* onto itself.

We consider the case where the origin corresponds to itself. No equation (17b) can occur and the permutation $\rho_{v'} \rightarrow \rho_v$ associates positive roots with positive roots.

The only roots ρ for which $\rho = 0$ forms a wall of W^* are the first ℓ roots, (5), $\varphi^1, \varphi^2, \dots, \varphi^\ell$, and their negatives. Hence our transformation (17a) must effect a permutation of $\varphi^1, \varphi^2, \dots, \varphi^\ell$

$$(18) \quad \psi^v = \varphi^{v'}$$

We want to show now that this permutation is the identity $\psi^v = \varphi^v$.

Suppose that we have U , (11), such that (18) is a permutation $\pi \neq 1$. Choose the numeration of the roots such that the transformation (17a) permutes the roots $\rho_{k+1}, \rho_{k+2}, \dots, \rho_{k+n}$ cyclically $(k+1)' = k+2, (k+2)' = (k+3), \dots, (k+n)' = k+1$. Then U_π will break up into constituents corresponding to the different sequences $\rho_{k+1}, \dots, \rho_{k+n}$, and such a constituent will be of the form

$$\left\| \begin{array}{cccccc} 0 & u_1 & 0 & \dots & 0 \\ 0 & 0 & u_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{n-1} \\ u_n & 0 & 0 & \dots & 0 \end{array} \right\|.$$

The characteristic roots of this matrix are the different values of $\sqrt[n]{u_1 u_2 \dots u_n}$.

We may replace U by hU , h main element, without disturbing (11). Then u_v has to be replaced by $e(\rho_{k+v})u_v$ and $u_1 \cdot u_2 \cdot \dots \cdot u_n$ by $e(\rho_{k+1} + \rho_{k+2} + \dots + \rho_{k+n}) \cdot u_1 \cdot u_2 \cdot \dots \cdot u_n$. Here the factor is different from 1 if h is properly chosen. In the other case we should have

in W^* . For $\varphi^1 = \varphi^2 = \dots = \varphi^\lambda = 0$ we get $N = 0$. Now

$\rho_{k+1}, \rho_{k+2}, \dots, \rho_{k+n}$ are all positive or all negative. Hence all the coefficients occurring are either all positive or all negative. Hence

$$\rho_{k+1} + \dots + \rho_{k+n} \neq 0.$$

We can choose the element h such that $e^{(\rho_{k+1} + \dots + \rho_{k+n})u_1 \dots u_{n-1}}$

and we can do that simultaneously for all the sequences $\rho_{k+1}, \dots, \rho_{k+n}$. Then 1 is not a characteristic root of U_2 .

Every element U of A_μ is conjugate to a main element. Hence U has λ characteristic roots 1. In our case U_1 in (13) must have λ roots 1. Moreover, U_1 of degree λ is similar to a diagonal matrix. Hence, $U_1 = 1$ is the unit matrix of degree λ . According to (13) and the definition of the adjoint group we have

$$Uh_1U^{-1} = h_1, \dots, Uh_\lambda U^{-1} = h_\lambda$$

It follows that $UhU^{-1} = h$ for all elements of $\mathcal{J}$ and hence $UhU^{-1} = h$ for all main elements h . Then (15) gives $\psi^\lambda = \varphi^\lambda$, q.e.d.

We consider now the general case where the mapping $(\varphi^1, \dots, \varphi^\lambda) \rightarrow (\psi^1, \dots, \psi^\lambda)$ (17a) and (17b), does not leave the origin θ invariant. Then θ is mapped on a point θ' for which all the roots have integral values 0, +1 or -1. We know however two classes of linear transformations of the $\varphi^1, \dots, \varphi^\lambda$ -space which transform every $(\varphi^1, \dots, \varphi^\lambda)$ into a point $(\psi^1, \dots, \psi^\lambda)$ such that $h(\varphi^1, \dots, \varphi^\lambda)$ and $h(\psi^1, \dots, \psi^\lambda)$ are conjugate in A_μ . I. the transformations S of (S); II. the transformations

$$\psi^\lambda = \varphi^\lambda + m^\lambda \quad m^\lambda, \text{ integer}$$

Here $h(\varphi^1, \dots, \varphi^\lambda) = h(\psi^1, \dots, \psi^\lambda)$. We denote by $\tilde{S}$ the group generated by the transformations I and II. We can find a transformation II which transforms θ' into θ . Then W^* is transformed into a part of a chamber W_1 . By

a transformation I we can carry W_1 over into W . We obtain a transformation $\tilde{S}$ which transforms W^* into itself and $\mathcal{O}'$ into $\mathcal{O}$. If we perform $\tilde{S}$ after

the transformation given by (17a) and (17b), we obtain a transformation

$(\varphi^1, \dots, \varphi^{\lambda}) \rightarrow (\chi^1, \dots, \chi^{\lambda})$ of W^* into itself, which leaves $\mathcal{O}$ invariant and which has the property: $\tilde{U}h(\varphi^1, \dots, \varphi^{\lambda})\tilde{U}^{-1} = h(\chi^1, \dots, \chi^{\lambda})$

with a fixed $\tilde{U}$ in A_{μ} . Our consideration above yields $\varphi^{\nu} = \chi^{\nu}$. It follows that $\tilde{S}^{-1}$ transforms $(\varphi^1, \dots, \varphi^{\lambda})$ into $(\psi^1, \dots, \psi^{\lambda})$. In particular, $\tilde{S}^{-1}$ transforms $(\varphi_0^1, \dots, \varphi_0^{\lambda})$ into $(\psi_0^1, \dots, \psi_0^{\lambda})$.

Until now we have assumed that the roots have distinct values (mod 1)

in $\varphi_0^1, \dots, \varphi_0^{\lambda}$. Suppose now that this is not true. Let

$Uh(\varphi_0^1, \dots, \varphi_0^{\lambda})U^{-1} = h(\psi_0^1, \dots, \psi_0^{\lambda})$ as in (11). The method on pp. 108

-109 permits us to represent the elements $h(\psi^1, \dots, \psi^{\lambda})$, $((\psi^1, \dots, \psi^{\lambda})$

in the neighborhood of $(\psi_0^1, \dots, \psi_0^{\lambda})$ in the form $U_1h(\varphi^1, \dots, \varphi^{\lambda})U_1^{-1}$

where U_1 is in the neighborhood of U and the point $(\varphi^1, \dots, \varphi^{\lambda})$ lies in the neighborhood of $(\varphi_0^1, \dots, \varphi_0^{\lambda})$. In particular, we may choose $(\psi^1, \dots, \psi^{\lambda})$

such that the values of the roots are distinct (mod 1). Then the same is true

for $(\varphi^1, \dots, \varphi^{\lambda})$. We obtain therefore that $(\varphi^1, \dots, \varphi^{\lambda})$ and

$(\psi^1, \dots, \psi^{\lambda})$ are connected by a transformation of $(\tilde{S})$. If $(\psi^1, \dots, \psi^{\lambda})$

tends to $(\psi_0^1, \dots, \psi_0^{\lambda})$ we see that the same is true for $(\varphi_0^1, \dots, \varphi_0^{\lambda})$ and

$(\varphi_0^1, \dots, \varphi_0^{\lambda})$.

Theorem 2. If $(\varphi^1, \dots, \varphi^{\lambda})$ and $(\psi^1, \dots, \psi^{\lambda})$ are two points of W^* , then $h(\varphi^1, \dots, \varphi^{\lambda})$ and $h(\psi^1, \dots, \psi^{\lambda})$ are conjugate if and only if $(\varphi^1, \dots, \varphi^{\lambda})$ and $(\psi^1, \dots, \psi^{\lambda})$ are connected by a transformation of $(\tilde{S})$. (For the definition of $(\tilde{S})$ cf. Theorem 4.)

If $(\varphi_0^1, \dots, \varphi_0^{\lambda})$ is any point, we may transform it into a point of W^* by a transformation of $(\tilde{S})$. If, namely, W^{**} is the part of the space formed

by the hyperplanes

$$\alpha = 0, \pm 1, \pm 2, \dots \quad \text{for all the roots } \alpha$$

which contains $(\varphi_0^1, \dots, \varphi_0^l)$ it follows from the theorem on p. 111 that any element of $A_{\mathcal{M}}$ may be written in the form $Uh(\varphi^1, \dots, \varphi^l)U^{-1}$, $(\varphi^1, \dots, \varphi^l)$ in W^{**} , U in $A_{\mathcal{M}}$. If we apply this representation to the identity of $A_{\mathcal{M}}$, the point $\theta' = (\varphi^1, \dots, \varphi^l)$ must have the property that all the roots are integers there. θ' lies necessarily on the boundary of W^{**} . By a transformation II we may transform θ' into θ . We can perform a transformation I hereafter such that $W^{**} \rightarrow W^*$. It follows now, more generally, that

Theorem 3. The main elements $h(\varphi^1, \dots, \varphi^l)$ and $h(\psi^1, \dots, \psi^l)$ are conjugate if and only if $(\varphi^1, \dots, \varphi^l)$ and $(\psi^1, \dots, \psi^l)$ are connected by a transformation of $\widetilde{(S)}$.

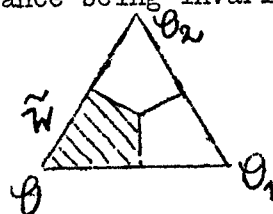
At the same time we have proved

Theorem 4. The elements of the group $\widetilde{(S)}$ can be written in the form PQ where Q belongs to (S) and P is of the form

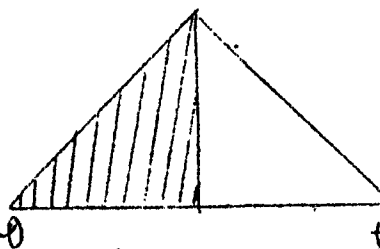
$$\varphi^r \rightarrow \psi^r = \varphi^r + m^r$$

Here $(m^1, \dots, m^l)$ underlies the condition that all the roots have integral values in this point.

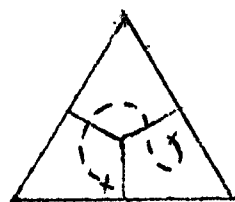
It is easy now to construct a fundamental region $\widetilde{W}$ for the group $\widetilde{(S)}$. Let $\theta, \theta_1, \theta_2, \dots, \theta_{f-1}$ be the vertices of W^* in which all the roots have integral values. Let $\widetilde{W}_\nu$ be the part of W^* for the points R of which the distance $R\theta^\nu$ is not greater than $R\theta^\mu$, $(\mu \neq \nu)$. Then W^* is composed of $\widetilde{W}, \widetilde{W}_1, \dots, \widetilde{W}_{f-1}$. The transformation $\widetilde{S}_\nu$ of $\widetilde{(S)}$ which transforms θ into θ_ν and W^* into itself, is uniquely determined. The distance being invariant under the transformations of $\widetilde{(S)}$, $\widetilde{W}$ is transformed into $\widetilde{W}_\nu$. Hence $\widetilde{W}$ is a fundamental region.



(In the case of the second figure, not all the roots are integers in the upper vertex of W^* .)



These theorems answer our questions for the case of the group A_m . We deal now with the "true" group B_m . Let t_0^* and t_1^* be two elements of B_m corresponding to the same element t_0 of A_m and let γ^* be a path from t_0^* to t_1^* on B_m and γ the corresponding path on A_m . Then γ is closed. We have $t_0^* = t_1^*$ if and only if γ can be shrunk to a point on A_m . We may assume without restriction that the elements t^* on γ^* are regular, with the possible exception of t_0^* and t_1^* . According to the theorem on p. III every t^* on γ^* can be written in the form $u^* h^*(\varphi^1, \dots, \varphi^{\lambda}) u^{*-1}$, where $(\varphi^1, \dots, \varphi^{\lambda})$ is a point of W^* and u^* belongs to B_m . We may assume that $(\varphi^1, \dots, \varphi^{\lambda})$ and u^* depend continuously on t^* varying on γ^* . Let $R_0 = (\varphi_0^1, \dots, \varphi_0^{\lambda})$ and $R_1 = (\varphi_1^1, \dots, \varphi_1^{\lambda})$ be the points corresponding to t_0^* and t_1^* . Then $h(\varphi_0^1, \dots, \varphi_0^{\lambda})$ and $h(\varphi_1^1, \dots, \varphi_1^{\lambda})$ are conjugate in A_m and therefore the points R_0 and R_1 are connected by a transformation of $(\tilde{S})$. The point $(\varphi^1, \dots, \varphi^{\lambda}) = R$ describes a path leading from R_0 to R_1 . We distinguish two cases: 1) R_0 and R_1 lie in the same $\tilde{W}_v$. Then $R_0 = R_1$. The same argument as on p. 43 and on p. 110 shows that γ can be shrunk to a point and we have $t_0^* = t_1^*$. 2) R_0 is lying in $\tilde{W}_v$ and R_1 in $\tilde{W}_\mu$, $v \neq \mu$. Suppose that γ can be shrunk to a point. We may assume that t_0 is fixed during this deformation. Then R_0 and R_1 will be fixed also and we obtain a contradiction. Hence t_0^* and t_1^* are different in this case. We have



Theorem 5. The number of sheets of B_m considered as universal covering manifold of A_m is equal to the number f of vertices $\theta, \theta_1, \dots, \theta_{f-1}$ of W^* in which all the roots are integers. Two different main elements

$h^*(\varphi^1, \dots, \varphi^l)$ and $h^*(\psi^1, \dots, \psi^l)$, $(\varphi^1, \dots, \varphi^l)$ and $(\psi^1, \dots, \psi^l)$ in W^* are not conjugate in $\mathcal{C}_M$. Every element of $\mathcal{C}_M$ is conjugate to a main element $h^*(\varphi^1, \dots, \varphi^l)$, $(\varphi^1, \dots, \varphi^l)$ in W^* .

The points of W^* represent therefore the classes of conjugate elements in $\mathcal{C}_M$. The f elements corresponding to $\theta, \theta_1, \dots, \theta_{f-1}$ are the elements of $\mathcal{C}_M$ which are represented by the identity in A_M . They form an invariant subgroup belonging to the centrum of $\mathcal{C}_M$.

We may replace W^* in our theorem by every other region W^{**} which is cut out of the space by the hyperplanes $\alpha = 0, \alpha = \pm 1, \alpha = \pm 2, \dots$ where α ranges over all roots. In particular, there is exactly one vertex of W^{**} which represents the identity of $\mathcal{C}_M$. These points representing the identity form a lattice $\mathcal{R}'$. If, namely,

$$h^*(\varphi^1, \dots, \varphi^l) = 1, \quad h^*(\psi^1, \dots, \psi^l) = 1$$

then $h^*(\varphi^1 \pm \psi^1, \dots, \varphi^l \pm \psi^l) = 1$. This lattice $\mathcal{R}'$ is a sublattice of the lattice $\mathcal{R}$ consisting of the points which correspond to the identity in A_M . These are the points where all the roots have integral values. We will characterize $\mathcal{R}'$ in a similar manner.

The same argument as in the case of A_M shows

Theorem 6. Let (S^*) be the group of all transformations PQ where

Q belongs to (S) and P is a translation

$$\varphi^r \rightarrow \varphi^r + n^r$$

where $(n^1, \dots, n^l)$ is a point of the lattice $\mathcal{R}'$. * The main elements

* For the definition of $\mathcal{R}'$ cf. Theorems 7, 7' or 7".

$h^*(\varphi^1, \dots, \varphi^l)$ and $h^*(\psi^1, \dots, \psi^l)$ are conjugate in $\mathcal{C}_M$ if and only if the points $(\varphi^1, \dots, \varphi^l)$ and $(\psi^1, \dots, \psi^l)$ are connected by a transformation of (S^*) . W^* is a fundamental region for (S^*) .

Let α be a root. We may assume without restriction that $\alpha \neq 0$

forms a wall of W . In the other case we replace W by one of the other chambers.

According to (5) we have then $\alpha = \varphi^i$. We may assume $i = 1$,

$\alpha = \varphi^1$. Let R be the point $\varphi^1 = 1, \varphi^2 = 0, \dots, \varphi^l = 0$, and β the vector from O to R . The reflection in $\alpha = 0$ turns β into the vector

$$(19) \quad \beta' = \beta - \frac{2(\alpha|\beta)}{(\alpha|\alpha)}\alpha = \beta - \frac{2\alpha}{(\alpha|\alpha)}$$

The element $h(1, 0, \dots, 0)$ of $A_{\mathcal{R}}$ is the identity since in $(1, 0, \dots, 0)$ all

the roots are integers. The element $h^*(1, 0, \dots, 0)$ is one of the elements

$\mathcal{E}_v$ of $\mathcal{C}_u$ which are represented by the identity in $A_{\mathcal{R}}$

Since $\beta = (\varphi^1, \dots, \varphi^l)$ and $\beta' = (\psi^1, \dots, \psi^l)$ are connect-

ed by a transformation of (S) , the elements $h^*(\varphi^1, \dots, \varphi^l) = \mathcal{E}_v$ and

$h^*(\psi^1, \dots, \psi^l)$ are conjugate in $\mathcal{C}_u$ (cf. Theorem 1). But $\mathcal{E}_v$ belongs

to the centrum of $\mathcal{C}_u$. Hence

$$(20) \quad \begin{aligned} h^*(\varphi^1, \dots, \varphi^l) &= h^*(\psi^1, \dots, \psi^l) \\ h^*(\varphi^1 - \psi^1, \dots, \varphi^l - \psi^l) &= 1 \end{aligned}$$

But it follows from (19) that $(\varphi^1 - \psi^1, \dots, \varphi^l - \psi^l) = \beta - \beta' = \frac{2\alpha}{(\alpha|\alpha)}$

Then we obtain from (20) that the vector $\frac{2\alpha}{(\alpha|\alpha)}$ belongs to the lattice $\mathcal{R}'$.

This holds for every root α . We are going to prove

Theorem 7. The lattice $\mathcal{R}'$ is spanned by the vectors $\frac{2\alpha}{(\alpha|\alpha)}$ where

α ranges over the roots.

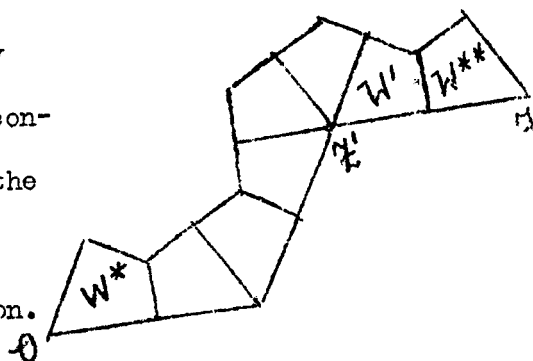
Proof. Let $Z = (\varphi^1, \dots, \varphi^l)$ be a point such that $h^*(\varphi^1, \dots, \varphi^l) = 1$. We have to prove that $(\varphi^1, \dots, \varphi^l)$ is a linear combination of the

$\frac{2\alpha}{(\alpha|\alpha)}$ with integral coefficients. Z is a vertex of one of the regions W^{**} cut out of the space by

the hyperplanes $\alpha = 0, \pm 1, \pm 2, \dots$. We may connect W^* and W^{**} by a "chain" of such regions of the

following kind: Two consecutive regions have an

$(l-1)$ -dimensional part of their boundary in common.



Let W' be the region preceding W^{**} in this chain. There is one vertex

Z' of W' corresponding to the identity of $\mathcal{G}_u$. We assume that we have proved already that the vector $\vec{OZ'}$ is an integral combination of vectors $\frac{2\alpha}{(\alpha\alpha)}$ and we will show the same for the vector $\vec{OZ}$.

Let the common part of the boundary of W' and W^{**} lie on the hyperplane $\alpha = k$, where α is a root and k an integer. The reflection in $\alpha = k$ may be composed of translations of the type

$$\psi^v = \varphi^v + m_v, \quad m_v \text{ integers}$$

and the reflection in $\alpha = 0$. It is clear that regions of our type are transformed into such regions again. In particular, W' is transformed into W^{**} .

The point Z' is mapped on a point Z'' of W^{**} given by

$$(21) \quad \vec{OZ''} = \vec{OZ'} + \frac{2(\mathcal{R} - (\alpha, \vec{OZ'}))}{(\alpha\alpha)} \cdot \alpha$$

The value of α in Z' , i.e. $(\alpha, \vec{OZ'})$ is an integer since Z' corresponds to the identity of $\mathcal{A}_u$, even of $\mathcal{G}_u$. From (21) it follows that the vector $\vec{OZ''}$ is an integral multiple of $\frac{2\alpha}{(\alpha\alpha)}$. Since Z' and $\frac{2\alpha}{(\alpha\alpha)}$ belong to $\mathcal{R}'$, so does Z'' . Hence $Z'' = Z$. (21) shows that $\vec{OZ}$ is an integral linear combination of vectors $\frac{2\alpha}{(\alpha\alpha)}$, q.e.d.

In an appropriate coordinate system the vectors of $\mathcal{R}'$ and only these have integral components $(\varphi^1, \varphi^2, \dots, \varphi^r)$. A linear form

$\Phi = m_1 \varphi^1 + m_2 \varphi^2 + \dots + m_r \varphi^r$ has been called an integral form in Chapter 2, sec. 4, if its values for $(\varphi^1, \dots, \varphi^r) = \frac{2\alpha}{(\alpha\alpha)}$ were integers. In

our new coordinate system, the coefficients of Φ are integers since Φ has an integral value for $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, ..., $(0, \dots, 0, 1)$. Con-

versely, Φ is an integral form if the coefficients are integers. Therefore

the coordinate system is chosen now according to the convention in Chap. 2, sec. 4

p. 93. We have

Theorem 7'. If the coordinate system is chosen such that the integral

forms $\Phi = m_1 \varphi^1 + \dots + m_l \varphi^l$ and only the integral forms have integral coefficients m_v , then the lattice $\mathcal{R}'$ consists of all points $(\varphi^1, \dots, \varphi^l)$ with integral coordinates φ^v

$$h^*(\varphi^1, \dots, \varphi^l) = 1 \text{ if and only if } \varphi^1 \equiv \varphi^2 \equiv \dots \equiv \varphi^l \equiv 0 \pmod{1}$$

$$h^*(\varphi^1, \dots, \varphi^l) = h^*(\psi^1, \dots, \psi^l) \text{ if and only if } \varphi^1 \equiv \psi^1, \varphi^2 \equiv \psi^2, \dots, \varphi^l \equiv \psi^l \pmod{1}$$

The last fact follows from the remark that $h^*(\varphi^1, \dots, \varphi^l) =$

$$= h^*(\psi^1, \dots, \psi^l) \text{ if and only if } h^*(\varphi^1 - \psi^1, \dots, \varphi^l - \psi^l) = 1.$$

We may formulate our result in the following form independent of the choice of the coordinate system:

Theorem 7''. $\mathcal{R}'$ consists of all points where all integral forms have integral values.

These results complete the considerations on p. 119. By using the completeness of the system of characters we obtain the theorem that all expressions (97), Chap. 2, are characters if we have $\Phi = m_1 \varphi^1 + \dots + m_l \varphi^l \geq_S \Psi$ for all S in (S) .

We deal now with the case where the given group $\mathcal{G}$ is simple.

Theorem 8. The region W^* is a simplex in the case of a simple group

$\mathcal{G}$. It is given by

$$(22) \quad 0 < \overset{(1)}{\alpha}, 0 < \overset{(2)}{\alpha}, \dots, 0 < \overset{(l)}{\alpha}; \quad \beta < 1$$

where $\overset{(1)}{\alpha}, \overset{(2)}{\alpha}, \dots, \overset{(l)}{\alpha}$ and β are certain roots.

Proof. We choose the coordinate system again in such a manner that the walls of W are given by

$$\overset{(1)}{\alpha} = \varphi^1 = 0, \quad \overset{(2)}{\alpha} = \varphi^2 = 0, \dots, \overset{(l)}{\alpha} = \varphi^l = 0.$$

Let $\beta = \beta_1 \varphi^1 + \dots + \beta_l \varphi^l$ be the highest root of $\mathcal{G}$. If

$\gamma = \gamma_1 \varphi^1 + \dots + \gamma_\lambda \varphi^\lambda$ is any positive root, we want to show that $\beta_i \geq \gamma_i$. It follows then that $\gamma \leq \beta$ in every point $(\varphi^1, \dots, \varphi^\lambda)$ of $\mathcal{H}$. Hence the domain (22) has no point in common with $\gamma = k$, k integer. Therefore (22) defines W^* .

We have $\beta_i \geq \gamma_i$, β being the highest root. If $\beta - \gamma = (\beta_1 - \gamma_1) \varphi^1 + \dots + (\beta_\lambda - \gamma_\lambda) \varphi^\lambda$ is a root, it is positive and all the $\beta_i - \gamma_i$ are ≥ 0 . We may assume therefore that $\beta - \gamma$ is not a root. From (1), (2), (3) ^(p. 165) it follows then that $\mu_{\gamma\beta} = \frac{2(\gamma, \beta)}{(\gamma, \gamma)} \leq 0$. Hence $(\gamma, \beta) \leq 0$.

$$(23) \quad \gamma^1 \beta_1 + \dots + \gamma^\lambda \beta_\lambda \leq 0$$

Suppose now that $\gamma_j > \beta_j$. We may assume that all the γ^i are ≥ 0 . We obtain, namely, by reflection of γ in the hyperplane $\alpha^{(i)} = 0$, that

$$\gamma' = \gamma - \frac{2(\gamma, \alpha^{(i)})}{(\alpha^{(i)}, \alpha^{(i)})} \alpha^{(i)} = \gamma - \frac{2\gamma^i}{g^{ii}} \varphi^i$$

is a root too. Here all the coefficients are the same as in γ , only the coefficient of φ^i has increased if $\gamma^i < 0$. We repeat this process as many times as possible and obtain finally a root which we denote by γ again such that still $\gamma_j > \beta_j$ and that all $\gamma^i \geq 0$. The difference $\beta - \gamma$ cannot be a root here because of $\gamma_j > \beta_j$, and thus (23) is valid. All the β_i are ≥ 0 , β being a positive root. Hence either $\beta_i = 0$ or $\gamma^i = 0$ according to (23). Suppose that

$$\beta_{v_1} > 0, \beta_{v_2} > 0, \dots, \beta_{v_p} > 0; \beta_{v_{p+1}} = 0, \dots, \beta_{v_e} = 0$$

We have $p \geq 1$, but $p \leq \lambda - 1$. (In the case $p = \lambda$, all the γ^i would be equal to 0.)

The β^i are all ≥ 0 because $\beta - \frac{2(\beta, \alpha^{(i)})}{(\alpha^{(i)}, \alpha^{(i)})} \alpha^{(i)} = \beta - \frac{2\beta^i}{g^{ii}} \varphi^i$

would be a root higher than β if $\beta^i < 0$. We have

$$(24) \quad \beta^i = \sum g^{ik} \beta_k = g^{i, v_1} \beta_{v_1} + g^{i, v_2} \beta_{v_2} + \dots + g^{i, v_p} \beta_{v_p}$$

We choose for i one of the indices $\gamma_{p+1}, \gamma_{p+2}, \dots, \gamma_\lambda$. Then (cf. (8), p. 162)
 $g^{i, \gamma_1} \leq 0, g^{i, \gamma_2} \leq 0, \dots, g^{i, \gamma_p} \leq 0$. If one of them should be smaller than 0, the right-hand side of (24) would be smaller than 0, whereas $\beta^i \geq 0$. Hence

$$g^{ik} = 0 \text{ for } \begin{matrix} i = \gamma_{p+1}, \gamma_{p+2}, \dots, \gamma_\lambda \\ k = \gamma_1, \gamma_2, \dots, \gamma_p \end{matrix}$$

It follows then that $\alpha^{(i)}$ is orthogonal to $\alpha^{(k)}$ for the same values of i and k .

Let $\mathcal{L}_1$ be the subspace spanned by the $\alpha^{(i)}$ and $\mathcal{L}_2$ be the subspace spanned by the $\alpha^{(k)}$. Both are orthogonal. If we reflect a vector of $\mathcal{L}_1$ in a hyperplane $\alpha = 0$ or $\alpha = 0$ we always obtain a vector of $\mathcal{L}_1$. The same is true for

$\mathcal{L}_2$. Every root can be obtained by successive reflections in hyperplanes $\alpha = 0$ starting from a $\alpha^{(m)}$. Hence every root lies either in $\mathcal{L}_1$ or in $\mathcal{L}_2$.

According to the results on p. , $\mathfrak{g}$ is not simple. If $\mathfrak{g}$ is simple, the assumption $\gamma_j > \beta_j$ leads to a contradiction. This proves the theorem.

Let $\omega = \omega_1 \varphi^1 + \dots + \omega_\lambda \varphi^\lambda$ be a linear form such that

$\frac{2(\omega, \alpha^{(i)})}{(\alpha^{(i)}, \alpha^{(i)})}$ is an integer, $i = 1, 2, \dots, \lambda$. We state that $\frac{2(\omega, \alpha)}{(\alpha, \alpha)}$ is an integer for every root α , or, in other words, that ω is an integral form.

Every root $\alpha^{(i)}$ can be obtained from an $\alpha^{(k)}$ by a succession of reflections in hyperplanes $\alpha = 0$. Let us assume that we have already proved that $\frac{2(\omega, \gamma)}{(\gamma, \gamma)}$

is an integer for a root γ and that $\alpha^{(k)}$ is obtained from γ by the reflection in $\alpha = 0$:

$$\alpha = \gamma - 2 \frac{(\gamma, \alpha^{(k)})}{(\alpha^{(k)}, \alpha^{(k)})} \alpha^{(k)}$$

$$\frac{2(\omega, \alpha)}{(\alpha, \alpha)} = \frac{2(\omega, \gamma)}{(\alpha, \alpha)} - \frac{2(\gamma, \alpha^{(k)})}{(\alpha, \alpha)} \frac{2(\omega, \alpha^{(k)})}{(\alpha^{(k)}, \alpha^{(k)})}$$

Now $(\alpha, \alpha) = (\gamma, \gamma)$ since a reflection transforms γ into α . Then

$\frac{2(\omega, \gamma)}{(\alpha, \alpha)} = \frac{2(\omega, \gamma)}{(\gamma, \gamma)}$ is an integer according to our assumption with regard

to γ . $\frac{2(\gamma, \alpha^{(R)})}{(\alpha, \alpha)} = \frac{2(\gamma, \alpha)}{(\gamma, \gamma)}$ is an integer since γ and α are roots.

Finally, $\frac{2(\omega, \alpha^{(R)})}{(\alpha, \alpha)}$ has been assumed to be an integer. This proves our statement.

Consequently the integral forms are characterized by the fact that

$$\frac{2(\omega, \alpha^{(i)})}{(\alpha, \alpha)} = \frac{2\omega^i}{g^{ii}} = p_i$$

is an integer. We have, therefore,

$$(25) \quad \omega^i = \frac{g^{ii}}{2} p_i$$

If ω is dominant, $\omega \geq S\omega$ for all S in (S) , we have $\omega^i \geq 0$ (cf. p. 164),

and hence $p_i \geq 0$ and vice versa.

We form χ special such dominant integral forms $\omega = \Phi_1, \omega = \Phi_2,$

..., $\omega = \Phi_\chi$ by putting $p_1 = 1, p_2 = 0, \dots, p_\chi = 0$, then $p_1 = 0, p_2 = 1,$

$p_3 = 0, \dots, p_\chi = 0$, ..., and finally $p_1 = 0, \dots, p_{\chi-1} = 0, p_\chi = 1$. According

to (25) every dominant integral form ω is of the form

$$(26) \quad \omega = p_1 \Phi_1 + p_2 \Phi_2 + \dots + p_\chi \Phi_\chi$$

p_i integers ≥ 0 .

There exist irreducible representations $\Gamma_1, \Gamma_2, \dots, \Gamma_\chi$ of $\mathcal{G}_u$

(and of the not restricted group $\mathcal{G}$) such that Φ_v is the highest weight of

Γ_v . The highest term in the character of Γ_v then is $e(\Phi_v)$. Let

Γ be any irreducible representation, and ω its highest weight. Then

$\omega \geq S\omega$ for every S in (S) and ω is an integral form. Then ω possess a

representation (26). The same considerations as on pp. 23⁽⁻²⁵⁾ are possible. We

form

$$(27) \quad \mathcal{G} = \Gamma_1^{p_1} \times \Gamma_2^{p_2} \times \dots \times \Gamma_\chi^{p_\chi}.$$

The highest weight is ω according to (26). Consequently, Γ is an irreducible

constituent of $\mathcal{G}$ and can be obtained by the method on p. 24. We denote this

method of forming (27) and splitting up the constituent of highest weight as
Cartan multiplication. We have

Theorem 9. There exist λ irreducible representations $\Gamma_1, \Gamma_2, \dots, \Gamma_\lambda$ such that every irreducible representation Γ can be obtained by Cartan multiplication of the Γ_ν .

In the case of a simple group, the group (S^*) , (Theorem 6), can be generated by $\lambda+1$ reflections, the reflections in $\alpha^{(1)} = 0, \alpha^{(2)} = 0, \dots, \alpha^{(\lambda)} = 0$ and in $\beta = 1$ in the notation of Theorem 8 as is easily seen. Let $\mathcal{R}$ be the sum of all the regions SW^* , S in (S) . Then $\mathcal{R}$ is a fundamental region for the subgroup $(\mathcal{P})$ of the translations $\mathcal{P}$ in (S^*) . We compute the (Cartesian) volume v of $\mathcal{R}$.

The vertices of W^* are given by

$$(28) \quad (0, 0, \dots, 0); \left(\frac{1}{\beta_1}, 0, \dots, 0\right); \left(0, \frac{1}{\beta_2}, 0, \dots, 0\right); \dots; \left(0, \dots, 0, \frac{1}{\beta_\lambda}\right)$$

The volume of W^* is, therefore, given by $c \cdot \frac{1}{\lambda!} \frac{1}{\beta_1} \frac{1}{\beta_2} \dots \frac{1}{\beta_\lambda}$ where c is a factor depending on the g^{ik} . Hence

$$(29) \quad v = c \cdot \frac{s}{\lambda!} \frac{1}{\beta_1} \frac{1}{\beta_2} \dots \frac{1}{\beta_\lambda}$$

where s is the order of (S) .

The number f of sheets of $\mathcal{C}_n$ over A_n is the number of vertices (28) with integral coordinates (Theorem 5). Hence $f-1$ is the number of β_ν which are equal to 1. The proof of Theorem 8 shows that $f-1$ is the number of φ^ν whose coefficient in every root is $= 0, +1$ or -1 .

We consider now a part $\mathcal{Q}$ of the space which is composed of k regions congruent with $\mathcal{R}$ under translations $\mathcal{P}$. We choose $\mathcal{Q}$ such that an arbitrarily large sphere about the origin belongs to $\mathcal{Q}$. The volume of $\mathcal{Q}$ is kv .

$\mathcal{R}$ contains exactly one point of $\mathcal{R}'$, namely the origin. Hence $\mathcal{Q}$ contains k such points. The lattice of all points with

$\varphi^1 \equiv 0, \varphi^2 \equiv 0, \dots, \varphi^l \equiv 0 \pmod{1}$ splits up into f sublattices. One of them is $\mathcal{R}'$. The other $f-1$ belong to the $f-1$ elements $\neq 1$ in $\mathcal{C}_u$ which are obtained from $\mathcal{R}'$ by translation $\psi^1 = \varphi^1 + m^1, \dots, \psi^l = \varphi^l + m^l, m^v$ integers. No two of the f sublattices have a point in common. It follows that the volume of a fundamental parallelepipedon of $\mathcal{R}'$ is equal to $c \cdot f$ where c is the same constant as in (29). The volume of $\mathcal{T}$ then is approximately $c \cdot f \cdot k$ since $\mathcal{T}$ contains k points of $\mathcal{R}'$. We let $k, \mathcal{T}$ increase. In the limit we obtain $v = c \cdot f$. According to (29) we have

$$\frac{S}{l!} \frac{1}{\beta_1} \frac{1}{\beta_2} \dots \frac{1}{\beta_l} = f$$

Theorem 10. If W^* is given by

$$0 < \varphi^1, 0 < \varphi^2, \dots, 0 < \varphi^l, \beta = \beta_1 \varphi^1 + \dots + \beta_l \varphi^l < 1$$

we have

$$S = l! \beta_1 \beta_2 \dots \beta_l \cdot f$$

where S is the order of (S) and f is the number of β_v which are equal to 1. (i. e. the number of sheets of $\mathcal{C}_u$ over A_u).

6. Discrete groups generated by reflections (Notes by Dr. H.S.M.Coxeter)

To every semi-simple
simple Lie group there corresponds a group generated by reflections whose fundamental region is finite
a simplex in Euclidean space.

Our ultimate object is to prove the converse, i.e. that the discrete groups which correspond to these Lie groups are the only discrete groups of this kind.

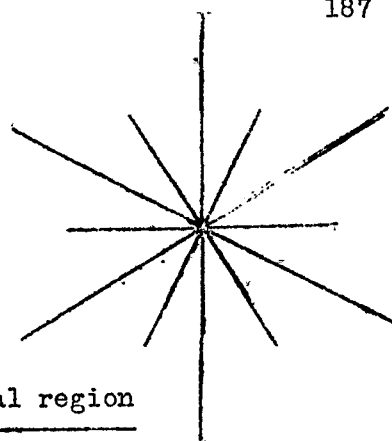
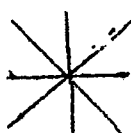
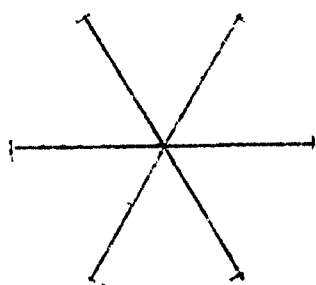
We shall first consider special cases in few dimensions, in order to become familiar with groups generated by reflections.

Each of the vector diagrams enumerated by van der Waerden (Math. Zeitschr., 37) is invariant under a finite group S generated by reflections in hyperplanes perpendicular to the vectors; but it sometimes happens that two of them (corresponding to different Lie groups) have the same group S , differing merely in the lengths of the vectors. Moreover, the groups S are not the only finite groups generated by reflections. The vectors have to satisfy arithmetical conditions which make S leave a certain lattice invariant. It is desirable, therefore, to consider the infinite discrete group derived from S by adding the lattice-translations.

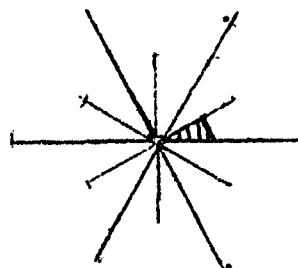
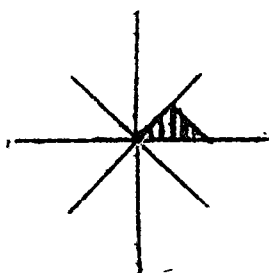
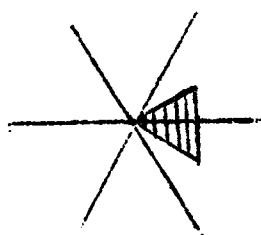
If α is a vector of the diagram, we have a translation $\frac{2\alpha}{(\alpha, \alpha)}$. Since S contains the reflection in the hyperplane through the origin perpendicular to α , the translation could be obtained by combining this reflection with the reflection in a parallel hyperplane, namely the perpendicular bisector of the vector $\frac{2\alpha}{(\alpha, \alpha)}$ from the origin. The infinite discrete group is thus generated by reflections.

Take for instance the two-dimensional cases.

The vectors α



The vectors $\frac{2\alpha}{(\alpha\alpha)}$ and the fundamental region



The infinite discrete group is generated by reflections in the sides of the fundamental region, which is a triangle of angles

$$(\pi/3, \pi/3, \pi/3), (\pi/2, \pi/4, \pi/4), (\pi/2, \pi/3, \pi/6)$$

in the three cases.

The simplest g.g.r. (group generated by reflections) is that generated by one reflection. We represent it by a single dot $\bullet$. Abstractly, it is the group of order 2. Considered in Euclidean 1-space, it is generated by reflection in a point, and the f.r. (fundamental region) is the half-line.

When there are two generating reflections, the group can be represented in 2 dimensions, reflecting in lines. When the angle between 2 lines is incommensurable with π , the group generated by reflections in them is not discrete. When the angle is commensurable with π , it can be taken to be just a submultiple of π , since some multiple of $\frac{d\pi}{n}$ must differ from $\frac{\pi}{n}$ by a multiple of π .

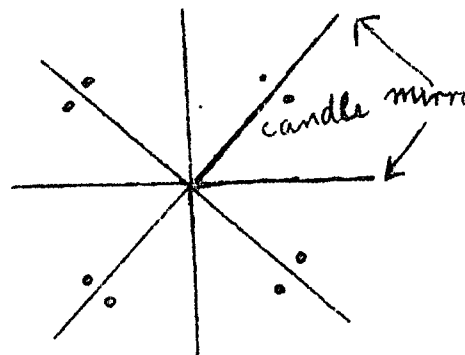
When the angle is a submultiple of π , the situation can be seen very clearly by taking actual mirrors, with a lighted candle for object. For angle $\frac{\pi}{n}$, the images (and object) form an alternate-sided $2n$ -gon,

and the group, of order $2n$, has an angular (fundamental region) f.r. $\angle$. We call it $\bullet \xrightarrow{n} \bullet$ when $n > 2$,

but merely $\bullet \bullet$ when $n = 2$ since it is

then the direct product of two groups $\bullet \bullet$.

(Abstractly, $\bullet \bullet$ is the Vierergruppe.)

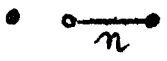
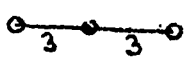
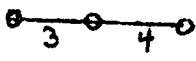
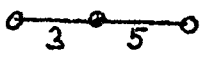


As the angle diminishes, the order of the group (= number of images) increases indefinitely. It is natural to let $\bullet \xrightarrow{\infty} \bullet$ denote the group generated by reflections in two parallel mirrors. This can be represented in space of 1 dimension, reflecting in two points. The f.r. is a finite segment of the line. (The corresponding Lie group is the rotation group in three dimensions.)

The situation can still be observed in a practical manner when a third mirror is introduced. When the planes of the three mirrors intersect, all the images lie on a sphere whose center is the point of intersection, and the planes cut out a spherical polygon. Regarding the group as operating on this sphere, instead of on the whole space, the angular f.r. is replaced by this spherical polygon. By the same argument as in the case of 2 mirrors, all the angles must be submultiples of π . If the f.r. is a triangle of angles $\frac{\pi}{l}, \frac{\pi}{m}, \frac{\pi}{n}$ we must have

$$\frac{1}{l} + \frac{1}{m} + \frac{1}{n} > 1.$$

The solutions of this inequality, and the symbols for the corresponding groups, are as follows:

2,	2,	n		(... when n = 2)
2,	3,	3		
2,	3,	4		
2,	3,	5		

In these symbols, as before, each dot represents a mirror, and two dots not (directly) linked represent perpendicular mirrors.

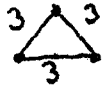
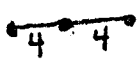
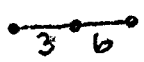
Since the area of the triangle is $\pi(1/2 + 1/m + 1/n - 1)$ and of the whole sphere 4, the order of the group is

$$\frac{4}{1/2 + 1/m + 1/n - 1}$$

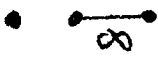
On the other hand, when the three planes do not intersect, the images all lie in one plane, and we have an infinite group in two dimensions. (This can be seen very nicely by means of three vertical mirrors, forming the lateral faces of a triangular prism.) If the f.r. is a triangle of angles $\pi/2, \pi/m, \pi/n$ we now have

$$\frac{1}{2} + \frac{1}{m} + \frac{1}{n} = 1$$

with three solutions

3,	3,	3	
2,	4,	4	
2,	3,	6	

(Cf. van der Waerden's three diagrams, described above. His proof that only angles $\pi/2, \pi/3, \pi/4, \pi/6$ can occur depends on the arithmetical condition that $\frac{2(\alpha, \beta)}{(\alpha, \alpha)}$ and $\frac{2(\alpha, \beta)}{(\beta, \beta)}$ must be integers, where α, β are two of the vectors. The product of these numbers is $4 \cos^2 \hat{\alpha, \beta}$, which can only be 0, 1, 2, 3.)

The group  is not interesting, since it is infinite

without having a finite f.r.

We pass now to the case of four mirrors. If they form a spherical or plane quadrangle of angles $\frac{\pi}{k}, \frac{\pi}{l}, \frac{\pi}{m}, \frac{\pi}{n}$ we must have

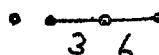
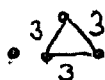
$$\frac{1}{k} + \frac{1}{l} + \frac{1}{m} + \frac{1}{n} \geq 2.$$

The only solution is $(2, 2, 2, 2)$, corresponding to the infinite group:



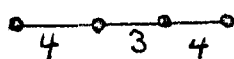
(reflections in the four sides of a rectangle, or in the four walls of a room).

The infinite groups



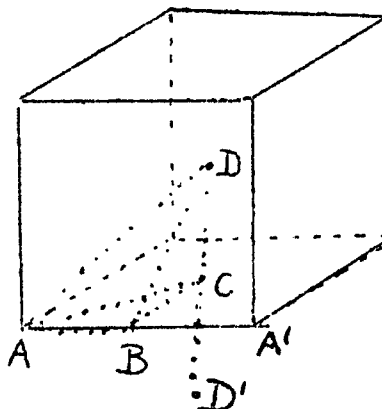
have infinite f.r.'s, and so are not interesting.

The complete symmetry-group of the ordinary cubic lattice is

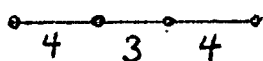


since the f.r.,

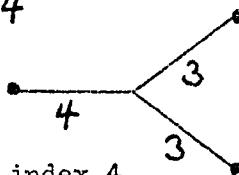
is a tetrahedron having one vertex (A) at a vertex of one of the cubes, one (B) at the center of an edge, one (C) at the center of a square face, and one (D) at the center of the whole cube.



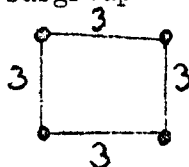
Let A' be another vertex of the cube, chosen so that B is the mid-point of AA' , and let D' be the center of another cube, chosen so that C is the mid-point of DD' . Then



has a subgroup of index 2

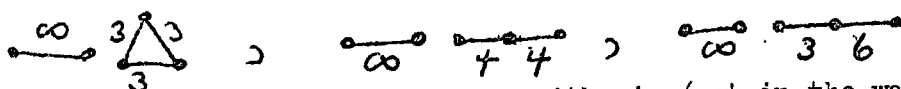


whose f.r. is $AA'CD$, and a subgroup of index 4

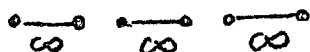


whose f.r. is $AA'DD'$.

With five mirrors we can generate the groups



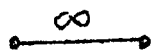
whose f.r.'s are triangular prisms; and with six (set in the walls, ceiling and floor of a room), the group



whose f.r. is a rectangular parallelepiped.

We shall see later that those discrete groups which correspond to simple Lie groups are just those infinite g.g.r.'s whose symbols are connected graphs. By juxtaposing several such graphs, we obtain the symbol for the direct product of the corresponding g.g.r.'s, and the direct sum of the corresponding Lie groups. (It is known that every semi-simple group is a direct sum of simple groups.) Finally, the fundamental region is a topological product (of simplexes lying in absolutely perpendicular spaces).

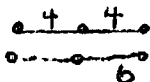
So far, we have considered seven connected graphs representing infinite g.g.r.'s. The corresponding simple groups are as follows:



$$L_2 = C_2 = O_3$$



$$L_3$$

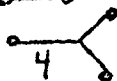


$$C_4 = O_5$$

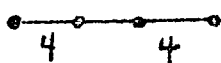
(van der Waerden's G_2)



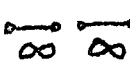
$$L_4 = O_6$$



$$O_7$$

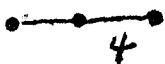


$$C_6$$

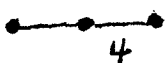
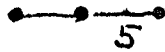
(Wherever a link is unmarked, we are to understand the mark 3.) L_n means the unimodular linear group, C_n (n even) the complex group, and O_n the orthogonal group. O_4 is only semi-simple, corresponding to the disconnected graph 

Incidentally, we have found several finite g.g.r.'s, which are listed

below, with their orders:

		2
	( when $n = 2$)	$2n$
	( when $n = 2$)	$4n$
		24
		48
		120

The last three of these are the complete symmetry groups of the regular solids.

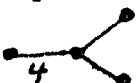
 is the symmetric group of degree 4, permuting the four vertices of the regular tetrahedron.  is the direct product of this with the group of order 2 (generated by the central inversion. We can inscribe in the cube two tetrahedra which are interchanged by the central inversion).  is the direct product of the alternating group of degree 5 with the group of order 2. (The five objects that are permuted can be taken to be five tetrahedra inscribed in the dodecahedron.)

The presence of the central inversion in these last two cases illustrates the following theorem:

If the f.r. of a finite g.g.r. contains no intrinsic symmetry, then the central inversion belongs to the group.

For, the central inversion is then the only operation that can transform the f.r. into its antipodal region.

Note that the group , whose f.r. is isosceles, does not contain the central inversion.

Any subset of the generating reflections of a g.g.r. generates a subgroup whose graph is part of the graph that represents the whole group. E.g., by omitting the generating reflections of , one at a time, we obtain the finite subgroups



Cartan has proved that the f.r. of a finite g.g.r. is a spherical simplex (cf. sec. 4). We shall give a geometrical proof (by induction) of the same theorem.

Any finite group of congruent transformations in Euclidean space leaves one point invariant, namely the centroid of all transforms of a point of general position. A finite g.g.r., therefore, is generated by reflections in a number of concurrent hyperplanes. We can assume this number to be at least as great as the number of dimensions of the Euclidean space, since otherwise the group would operate essentially in a space of fewer dimensions. Under the operations of the group, these hyperplanes are transformed into a larger set of hyperplanes (all passing through the same fixed point) which divide up the whole space into a number of congruent angular regions ("chambers"). By drawing a sphere (say of unit radius) around the fixed point, we replace these angular regions by spherical polytopes. By Witt's topological argument (p. 159/160), such a polytope is a fundamental region for the g.g.r. (This result depends on the simple connectivity of the spherical space. It holds also for many other kinds of space, even multiply-connected, by not for elliptic space.)

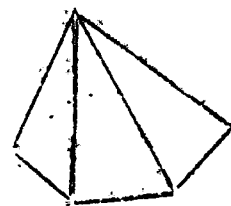
The f.r. can be any spherical polytope all of whose dihedral angles are submultiples of π (namely $\frac{\pi}{2}$ or $\frac{\pi}{3}$ or $\frac{\pi}{4}$ etc.). In order to prove that this must be a simplex, it is sufficient to know that the dihedral angles are $\leq \frac{\pi}{2}$

Any spherical polytope whose dihedral angles are all non-obtuse is a

simplex.

In spherical space of 2 dimensions (i.e. on an ordinary sphere), the angle-sum of a triangle is always greater than

An n -gon can be divided into $n-2$ triangles in such a way that the sum of the angles of the n -gon is equal to the sum of the angles of all the triangles. This angle sum, then, is greater than $(n-2)\pi$.



Hence the average angle of the n -gon is $> \frac{(n-2)\pi}{n}$. But the average angle must be $\leq \frac{\pi}{2}$. Hence

$$1 - \frac{2}{n} < \frac{1}{2}$$

i.e.

$$n < 4.$$

This proves that the spherical polygon is a triangle.

In order to extend this result to more dimensions, we require the following important lemma:

If all the angles of a spherical triangle are $\leq \frac{\pi}{2}$, then all the sides are $\leq \frac{\pi}{2}$.

This follows from the formula for a side in terms of the three angles:

$$\cos c = \frac{\cos A \cos B + \cos C}{\sin A \sin B}.$$

Consider now a polyhedron (in spherical 3-space) whose dihedral angles are non-obtuse. By drawing a sphere around any vertex, we obtain a spherical polygon whose angles occur among the dihedral angles of the polyhedron. Since this polygon has been proved to be a triangle, the polyhedron must have trihedral vertices. Moreover, the sides of this triangle (regarded as lying on a sphere of unit radius) are face-angles of the polyhedron. By the lemma, these are likewise non-obtuse; so the faces are triangles. But a polyhedron with triangular faces and trihedral vertices can only be a simplex. Thus the theorem is proved for three dimensions.

The inductive proof for n dimensions is quite similar. By assuming the result for $n-1$ dimensions, we show that the polytope cut out on a sphere drawn around any vertex is a simplex. I.e., the vertices are simplicial. The section by a 3-space absolutely perpendicular to the space of any $(n-3)$ -dimensional element is a trihedral angle, whose dihedral angles are dihedral angles of the whole polytope, while its face-angles are dihedral angles of bounding polytopes (of $n-1$ dimensions). By the lemma, these last angles are non-obtuse. By the inductive assumption, the bounding polytopes are simplexes. We have already shown that only n of these meet at each vertex. Hence the whole polytope is a simplex.

In order to enumerate the spherical and Euclidean simplexes whose dihedral angles are submultiples of π , we need an extension of the above lemma.

If all the dihedral angles of a spherical simplex are $\leq \frac{\pi}{2}$, then all the edges are $\leq \frac{\pi}{2}$.

This is easily proved by induction. The simplex of $n-1$ dimensions, obtained by drawing a sphere around one vertex, has non-obtuse edges by the inductive assumption. But these edges are angles of triangles in the original simplex. Hence, by the lemma in its two-dimensional form, the edges (being sides of such triangles) are non-obtuse.

We proceed to express this result algebraically.

Consider the $(n+1)$ -dimensional Euclidean space in which the spherical simplex is cut out by $n+1$ hyperplanes. We may take these $n+1$ hyperplanes to define an oblique coordinate-system, with coordinates $(x_1, x_2, \dots, x_{n+1})$. We may take the perpendiculars to these hyperplanes through the origin to define another coordinate system, with coordinates $(x^1, x^2, \dots, x^{n+1})$, the positive direction of the x^1 axis being that for which $x_1 > 0$. We can regard

$(x_1, x_2, \dots, x_{n+1})$ and $(x^1, x^2, \dots, x^{n+1})$ as covariant and contravariant vectors, such that

$$x_i = \sum g_{ij} x^j, \quad x^i = \sum g^{ij} x_j.$$

By a suitable choice of units, the components g_{ij} ($i \neq j$) of the covariant metric tensor are just the cosines of the edges of the polar simplex; i.e.,

$$g_{ii} = 1, \quad g_{ji} = g_{ij} = -\cos \frac{\pi}{k_{ij}} \quad (i < j).$$

The contravariant tensor then has components of the form

$$g^{ii} = (\lambda_i)^2, \quad g^{ji} = g^{ij} = \lambda_i \lambda_j \cos \alpha_{ij} \quad (i < j)$$

where α_{ij} is an edge of the original simplex, and the λ are positive multipliers.

The extended lemma tells us that $\alpha_{ij} \leq \frac{\pi}{2}$, whence

$$g^{ij} \geq 0.$$

Let Δ denote the determinant, g_{ij} . This being the determinant of a positive definite form, we have

$$\Delta > 0.$$

Hence

$$\frac{\partial \Delta}{\partial g_{ij}} = \Delta g^{ij} \geq 0,$$

i.e.

$$\frac{\partial \Delta}{\partial (\cos \frac{\pi}{k_{ij}})} \leq 0.$$

Thus Δ diminishes for any discrete increment of $\cos \frac{\pi}{k_{ij}}$, and consequently diminishes when we replace any one of the integers k_{ij} by the next higher integer. Moreover, Δ is a quadratic function of $\cos \frac{\pi}{k_{ij}}$ in which the coefficient of $\cos^2(\frac{\pi}{k_{ij}})$ is derived from Δ by omitting the i^{th} and j^{th} rows and columns and reversing the sign. This coefficient is negative, since the derived determinant is the Δ of an analogous simplex in two fewer dimensions. Hence, when k_{ij} is increased to such an extent as to make $\Delta \leq 0$, further increment will certainly leave $\Delta < 0$.

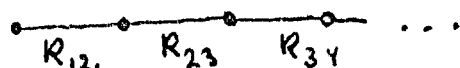
It is clear that the limiting cases where $\Delta = 0$ are just the cases where the simplex is Euclidean instead of spherical. These, of course, give the infinite groups that we seek.

When a group, or its fundamental region, is represented by a graph, the sub-graph derived by omitting one of the dots represents a vertex angle of the fundamental region. Regarded as a simplex of one fewer dimensions, this must be definitely spherical, even when the whole fundamental region is Euclidean. Thus we can never obtain an admissible graph by adding new links to a graph that represents a Euclidean simplex.

Suppose we know that a certain graph represents a proper spherical simplex. By increasing one of the marks k_{ij} we obtain new simplexes which remain spherical as long as Δ remains positive. But as soon as Δ becomes zero or negative, further increase of k_{ij} is useless.

A graph consisting of several separate pieces represents a spherical simplex if and only if each of the pieces represents a spherical simplex. For, in such a case we have, in Euclidean $(m+n)$ -space, $m+n$ concurrent hyperplanes of which the first m are all perpendicular to all the last n . The first m hyperplanes generate a group which operates essentially in an m -space (namely the m -space common to the last n), and vico versa. Thus, in the enumeration, it is sufficient to consider connected graphs.

The simplest type of graph is the simple chain



In this case

$$\Delta = \begin{vmatrix} 1 & -\cos \pi/R_{12} & 0 & \dots & 0 \\ -\cos \pi/R_{12} & 1 & -\cos \pi/R_{23} & \dots & 0 \\ 0 & -\cos \pi/R_{23} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}$$

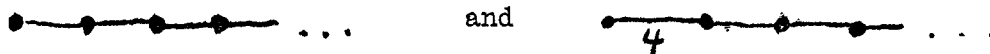
and we have to determine the admissible values of $k_{12}, k_{23}, \dots$. We note first that $\Delta > 0$ when

$$k_{12} = k_{23} = k_{34} = \dots = 3$$

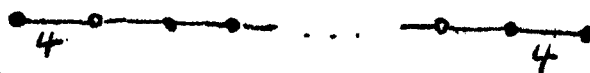
and when

$$k_{12} = 4, k_{23} = k_{34} = \dots = 3,$$

and that $\Delta = 0$ when both the first link and the last are marked 4. Thus



are finite groups, while

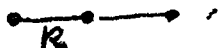


is infinite (Euclidean). E.g.

is infinite.

For the rest, we have to consider the cases of few dimensions. 

is finite for all values of k , since $\Delta = \sin^2 \pi/R$. In the case of



Δ remains positive when $k = 5$, but vanishes when $k = 6$.

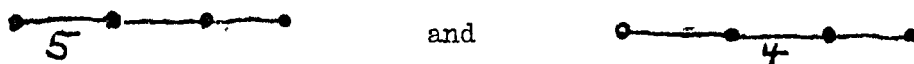
In the case of

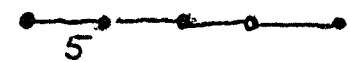


Δ is again positive when $k = 5$, but it is negative when $k = 6$. In the case of



$\Delta > 0$ when $k = 4$, but $\Delta < 0$ when $k = 5$. We thus have two more finite groups:



Now,  makes $\Delta < 0$, so the only remaining possibility is



which makes Δ vanish, giving one more infinite group.

The next simplest graph is the polygon



or



or

...

In these cases, when the links are unmarked,

$$\Delta = \begin{vmatrix} 1 & -\frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{1}{2} & 0 & 0 & \dots & -\frac{1}{2} & 1 \end{vmatrix} = 0$$

Thus we have another infinite family of infinite groups. It follows that none of the links may be marked, and that no graph may contain a circuit of links. I.e., apart from these polygons of unmarked links, every admissible graph is a tree.

If there is just one dot where three links occur, we may first consider the case when three chains, of n , p , q (unmarked) links respectively, emanate from this one dot. E.g., when $n = p = q = 1$, we have



It can be proved by induction that

$$\Delta = (n + p + q - npq + 2)/2^{n+p+q+1}.$$

Hence the group is finite when

$$n + p + q + 2 > npq$$

and infinite when $n + p + q + 2 = npq$.

The actual cases are as follows.

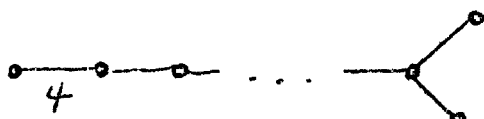
Finite groups

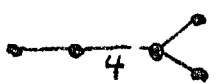
n	p	q
n	1	1
2	2	1
3	2	1
4	2	1

Infinite groups

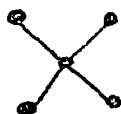
n	p	q
2	2	2
3	3	1
5	2	1

We next ask whether any link of these graphs can be marked 4. in the case of



Δ vanishes. Thus we have another family of infinite groups but no further possibilities. For example,  would consist of  with an added link.

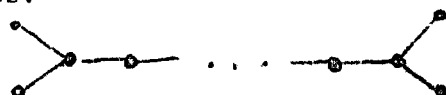
When four links meet at one dot, the simplest possibility is the only one:



is an infinite group, since

$$\Delta = \begin{vmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 1 \end{vmatrix} = 0$$

Finally, when there are two distinct branchings, we have a fourth family of infinite groups:



Since $\Delta = 0$ here, we cannot have any further branching, nor any marked links. The enumeration is now complete.

Collecting results, we have the following table:

Simple Continuous Groups	Infinite Discrete Groups	Finite Groups (S)	χ	h	g
L_2			1	2	2
L_3			2	3	6
$L_{\chi+1}$	(Graph: $(\chi+1)$ -gon)	(χ dots in a simple chain)	χ	$\chi+1$	$(\chi+1)!$
O_8			4	6	192
O_{10}			5	8	1920
$O_{2\chi}$			χ	$2(\chi-1)$	$2^{\chi-1}\chi!$
C_4			2	4	8
C_6			3	6	48
O_7			4	8	384
C_8			4	8	384
O_9			4	8	384
$C_{2\chi}$			χ	2χ	$2^{\chi}\chi!$
$O_{2\chi+1}$			χ	2χ	$2^{\chi}\chi!$
(G_2)			2	6	12
(F_4)			4	12	1152
(E_6)			6	12	51840
(E_7)			7	18	2903040
(E_8)			8	30	696729600
$---$			2	$2k$	$2k$
$---$			3	10	120
$---$			4	30	14400

Our infinite and finite groups are listed in the second and third columns. We have arranged these columns so that each finite group (save the last three) is a subgroup of the adjacent infinite group (or groups), namely the symmetry group S of the "vector diagram". The first column gives the corresponding Lie groups. (The symbols in parentheses are in van der Waerden's notation.) χ is the number of (Euclidean) dimensions, which is also the number of generating reflections for the finite group (i.e., the number of dots in the graph for the finite group). The last column gives the order of the finite group (the s of sec. 5). We still have to define h .

It can be proved by induction that, as long as the graph is a tree (as it is in all the finite cases), the $\chi!$ continued products of the generating reflections are all conjugate operations. We define h as the period of any one of these operations. For an infinite group this is infinite (even when the graph is a polygon, in which case the various products may not all be conjugate).

It can also be proved (by detailed consideration of the separate cases) that the total number of reflections in the finite group is $\frac{1}{2}\chi h$, so that the number of vectors in the vector diagram is χh , and the number of parameters of the continuous group is

$$\chi(h+1).$$

(We shall give a third property of h later on.)

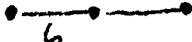
The theorem that enables us to associate a particular finite group with each infinite group is the following:

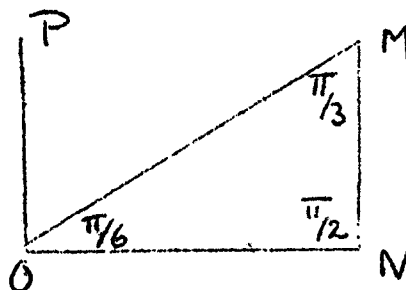
When we regard the reflecting hyperplanes of an infinite group generated by reflections as falling into families of parallel hyperplanes, the fundamental region has at least one vertex that lies in one hyperplane of every family.

In Annals of Math. 35 (1934), p. 591, I proved that the most general finite fundamental region for a group generated by reflections is a rectangular

product of simplexes. Therefore it will be sufficient to consider the case when the fundamental region is just a simplex, in which case what we have to prove is that this simplex has at least one bounding hyperplane that is parallel to one of the derived hyperplanes through the opposite vertex. (By removing such a bounding hyperplane, we obtain the special finite subgroup of the infinite group. This is Weyl's group (S), since the families of parallel hyperplanes correspond to the pairs of opposite vectors of the vector diagram.)

For each of the four infinite families of groups, we shall be content to take $\chi = 5$ or 6, since the situation for other values of χ will then be quite evident. Empirically, we can find equations for the bounding hyperplanes of the fundamental region in terms of Cartesian coordinates, χ , $\chi+1$, or (in one case) $\chi+2$, in number. (The extra coordinates are introduced for the sake of elegance.) In each case we first find $\chi+1$ hyperplanes through the origin, inclined at the proper angles, and then translate one of these hyperplanes away from the origin (by adding 1 to one side of the equation), so as to make a proper simplex. What we shall prove is that the reflections in the remaining χ hyperplanes suffice to restore this special hyperplane in its original position. E.g.

in the case of , with fundamental region MNO, the line OP, parallel to NM, can be derived from OM by a rotation about O that is a combination of reflections in OM and ON. Thus the special finite subgroup is

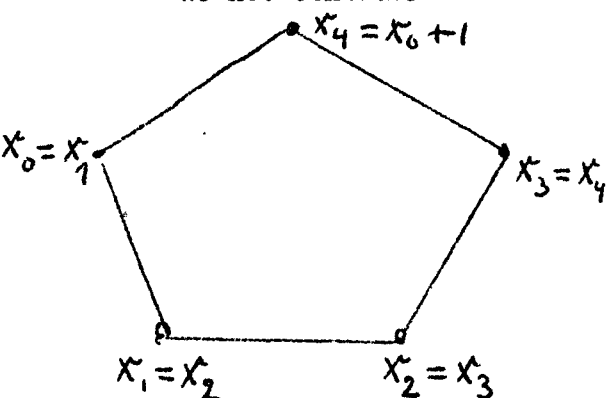


 (derived by omitting the final dot), which is generated by reflections in OM and ON. There is, on the other hand, no reflecting line through M parallel to OM. But when the fundamental region has intrinsic symmetry, as in the case of , the special hyperplane can be chosen in several ways, and we select any particular one for convenience.

The following obvious facts will be used repeatedly.

- (1) The reflection in the hyperplane $x_i = x_j$ is the transposition $(i j)$.
- (2) The transpositions $(1 2), (2 3), \dots, (n-1, n)$ generate the symmetric group on the n suffixes, and so enable us to replace any of these suffixes by any other (so long as we avoid repetition).

We not consider the actual cases, in turn.

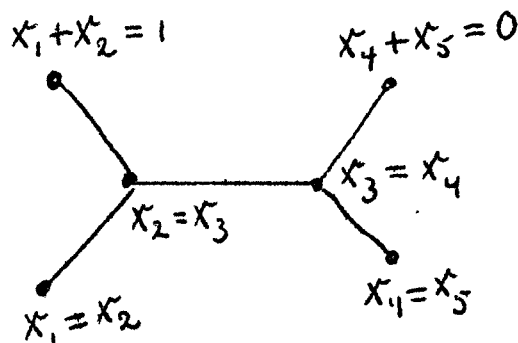


Here the special hyperplane is

$$x_4 = x_0 + 1, \text{ since the transpositions } (0 1), (1 2), (2 3), (3 4)$$

enable us to derive the hyperplane $x_4 = x_0$ from any one of the other four. (All these hyperplanes are perpendicular to $\sum x = 0$:

so the group really operates in four dimensions, not five.)

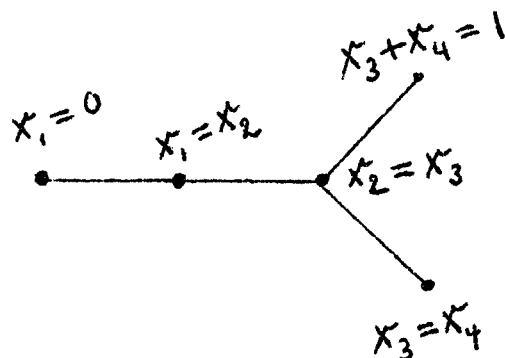


Here the special hyperplane is

$$x_1 + x_2 = 1, \text{ since the transpositions } (1 2), (2 3), (3 4), (4 5)$$

enable us to derive $x_1 + x_2 = 0$ from $x_4 + x_5 = 0$. (Of course we could equally well have made any other of the four outer

dots represent the special hyperplane.)

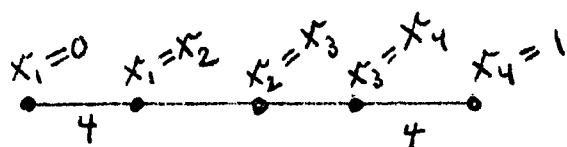


Special hyperplane $x_3 + x_4 = 1$, since the transpositions

$$(1 2), (2 3), (3 4)$$

enable us to derive $x_3 + x_4 = 0$ from $x_1 + x_2 = 0$, which is the image of

$$x_1 = x_2 \text{ in } x_1 = 0.$$



Special hyperplane $x_4 = 1$.

($x_1 = x_2 + 1$ special.)

($x_3 = x_4 + 1$ special.)

Here we introduce two extraneous coordinates and make all our hyperplanes perpendicular to $x_0 + x_7 = 0$ and to $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 0$. The reflection in the complicated hyperplane may be called $[0123 \cdot 4567]$;

it behaves like the "bifid substitution"

that arises in connection with the 28 bitangents of the plane quartic curve.

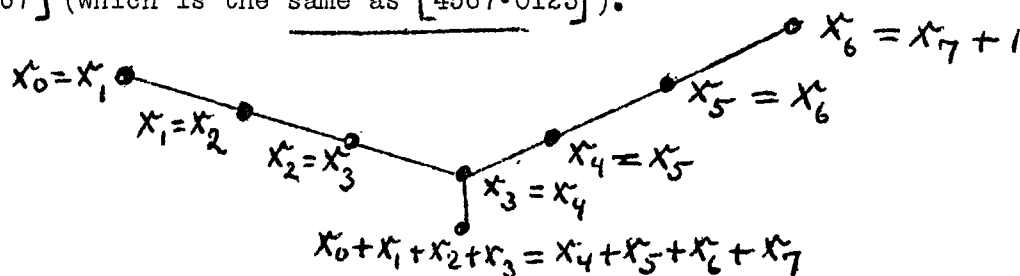
Since

$$((07) [0123 \cdot 4567])^3 = 1,$$

we have

$$\begin{aligned} [0123 \cdot 4567] (07) [0123 \cdot 4567] &= (07) [0123 \cdot 4567] (07) \\ &= [7123 \cdot 4560] \\ &= (14)(25)(36) [4567 \cdot 0123] (14)(25)(36). \end{aligned}$$

Thus the symmetric group on 1, 2, 3, 4, 5, 6 enables us to express (07) in terms of $[0123 \cdot 4567]$ (which is the same as $[4567 \cdot 0123]$).



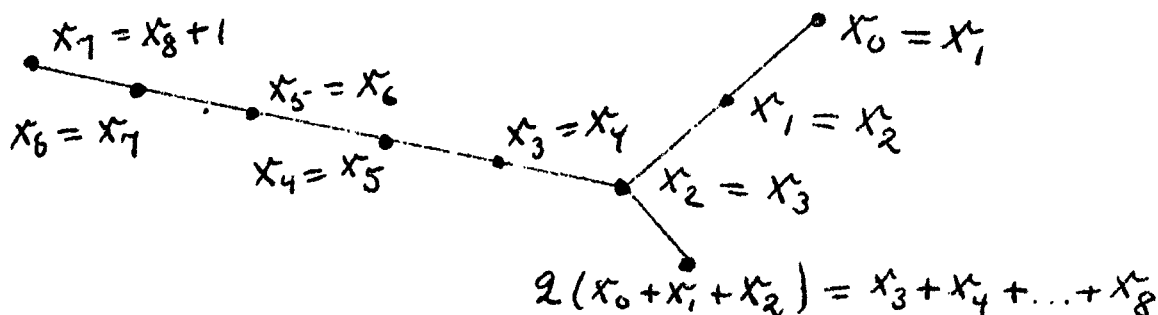
Here we observe that

$$(67) = [0\ 1\ 2\ 6 \cdot 4\ 5\ 3\ 7] [0\ 1\ 2\ 7 \cdot 4\ 5\ 3\ 6] [0\ 1\ 2\ 6 \cdot 4\ 5\ 3\ 7]$$

where $[0\ 1\ 2\ 6 \cdot 4\ 5\ 3\ 7] = (3\ 6) [0\ 1\ 2\ 3 \cdot 4\ 5\ 6\ 7] (3\ 6)$

and $[0\ 1\ 2\ 7 \cdot 4\ 5\ 3\ 6] = (0\ 4)(1\ 5)(2\ 6) [4\ 5\ 6\ 7 \cdot 0\ 1\ 2\ 3] (0\ 4)(1\ 5)(2\ 6)$

Hence the symmetric group on 0, 1, 2, 3, 4, 5, 6 enables us to express (6 7) in terms of $[0\ 1\ 2\ 3 \cdot 4\ 5\ 6\ 7]$.



Calling the complicated reflection V_{012} , we observe that

$$\begin{aligned} (7\ 8) &= V_{017} V_{018} V_{017} \\ &= V_{017} V_{234} V_{567} V_{234} V_{017}. \end{aligned}$$

All these V's can be derived from V_{012} by means of the symmetric group on 0, 1, 2, 3, 4, 5, 6, 7.

Let us now transform to oblique axes, with coordinates $\varphi_1, \varphi_2, \dots, \varphi_\lambda$,

in such a way that the bounding hyperplanes of the Euclidean simplex are

$$\varphi_1 = 0, \dots, \varphi_\lambda = 0,$$

and

$$a_1 \varphi_1 + \dots + a_\lambda \varphi_\lambda = 1.$$

We shall normalize these coordinates in such a way that the nearest hyperplanes parallel to

$$\varphi_1 = 0, \dots, \varphi_\lambda = 0$$

are precisely

$$\varphi_1 = \pm 1, \dots, \varphi_\lambda = \pm 1.$$

In other words, the totality of reflecting hyperplanes parallel to $\varphi_i = 0$ shall be $\varphi_i = k$, where k runs through all integer values (positive, zero and negative).

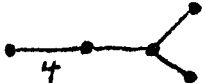
Whenever all the links in the graph are unmarked, this normalization is easily achieved, since then all the generating reflections are conjugate. (In fact, if P, Q are two reflections, represented by dots that are joined by an unmarked link, we have $PQP = QPQ$. Repeated application of this principle proves our result.) In this case the φ 's, when expressed in terms of the x 's, all have the same sum of squares of coefficients. For the sake of symmetry, we define

$$a_0 = 1, \quad \varphi_0 = 1 - a_1 \varphi_1 - \dots - a_\chi \varphi_\chi,$$

so that

$$a_0 \varphi_0 + a_1 \varphi_1 + \dots + a_\chi \varphi_\chi = 1.$$

We are now using homogeneous coordinates based on the fundamental region.

Let λ denote the number of ways in which the "special" hyperplane can be chosen; e.g. 4 for  or , 2 for . Then it turns out that λ of the a 's are equal to 1, while the remaining $\chi + 1 - \lambda$ are greater integers. Professor Weyl (who writes f for λ) has proved (sec. 5, Theorem 10) that the order of the corresponding finite group is

$$g = \lambda \cdot \prod a_i$$

(Since $a_0 = 1$, it does not matter whether we do or do not include a_0 in the product of a 's.) A new result, which has not yet been explained in general terms, is that

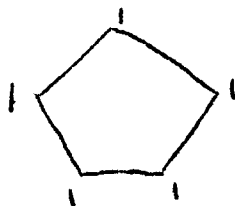
$$h = a_0 + a_1 + \dots + a_\chi.$$

This may be stated as follows:

The number of parameters of a simple continuous group is equal to

$$\chi(1 + a_0 + a_1 + \dots + a_\chi).$$

We proceed to consider the individual cases, marking the dots of each graph with the values of the relevant a 's.



$$\lambda = \chi + 1.$$

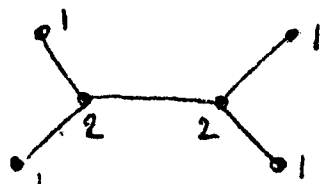
$$\prod a = 1,$$

$$g = (\chi + 1)!$$

$$\varphi_i = x_{i-1} - x_i \quad (i = 1, \dots, \chi),$$

$$x_0 - x_\chi = \varphi_1 + \varphi_2 + \dots + \varphi_\chi.$$

$$h = \chi + 1,$$



$$\lambda = 4,$$

$$\prod a = 2^{\chi-3},$$

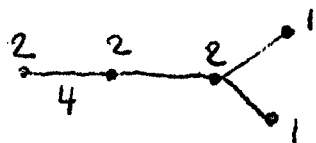
$$g = 2^{\chi-1} \chi!$$

$$\varphi_i = x_i - x_{i+1} \quad (i = 1, \dots, \chi-1),$$

$$\varphi_\chi = x_{\chi-1} + x_\chi.$$

$$x_1 + x_2 = \varphi_1 + 2\varphi_2 + 2\varphi_3 + \dots + \varphi_{\chi-1} + \varphi_\chi.$$

$$h = 2\chi - 2,$$



$$\lambda = 2,$$

$$\prod a = 2^{\chi-1},$$

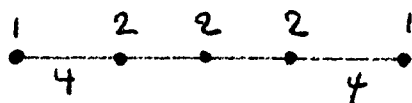
$$g = 2^\chi \chi!$$

$$\varphi_1 = x_1,^* \quad \varphi_i = x_i - x_{i-1} \quad (i = 2, \dots, \chi),$$

$$x_\chi + x_{\chi-1} = 2\varphi_1 + 2\varphi_2 + \dots + 2\varphi_{\chi-1} + \varphi_\chi.$$

$$h = 2\chi,$$

* $x_1 = 1$ is the image of $x_2 = 0$ in $x_1 + x_2 = 1$.

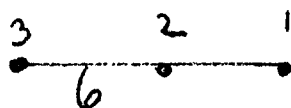


(h , λ and g as in the previous case.)

$$\varphi_1 = x_1, \quad \varphi_i = \frac{x_i - x_{i-1}}{2} \quad (i = 2, \dots, \chi),^{**}$$

$$x_\chi = \varphi_1 + 2\varphi_2 + \dots + 2\varphi_\chi.$$

** $x_2 - x_1 = 2$ is the image of $x_1 + x_2 = 0$ in $x_2 = 1$.



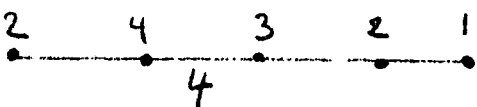
$$\lambda = 1,$$

$$g = 3 \cdot 2 \cdot 2! = 12.$$

$$\varphi_1 = \frac{2x_0 - x_1 - x_2}{3}, \quad \varphi_2 = x_1 - x_0.$$

$$x_1 - x_2 = 3\varphi_1 + 2\varphi_2.$$

$$h = 1 + 3 + 2 = 6$$



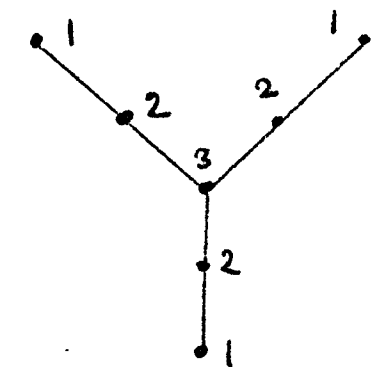
$$\varphi_1 = \frac{x_1 + x_2 + x_3 + x_4}{2}, \varphi_2 = -x_1, \varphi_3 = x_1 - x_2, \\ \varphi_4 = x_2 - x_3.$$

$$x_4 - x_3 = 2\varphi_1 + 4\varphi_2 + 3\varphi_3 + 2\varphi_4.$$

$$h = 1 + 2 + 4 + 3 + 2 = 12,$$

$$\lambda = 1,$$

$$g = 2 \cdot 4! \cdot 3 \cdot 2 \cdot 4! = 1152.$$



$$\varphi_i = x_{i+1} - x_i \quad (i = 1, \dots, 5)$$

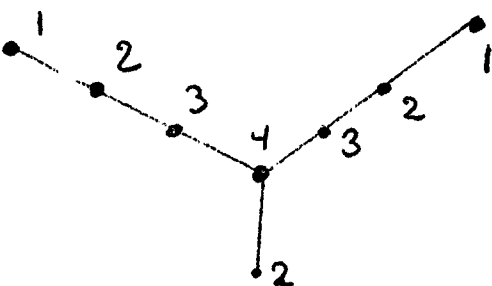
$$\varphi_6 = \frac{x_0 + x_1 + x_2 + x_3 - x_4 - x_5 - x_6 - x_7}{2}.$$

$$x_0 - x_7 = \varphi_1 + 2\varphi_2 + 3\varphi_3 + 2\varphi_4 + \varphi_5 + 2\varphi_6.$$

$$h = 3 + 3 \cdot 2 + 3 = 12,$$

$$\lambda = 3, \quad g = 3 \cdot 2^3 \cdot 3 \cdot 6! = 72 \cdot 6!$$

$$= 51840.$$



$$\varphi_i = x_i - x_{i-1} \quad (i = 1, \dots, 6),$$

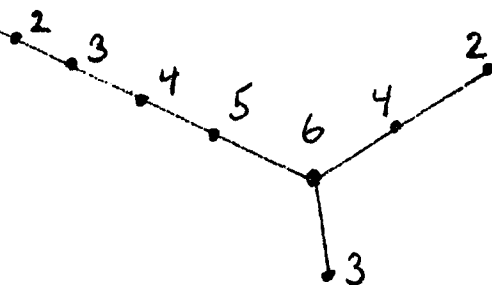
$$\varphi_7 = \frac{x_0 + x_1 + x_2 + x_3 - x_4 - x_5 - x_6 - x_7}{2}.$$

$$x_6 - x_7 = \varphi_1 + 2\varphi_2 + 3\varphi_3 + 4\varphi_4 + 3\varphi_5 + 2\varphi_6 + 2\varphi_7.$$

$$h = 2 + 3 \cdot 2 + 2 \cdot 3 + 4 = 18,$$

$$\lambda = 2,$$

$$g = 2 \cdot 2^3 \cdot 3^2 \cdot 4 \cdot 7! = 8 \cdot 9! = 2903040.$$



$$\varphi_i = x_i - x_{i-1} \quad (i = 1, \dots, 7),$$

$$\varphi_8 = \frac{2(x_0 + x_1 + x_2) - x_3 - \dots - x_8}{3}$$

$$x_7 - x_8 = 2\varphi_1 + 4\varphi_2 + 6\varphi_3 + 5\varphi_4 + 4\varphi_5 + 3\varphi_6 + 2\varphi_7 + 3\varphi_8.$$

$$h = 1 + 2 \cdot 2 + 2 \cdot 3 + 2 \cdot 4 + 5 + 6 = 30,$$

$$\lambda = 1,$$

$$g = 2^2 \cdot 3^2 \cdot 4^2 \cdot 5 \cdot 6 \cdot 8! = 192 \cdot 10!$$

$$= 696729600.$$

Let M denote the matrix of coefficients of the x 's in the expressions for $\varphi_0, \varphi_1, \dots, \varphi_\chi$, so that

$$(a_0, a_1, \dots, a_\chi)M = (0, 0, \dots, 0).$$

In the cases when all links in the graph are unmarked, let S denote the sum of squares of coefficients in each φ_i , and M' the transposed of M . Then

$$(0, 0, \dots, 0) = (a_0, a_1, \dots, a_\chi) \frac{MM'}{S} = \begin{pmatrix} 1 & -\cos \pi/R_{01} & \dots & -\cos \pi/R_{0\chi} \\ -\cos \pi/R_{01} & 1 & \dots & -\cos \pi/R_{1\chi} \\ \dots & \dots & \dots & \dots \\ -\cos \pi/R_{0\chi} & -\cos \pi/R_{1\chi} & \dots & 1 \end{pmatrix} \cdot (a_0, a_1, \dots)$$

Here each $k_{ij} = 3$ or 2 according as the i^{th} and j^{th} dots in the graph are or are not linked. The a 's can accordingly be written down successively by a very simple rule:

$$2a_i = a_j + a'_k + \dots$$

whenever the i^{th} dot is linked to the j^{th} , k^{th} , etc. E.g., in the last case that we considered,

$$2 \cdot 1 = \underline{2}, \quad 2 \cdot 2 = 1 + \underline{3}, \quad 2 \cdot 3 = 2 + \underline{4}, \quad 2 \cdot 4 = 3 + \underline{5}, \quad 2 \cdot 5 = 4 + \underline{6}, \quad 2 \cdot 3 = 6, \quad 2 \cdot 6 = 5 + \underline{4} + 3,$$

$$2 \cdot 4 = 6 + \underline{2}, \quad 2 \cdot 2 = 4.$$

DuVal* has shown that each graph of unmarked links that represents a

* Proc. Camb. Phil. Soc. 30 (1934), pp. 453-465, 483-491. See especially p. 456.

finite group, also represents the tree of neighborhoods of a singular point on an algebraic surface. The numbers $a_1, a_2, \dots, a_\chi$ are precisely the coefficients of the partial neighborhoods in the expression for the total neighborhood of the singular point; e.g.

$$6 \cdot 0 + 5 N_1 + 4 (N_2 + P_1) + 3 (N_3 + Q) + 2 (N_4 + P_2).$$