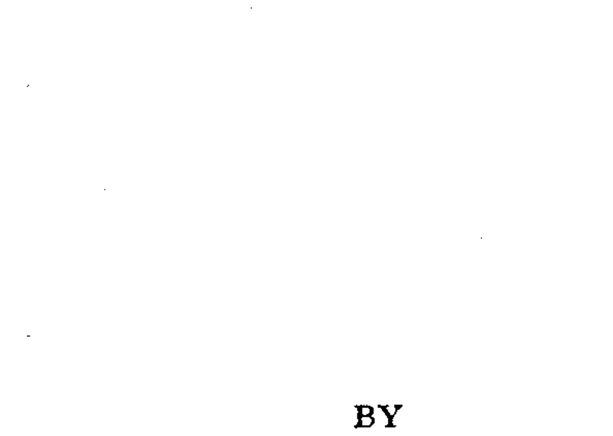
MATSCIENCE REPORT 25

LECTURES ON
LOCAL LIE GROUPS AND THEIR REPRESENTATIONS

BY

L. O. RAIFEARTAIGH,
Visiting Member, MATSCIENCE



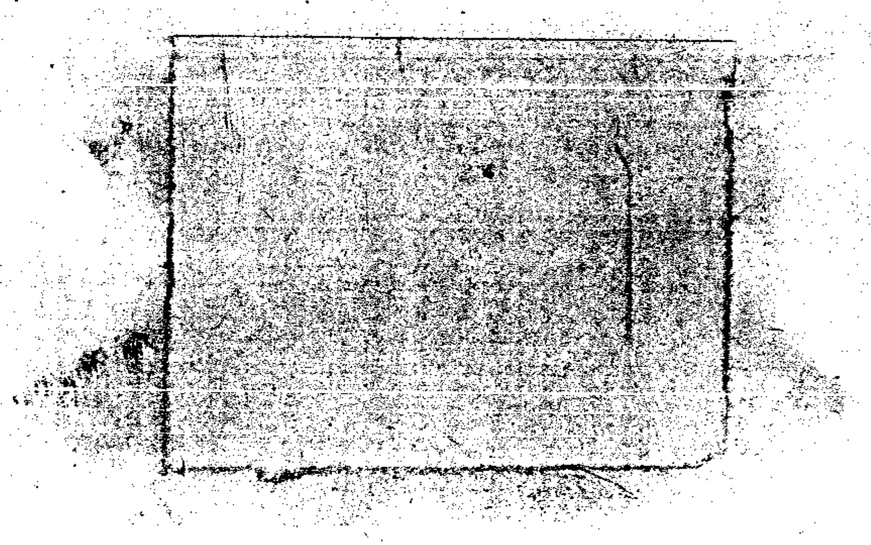
THE INSTITUTE OF MATHEMATICAL SCIENCES, MADRAS-20, INDIA.

MATSCIENCE REPORT 25

THE INSTITUTE OF MATHEMATICAL SCIENCES
MADRAS - 4 (India)

LECTURES ON LOCAL LIE GROUPS AND THEIR REPRESENTATIONS.

By
L.O'Raifeartaigh+
Visiting Member (December 1963 - February 1964),
MATSCIENCE, Madras.



+ Permanent Address: Dublin Institute for Advanced Studies,
DUBLIN, Ireland.

INTRODUCTION

The following series of 24 lectures delivered at the Madras Institute of Mathematical Sciences was aimed at providing a fairly extensive working knowledge of the theory of local Lie groups and their representations for physicists..

The series falls roughly into 3 parts. Part I cover the general theory of abstract local Lie groups, and the associated transformation groups. It ends with the introduction of the Lie algebra

$$[X_\mu , X_\nu] = C^{\lambda}_{\mu\nu} X_\lambda \qquad (1)$$

The mathematics used in this part does not go beyond that used in ordinary functional analysis.

Part II starts with the relation (1), and for the case of semi-simple Lie groups (defined in Part I), it deals with the construction of their irreducible representations. The mathematics used in this section does not go beyond the theory of linear operators (matrices) on vector and metric spaces, and should be familiar to every student of quantum mechanics.

Part III deals with a particular problem in the representation theory of Part II, namely, the problem of constructing all the independent invariants for any given Lie group. This part has been included in the lecture notes, although, in fact, time did not permit it to be given as a set of lectures.

It is clear from the last paragraph that the mathematical language used in these notes is somewhat elementary from the ''professional'' group theoreticians point of view. However, the

use of such elementary mathematics is deliberate, and is due to the fact that the course is aimed primarily at physicists, who do not always have the time necessary to make themselves familiar with the more sophisticated ideas and notations of modern group theory.

A more serious criticism of the present series is that, in order to simplify the presentation, the results used at one or two crucial points in the series, have been quoted rather than proved. However, all the results quoted can be proved using elementary mathematical methods (see, for example, ref$^{(1)}$), and the omission of the proof is only in the interest of simplicity. I have also tried to keep the number of unproven results down to a bare minimum.

Throughout the lecture notes, questions which were asked by members of the audience during or after the lectures have been included as part of the text, since these questions and the discussions which they inaugurated seemed to me to be the most important part of the whole series. Some of the questions have been included in a simple Q:A (Question and Answer) form, although it is rarely necessary to say that the answers to the questions did not come as promptly as would appear from the text. Other questions were such that it was more convenient to include them implicitly in the text. In fact, it is not too much to say that the final form of these notes was largely determined by the questions asked at the lectures, and I should like to take this opportunity of thanking all those who attended the lectures for their kind and persevering attention, and for the many fruitful

discussions which followed their various questions and remarks.

The sources from which most of the results presented in this series have been drawn should also be mentioned. These consist of the short list of standard works given on the next page. These are recommended to anybody who wishes to go more deeply into the theory than I have gone here.

Finally, I should like to express my thanks to Professor Alladi Ramakrishnan who invited me to come to Matscience and who suggested that I give this series of lectures. My thanks are also due to the Governing Board of the Dublin Institute for Advanced Studies for allowing me leave of absence to visit Matscience.

Lochlainn O'Raifeartaigh,
1--3--1964.

Standrad Works

1. Hermann Boerner;
   Representations of Groups,
   North Holland Publishing Company,
   Amsterdam, 1963.

2. Nathan Jacobson;
   Lie Algebras,
   Interscience Publishers,
   New York, 1962.

3. E.B.Dynkin,
   American Mathematical Society,
   Trans. I, 9.

4. G. Racah,
   Princeton Lecture Notes,
   CERN 61-8 (1961).

5. Weyl, H.,
   Theory of Groups and Quantum Mechanics,
   Princeton University Press,
   Princeton, 1931.

6. Weyl, H.,
   The Classical Groups,
   Princeton University Press,
   Princeton, 1946.

7. Morton. Hamermesh,
   Group Theory and its application to Physical problems,
   Addison-Wesley Publishing Company, INC,
   U.S.A., 1962.

8. Pontrjagin, L.S.,
   Continuous Groups,
   Moscow, 1938.

9. Pontrjagin, L.S.,
   Topological Groups,
   Princeton 1946.
   (Princeton Math. Series, Noh).

CONTENTS

Part I.

General theory of Groups ..	Lecture 1	
General theory of local Lie Groups	Lectures 2 - 7	
Transformation associated with local Lie groups	Lectures 7	"a" Series
Linear transformations ..	Lecture 8	"a" Series
Lie algebras ..	Lectures 8-9	"a" Series

Part II.

Reduction of semi-simple Lie Algebras to their canonical form	Lectures 10-13	"b" Series
Root diagrams ..	Lectures 14-15	"b" Series
Classification of semi-simple Lie algebras ..	Lectures 15-16	"b" Series
Weight diagrams ..	Lectures 17-18	
Representations: C_ℓ ..	Lecture 19	"c" series
A_ℓ ..	Lecture 20	"c" series
B_ℓ ..	Lecture 21	"c" series
D_ℓ ..	Lecture 22	"c" series

Part III.

Construction of the invariants for the simple Lie algebras.		

Definition:

Let $G = \{\ldots g \ldots\}$ be any set of elements g. Let $g \cdot g'$ denote any operation which is defined for every pair g and g' in G. If $g \cdot g'$ is again a member of G i.e., if

$$g \cdot g' = g'' \in G, \qquad (1.1)$$

then G is said to be closed under the operation $g \cdot g'$.

The set G are said to form a group with respect to any operation $g \cdot g'$, under which G is closed, if the operation satisfies also the following three subsidiary conditions:

(1) It is associative i.e.,

$$g \cdot (g' \cdot g'') = (g \cdot g') \cdot g'', \text{ all } g, g', g'' \in G. \qquad (1.2)$$

(2) There exists a (unit) element $g_0 \in G$ such that

$$g_0 \cdot g = g \cdot g_0 = g, \text{ all } g \in G. \qquad (1.3)$$

(3) For each $g \in G$ there exists an (inverse) element $g^{-1} \in G$ such that

$$g \cdot g^{-1} = g^{-1} \cdot g = g_0, \quad \text{all} \quad g \in G \qquad (1.4)$$

The number of elements in a group may be finite, denumerably infinite or non-denumerably infinite. If it is finite, the number of elements is called the order of the group.

Remarks: A group consists therefore of one set of elements and one operation. The operation is often called 'group multiplication' though it may have nothing to do with ordinary multiplication. Note that it is necessary to say with respect to which operation a set forms a group,

since any given set might form a group with respect to one operation but not with respect to another. For example, the set of four complex numbers 1, i, -1, -i form a group with respect to ordinary multiplication, but with respect to ordinary addition the set is not even closed

With regard to the subsidiary conditions, we note first that if the unit element exists at all, it is unique, since if g_0 and g_0' are two unit elements, then, by definition,

$$g_0 = g_0 \cdot g_0' = g_0' . \tag{1.5}$$

Secondly, we note that for the inverse elements to exist, the unit element must already exist. If all of these exist and if the associative law (2) is satisfied, then the inverse elements are unique. For let g' and g'' be inverse to g. Then

$$g' = g' \cdot (g \cdot g'') = (g' \cdot g) \cdot g'' = g'' . \tag{1.6}$$

From the uniqueness of the inverse it follows, in particular, that

$$(g \cdot g')^{-1} = g'^{-1} \cdot g^{-1} . \tag{1.7}$$

Q: Can you give an example of a non-associative operation?

A: Yes. A non-associative operation which we shall be meeting later in the theory of Lie algebras is the operation of commutation, i.e.,

$$A \cdot B = [A, B] = AB - BA , \tag{A.1}$$

for matrices. This is non-associative since

$$[[A,B],C] \neq [A,[B,C]] , \tag{A.2}$$

in general. In fact, the associative law is replaced for this operation by the Jacobi identity

$$\left[C,\left[A,B\right]\right]+\left[B,\left[C,A\right]\right]+\left[A,\left[B,C\right]\right]=0\ . \tag{A.3}$$

A set of elements and an operation which satisfy all the group conditions except the inverse condition (3) are said to form a semi-group.

Finally, we note that it is not necessary for a group operation to be commutative i.e., to satisfy the relations

$$g\cdot g'=g'\cdot g\ ,\ \text{all}\ \ g,g'\in G\ . \tag{1.8}$$

If it does not satisfy these relations the group is said to be abelian. It is interesting to note that any finite group of order $N\leq 5$ must be abelian.

Examples: (1) Consider the set of integers 1, 2, ... 9. Under ordinary addition this set is not closed, since, for example, 6 + 5 = 11. Thus under ordinary addition this set cannot form a group. Consider, however, the set of all positive integers. This set is certainly closed under addition. Furthermore, addition satisfies the associative law $a+(b+c)=(a+b)+c$ and even the commutative law $a+b=b+a$. Nevertheless, the set of all positive integers does not form a group with respect to addition, since it contains no unit element i.e., there is no positive integer a such that

$$a+b=b+a=b \tag{1.9}$$

for all positive integers b. This defect can be overcome by

widening our set to include zero. However, by doing this we obtain only a semi-group, because the ''inverse'' condition (3) is still not satisfied i.e., given any non-negative integer a, there exists no non-negative integer b such that

$$a + b = 0 \quad , \tag{1.10}$$

(except for $a = 0$). However, if we widen our set still further to include also the negative integers, then we obtain a group. In other words, the set of all integers (positive, negative and zero) forms a group under ordinary addition. The set of all non-negative integers (or the set of all non-positive integers) forms a semi-group.

(2) Consider the set of all $(n \times n)$ matrices. Under the operation of matrix multiplication this set is closed, since the product of 2 $(n \times n)$ matrices is again an $(n \times n)$ matrix. Further matrix multiplication satisfies the associative (though not, in general, the commutative) law, and the unit $(n \times n)$ matrix plays the role of a unit element. Thus this set forms a semi-group under matrix multiplication. It does not form a group, however, because the inverse of an $(n \times n)$ matrix will respect to matrix multiplication does not exist unless the matrix happens to be non-singular.

Suppose, however, we restrict our original set to include only all non-singular $(n \times n)$ matrices. Then, as is easily verified, we obtain a group. Similarly if we consider the set of all <u>unimodular</u> $(n \times n)$ matrices i.e., all $(n \times n)$ matrices with determinant = 1, and the operation of matrix multiplication, we

obtain a group (the product of 2 unimodular matrices and the inverse of a unimodular matrix are again unimodular matrices, as is the unit matrix). One easily verifies that the same holds for the set of all unitary $(n \times n)$ matrices, all unitary unimodular $(n \times n)$ matrices, and the set of all real non-singular $(n \times n)$ matrices. However, the set of all traceless $(n \times n)$ matrices do not form a group with respect to matrix multiplication, since under this operation the set is not even closed (the product of two traceless matrices is not always traceless). The same is true of the set of all hermitian $(n \times n)$ matrices.

With respect to the operation of matrix addition, however, the position is reversed. The set of all $(n \times n)$ matrices, the set of all traceless $(n \times n)$ matrices and the set of all hermitian $(n \times n)$ matrices form groups. But the set of all unimodular matrices, all unitary matrices, all unitary unimodular matrices and all real non-singular matrices do not.

(3) In the above two examples we have mentioned only groups which have an infinite number of elements (denumerably infinite in the first case, non-denumerably infinite in the second). But, of course, there are many important groups with a finite number of elements. An example of such a group is:
The group of perturbations of n objects. The set here is the set of $n!$ possible permutations of the n objects, and the group operation is simply the operation of following one permutation by another. The order of the group is obviously $n!$ and it is non-abelian for $n > 2$. In particular it is non-abelian for

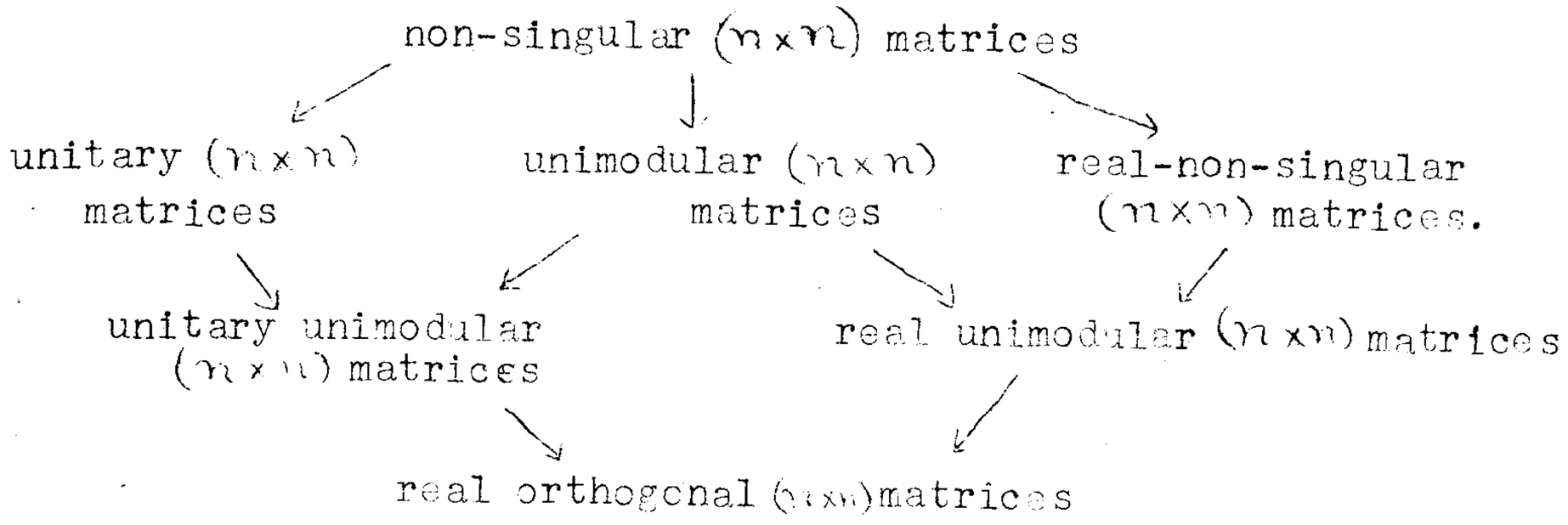
$n = 3$, and this furnishes an example of a non-abelian group of the lowest possible order, namely 6 (= 3!).

Subgroups: If a subset of the set of elements of a group is itself a group (with respect to the group operation of the original group) it is called a subgroup. Note that all subgroups contain the unit element of the group and so overlap.

Examples of subgroups:

(1) Consider the group of all integers (with respect to ordinary addition) mentioned above. It is easy to verify that a subgroup of this is the set of all even integers (the addition of two even integers and the inverse (negative) of an even integer are again even integers, as is the unit element, zero).

(2) Consider the group of non-singular $(n \times n)$ matrices mentioned above. It is easy to verify that the other matrix groups mentioned, namely, the group of unitary $(n \times n)$ matrices, the group of unimodular $(n \times n)$ matrices and the group of real non-singular $(n \times n)$ matrices are all subgroups of this group. Furthermore, the group of unitary unimodular $(n \times n)$ matrices is a common subgroup of the first two of these, while the group of real unimodular matrices is a common subgroup of the second two. Finally, it is easy to verify that the group of real orthogonal $(n \times n)$ matrices is a common subgroup of all of these groups. Schematically:



(3) The group of permutations of any $m < n$ objects is clearly a subgroup of the group of permutations of n objects.

We shall now consider two particular types of subgroups.

(a) Abelian subgroups:

If a subgroup of a group is an abelian group it is said to be an abelian subgroup. It is clear that any subgroup of an abelian group is an abelian subgroup. However, a non-abelian group may contain an abelian subgroup. For example, the group of permutations of 2 objects is an abelian subgroup of the non-abelian group of permutations of 3 objects. Similarly the group of rotations around the z-axis in 3-space is an abelian subgroup of the non-abelian group of all rotations in 3-space.

(b) Invariant subgroups:

Let S be a subgroup of any group G i.e., if

$$g_s, g_{s'} \in S,$$

then

$$g_S'' = g_S \cdot g_S' \in S \tag{1.11}$$

If, in addition, for any $g_S \in S$, and any $g \in G$,

$$g\, g_S\, g^{-1} = g_{S'} \in S. \tag{1.12}$$

then the subgroup is said to be an invariant subgroup. Naturally, for every subgroup (1.12) holds for $g \in S$ (it is for such a sufficient) g, a weaker consequence of (1.11) but what characterizes an invariant subgroup is that (1.12) holds for all $g \notin S$ also. The element $g\, g_S\, g^{-1}$ in (1.12) is, of course, not necessarily equal to g_S itself, but only to some $g_S' \in S$.

Examples: (1) The group of even integer is an invariant subgroup of the group of all integers discussed above. This is because in this case

$$g\, g_S\, g^{-1} = m + 2n - m = 2n = g_S \tag{1.13}$$

where m and n are integers. Note that here,

$$g_S' = g_S.$$

This is a particular case of the general result that any subgroup of an abelian group is invariant, since for an abelian group

$$g\, g_S\, g^{-1} = g g^{-1} g_S = g_e g_S = g_S. \tag{1.14}$$

Note, however, that for (1.14) to be valid the full group must be abelian. If only the subgroup is abelian, it is not necessarily an invariant subgroup.

(2) We saw earlier that the group of all non-singular $(n \times n)$ matrices had the following three subgroups:

non-singular $(n \times n)$ matrices

unitary $(n \times n)$ matrices | <u>unimodular $(n \times n)$ matrices</u> | real non-singular $(n \times n)$ matrices.

... and we might ask which, if any, of these are invariant subgroups. The answer is that the group of unimodular $(n \times n)$ matrices is an invariant subgroup, while the other two are not.

<u>Proof:</u>

Let S be a unimodular matrix, and A any non-singular matrix. Then

$$\det (A S A^{-1}) = \det S = 1 , \tag{1.15}$$

So that $A S A^{-1}$ is unimodular. Thus the group of unimodular $(n \times n)$ matrices is an invariant subgroup.

On the other hand, if U is a unitary matrix, and A any non-singular matrix

$$A U A^{-1} \tag{1.16}$$

is not necessarily unitary. Thus the group of unitary $(n \times n)$ matrices is not an invariant subgroup.

Similarly, if R is a real unimodular matrix, and A is any non-singular matrix

$$A R A^{-1} \tag{1.17}$$

is not necessarily real. Thus the group of real unimodular matrices is not an invariant subgroup.

Having defined the abelian and invariant subgroups in the above way, we now use them to classify the various groups, as follows:

Simple Groups: Suppose that a group contains no invariant subgroups. It is then said to be simple.

Note that a simple group may contain any number of subgroups and even abelian subgroups.

Semi-simple Groups: Suppose that a group contains no abelian invariant subgroups. It is then said to be semi-simple.

Note that a semi-simple group may contain any number of subgroups, any number of abelian subgroups provided that they are not invariant, and any number of invariant subgroups provided that they are not abelian.

Clearly any simple group is automatically semi-simple. If a group is semi-simple without being simple it must contain at least one non-abelian invariant subgroup.

Later we shall see the very important role which is played by the semi-simple groups in the theory of Lie groups.

Lecture 2.

In this lecture and the following ones, we shall consider Lie groups only. We introduce them as follows:

Definition: If a group G is such that its elements g can be specified uniquely by the values of a finite set of parameters $\{x_1, x_2, \ldots x_r\}$ each of which has a continuous range of values, the group is set to be a Lie group. (subject to certain conditions of continuity, differentiability and connectedness to be discussed below.)

Thus for a Lie group

$$\{\ldots g \ldots\} = \{\ldots g_{(x_1, x_2, \ldots x_r)} \ldots\} = \{\ldots g_{(x_\alpha)} \ldots\} \quad (2.1)$$
$$= \{\ldots g_{(x)} \ldots\}, \quad \alpha = 1, \ldots r$$

The number of elements in a Lie group is, of course, infinite, and even non-denumerably infinite. Nevertheless, the Lie groups are often called finite continuous groups, to distinguish them from groups in which the number of parameters is also infinite. Similarly, the number of parameters which specify the elements uniquely, namely r, is called the order of the group. Note that we assume that we have a unique paramatrization (u. p.) i.e.,

$$g_{(x)} = g_{(y)} \longleftrightarrow x = y. \quad (2.2)$$

The multiplication law

$$g_{(x)} \cdot g_{(y)} = g_{(z)}$$

for a Lie group is most conveniently expressed by a relation of the kind

$$z_\alpha = \phi_\alpha (x_\beta, y_\gamma), \quad \alpha, \beta, \gamma = 1, \ldots, r \quad (2.3)$$

or, more briefly, by

$$z = \phi(x, y) \tag{2.4}$$

On account of the unique parametrization, the functions ϕ_α characterize the group completely, playing for the Lie group the same role as is played by the multiplication table for a finite group.

At this stage it might not be out of place to consider an example. Consider the group of all non-singular 2 x 2 matrices (group operation = matrix multiplication). An element $g_{(x)}$ of this group is specified by the four parameters in

$$g_{(x)} = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \quad (x_1 x_4 - x_2 x_3) \neq 0 . \tag{2.5}$$

The group multiplication

$$g_{(x)} \cdot g_{(y)} = g_{(z)}$$

becomes in this case

$$\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}, \tag{2.6}$$

from which

$$\begin{aligned} z_1 &= \phi_1(x_\alpha, y_\beta) = (x_1 y_1 + x_2 y_3), \\ z_2 &= \phi_2(x_\alpha, y_\beta) = (x_1 y_2 + x_2 y_4), \end{aligned} \tag{2.7}$$

etc.

Going back to the general Lie groups, we introduce the convention that the parameters x_α are so defined that to the unit element g_0 corresponds the value 0 for each parameter, i.e.

$$g_0 = g_{(0,0,\ldots 0)} = g_{(0)} . \tag{2.8}$$

Great simplification is achieved in the study of Lie groups if one confines oneself to local Lie groups, i.e., to those for which the parameters x_α are in the neighbourhood of the origin. This we shall do in these lectures, which may therefore be considered as lectures on local Lie groups. Such groups are automatically simply connected, the only problems we need worry about are the continuity and analyticity properties of the $\phi^i{}_{jk}$. We shall assume that all the $\phi^i{}_{jk}$ are analytic (i.e. differentiable to all orders) in all their $2r$ arguments in the neighbourhood of the origin. This is not a necessary assumption -- we shall indicate later how one could get away with assuming only differentiability to third order, and, in fact, I think it can now be shown that one need assume only continuity -- but it is a convenient assumption for our purposes.

It might be worth remarking at this stage that since we can talk about x being in the neighbourhood of the origin, we can talk about $g_{(x)}$ being in the neighbourhood of $g_{(e)}$, namely by defining it to be so when x is in the neighbourhood of 0. This means that the set $\{\cdots g_{(x)} \cdots\}$ has more than just the group property, it has a topological property as well. Thus a Lie group is not merely a special kind of group, it is a special kind of topological group.

Q: Do you assume that the parameters x_α are real?

A: No. It is usual to work with real parameters, but for local groups all of the results we shall prove go through for complex parameters also, as far as I know. If we meet one which does not we shall specialize to real parameters. Until then we shall keep an open mind.

What we now wish to do is to investigate the nature of the restrictions placed on the functions ϕ , by the three subsidiary group conditions (1.1), (1.2) and (1.3) . It will turn out that these conditions are very strong, so strong that all possible local Lie groups can be classified in a denumerable way. The three subsidiary group conditions are expressed, of course, in terms of the group elements g . Our problem will be to translate them into conditions on the ϕ's . For this translation from g - language into ϕ -language the assumption (2.2) concerning the unique parametrization (u.p) will play a fundamental role. We consider the conditions in the following order:

(1) Existence of g_0 . We have, from (1.1), and the convention (2.8) above,

$$g_{(x)} \cdot g_{(0)} = g_{(0)} \cdot g_{(x)} = g_{(x)} , \tag{2.9}$$

or

$$g_{\phi(x,0)} = g_{\phi(0,x)} = g_{(x)} . \tag{2.10}$$

Hence, by the u.p. assumption

$$\phi(x,0) = \phi(0,x) = x . \tag{2.11}$$

This is what the existence of g_0 implies in ϕ- language. To see what this means let us expand $\phi(x,y)$ in terms of x and y . The most general form for ϕ would be

$$\begin{aligned}\phi(x,y) = a_0 + a x + b y + d x^2 + c x y + e y^2 \\ + k x^3 + f x^2 y + g x y^2 + h y^3 + \cdots\end{aligned} \tag{2.12}$$

where we have suppressed the subscripts. What (2.11) says is

that this reduces to $\left(\text{with } a_0 = 0,\ a = b = 1\right)$

$$\phi(x,y) = x + y + c\,xy + f\,x^2 y + g\,xy^2 + \cdots + g^{\alpha}_{\beta\gamma\delta}\,x_\beta\,y_\gamma\,y_\delta + \cdots \tag{2.13}$$

or, more precisely,

$$\phi_\alpha(x_\beta, y_\gamma) = x_\alpha + y_\alpha + c^{\alpha}_{\beta\gamma}\,x_\beta y_\gamma + f^{\alpha}_{\beta\gamma\delta}\,x_\beta x_\gamma y_\delta \tag{2.14}$$

(2) The existence of the inverse. We have, from (1.2), for any given $g(x)$, a $g_{(x')}$ exists such that

$$g_{(x)} \cdot g_{(x')} = g_{(x')}\, g_{(x)} = g_0, \tag{2.15}$$

or

$$g_{[\phi(x,x')]} = g_{[\phi(x',x)]} = g_0 \tag{2.16}$$

Translating this into ϕ-language, we see that it implies that, for each x, the equations

$$\phi(x,x') = 0 \tag{2.17}$$

$$\phi(x',x) = 0 \tag{2.18}$$

should be solvable for x'. From (2.14), it is easy to see (by successive approximation in powers of x) that each of these equations can, in fact, be solved for x' in the neighbourhood of the origin.

Q: That each equation can be solved imposes no conditions, therefore. But does not the fact the solution of both equations (2.17), (2.18) must be the same x' impose same conditions?

A: Yes. You are completely right. To order x^2, we get no conditions from the compatibility of the two equations in (2.17), but for higher orders we do. We shall not have to discuss these conditions however, because they are weaker than, and are automatically fulfilled on account of, our next and last subsidiary condition, namely associativity. In fact, suppose x' is a solution of (2.17) and x'' a solution of (2.18), and that we have associativity. Then

$$g_{(x'')} = g_{(x'')} g_{(0)} = g_{(x'')}\left(g_{(x)} g_{(x')}\right) = \left(g_{(x'')} g_{(x)}\right) g_{(x')} \qquad (2.19)$$

from which, by the u.p., $= g_{(\cdot)} g_{(x')} = g_{(x')}$,

$$x'' = x \qquad (2.20)$$

(3) The associativity condition. From (1.3), we have

$$g_{(x)}\left(g_{(y)} g_{(z)}\right) = \left(g_{(x)} g_{(y)}\right) g_{(z)}\,, \qquad (2.21)$$

or

$$g_{(x)}\, g_{[\phi(yz)]} = g_{[\phi(xy)]}\, g_{(z)}\,, \qquad (2.22)$$

or

$$g_{\phi[x,\,\phi(y,z)]} = g_{\phi[\phi(x,y),\,z]}\,, \qquad (2.23)$$

from which by the u.p.

$$\phi[x, \phi(y,z)] = \phi[\phi(x,y), z]. \qquad (2.24)$$

This is the associativity condition in ϕ-language. As we shall see, it is an extremely powerful condition. In fact it is this condition which gives the ϕ's a definite structure.

Before proceeding to the general investigation of (2.24), we should like to illustrate just how powerful the restriction (2.24) is by considering first the case of a one parameter Lie group. In this case, the x, y, z in (2.24) need no subscript and the equation can be taken literally as an equation in three variables. Differentiating it with respect to the middle variable y, we obtain

$$\frac{\partial \phi[x, \phi(y,z)]}{\partial \phi(y,z)} \frac{\partial \phi(y,z)}{\partial y} = \frac{\partial \phi[\phi(x,y),z]}{\partial \phi(x,y)} \frac{\partial \phi(x,y)}{\partial y}. \quad (2.25)$$

Setting the same variable y equal to zero, this reduces to

$$\frac{\partial \phi(x,z)}{\partial z} \psi(z) = \frac{\partial \phi(x,z)}{\partial x} \theta(x), \quad (2.26)$$

where we have used (2.11) and written

$$\left[\frac{\partial \phi(y,z)}{\partial y}\right]_{y=0} = \psi(z) \; ; \left[\frac{\partial \phi(x,y)}{\partial y}\right]_{y=0} = \theta(x), \quad (2.27)$$

to emphasize the fact that these are functions of z only, and x only, respectively. Thus

$$\frac{\partial \phi(x,z)}{\partial z} \frac{dz}{dZ} = \frac{\partial \phi(x,z)}{\partial x} \frac{dx}{dX} \quad (2.28)$$

where we define $Z(z)$ and $X(x)$ by the differential equations

$$\frac{dZ(z)}{dz} = \frac{1}{\psi(z)} \; , \quad \frac{dX(x)}{dx} = \frac{1}{\theta(x)} \quad (2.29)$$

and boundary conditions

$$Z(0) = 0 \; , \qquad X(0) = 0 \quad (2.30)$$

Q: How do you know that $\psi(z) \neq 0$? If $\psi(z) = 0$, this is not possible.

A: From (2.27),

$$\psi(o) = \left[\frac{\partial \phi(y,\beta)}{\partial y}\right]_{\substack{y=o \\ \beta=o}} = \left[\frac{\partial \phi(y,o)}{\partial y}\right]_{y=o} = \left[\frac{\partial y}{\partial y}\right]_{y=o} = 1 \quad (2.31)$$

Hence, by the analyticity assumption, $\psi(\beta) \simeq 1$ in the neighbourhood of the origin.

Returning to our equations, we see that (2.8) now becomes

$$\frac{\partial \phi(x,\beta)}{\partial Z} = \frac{\partial \phi(x,\beta)}{\partial X} \quad , \quad (2.32)$$

or

$$\phi(x,\beta) = F\left(Z(\beta) + X(x)\right) . \quad (2.33)$$

But now

$$x = \phi(x,o) = F\left(Z(o) + X(x)\right) = F\left(X(x)\right) . \quad (2.34)$$

Hence X is the "inverse" function to F, which we write F^{-1}. For example if

$$F(x) = x\,(1+x) ,$$

$$F^{-1}(x) = x\left[(1-x) + 2x^2 + \cdots\right] . \quad (2.35)$$

Similarly $Z(\beta) = F^{-1}(\beta)$. Thus Z is the same function as X.

Hence, finally,

$$\phi(x\beta) = F\left(F^{-1}(x) + F^{-1}(\beta)\right) \quad (2.36)$$

This ϕ must be of this very special form. How special this is can be seen by writing (2.36) in the form

$$F^{-1}(\phi) = F^{-1}(x) + F^{-1}(\beta) \quad (2.37)$$

and changing the parameter from x to $x' = F^{-1}(x)$. With this parameter we have

$$\phi'(x', y') = x' + y' \tag{2.38}$$

Thus, apart from a change in the parameter, every one parameter local Lie group has the same very simple multiplication function $\phi(x, y)$, namely,

$$\phi(x, y) = x + y \tag{2.39}$$

Q: Does not the fact that you can use a new parameter $x' = F^{-1}(x)$ contradict the assumption concerning the unique parametrization?

A: No. The unique parametrization assumption says only that there will be a one-one correspondence between $g(x)$ and x . We can change to any $x' = \psi(x)$, provided only that there is a one-one correspondence between x' and x . Then we have, schematically,

$$g_{(x)} = g_{(y)} \longleftrightarrow x = y \longleftrightarrow x' = y' \tag{2.40}$$

Q: How do you know that your transformation $x' = F^{-1}(x)$ is one-one? Also how do you know that the Jacobian J of the transfomation is non-zero and finite?

A: Both these questions are answered if we can show that $J(x)$ is non-zero and finite. But we have

$$J(x) = \frac{d F^{-1}(x)}{d x} = \frac{d X(x)}{d x} = \frac{1}{\psi(y)} \; , \tag{2.41}$$

from (2.29), and we have already seen that in the neighbourhood of zero $\psi(y) \simeq 1$.

Returning to (2.39), we should like to make 3 remarks:

(1) We have seen that every one parameter local Lie group is isomorphic to the group with multiplication law (2.39). This is called the translation group. Thus there is, in fact, only one one-parameter local Lie group, a result which illustrates the powerful nature of the associativity condition (2.24). In passing we might remark that this result holds only locally. For example, the translation group, and the group of rotations around one axis are both one-parameter Lie groups but they are not isomorphic in the large.

(2) From (2.39)

$$g_{(x)}\, g_{(y)} = g_{[\phi(x,y)]} = g_{(x+y)} = g_{(y+x)} = g_{[\phi(y,x)]} \qquad (2.42)$$
$$= g_{(y)}\, g_{(x)}.$$

In other words, the one parameter local Lie group is abelian.

(3) We see from (2.39) and (2.36), that the $\phi(x,y)$ for [illegible] is not unique. It is undetermined up to a change induced by a change in the parameter. This emphasizes the point that in investigating the nature of the ϕ's for the general local Lie group, we should not expect them to be determined uniquely, but to be arbitrary at least up to a change induced by any one-one transformation in parameter space.

This completes our study of the one-parameter local Lie group. In the next lectures we shall investigate the conditions placed on the ϕ's by the associativity condition (2.24) for the general local Lie group.

Lecture 3.

In the last lecture we introduced the idea of a Lie group as a set of elements $g_{(x)}$ which are labelled by a finite number r of parameters x_α, each of which can take any value in the neighbourhood of zero. We have unique parametrization (U.P) i.e., that

$$g_{(x)} = g_{(y)} \longleftrightarrow x = y \tag{3.1}$$

and the group operation is specified uniquely by the set of functions

$$x''_\alpha = \phi_\alpha (x_\beta, x'_\gamma) \tag{3.2}$$

which determine it according to the rule

$$g_{(x)}\ g_{(x')} = g_{[x''=\phi(x,x')]}. \tag{3.3}$$

On examining what conditions for the ϕ's, were implied by the three subsidiary group conditions for the g's, we found that

(1) The existence of the unit element g_0 with the convention $g_0 = g_{[0]}$, implied that ϕ should be of the form

$$\phi_\alpha (x_\beta, x'_\gamma) = x_\alpha + x'_\alpha + c^\alpha_{\beta\gamma} x_\beta x'_\gamma + f^\alpha_{\beta\gamma\delta} x_\beta x_\gamma x'_\delta + f^\alpha_{\beta\gamma\delta} x_\beta x'_\gamma x'_\delta + \ldots \tag{3.4}$$

(2) The existence of an inverse such that $g_{(x)} g_{(x')} = g_{(0)}$ and $g_{(x')} g_{(x)} = g_{(0)}$ implied that

$$\phi(x, x') = 0 \quad \text{and} \quad \phi(x', x) = 0, \tag{3.5}$$

should be solvable for x' . Equation (3.4) already guarantees that each of (3.5) is solvable separately. The condition that the solution should be the same x' gives further conditions which will be discussed later.

(3) The associativity condition

$$g_{(x)}\left[g_{(y)}\cdot g_{(z)}\right]=\left[g_{(x)}\,g_{(y)}\right]g_{(z)} \tag{3.6}$$

implies that ϕ satisfies the condition

$$\phi\left(x,\phi(y,z)\right)=\phi\left(\phi(x,y),z\right) \tag{3.7}$$

This condition is extremely strong. To illustrate this we showed that for a one parameter group it means that the group is isomorphic to the translation group (locally). We also saw that we can in any case expect the ϕ's to be defined only up to a change in parametrization. In this lecture, we shall investigate the meaning of the condition (3.7) in the general case. The idea will be the same as in the one-parameter case, namely, to express the condition in terms of a differential equation, and proceed from there. However, the differential equation we require is more complicated than in the one parameter case, and to obtain it we have to proceed in the following somewhat unnatural way: We have,

$$g_{(x)}\;g_{(y)}=\;g_{[\phi(x,y)]} \tag{3.8}$$

Then

$$g_{(x+dx)}\;g_{(y)}=\;g_{[\phi(x+dx,\,y)]}\,. \tag{3.9}$$

Hence, since by the associativity condition $\left[g_{(x)} g_{(y)}\right]^{-1} = g^{-1}_{(y)}\, g^{-1}_{(x)}$, (lecture I), we have

$$g_{[\phi(x+dx,y)]}\, g^{-1}_{[\phi(x,y)]} = g_{(x+dx)}\, g_{(y)} \left[g_{(x)}\, g_{(y)}\right]^{-1} \qquad (3.10)$$

$$= g_{(x+dx)}\, g_{(y)}\, g^{-1}_{(y)}\, g^{-1}_{(x)} = g_{(x+dx)}\, g^{-1}_{(x)}$$

Note that in the second step in (3.10) we have again used the associativity condition. Translating this into ϕ- language, we have

$$\phi\left[\phi(x+dx,y), \phi^{-1}(x,y)\right] = \phi\left(x+dx, x^{-1}\right). \qquad (3.11)$$

[We denote x', the inverse point to x, such that $\phi(x,x')=0$ by x^{-1}. This notation is loose, but it makes equations such as (3.11) more readable, and will, we hope, cause no confusion. $x^{-1} \neq \frac{1}{x}$, of course.]

Expanding both sides of this equation in dx and taking only the coefficient of dx we have

$$\frac{\partial \phi\left[\phi(x,y), \phi^{-1}(x,y)\right]}{\partial \phi(x,y)} \frac{\partial \phi(x,y)}{\partial x} = \frac{\partial \phi}{\partial x}\left(x, x^{-1}\right) \qquad (3.12)$$

Let

$$V^{\alpha}_{\beta}(x) = \frac{\partial \phi^{\alpha}(x,x^{-1})}{\partial x_{\beta}} \equiv \left[\frac{\partial \phi^{\alpha}(x,y)}{\partial x_{\beta}}\right]_{y=x^{-1}} \qquad (3.13)$$

We note that this quantity occurs twice in (3.12), once with x as argument and once with ϕ. Hence we can write (3.12) as

$$V(\phi)\, \frac{\partial \phi}{\partial x}(x,y) = V(x). \qquad (3.14)$$

This equation is the required differential equation for ϕ . However, as we shall be discussing integrability conditions for it, it is more convenient to unite it in a slightly different form. This we do by introducing the quantities

$$u^{\alpha}_{\beta}(x) = \left[\frac{\partial \phi^{\alpha}(y,x)}{\partial y_{\beta}}\right]_{y=0} \qquad (3.15)$$

These are the analogues of the $\psi(x)$, and $\theta(x)$ we had for the one-parameter group. The usefulness of these quantities is that with respect to the $V's$ of (3.12), they have the property

$$V^{\alpha}_{\beta}(x)\, u^{\beta}_{\gamma}(x) = \delta_{\alpha\gamma} \,. \qquad (3.16)$$

This relation follows from (3.12) by setting $x = 0$ in that equation. From (3.16) follows also the inverse relation

$$u^{\alpha}_{\beta}(x)\, V^{\beta}_{\gamma}(x) = \delta_{\alpha\gamma} \,. \qquad (3.17)$$

Note that if we regard $V(x)$ as a matrix, $u(x)$ is just the inverse matrix.

Using these quantities it is easy to see that (3.14) can be written in the form

$$\frac{\partial \phi_{\alpha}(x\,y)}{\partial x_{\beta}} = u^{\alpha}_{\gamma}(\phi)\, V^{\gamma}_{\beta}(x) \,. \qquad (3.18)$$

This will be our basic equation.

For (3.18) and (3.14) we have, of course, the initial conditions

$$\phi_{\alpha}(0,y) = y_{\alpha} \,. \qquad (3.19)$$

In terms of the expansion (3.4), the auxiliary quantities u and V which, we have introduced may be written

$$u^{\alpha}_{\beta}(x) = \delta_{\alpha\beta} + c^{\alpha}_{\beta\gamma} x_{\gamma} + f^{\alpha}_{\beta\gamma\delta} x^{\gamma} x^{\delta} + \cdots \qquad (3.20)$$

and

$$V^{\alpha}_{\beta}(x) = \delta_{\alpha\beta} + c^{\alpha}_{\beta\gamma} x_{\gamma} + \cdots \qquad (3.21)$$

Lecture 4.

In this lecture we shall proceed with the investigation of our basic equation (or rather equations)

$$\frac{\partial \phi(xy)}{\partial x} = u(\phi)\, v(x) , \qquad (4.1)$$

with boundary conditions

$$\phi(0,y) = y . \qquad (4.2)$$

This equation must hold if the $\phi(xy)$ are the transformation functions of a Lie group. Thus if the $\phi(xy)$ are the transformation functions for a Lie group, the

$$u(x) = \left[\frac{\partial \phi(yx)}{\partial y}\right]_{y=0} \qquad (4.3)$$

must be such that (4.1) are integrable. Note that we need not discuss the $V(x)$ since these are determined by the $u(x)$ from the (matrix relation

$$u(x)V(x) = V(x)u(x) = 1 . \qquad (4.4)$$

What we now want to show is that, conversely, if the $u(x)$ are such that the equations (4.1) are integrable then the ϕ's which we obtain are the transformation functions ϕ of a Lie group. Later we shall be discussing the conditions under which the $u(x)$ are such that (4.1) is integrable.

Suppose, therefore, that (4.1) is integrable. From the ordinary theory of differential equations we see that if $u(x)$ and $v(\phi)$ are known then (4.1) and (4.2) determine the $\phi(xy)$ uniquely. Now we ask: What conditions must the ϕ satisfy in order to be the transformation functions of a Lie group ? The answer is:

(1) they must satisfy the condition for the existence of the unit element i.e.,

$$\phi(x,0) = \phi(0,x) = x \qquad (4.5)$$

It is easy to verify that both of (4.5) are indeed implied by (4.1) and (4.2).

(2) they must satisfy the condition for the existence of the inverse. However, we have seen above that this will always be satisfied if (1) and the associativity condition are satisfied.

(3) they must satisfy the associativity condition

$$\phi\left(x,\phi(y,z)\right) = \phi\left(\phi(xy),z\right) \qquad (4.6)$$

The proof that they do indeed satisfy this condition also is as follows: Let us regard $\phi\left(x,\phi(yz)\right)$ and $\phi\left(\phi(x,y)z\right)$ as two functions of x, $A(x)$ and $B(x)$ respectively. Then

we prove that

$$A(x) = B(x) , \tag{4.7}$$

by proving that $A(x)$ and $B(x)$ satisfy the same differential equation, with the same initial condition.

First, the initial condition: we have from (4.5)

$$A(0) = \phi(0, \phi(y,\beta)) = \phi(y,\beta) = \phi(\phi(0,y),\beta) = B(0) \tag{4.8}$$

Next, the differential equation: from (4.1)

$$\begin{aligned} \frac{\partial A(x)}{\partial x} = \frac{\partial \phi(x,\phi(y,\beta))}{\partial x} &= v\left(\phi(x,\phi(y,\beta))\right) u(x) \\ &= v\left(A(x)\right) u(x) , \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} \frac{\partial B(x)}{\partial x} = \frac{\partial \phi(\phi(xy),\beta)}{\partial x} &= \frac{\partial \phi(\phi(xy),\beta)}{\partial \phi(xy)} \frac{\partial \phi(xy)}{\partial x} \\ &= v\left(\phi(\phi(x,y),\beta)\right) u\left(\phi(xy)\right) v\left(\phi(xy)\right) u(x) \\ &= v\left(\phi(\phi(xy),\beta)\right) u(x) \\ &= v\left(B(x)\right) u(x) \qquad \text{Q.E.D.} \end{aligned} \tag{4.10}$$

Thus every set of $u(x)$ which make (4.1) integrable generates a set of ϕ's which are the multiplication functions for a local Lie group. In other words, we may say that every such set of ϕ's generates a local Lie group. The problem now will be to find out what $u(x)$'s are allowed by the condition that (4.1) should be integrable. This problem we shall tackle in the next lecture.

Note by the way, that our problem has now been reduced to discussing the functions of <u>one</u> variable (or set of variables) x , instead of the functions of two variables $\phi(xy)$.

Lecture 5.

In the last lecture we derived the equation

$$\frac{\partial \phi_\alpha(xy)}{\partial x_\beta} = u_\sigma^\alpha(\phi)\, v_\beta^\sigma(x), \qquad (5.1)$$

$$\phi_\alpha(o\,y) = y_\alpha \qquad (5.2)$$

for the transformation functions $\phi_\alpha(xy)$ of a Lie group. We showed also that conversely, if we could find a set of $u's$ such that (5.1) was integrable, then the $\phi's$ obtained would be the transformation functions $\phi_\alpha(xy)$ of a local Lie group. The functions V are, of course, defined in terms of the $u's$ by the relation

$$V_\sigma^\alpha(x)\, u_\gamma^\sigma(x) = \delta_{\alpha\gamma} \qquad (5.3)$$

What we now wish to do is to find out what $u's$ will make (5.1) integrable. To do this we write down the integrability conditions for the system (5.1), which read

$$\frac{\partial}{\partial x_\gamma} u_\sigma^\alpha(\phi)\, v_\beta^\sigma(x) = \frac{\partial}{\partial x_\beta} u_\sigma^\alpha(\phi)\, v_\gamma^\sigma(x), \qquad (5.4)$$

or

$$\frac{\partial u_\sigma^\alpha(\phi)}{\partial \phi_\lambda}\frac{\partial \phi_\lambda}{\partial x_\gamma} v_\beta^\sigma(x) + u_\sigma^\alpha(\phi)\frac{\partial}{\partial x_\gamma} v_\beta^\sigma(x) = \frac{\partial u_\sigma^\alpha(\phi)}{\partial \phi_\lambda}\frac{\partial \phi_\lambda}{\partial x_\beta} v_\gamma^\sigma(x) + u_\sigma^\alpha(\phi)\frac{\partial v_\gamma^\sigma(x)}{\partial x_\beta} \qquad (5.5)$$

Using the original equation (5.1) in this, we obtain

$$\frac{\partial u_\sigma^\alpha(\phi)}{\partial \phi_\lambda} v_\mu^\lambda(\phi)\, u_\gamma^\mu(x)\, v_\beta^\sigma(x) + u_\sigma^\alpha(\phi)\frac{\partial v_\beta^\sigma(x)}{\partial x_\gamma} = \frac{\partial u_\sigma^\alpha(\phi)}{\partial \phi_\lambda} v_\mu^\lambda(\phi)\, u_\beta^\mu(x)\, v_\gamma^\sigma(x) + u_\sigma^\alpha(\phi)\frac{\partial v_\gamma^\sigma(x)}{\partial x_\beta}. \qquad (5.6)$$

By using (5.3), we can carry out some manipulations on this equation so that it finally reads

$$\left[\frac{\partial u_\beta^\sigma(x)}{\partial x_\gamma} - \frac{\partial u_\gamma^\sigma(x)}{\partial x_\beta}\right] V_\mu^\beta(x) V_\tau^\gamma(x) = \left[\frac{\partial V_\tau^\alpha(\phi)}{\partial \phi_\lambda} V_\mu^\lambda(\phi) - \frac{\partial V_\mu^\alpha(\phi)}{\partial \phi_\lambda} V_\tau^\lambda(\phi)\right] u_\alpha^\sigma \tag{5.7}$$

the idea in bringing it to this form being that the left hand side is a function of x only, and the right hand side a function of $\phi(xy)$. We immediately make use of this property by choosing the variable y to be x^{-1}. Then $\phi(xy) = 0$ and the right hand side of (5.8) is a constant which we denote by $\hat{c}_{\mu\tau}^\sigma$. Thus, we obtain

$$\left[\frac{\partial u_\beta^\sigma(x)}{\partial x_\gamma} - \frac{\partial u_\gamma^\sigma(x)}{\partial x_\beta}\right] V_\mu^\beta(x) V_\tau^\gamma(x) = \hat{c}_{\mu\tau}^\sigma = \left[\left(\frac{\partial V_\mu^\alpha(\phi)}{\partial \phi_\lambda}\right)_{\phi=0} V_\tau^\lambda(0) - \left(\frac{\partial V_\tau^\alpha(\phi)}{\partial \phi_\lambda}\right) V_\mu^\lambda(0)\right] u_\alpha^\sigma(0). \tag{5.8}$$

where the c's are constant. Using (5.3) we can transform this into

$$\frac{\partial u_\beta^\sigma(x)}{\partial x_\gamma} - \frac{\partial u_\gamma^\sigma(x)}{\partial x_\beta} = \hat{c}_{\mu\tau}^\sigma u_\beta^\mu(x) u_\gamma^\tau(x), \tag{5.9}$$

which is an equation in terms of the u's. Thus (5.9) is a necessary condition for (5.1) to be integrable. But it is also sufficient! To see this, let us first write (5.9) in terms of the V's by using the relation (5.3). By direct substitution of (5.3) in (5.9), we obtain (after some manipulation !)

$$\left[\frac{\partial V_\mu^\alpha(x)}{\partial x_\lambda} V_\tau^\lambda(x) - \frac{\partial V_\tau^\alpha(x)}{\partial x_\lambda} V_\mu^\lambda(x)\right] u_\alpha^\sigma(x) = \hat{c}_{\tau\mu}^\sigma \tag{5.10}$$

Thus if (5.9) is satisfied, (5.7) is satisfied, and since (5.7) is merely another form of the integrability condition (5.4), the latter condition is satisfied.

We see therefore that if we are given a Lie group, we can construct from its $\phi(x,y)$ functions, the corresponding $u(x)=\left[\frac{\partial \phi(x,y)}{\partial y}\right]_{y=0}$ and these must satisfy (5.8). Conversely, if we are given a set of u's satisfying (5.8), we can define the V functions by (5.3), write down the equations (5.1), (5.2) and since this will then be integrable, obtain from a unique set of ϕ's. From the discussion of the previous lecture, these will be the ϕ functions of a local Lie group.

Our problem reduces, therefore, to finding out what kind of u's are defined by (5.9). In other words, we now have to find the integrability conditions for (5.9). The integrability conditions for (5.9) will yield another set of equations, but the process stops with these, because as we shall see, they reduce to purely algebraic relations for the $\hat{C}$.

Q: Would it be true to say that the process stops at the next step because, whereas (5.1) is an equation for the ϕ in terms of the u's, (5.9) is an equation for the u in terms of itself?

A: Yes, I think so.

Before going on to write down the integrability conditions for (5.9), however, let me just mention an analogy between (5.9) and an equation with which we are all familiar, namely, the equation which defines the vector potential $A_\mu(x)$ in the classical theory of electricity. We have in that theory an equation

$$\left[A_{\mu,\nu}(x)-A_{\nu,\mu}(x)\right] = F_{\mu\nu}(x) . \qquad (5.11)$$

This equation for $A_\mu(x)$ has two properties which I should like to mention:

(1) It is integrable, if, and only if, $F_{\mu\nu}(x)$ satisfies the two conditions

$$F_{\mu\nu} = -F_{\nu\mu} \,, \tag{5.12}$$

and

$$F_{\mu\nu,\lambda} + F_{\nu\lambda,\mu} + F_{\lambda\mu,\nu} = 0 \,. \tag{5.13}$$

(The second expresses the fact that 'divcurl $\equiv$ 0').

(2) If these conditions are satisfied, $A_\mu(x)$ exists, but it is not unique. In fact, as is well-known, $A_\mu(x)$ is arbitrary up to the gauge-transformation

$$A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \Lambda(x) \,, \tag{5.14}$$

where $\Lambda(x)$ is an arbitrary function. [This, in turn, is an expression of the fact that ''curlgrad $\equiv$ 0''.]

The analogy between (5.11) and our equation (5.9) for the u, is obvious. The only difference between the two is that (1) in our case we have a set of ''potentials'' (one for each α) instead of just one, and (2) that in our case $F_{\mu\nu}$ is not an ''external'' quantity but is defined in terms of the u's themselves.

The integrability conditions for (5.9) are just the same as for $A_\mu(x)$, namely (5.12) and (5.13). With

$$F^{\alpha}_{\mu\nu} = \hat{C}^{\alpha}_{\sigma\tau}\, u^{\sigma}_{\mu}\, u^{\tau}_{\nu} \,, \tag{5.15}$$

these become,

$$\hat{C}^{\alpha}_{\sigma\tau} = -\hat{C}^{\alpha}_{\tau\sigma} , \tag{5.16}$$

and

$$\sum_{\text{cyclic in } \mu,\nu,\lambda} \hat{C}^{\alpha}_{\sigma\tau}\left(u^{\sigma}_{\mu} u^{\tau}_{\nu}\right)_{\lambda} = 0 . \tag{5.17}$$

Expanding the latter equation we have

$$\hat{C}^{\alpha}_{\sigma\tau} u^{\sigma}_{\mu} u^{\tau}_{\nu,\lambda} + \hat{C}^{\alpha}_{\sigma\tau} u^{\tau}_{\nu} u^{\sigma}_{\mu\lambda} \quad + 4 \text{ similar terms} = 0. \tag{5.18}$$

Rearranging the terms and using (5.16) we obtain

$$\hat{C}^{\alpha}_{\sigma\tau} u^{\sigma}_{\mu}\left(u^{\tau}_{\nu,\lambda} - u^{\tau}_{\lambda,\nu}\right) \quad + 2 \text{ similar terms} = 0. \tag{5.19}$$

But now we can make use of equation (5.9) itself, to get

$$\hat{C}^{\alpha}_{\sigma\tau}\left(C^{\tau}_{\alpha\beta} u^{\alpha}_{\nu} u^{\beta}_{\lambda}\right) u^{\sigma}_{\mu} \quad + 2 \text{ similar terms} = 0. \tag{5.20}$$

or

$$\left[\hat{C}^{\alpha}_{\alpha\tau}\hat{C}^{\tau}_{\beta\gamma} + \hat{C}^{\alpha}_{\beta\tau}\hat{C}^{\tau}_{\gamma\alpha} + \hat{C}^{\alpha}_{\alpha\tau}\hat{C}^{\tau}_{\beta\gamma}\right] u^{\alpha}_{\mu} u^{\beta}_{\nu} u^{\gamma}_{\lambda} = 0. \tag{5.21}$$

By putting $x = 0$ in this equation, we get

$$\hat{C}^{\alpha}_{\mu\tau}\hat{C}^{\tau}_{\nu\lambda} + \hat{C}^{\alpha}_{\lambda\tau}\hat{C}^{\tau}_{\mu\nu} + \hat{C}^{\alpha}_{\nu\tau}\hat{C}^{\tau}_{\lambda\mu} = 0 \tag{5.22}$$

Since $u^{\alpha}_{\mu}(0) = \delta^{\alpha}_{\mu}$. This equation (known as the Jacobi 'identity') is therefore a necessary condition for (5.9) to be integrable. It is (together with (5.16)) also a sufficient condition, however, since if (5.22) is satisfied, so is (5.21) and the latter is no more than a simplified version of the integrability condition (5.17).

We have now reached the end of the line. We have seen that for any Lie group, we can define a set of u's in terms of the ϕ's, and from the u's determine, according to (5.8) a set of constants $\hat{C}$ which must satisfy the conditions (5.16) (anti-symmetry condition) and (5.22) (Jacobi identity). We now see that, conversely, if we are given a set of constants $C^{\alpha}_{\beta\gamma}$, $\alpha, \beta, \gamma = 1 \ldots\ldots n$, any n, we write down the equation (5.9) and solve it for u. Having got our u, we can define V in terms of u in the usual way $(uV = Vu = 1)$ and write down the equation (5.1). The ϕ obtained from this equation (which is integrable since the u's satisfy (5.9) by definition) will be the ϕ function for a local Lie group. Thus we have the result: To every local Lie group corresponds a set of constants $\hat{C}^{\alpha}_{\beta\gamma}$ satisfying (5.16) and (5.22). Conversely to every set of constants satisfying (5.16) and (5.22) corresponds a local Lie group. The C's are called the structure constants of the group.

The only question is: is the correspondence one-to-one? One might suspect that it is not, because, from the analogy with the vector-potential, we can see that the u's defined by (5.9) for a given set of $\hat{C}$'s, are not unique. However, it can be shown that the arbitrariness in the u's corresponds exactly to the arbitrariness in the choice of parameters, and this, we know, is always present. In the next lecture I hope to give some indication of why this is so. For the moment, we shall simply say that, in spite of the arbitrariness in the u's, there is, in fact, a one-to-one correspondence between the sets of $\hat{C}$'s and the local Lie groups.

I shall conclude this lecture by evaluating the structure constants $\hat{c}$'s in terms of the coefficients in the expansion of $\phi(x,y)$. i.e.

$$\phi_\alpha(x,y) = x_\alpha + y_\alpha + c^\alpha_{\beta\gamma} x_\beta y_\gamma + f x^2 y + g x y^2 + \cdots\cdots \quad (5.23)$$

Thus,

$$u^\alpha_\beta(x) = \left[\frac{\partial \phi_\alpha(x,y)}{\partial y_\beta}\right]_{y=0} = \delta_{\alpha\beta} + c^\alpha_{\beta\gamma} x_\gamma + f x^2 + \cdots \quad (5.24)$$

and its ''inverse'',

$$v^\alpha_\beta(x) = \delta_{\alpha\beta} - c^\alpha_{\beta\gamma} x_\gamma + f' x^2 + \cdots\cdots \quad (5.25)$$

Thus, by the definition (5.8)

$$\begin{aligned} \hat{c}^\alpha_{\beta\gamma} &= -\left[\left[\frac{\partial v^\sigma_\beta(\phi)}{\partial \phi_\lambda}\right]_{\phi=0} v^\lambda_\gamma(0) - \left[\frac{\partial v^\sigma_\gamma(\phi)}{\partial \phi_\lambda}\right]_{\phi=0} v^\lambda_\beta(0)\right] u^\alpha_\sigma(0) \\ &= -\left[-c^\sigma_{\beta\gamma}\delta^\lambda_\gamma - \left[-c^\sigma_{\gamma\lambda}\right]\delta^\lambda_\beta\right]\delta_{\alpha\sigma} \quad (5.26) \\ &= c^\alpha_{\beta\gamma} - c^\alpha_{\gamma\beta} \,. \end{aligned}$$

Thus, the anti-symmetric parts of the second order coefficients in (5.23) determine the local group completely.

Lecture 6.

In the last lecture we showed that from the ϕ-functions of a local Lie group we could obtain a set of functions $u^\alpha_\beta(x)$ from which, in turn we could obtain a set of <u>structure constants</u>

$C^{\lambda}_{\alpha\beta}$ which satisfy the two relations

$$C^{\lambda}_{\alpha\beta} = - C^{\lambda}_{\beta\alpha} \,, \tag{6.1}$$

$$C^{\alpha}_{\beta\lambda}\, C^{\lambda}_{\gamma\delta} + C^{\alpha}_{\gamma\lambda}\, C^{\lambda}_{\delta\beta} + C^{\alpha}_{\delta\lambda}\, C^{\lambda}_{\beta\gamma} = 0 \,,$$

and, conversely, that given any set of constants satisfying (6.1) we could construct the ϕ -functions of a local Lie group. In this lecture we wish to discuss the uniqueness of the relation between the ϕ's and the C's .

We start by noting that the ϕ's corresponding to any Lie group are not themselves unique, since we can always make a change of parameter $x \longrightarrow x' = f(x)$ and this will change the functional form of ϕ . However, by definition, this is the only arbitrariness allowed in the ϕ . If we write the equation $x' = f(x)$ in the form

$$x'_{\alpha} = a^{\alpha}_{\beta}\, x_{\beta} + a^{(2)\alpha}_{\beta\gamma}\, x_{\beta} x_{\gamma} + a^{(3)\alpha}_{\beta\gamma\delta}\, x_{\beta} x_{\gamma} x_{\delta} + \cdots + a^{(n)\alpha}\, x^{n} + \cdots \tag{6.2}$$

which is the most general relation which preserves the convention $g_0 = g_{[0]}$, we see that any such transformation can be considered as the product of two transformations of the following kind:

(1)

$$x'_{\alpha} = x_{\alpha} + a^{(2)\alpha}_{\beta\gamma}\, x_{\beta} x_{\gamma} + a^{(3)\alpha}_{\beta\gamma\delta}\, x_{\beta} x_{\gamma} x_{\delta} + \cdots + a^{(n)\alpha}\, x^{n} + \cdots \,, \tag{6.3}$$

and

(2)
$$x'_{\alpha} = a^{\beta}_{\alpha}\, x_{\beta} \tag{6.4}$$

It is convenient to treat these two types of transformation separately.

To investigate what happens when we make a transformation of the first kind above we expand $\phi(x,y)$ in the form

$$\phi(x,y) = x+y+cxy+\left(g_3^{(3)}x^2y+f_3xy^2\right) + \cdots + \left(g_n^{(n)}x^{n-1}y+\cdots\right)+\cdots\cdots \tag{6.5}$$

Then we have from (6.3)

$$\begin{aligned}\phi' &= \phi + a^{(2)}\phi^2+\cdots\\ &= x+y+cxy+a^{(2)}(x+y)^2+\cdots\\ &= \left[x'-a^{(2)}x'^2\right]+\left[y'-a^{(2)}y'^2\right]+cx'y'+a^{(2)}(x'+y')^2\\ &= x'+y'+\left[c+2a^{(2)}\right]x'y'+\cdots,\\ &= x'+y'+c'x'y'+\cdots,\end{aligned} \tag{6.6}$$

where

$$c' = c+2a^{(2)}. \tag{6.7}$$

Thus, as mentioned above, ϕ is not the same function of x' and y' as ϕ is of x and y. Note, however, that, to second order, the property

$$\phi(x,0) = \phi(0x) = x, \tag{6.8}$$

is preserved. It is not difficult to verify that this holds to all orders.

We see from (6.7) that the c's are not invariant under our transformation. The same will be true, of course, for g, f ..... It is easy to see from this that the u's defined by

$$u^{\alpha}_{\beta}(x) = \left[\frac{\partial \phi_{\alpha}(x,y)}{\partial y}\right]_{y=0}$$

$$= \delta^{\alpha}_{\beta} + c^{\alpha}_{\beta\gamma}\, x^{\gamma} + g^{(3)\alpha}_{\beta\gamma\delta}\, x^{\gamma} x^{\delta} + \cdots + g^{(n)\alpha}\, x^{n-1} + \cdots \qquad (6.9)$$

will not remain invariant, i.e., the u's will not be the same functions of the x's as the u are of the x's. However, if we now consider the structure constants $\hat{c}_{\alpha\beta}$, we find that

$$\hat{c}'_{\alpha\beta} = c'_{\alpha\beta} - c'_{\beta\alpha} = c_{\alpha\beta} - c_{\beta\alpha} + 2\left(a_{\alpha\beta} - a_{\beta\alpha}\right)$$
$$= \left(c_{\alpha\beta} - c_{\beta\alpha}\right) = \hat{c}_{\alpha\beta}, \qquad (6.10)$$

because the $a_{\alpha\beta}$ are by definition, (3), completely symmetric. Thus, under a transformation of the first kind, the structure constants remain invariant.

Consider now the equation for the u's

$$\frac{\partial u^{\alpha}_{\beta}}{\partial x_{\gamma}} - \frac{\partial u^{\alpha}_{\gamma}}{\partial x_{\beta}} = c^{\alpha}_{\lambda\mu}\, u^{\lambda}_{\gamma}\, u^{\mu}_{\beta}. \qquad (6.11)$$

If for a given Lie group for which the c satisfy (6.11) we change the parameter x, we obtain

$$\frac{\partial u'^{\alpha}_{\beta}(x')}{\partial x'_{\gamma}} - \frac{\partial u'^{\alpha}_{\gamma}(x')}{\partial x'_{\beta}} = c^{\alpha}_{\lambda\mu}\, u'^{\lambda}_{\gamma}(x')\, u'^{\mu}_{\beta}(x'), \qquad (6.12)$$

which shows that the u's satisfy the same equation as the u's. [Note that C is the same in (6.11) and (6.12).] But we saw above that the u's are not the same functions as the u's. We see, therefore, that the equation (6.11) does not determine the u's uniquely. In fact, we can generate a whole family of solutions from any given u by constructing for that u the corresponding (unique) ϕ, changing the parameters by transformations of the first kind, and defining new u's in terms of the new ϕ's, according to (6.9).

The question now is: does this family, of $u's$ include all the possible $u's$ which are solutions of (6.11) ? The answer is ''yes''. This is somewhat difficult to prove, and I shall not prove it here. But it has the important consequence that we can now say: The equations (6.11) for the $u's$ do not define a unique set of $u's$, and hence do not define a unique set of $\phi's$. But they do define a unique local Lie group because the $\phi's$ obtained are such that they differ from each other only by a change of parameter. Thus, any set of $C's$ satisfying (6.1) determine a local Lie group uniquely.

This practically completes our discussion of the first kind of transformation, but before going on to the second kind, I should like first to mention how the freedom of changing parameter can be used to simplify ϕ a little. We recall that for the case of a one-parameter group, we could reduce ϕ to the simple function $\phi(x,y) = x + y$ by means of a parameter transformation. What we are now going to do will be the analogue of this for the general case.

We begin by noting from (6.7) that if we choose $a^{(1)}_{\alpha\beta}$ which is completely at our disposal, to be

$$a^{(1)}_{\alpha\beta} = -\frac{1}{4}\left(C_{\alpha\beta} + C_{\beta\alpha}\right). \qquad (6.13)$$

we make the symmetric part of $C_{\alpha\beta}$ zero. Further this determines $a^{(2)}_{\alpha\beta}$ completely.

In a similar way, by choosing $a^{(3)}_{\alpha\beta\gamma}, a^{(4)}_{\alpha\beta\gamma\delta}, \ldots, a^{(n)}$, ... , ... suitably, we can transform to zero the

completely symmetric parts of $g^{(3)}_{\alpha\beta\gamma}$, $g^{(4)}_{\alpha\beta\gamma\delta}$, ... $g^{(n)}$, and this determines $a^{(3)}_{\alpha\beta\delta}$, $a^{(4)}$, ... $a^{(n)}$, completely. The proof of this (by induction) is not difficult, but we shall not give it here.

If we do choose the $a^{(2)}$, $a^{(3)}$, ... in this way, the parameters x which we obtain are called normal parameters. The normal parameters are unique as far as transformations of the first kind are concerned, and we shall now derive a very simple definition of them. To derive it, we go back to (6.9) and note that if we multiply the right hand side of that equation by x_β only the completely symmetric parts of c, $g^{(3)}$, $g^{(4)}$, ... survive. If these are zero, we get

$$u^\alpha_\beta(x)\, x_\beta = x_\alpha \,, \tag{6.14}$$

which is the required definition.

Q: Can one say from (6.14) that $u^\alpha_\beta(x)$ is a unit operator of some kind in parameter space?

A: Not really. For example, for $y_\alpha \neq x_\alpha$,

$$u^\alpha_\beta(x)\, y_\beta \neq y_\alpha \,. \tag{6.15}$$

Note, by the way, that since $Vu = uV = 1$,

$$V^\alpha_\beta(x)\, x_\beta = x_\alpha \,, \tag{6.14a}$$

as an alternative definition of the normal parameters.

We should now like to consider the transformations of the second kind. These are much easier to handle, because they are linear. First of all, we note that to preserve the unique

parametrization, we must have

$$\det |A| \neq 0 \tag{6.16}$$

where we are regarding this kind of transformation as a matrix transformation

$$x' = Ax. \tag{6.17}$$

Since ϕ is itself an x, we have

$$\phi' = A\phi, \tag{6.18}$$

and on expanding this we have

$$\begin{aligned}
\phi' = A\phi &= A\{x + y + cxy + g_3 x^2 y + \cdots\cdots\} \\
&= A\{A^{-1}x' + A^{-1}y' + cA^{-1}A^{-1}x'y' + g_3 A^{-1}A^{-1}A^{-1}x'^2 y' \\
&\qquad + \cdots\} \\
&= x' + y' + AcA^{-1}A^{-1}x'y' + Ag_3 A^{-1}A^{-1}A^{-1}x'^2 y' \\
&\qquad + \cdots\} \\
&= x' + y' + c'x'y' + g^{(3)\prime}x'^2 y' + \cdots\cdots,
\end{aligned} \tag{6.19}$$

where

$$\begin{aligned}
c'^{\alpha}_{\beta\gamma} &= AcA^{-1}A^{-1} = a^{\alpha}_{\mu} a^{-1\lambda}_{\beta} a^{-1\sigma}_{\gamma} c^{\mu}_{\lambda\sigma}, \\
g^{(3)\prime\alpha}_{\beta\gamma\delta} &= Ag^{(3)}A^{-1}A^{-1}A^{-1} = a^{\alpha}_{\mu} a^{-1\lambda}_{\beta} a^{-1\sigma}_{\gamma} a^{-1\tau}_{\delta} g^{(3)\mu}_{\lambda\sigma\tau},
\end{aligned} \tag{6.20}$$

etc. We see that the property

$$\phi(x,0) = \phi(0,x) = x \tag{6.21}$$

is again preserved.

The transformation property of the u^{α}_{β} [illegible]

$$u' = u'^{\alpha}_{\beta}(x) = a^{\alpha}_{\lambda} a^{-1\mu}_{\beta} u^{\lambda}_{\mu}(x) = AuA^{-1}, \tag{6.22}$$

follows immediately from the definition of C (6.9).

Thus, under a transformation of the second kind also, the ϕ's and C's change their functional form. But in contrast to the first kind of transformation, in the present case the structure constants also change. This follows immediately from (6.20) since

$$\hat{C}' = \hat{C}'^{\alpha}_{\beta\gamma} = C'^{\alpha}_{\beta\gamma} - C'^{\alpha}_{\gamma\beta} = a^{\alpha}_{\lambda} a^{-1\lambda}_{\beta} a^{-1\sigma}_{\gamma} \left(C^{\lambda}_{\mu\sigma} - C^{\lambda}_{\sigma\mu} \right)$$

$$= a^{\alpha}_{\lambda} a^{-1\mu}_{\beta} a^{-1\sigma}_{\gamma} \hat{C}^{\lambda}_{\mu\sigma} = A \hat{C} A^{-1} A^{-1} \qquad (6.23)$$

Thus the C's corresponding to a given Lie group are not unique. On the other hand, since we have now considered the most general parameter transformations possible, (6.23) represents at the same time the maximum arbitrariness in the definition of the $C^{\alpha}_{\beta\gamma}$. It is obvious that we also have the following sort of converse result, namely, that if any two C's are related by a transformation of the kind (6.22) they determine the same local Lie group, since the ϕ's determined by them will differ only by a parameter charge of the second kind with the given A (plus, perhaps, a parameter change of the first kind).

Finally, we note that under a parameter transformation of the second kind the normal coordinate condition (6.14) is preserved. Thus, a set of normal parameters, far from being unique with respect to the second kind of transformation, transform into another set of normal parameters under such a transformation.

Collecting all the results of the present lecture together, we may, therefore, write down the following scheme:

Local Lie group	$\longleftrightarrow$	set of ϕ's related to each other by parameter transformations of the 1st and 2nd kind i.e., all parameter transformations.	$\longleftrightarrow$	Set of all a's related to each other by transformations of the kind ( ) and ( )	$\longleftrightarrow$	Set of all c's, related to each other by transformations $c' = AcA^{-1}A^{-1}$ as in (6.23) (6.24)

Here, " $\longleftrightarrow$ " means: "there is a one-to-one correspondence between".

Lecture 7

So for, we have considered Lie groups in the abstract only, that is to say, we have considered the $\phi(x,y)$ functions to be given without mentioning from where they come. What we now wish to discuss is the connection between the Lie groups and groups of continuous transformations. The connection is very intimate, and, in fact, it was from considering groups of transformations that Lie was led to the concept of the Lie groups themselves. In the present lecture we shall discuss the general transfomations associated with Lie groups, and in the next lecture we shall specialize to the case of linear transfomations (or representations).

We shall introduce the idea of a transformation associated with a Lie group by means of an example. We saw earlier that the set of all matrices

$$g_{[x]} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \qquad (x_1x_4 - x_2x_3) \neq 0, \tag{7.1}$$

form a Lie group under the operation of matrix multiplication i.e.

$$g_{[x']}\, g_{[x]} = \begin{bmatrix} x_1' & x_2' \\ x_3' & x_4' \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} x_1'' & x_2'' \\ x_3'' & x_4'' \end{bmatrix} \tag{7.2}$$

$$= g_{[x'']},$$

where

$$x_i'' = \phi_i(x_j', x_k) \qquad i,j,k = 1,2,3,4 \tag{7.3}$$

But a matrix can always be thought of as a linear transformation of a vector space. In the above case, such transformations would be

$$\begin{pmatrix} y \\ \eta \end{pmatrix} \longrightarrow \begin{pmatrix} y' \\ \eta' \end{pmatrix} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \begin{pmatrix} y \\ \eta \end{pmatrix}, \tag{7.4}$$

where (y, η) is a 2-vector.

If we now apply two successive transformations of this kind, we obtain the vector

$$\begin{pmatrix} y'' \\ \eta'' \end{pmatrix} = \begin{bmatrix} x_1' & x_2' \\ x_3' & x_4' \end{bmatrix} \begin{pmatrix} y' \\ \eta' \end{pmatrix} = \begin{bmatrix} x_1' & x_2' \\ x_3' & x_4' \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \begin{pmatrix} y \\ \eta \end{pmatrix} \tag{7.5}$$

$$= \begin{bmatrix} x_1'' & x_2'' \\ x_3'' & x_4'' \end{bmatrix} \begin{pmatrix} y \\ \eta \end{pmatrix}$$

That is to say, we obtain the vector (y'', η'') which could be obtained by a single transformation of this kind (with suitable values x'' for the parameters).

We now generalize this idea. Let $y_i' = f_i(y_j, x_\alpha)$ be a set of transformations depending on the finite set of continuous parameters x_α, $\alpha = 1, \ldots r$, and transforming the finite set of n variables y_j into the finite set of n variables y_i' $(i, j = 1, \ldots n)$ ($n \neq r$ in general) in such a way that the following conditions are satisfied:

(1) Two successive transformations of this kind are equivalent to a single transformation(with certain other values of the parameters). i.e. If

$$y' = f(y, x), \tag{7.6}$$

and

$$y'' = f(y', x'),$$

then

$$y'' = f\left(f(y, x), x'\right) = f\left(y, \phi(x, x')\right), \tag{7.7}$$

where the ϕ's are a suitable set of parameters, depending on x and x'.

(2) The set of transformations includes the identity transformation $y_i' = y_i$. By convention the parameters are so chosen that the identity transformation is obtained by setting $x_\alpha = 0$

$$y_i = f_i(y_k, 0) \tag{7.8}$$

(3) For every transformation

$$y_i' = f_i(y_k, x_\alpha) \tag{7.9}$$

the inverse transformation exists, i.e. there exists y_α such that

$$y_i = f_i(y_k', y_\alpha) \tag{7.10}$$

In addition we assume that the parameterization is unique i.e.

$$x_\alpha = \hat{x}_\alpha \longleftrightarrow f_i(y_j, x_\alpha) = f_i(y_j, \hat{x}_\alpha), \qquad (7.11)$$

and that the f_i are analytic in y and x in the neighbourhood of $x = 0$, all y (Actually one can get away with much weaker assumptions concerning differentiability but for the sake of simplicity we shall assume full analyticity).

If these conditions are all fulfilled we shall say that the f_i form a group of continuous transformations. We shall now show that to every group of continuous transformations corresponds a unique local Lie group. In fact we shall show that the $\phi(x, x')$ defined in (1.7) are the ϕ-functions of a local Lie group. Conversely, we shall show that to each local Lie group there corresponds a number of groups of transformations $f(y, x)$.

Q:- must the transformations f_i be linear?

No. For example, the set of transformations of one variable $y' = \left(\frac{x_1 y + x_2}{x_3 y + x_4}\right)$ satisfy all our conditions and are not linear. However, the linear transformations are a very important special case, and will be discussed in the next lecture.

To show that the $\phi(x, x')$ of (1.7) satisfy all the conditions necessary for them to be the ϕ-functions of a Lie group we proceed as follows: First we prove that the ϕ's satisfy the associativity condition.

Proof: Let

$$\begin{aligned} y' &= f(y,x), \\ y'' &= f(y',x'), \\ y''' &= f(y'',x''), \end{aligned} \tag{7.12}$$

Then

$$\begin{aligned} f(y,\phi(x'\,\phi(x''x'''))) &= f(f(y,x'),\ \phi(x''x''')) \\ &= f[f(f(y,x'),x''),x'''] \\ &= f[f(y,\phi(x'x''),x''')] \\ &= f(y,\phi(\phi(x'x''),x''')), \end{aligned} \tag{7.13}$$

and so, by the unique parameterization

$$\phi(x'\,\phi(x'',x''')) = \phi(\phi(x',x''),x'''), \tag{7.14}$$

as required, Note that we did not have to postulate explicitly an associativity condition for the f's in order to get this result. The assumption is contained implicitly in the definition of the f's.

The next step is to show that the ϕ's satisfy the condition for the existence of the unit element. We have from

(7.8)

$$f(y,x) = f(f(y,x),0) = f(y,\phi(x,0)), \tag{7.15}$$

and

$$f(y,x) = f\left(f(y,0),x\right) = f\left(y\ \phi(0,x)\right), \tag{7.16}$$

from which, by the unique parameterization,

$$x = \phi(x,0) = \phi(0,x), \tag{7.17}$$

as required.

Finally, we prove that the ϕ's satisfy the condition for the existence of the inverse. In fact, from (7.8), (7.9) and (7.10)

$$f(\zeta,0) = \zeta = f(\gamma,\eta) = f\left(f(\gamma,x),\eta\right)$$
$$= f\left(y,\phi(x,y)\right),$$

from which, by the unique parameterization,

$$0 = \phi(x,y), \tag{7.18}$$

as required. We have thus shown that to every set of transformations of the above kind corresponds a unique ϕ, which is the ϕ-function of a local Lie group. Hence to every such transformation corresponds a local Lie group.

To prove the converse, that given any local Lie group we can construct transformations of the above kind, to which this local Lie group then corresponds, one proceeds as follows: Let $\phi(x,y)$ be the ϕ-functions for the local Lie group $u^{\alpha}_{\beta}(x)$ the corresponding u-functions, and $C^{\alpha}_{\beta\gamma}$, $\alpha,\beta,\gamma = 1,\ldots,r$, the structure constants satisfying

$$C^{\alpha}_{\beta\gamma} = - C^{\alpha}_{\gamma\beta} , \tag{7.19}$$

$$C^{\alpha}_{\gamma\lambda} C^{\lambda}_{\beta\delta} + C^{\alpha}_{\delta\lambda} C^{\lambda}_{\gamma\beta} + C^{\alpha}_{\beta\lambda} C^{\lambda}_{\gamma\delta} = 0$$

The u^{α}_{β}'s satisfy the equation

$$\frac{\partial u^{\alpha}_{\beta}(x)}{\partial x_{\gamma}} - \frac{\partial u^{\alpha}_{\gamma}(x)}{\partial x_{\beta}} = C^{\alpha}_{\mu\nu} u^{\mu}_{\gamma}(x) u^{\nu}_{\beta}(x) , \tag{7.20}$$

and by defining in the usual way

$$V^{\alpha}_{\beta}(x) u^{\beta}_{\gamma}(x) = \delta^{\alpha}_{\gamma} , \tag{7.21}$$

we obtain for the V^{α}_{β}, the differential equation

$$V^{\alpha}_{\mu,\lambda}(x) V^{\lambda}_{\nu}(x) - V^{\alpha}_{\nu,\lambda}(x) V^{\lambda}_{\mu}(x) = C^{\beta}_{\mu\nu} V^{\alpha}_{\beta}(x) \tag{7.22}$$

So for, this is merely a recapitulation of what we have done for the abstract group in the previous lectures. One now introduces the transformations as follows: One takes any set of variables $\zeta_1, \zeta_2, \ldots \zeta_n$, where $n \neq r$ in general, and defines for them the nr functions $V^{i}_{\alpha}(\zeta)$ $i = 1, \ldots n$, $\alpha = 1, \ldots r$, by letting them satisfy the equation

$$V^{i}_{\mu,j}(y) V^{j}_{\nu}(y) - V^{i}_{\nu,j}(y) V^{j}_{\mu}(y) = C^{\beta}_{\mu\nu} V^{i}_{\beta}(y) \tag{7.23}$$

which is obtained from (7.22) by letting the (subscript changes) on the V's in that equation change to ζ, η and the argument of the V's change to ζ. The boundary conditions at $y = o$ are arbitrary.

Q: How do you know that this equation defines a set of $V'_\alpha(\zeta)$, i.e. that it is integrable?

A: To check that (7.23) actually defines a set of $V'_\alpha(\zeta)$ one must examine the integrability conditions for it. But, on examination, the integrability conditions turn out to be just the equations (7.19), and these are satisfied by hypothesis.

Q: What about the uniqueness of the $V'_\alpha(\zeta)$?

A: They are not all unique. First of all the $C^\alpha_{\beta\gamma}$ are not unique, but, as we saw earlier, defined only up to an $ACA^{-1}A^{-1}$ transformation. Secondly, even for given boundary conditions (7=23) does not define the V's uniquely. Finally, the boundary conditions are arbitrary.

Having obtained a set of $V'_\alpha(\zeta)$ from (7.23) one next writes down the equations

$$\frac{\partial f_i(y,x)}{\partial x_\alpha} = V'_\beta(f)\, u^\beta_\alpha(x) \tag{7.24}$$

with the boundary conditions

$$f_i(y,0) = f_i \tag{7.25}$$

This equation is integrable, because if one writes down for it the integrability conditions, one sees that they are satisfied by virtue of (7.23) and (7.20). Further, it is clear that for a given set of $V_\alpha^i(y)$, (7.24) and (7.25) determine $f_i(\xi,x)$ uniquely. What we now wish to show is that the $f_i(y,x)$ so obtained are the f_i of a group of transformations of the kind introduced above. To show this, we need only show that

$$f\left(f(y,x)\,x'\right) = f\left(y,\phi(x\,x')\right), \tag{7.26}$$

where ϕ is the ϕ-function for the given Lie group, since from (7.26) and the properties of the ϕ one can then easily deduce that f satisfies the conditions (2) and (3) above. We shall prove (7.26) by a now familiar technique, namely, by calling the left hand side $A(x')$ and the right hand side $B(x')$ and showing that A and B satisfy the same differential condition: Putting $x' = 0$ in (7.26) we have

$$A(0) = f\left(f(y,x),0\right) = f(y,x) = f\left(y,f(x,0)\right) \tag{7.27}$$
$$= B(0),$$

so that this is certainly satisfied. Secondly, from (1.22) we have

$$\frac{\partial A(x')}{\partial x'} = \frac{\partial}{\partial x'}\left(f(f(y,x),x')\right) = V\left(f\left(f(y,x),x'\right)\right)u(x')$$
$$= V\left(A(x')\right)u(x'),$$

while

$$\frac{\partial B(x')}{\partial x'} = \frac{\partial f\left(y, \phi(x x')\right)}{\partial \phi(x x')} \quad \frac{\partial \phi(x x')}{\partial x'}$$

$$= V\left(f(y, \phi(x x'))\right) u(\phi) V(\phi) u(x')$$

$$= V\left(B(x')\right) u(x'). \tag{7.28}$$

Thus (7.26) holds, and we have shown how to construct from a given local Lie group, corresponding groups of transformations.

To complete this lecture, I shall just mention two particular. transformations which one can construct. By choosing for the $V_\alpha^\beta(y)$ the $V_\alpha^\beta(x)$ of the group itself (eq.(7.22) with $x \to y$ one obtains for (7=24)

$$\frac{\partial f_\beta(\zeta, x)}{\partial x_\alpha} = V_\sigma^\beta(f)\, u_\alpha^\sigma(x) \tag{7.29}$$

which is just the equation for $\phi(x, \zeta)$ itself. In other words, we have

$$f(y, x) = \phi(x, \zeta), \tag{7.30}$$

as a particular transformation. Note that

$$f\left(f(y, x)\, x'\right) = \phi\left(x', \phi(x, \zeta)\right) \tag{7.31}$$

$$= \phi\left(\phi(x' x)\, \zeta\right) = f\left(\zeta, \phi(x x)\right)$$

and that the y in $\phi(x\ y)$ plays the role of the variable to be transformed. The group of transformations so obtained is called the first parameter group. In a similar way, we can regard the x in $\phi(x,y)$ as the variable to be transformed and define the transformation by

$$f(x,y) = \phi(x,y) .$$

from which

$$f\left(f(x,y),y'\right) = \phi\left(\phi(x,y)y'\right) = \phi\left(x\,\phi(yy')\right) = f\left(y\,\phi(yy')\right) .$$

This is called the second parameter group

Lecture 8

Let

$$y' = f(x,\alpha) \tag{8.1}$$

be a group of transformations in the space x, as discussed in the last lecture. It is easy to show from the condition

$$y'' = f(x',\alpha) = f\left(f(x,\alpha),\alpha\right) = f\left(x,\phi(\alpha\alpha)\right), \tag{8.2}$$

that the f's satisfy the differential equation

$$\frac{\partial f_i}{\partial \alpha_\alpha} = V_\sigma^i(f)\ u_\alpha^\sigma(\alpha) \tag{8.3}$$

where the u's are the u's of the Lie group defined by the $\phi(\alpha,y)$ and the $V_\sigma^i(x)$ are the components of the "velocity field" defined by

$$V_\sigma^i(y) = \left[\frac{\partial f_i(y,\alpha)}{\partial \alpha_\sigma}\right]_{\alpha=0} , \tag{8.4}$$

and satisfying the equation

$$V_\sigma^i(y)\frac{\partial V_\alpha^i(y)}{\partial y_j} - V_\alpha^j(\xi)\frac{\partial V_\sigma^i(\xi)}{\partial \xi_j} = C_{\sigma\alpha}^{\lambda} V_\lambda^i(\xi) \qquad (8.5)$$

Conversely, as we showed in the last lecture, any set of satisfying (8.5), if inserted in (8.3), will yield a set of transformations f .

In order to express the equation (2.5) for in a simpler way, one introduces a set of quantities know as the <u>infinitesimal generators</u> of the transformation f . This one does by noting that on the left hand side of (8.5) we have the differential operators

$$X_\alpha = V_\alpha^j(\xi)\frac{\partial}{\partial \xi_j} \qquad (8.6)$$

occurring. These operators X_α are the infinitesimal generators, and using them, it is easy to see that (8.5) can be expressed in the form

$$[X_\alpha, X_\beta] = C_{\alpha\beta}^{\gamma} X_\gamma \qquad (8.7)$$

The algebra with elements X_α, addition law defined in the ordinary way be $X_\alpha + X_\beta$ and multiplication law defined by the <u>commutator</u> $[X_\alpha, X_\beta]$ is called the Lie algebra of the Lie group defined by the $C_{\alpha\beta}^{\gamma}$. The algebra may be over the field of real or complex number. It is a non-associative algebra since

$$[[X_\alpha, X_\beta]X_\gamma] \neq [X_\alpha[X_\beta, X_\gamma]] .$$

Note that the infinitesimal generators X_α are associated with the transformation group rather than the Lie group itself. In fact, they derive their name from the fact that they generate the transformations in the sense that if we know the X_α we know the $V_\alpha^i(\xi)$ and the latter determine the f_j^i according to (8.3). The X_α also the following significance. Let $\psi(\xi)$ be any function of the variables ξ and let us calculate the variation of $\psi(\xi)$ due to the change induced in the ξ when we vary the x_α from 0 to dx_α. We have,

$$\frac{\partial \psi(\xi)}{\partial x_\alpha} = \frac{\partial \psi(f(\xi, x))}{\partial f_j} \frac{\partial f_j(\xi, x)}{\partial x_\alpha} \tag{8.8}$$

so that,

$$\left[\frac{\partial \psi(\xi)}{\partial x_\alpha}\right]_{x=0} = \frac{\partial \psi(\xi)}{\partial \xi_j} V_\alpha^j(\xi) = X_\alpha \psi(\xi) \tag{8.9}$$

Thus, $X_\alpha \psi(\xi)$ is the first variation in ψ under such a transformation.

This completes our discussion of the general transformations associated with a Lie group. We shall now discuss the particular case of linear transformations.

Linear transformations associated with a Lie group

Suppose now that a transformation of the kind (8.1) is linear in ξ i.e. we have

$$\xi_i' = f_i(\xi_j, x_\alpha) = A_{ij}(x_\alpha)\, \xi_j \tag{8.10}$$

Then we have the following results

(1) The condition

$$y^{\circ\prime\prime} = f(y', x') = f\big(f(y\,x)\,x'\big) = f\big(y\,,\,\phi(x\,x')\big) \tag{8.11}$$

becomes

$$y_i'' = A_{ij}(x')\,y_j' = A_{ij}(x')\,A_{jk}(x)\,y_k = A_{ik}\big(\phi(x\,x')\big)\,y_k \tag{8.12}$$

or

$$A_{ij}(x')\,A_{jk}(x) = A_{ik}\big(\phi(x\,x')\big) \tag{8.13}$$

Translating this into -language, we have, on account of the unique parameterization,

$$\begin{aligned} A_{ij}\big(g(x')\big)\,A_{jk}\big(g(x)\big) &= A_{ik}\big(g\,[\phi(x'x)]\big) \\ &= A_{ik}\big(g(x')\,g(x)\big) \end{aligned} \tag{8.14}$$

But this is just the condition for the matrices A_{ij} to form a representation of the group $g[x]$. Thus we have a result: The transformation matrices $A(x)$ of any linear transformation (8.10) associated with a Lie group, form a representation of that group.

(2) The velocity field $u_\alpha^i(\xi)$ in this case is given by

$$u_\alpha^i(\xi) = \left[\frac{\partial f_i(\xi, x)}{\partial x_\alpha}\right]_{x=0} = \left[\frac{\partial A_{ij}(x)}{\partial x_\alpha}\right]_{x=0} \xi_j \tag{8.15}$$

$$= T_{ij}^{\alpha}\,\xi_j$$

where the

$$I^{\alpha}_{ij} = \left[\frac{\partial A_{ij}(x)}{\partial x_{\alpha}}\right]_{x=0} \tag{8.16}$$

a set of numerical matrices.

(3) The equation (8.5) for the $V^{i}_{\alpha}(\xi)$ becomes

$$I^{\sigma}_{jk}\xi_{k} I^{\alpha}_{ij} - I^{\alpha}_{ik}\xi_{k} I^{\sigma}_{ij} = +C^{z}_{\alpha\sigma} I^{z}_{jk}\xi_{k} \tag{8.17}$$

from which we see that the numerical matrices I^{α} satisfy the Lie algebra equation

$$\left[I^{\sigma}, I^{\alpha}\right] = C^{z}_{\sigma\alpha} I^{z} \tag{8.18}$$

and so form a "representation" of this algebra. It is usual to call the I^{σ}_{ij} the infinitesimal generators of the transformations (8.10), although, according to our earlier definition it is really the quantities

$$I^{\sigma}_{ij}\,\xi_{i}\frac{\partial}{\partial \xi_{j}} \tag{8.19}$$

(which also satisfy (8.18) which are the infinitesimal generators. Thus we have the result: If a set of matrices $A(x)$ form a representation of a Lie group, the corresponding infinitesimal generators I^{α} of (8.16) form a "representation" of the Lie algebra. Conversely, if we have a set of numerical matrices I^{α} satisfying (8.18), then from them we can construct a representation $A(x)$ of the group. The construction will be given explicitly below. For the moment we note only that this result

mean that in order to discuss the representations of the local Lie groups we can confine ourselves to the "representations" of the local Lie algebras. The advantage of doing this is that any representation of the algebra consists of a finite number of numerical matrices, while the corresponding representation of the group itself consists of an infinite number of matrices depending on the parameters x_α .

Explicit construction of the group representations.

Let I^{σ}_{ij} be a set of numerical matrices satifying (2.18) we construct from them the quantities

$$V^{i}_{\alpha}(\xi) = I^{\sigma}_{ij}\,\xi_j \,, \tag{8.20}$$

and insert these in the equation (2.3) for the f_i obtaining

$$\frac{\partial f_i(\xi, x)}{\partial x_\alpha} = I^{\sigma}_{ij}\, f_j\, u^{\sigma}_{\alpha}(x) \tag{8.21}$$

Since we wish to construct a linear transformation f , however, we set in (8.21) $f = A\xi$ and obtain

$$\frac{\partial A_{ik}(x)}{\partial x_\alpha}\,\xi_k = I^{\sigma}_{ij}\, A_{jk}(x)\,\xi_k\, u^{\sigma}_{\alpha}(x) \tag{8.22}$$

or

$$\frac{\partial A(x)}{\partial x_\alpha} = I^{\sigma} A(x)\, u^{\sigma}_{\alpha}(x) \tag{8.23}$$

Let us now use a normal system a parameters, so that

$$u^{\sigma}_{\alpha}(x)\, x^{\alpha} = x^{\sigma} \tag{8.24}$$

Further, in the parameter space we now choose any fixed point $\overset{0}{x}$ [fig (8.1)], and draw the straight line from $\overset{0}{x}$ to the origin O. Let t be the distance from O measured along this line, with $t_0 = O\overset{0}{x}$. The points on the line can be parameterized according to

$$x_\sigma(\text{on line}) = l_\sigma t \tag{8.25}$$

where l_σ are the direction cosines of the line $O\overset{0}{x}$ at O. As we move along this line $A(x)$ will be a function of $x(t)$ and hence of t only. But using (8.23), (8.24), (8.25) we have,

$$\frac{dA(t)}{dt} = \frac{\partial A(x(t))}{\partial x_\alpha(t)} \frac{dx_\alpha(t)}{dt} = \underline{I}^\sigma A(t) u^\sigma_\alpha(x)\, l_\alpha = \underline{I}^\sigma A(t) l_\sigma \tag{8.26}$$

Integrating this equation, with the initial condition we obtain

$$A(t) = e^{l_\sigma I^\sigma t} \tag{8.27}$$

from which, on putting $t = t_0$, we have the relation

$$A(\overset{i}{x}) = e^{\overset{0}{x}_\sigma I^\sigma} \tag{8.28}$$

THE INSTITUTE OF MATHEMATICAL SCIENCES
LIBRARY
Date, ...
MADRAS.

Letting $\overset{\circ}{x}$ now be a running coordinate, we have finally

$$A(x) = e^{x_\sigma I^\sigma} \tag{8.29}$$

This formula gives the connexion between the $A(x)$ and I_σ explicitly. Note that because we know that the integrability conditions for (8.23) are satisfied, we would arrive at the same $A(x)$ even if we choose a different path $\overset{c}{x}$. However, it is not so easy to show explicitly that we obtain (8.29) if we use another path and that is why we use the straight line.

Q: Suppose you took some other curve. Why would the above argument not go through in the same way for this=

A: In (8.26) we have used the fact that

$$u^\sigma_\alpha(x)\frac{dx_\alpha}{dt} = u^\sigma_\alpha(x)\ell_\alpha = u^\sigma_\alpha(x)\frac{x_\alpha}{t} = \frac{x_\sigma}{t} = \ell_\sigma \tag{8.30}$$

which holds only if

$$x_\alpha = \ell_\alpha t \tag{8.31}$$

Of course, if we use a general curve $x_\alpha(t)$ we must eventually arrive at the same result (8.29), but to obtain it we shall have to use explicitly the various properties of the $u^\sigma_\alpha(x)$ found the first series of lectures. By using the straight line we avoid using them explicitly. They are used implicitly, of course, in (8.24).

Note also that the formula (8.29)holds only for normal coordinates. If we transform to any set of non-normal coordinates. If we transform to any set of non-normal coordinates, i.e. make a non-linear transformation in parameter space, (8.29) no longer holds.

Finally, I shall like to mention a particular representation which exists for every Lie group. It is called the adjoint representation and is constructed as follows. One takes the structure constants $C^{\lambda}_{\mu\nu}$ of the group and regards the quantity.

$$I_{\sigma} = C^{\lambda}_{\sigma\mu} \tag{8.32}$$

for each fixed σ as a matrix, with matrix indices λ and μ, in the sense that the product $I_{\sigma} I_{\tau}$ is defined to be the matrix.

$$I_{\sigma} I_{\tau} = C^{\lambda}_{\sigma\mu} C^{\mu}_{\tau\omega} \tag{8.33}$$

One then easily verifies that the factor identity for the can be written in the form

$$\left[I_{\sigma}, I_{\tau} \right] = C^{\lambda}_{\sigma\tau} I_{\lambda} \tag{8.34}$$

But this means that the matrices I_{σ} satisfy the condition which is necessary and sufficient for them to be the infinitesimal generators of a representation of the Lie group, as defined above. The representation constructed from these according to (8.32) is just the adjoint representation, Note that the matrices of this representation are r-dimensional $(i.e.\ r \times r)$.

Q: Earlier you mentioned a group of transformations connected with every Lie group, namely those obtained by letting the y in $\phi(x,y)$ be the quantity to be transformed and letting ϕ be the transformed quantity i.e.

$$y' = \phi(xy) = \phi(x,\xi) \tag{8.35}$$

Is this transformation connected with the adjoint representation you have just defined?

A: No. It is true that in both case the number of base elements in the transformation space is the same, namely γ . But whereas the adjoint representation is a <u>linear</u> transformation, by definition, the above transformation is not, since if we expand ϕ we have

$$y_\mu' = x_\mu + y_\mu + c^{\mu}_{\lambda\nu} x_\lambda y_\nu + f^{\mu}_{\lambda\nu\sigma} x_\lambda y_\nu y_\sigma + \cdots \tag{8.36}$$

and only in special trivial cases is this linear.

Q: Sometimes the representation which you have called the adjoint representation is called the regular representation. Why is this?

A: My own opinion is that, strictly the name regular representation is wrong. To understand how it has come to be used, however, let me consider for a moment the case of a <u>finite</u> group g_s, $s = 1 \cdots r$ For such a group one can construct a group algebra (over the real or complex numbers) simply by taking as the base elements of the algebra the group elements g_s . The group algebra is a vector space of γ dimensions and on this space one can define a set of

linear operators or matrices M_s by the relations

$$M_s g_t = g_u \left(= g_s g_t\right) \qquad s,t,u = 1 \ldots r \tag{8.37}$$

These M's form a representation of the finite group, because from the associativity condition

$$g_s(g_t g_u) = (g_s g_t) g_u \ , \tag{8.38}$$

for the group, it follows that

$$g_s g_t = g_u \rightarrow M_s M_t = M_u \tag{8.39}$$

This representation is called the <u>regular</u> representation of the finite group. It is of dimension r .

The point now is that one can generalize this idea in two ways for the case of a Lie group. The direct generalization, is simply to replace the indices $s, t, u = 1 \ldots r$ above, by their counterparts $[x], [y], [z]$ for a Lie group. One obtains in this way the <u>regular</u> representation of the Lie group. The only new feature is that the number of matrices $M_{[x]}$ is now infinite, and that their dimension is also infinite ($r \times$ continuum).

On the other hand the adjoint representation defined above is an indirect generalization of the above idea, because the base elements of the representation space of the adjoint representation are just the base elements I_σ of the Lie algebra. But it is only an <u>indirect</u> generalization because whereas for the

regular representation the base elements of the representation space are the elements of the group for the adjoint representation they are elements of the Lie algebra.

One way to avoid confusion might be to say that we have two representations:

(a) the regular representation of the Lie group. This is infinite dimension.

(b) the adjoint representation of the Lie group. ~~This is infinite dimension.~~

= the regular representation of the Lie algebra. This is of dimension r .

Lecture 9

In the previous series of lectures we saw that a Lie group is completely determined by its structure constants $C^{\lambda}_{\mu\nu}$. It follows that any particular properties of the group or its subgroups should be expressible in terms of the $C^{\lambda}_{\mu\nu}$. The purpose of this lecture is to express some properties of the Lie groups, such as, for example, the existence of certain types of subgroup, in this way. I shall also indicate how these properties are then reflected in the Lie algebra.

(a) Abelian groups.

For an abelian group we have

$$g_{[x]}\, g_{[y]} = g_{[y]}\, g_{[x]} \qquad (9.1)$$

In terms of the $\phi(x,y)$ this becomes

$$\phi(x\,y) = \phi(y\,x) \,, \qquad (9.2)$$

from which we see at once that

$$\hat{C}^{\lambda}_{\mu\nu} = C^{\lambda}_{\mu\nu} - C^{\lambda}_{\nu\mu} = 0 \qquad (9.3)$$

Thus, for an abelian Lie group all the structure constants are zero. The converse is also true, namely, if the structure constants of a Lie group are all zero, the group is abelian.

Proof: The equations ( )for the $u^{\lambda}_{\nu}(x)$ become in this case

$$\frac{\partial u^{\lambda}_{\nu}(x)}{\partial x_{\sigma}} - \frac{\partial u^{\lambda}_{\sigma}(x)}{\partial x_{\nu}} = 0 \qquad (9.4)$$

$$u^{\lambda}_{\mu}(0) = \delta^{\lambda}_{\mu} \,. \qquad (9.5)$$

A solution of these equation is obviously

$$u^{\lambda}_{\mu}(x) = 0 \qquad (9.6)$$

and if we insert this into the equations ( ) for the $\phi(xy)$ we obtain

$$\frac{\partial \phi^{\mu}(x\,y)}{\partial x_{\nu}} = \delta^{\mu}_{\nu} \,, \qquad (9.7)$$

from which

$$\phi^{\lambda}(o\,y) = y^{\lambda} \,, \qquad (9.8)$$

$$\phi^{\lambda}(x\,y) = x^{\lambda} + y^{\lambda} \,.$$

Thus

$$\phi^{\lambda}(x\,y) = x^{\lambda} + y^{\lambda} = y^{\lambda} + x^{\lambda} = \phi^{\lambda}(y\,x) \,, \qquad (9.10)$$

and the group is abelian, as required.

Q: $u^{\lambda}_{\mu}(x) = \delta^{\lambda}_{\mu}$ is <u>one</u> solution of (3.4) and (9.5). What happens if we choose a different solution.

A: As mentioned in the general theory, a change of solution of these equations amounts only to a change in parameter $x \to x' = f(x)$ in the $\phi(xy)$. Since

$$\phi'(x',y') = f(\phi(xy)) = f(\phi(yx)) = \phi'(y'x') \,, \tag{9.11}$$

this change will not affect the abelian property.

The Lie algebra associated with an abelian group is very trivial, since every element of it commutes with every other one i.e.

$$[X_\mu, X_\nu] = 0 \,, \qquad \text{all } \mu, \nu \,. \tag{9.12}$$

Such a Lie algebra is called an <u>abelian</u> Lie algebra.

(b) Subgroups.

We consider now the case of a subgroup of a Lie group of a Lie group. We can arrange the parameters of the group so that the subgroup (of order m, say) consists of those elements for which $x_\alpha = 0$, $\alpha = m+1 \ldots r$. Then by definition

$$g_{[x'']} = g_{[x']}\, g_{[x]} \tag{9.13}$$

is a member of the subgroup if $g_{[x']}$ and $g_{[x]}$ are members. In other words

$$x''_\alpha = 0 \tag{9.14}$$

if

$$x_\alpha' = x_\alpha = 0 \tag{9.15}$$

Translating this result into ϕ-language we have

$$\phi_\alpha(x, x') = 0 \,, \tag{9.16}$$

if

$$x_\alpha = x_\alpha' = 0 \tag{9.17}$$

By expanding ϕ in the usual way, we see that a particular consequence of this is that

$$\hat{C}^{\alpha}_{ij} = 0, \quad 1 \le i, j \le m, \; m+1 \le \alpha \le r \tag{9.18}$$

This expresses the existence of the subgroup in terms of the structure constants.

We can also show that, conversely, if we have a set of structure constants (i.e. a set of numbers satisfying the anti-symmetry condition and the Jacobi identity) for which (9.18) holds, then the Lie group constructed with these contains as a subgroup, the set of elements with parameters $x_\alpha = 0$.

Proof: For $x_\alpha = 0$ the equations ( ) for the $u^\alpha_i(x)$ reduce to

$$\frac{\partial u^\alpha_i(x)}{\partial x_j} - \frac{\partial u^\alpha_j(x)}{\partial x_i} = C^\alpha_{\lambda\mu} u^\lambda_i(x) u^\mu_j(x) \tag{9.19}$$

$$u^\alpha_i(0) = 0 \,, \tag{9.20}$$

since all the others involve a differentiation with respect to x_α. But a solution of (9.19) is clearly

$$u^\alpha_i(x) = 0 \,, \qquad x_\alpha = 0 \,, \tag{9.21}$$

since if we assume this, the left hand side of (3.19) is identically zero, while the right hand side can give a contribution only if both λ and μ take on values in the range $1 \ldots m$ and in this case the C is zero by hypothesis. Note however, that (9.21) holds only for $x_\alpha = 0$ since we have neglected the equations involving derivatives with respect to x_α and these are coupled to (9.19) in a non-trivial way.

From the equation (9.21), it is easy to see that also

$$V^\alpha_i(x) = \delta \quad , \qquad x_\alpha = 0 \tag{9.22}$$

If we now set up the equations ( ) to determine the $\phi^\mu(xy)$ corresponding to the given structure constants, we have for $x_\alpha = y_\alpha = 0$ as a subset of these, the equation

$$\frac{\partial \phi^\alpha(xy)}{\partial x_\beta} = V^\alpha_\tau(\phi) u^\tau_j(x) = V^\alpha_i(\phi) u^i_j(x) + \tag{9.23}$$
$$V^\alpha_\beta(\phi) u^\beta_j(x) = 0 ,$$

using (9.21) and (8.22). On the other hand, if $y_\alpha = 0$, the initial condition $\phi^\alpha(0,y) = y^\alpha$ becomes $\phi^\alpha(0y) = 0$,

(9.24)

and combining these two equations, we have

$$\phi^\alpha(xy) = 0 , \quad \text{if} \quad x_\alpha = y_\alpha = 0 , \tag{9.25}$$

as required. Thus the condition $\hat{C}^\alpha_{ij} = 0$ is both necessary and sufficient for the existence of the above subgroup.

The existence of the subgroup is reflected in the Lie algebra in a very simple way. From the condition $\hat{C}^{\alpha}_{ij} = 0$ we have, namely, the result that

$$[X_i, X_j] = C^{k}_{ij} X_k \qquad i, j, k = 1 \cdots m, \quad (9.26)$$

which means that the X_i form a subalgebra of the Lie algebra. Thus to each subgroup of a Lie group corresponds a subalgebra of the Lie algebra.

(c) Invariant subgroups.

Let us now consider the case where the subgroup defined in the last few paragraphs is an <u>invariant</u> subgroup. This means that not only is

$$g_{[x'']} = g_{[x']} \cdot g_{[x]} \quad (9.27)$$

a member of the subgroup if $x_\alpha' = x_\alpha = 0$ but also

$$g_{[x'']} = g_{[x']} \cdot g_{[x]} \cdot g^{-1}_{[x']} \quad (9.28)$$

is a member of $x_\alpha = 0$ (for <u>all</u> x'). Let us now try to express these equations in terms of the $\hat{C}'s$. First, we have in terms of the $\phi's$ from (9.28),

$$\phi_\alpha\left(\phi(x', x), x'^{-1}\right) = 0, \quad \text{for } x_\alpha = 0 \quad (9.29)$$

Expanding this equation we have

$$\phi_\alpha(x', x) + x'^{-1}_\alpha + C^{\alpha}_{\mu\nu}\, \phi_\mu(x', x)\, x'^{-1}_\nu + \cdots = 0 \quad (9.30)$$

On the other hand, expanding $\phi(x',x)$ and x^{-1}, we have

$$\phi_\alpha(x',x) = x_\alpha' + x_\alpha + C^\alpha_{\mu\nu} x_\mu' x_\nu + \cdots , \quad (9.31)$$

$$x_\alpha'^{-1} = -x_\alpha' + C^\alpha_{\mu\nu} x_\mu' x_\nu' + \cdots , \quad (9.32)$$

and on inserting these equations into (9.30) and keeping only terms of second order, we obtain

$$x_\alpha' + x_\alpha + C^\alpha_{\mu\nu} x_\mu' x_\nu' - x_\alpha' + C^\alpha_{\mu\nu} x_\mu' x_\nu' + C^\alpha_{\mu\nu}(x_\mu' + x_\mu)(-x_\nu') = 0 \quad (9.33)$$

or (since $x_\alpha = 0$ by hypothesis)

$$\hat{C}^\alpha_{\mu\nu} x_\mu' x_\nu = \left(C^\alpha_{\mu\nu} - C^\alpha_{\nu\mu}\right) x_\mu' x_\nu = 0 \quad (9.34)$$

Since x' is completely arbitrary, and x_ν is arbitrary except for the condition $x_\alpha = 0$ we obtain from this the condition

$$\hat{C}^\alpha_{i\nu} = 0 \qquad i = 1 \cdots m, \ \alpha = m+1 \cdots r, \ \nu = 1 \cdots r \quad (9.35)$$

This is the condition for the existence of an invariant subgroup. interms of the $\hat{C}$'s, and, as might be expected it is stronger than the condition for the existence of a non-invariant subgroup since here ν runs from 1 to r whereas for a non-invariant subgroup (9.35) would hold only for $\nu = 1 \cdots m$.

As might also be expected, we have here again a converse theorem, namely that if we have a set of structure constants which satisfy (9.35), the Lie group constructed from them will contain an invariant subgroup of order m.

Proof: With the $\hat{c}'s$ as in (9.35) the equations ( ) for the $u^{\mu}_{\nu}(x)$ become

$$\frac{\partial u^{\alpha}_{\nu}(x)}{\partial x_{\tau}} - \frac{\partial u^{\alpha}_{\tau}(x)}{\partial x_{\nu}} = c^{\alpha}_{\sigma\lambda} u^{\sigma}_{\tau} u^{\lambda}_{\nu} = c^{\alpha}_{\beta\gamma} u^{\beta}_{\tau} u^{\gamma}_{\nu} \quad (9.36)$$

together with equations for the $u^{i}_{\nu}(x)$ which involve the $u^{\alpha}_{\nu}(x)$ on the right hand side only. Thus the derivatives of the $u^{\alpha}_{\nu}(x)$ are expressed completely in terms of the $u^{\alpha}_{\nu}(x)$ themselves, $\alpha = m+1 \ldots r$. It is in this respect that the invariant subgroup differentiates itself from an ordinary subgroup. As a result a solution of (9.35) (consistent with the equations for the $u^{i}_{\nu}(x)$ and the initial conditions $u^{\mu}_{\nu}(0) = \delta^{\mu}_{\nu}$ is

$$u^{\alpha}_{k}(x) = 0 \quad (9.37)$$

$$\frac{\partial u^{\alpha}_{\beta}(x)}{\partial x_{k}} = 0 \quad (9.38)$$

for <u>all</u> x, not merely $x_{\alpha} = 0$. This can easily be verified by writing (9.35) for $\nu, \tau = 1 \ldots m$, $\nu = 1 \ldots m$, $\tau = m+1 \ldots r$ and $\nu, \tau = m+1 \ldots r$ separately.

If we now consider the quantity $A(x) = \phi_{\alpha}\left(\phi(x'x)x'\right)$ for $x_{\alpha} = 0$, x' arbitrary, we have

$$A(0) = \phi_{\alpha}\left(x', x'^{-1}\right) = 0 \quad (9.39)$$

$$\frac{\partial A(x)}{\partial x_i} = \frac{\partial \phi_\alpha\left(\phi(x'x)\ x'^{-1}\right)}{\partial x_i} = \frac{\partial \phi_\alpha\left(\phi(x'x)\,x'^{-1}\right)}{\partial \phi_\mu(x'x)} \frac{\partial \phi_\mu(x'x)}{\partial x_i} \quad (9.40)$$

$$= \frac{\partial \phi_\alpha(\phi, x'^{-1})}{\partial \phi_j} \frac{\partial \phi_j(x'x)}{\partial x_i} + \frac{\partial \phi_\alpha(\phi, x'^{-1})}{\partial \phi_\beta} \frac{\partial \phi_\beta(x', x)}{\partial x_i} \quad (9.41)$$

$$= 0$$

since

$$\frac{\partial \phi_\alpha(\phi, x'^{-1})}{\partial \phi_j} = V^\alpha_\tau\left(\phi(\phi, x'^{-1}) U^\tau_j(\phi)\right) \quad (9.42)$$

$$= V^\alpha_i\left(\phi(\phi, x'^{-1})\right) U^i_j(\phi) + V^\alpha_\beta\left(\phi(\phi, x'^{-1})\right) U^\beta_j(\phi) = 0$$

from (9.37) and the corresponding equation for and in a similar way, it can be shown that

$$\frac{\partial \phi_\beta(x'x)}{\partial x_i} = 0 \quad (9.43)$$

(one needs a slight modification of (9142) to obtain (9.43) because in (9.43) the differentiation is with respect to the second argument of ϕ). From these results it follows that

$$A(x) = \phi_\alpha\left(\phi(x'x)\,x'\right) = 0 \quad (9.44)$$

for $x_\alpha = 0$ and this is just the condition for the subgroup $x_\alpha = 0$ to be invariant.

In this way we see that the condition $\hat{C}^{\alpha}_{i\lambda} = 0$ $\alpha = m+1 \ldots r$, $i = 1 \ldots m$, $\lambda = 1 \ldots r$ is a necessary and sufficient condition for the existence of an invariant subgroup. For the corresponding Lie algebra we have

$$[X_i, X_\lambda] = C^{k}_{i\lambda} X_k \qquad (3.44)$$

(9.44)

in which case the X_i are said to form an invariant subalgebra of the Lie algebra X_μ .

(d) Abelian subgroups.

Let us next consider the case of an abelian (not necessarily invariant) subgroup. From the subgroup property we have

$$\hat{C}^{\alpha}_{ij} = 0 \quad , \qquad (9.45)$$

while from the abelian property we have

$$\hat{C}^{k}_{ij} = 0 \qquad (9.46)$$

Combining these two equations, we have

$$\hat{C}^{\mu}_{ij} = 0 \qquad (9.47)$$

As is easily seen from our earlier consideration regarding subgroups and abelian groups, this condition (9.47) is both a necessary and sufficient condition for the existence of an abelian subgroup. The corresponding Lie glgebra has clviously the abelian subalgebra

$$[X_i, X_j] = 0 \tag{9.48}$$

(e) Invariant abelian subgroups.

Finally, we consider the case of an <u>invariant abelian</u> subgroup we have the above condition

$$\hat{C}^{\mu}_{ij} = 0 \quad , \tag{9.49}$$

but in addition we have the invariant condition obtained earlier

$$\hat{C}^{\alpha}_{i\lambda} = 0 \tag{9.50}$$

These two conditions are characteristic for the existence of an invariant abelian subgroup because from our previous considerations regarding abelian and invariant subgroups it is easy to see that the two conditions (9.49) and (9.50) are not only necessary but also sufficient for the existence of an invariant abelian subgroup.

We can summarize the results just obtained in the following table:

Subgroup	$\hat{C}^{\alpha}_{ij} = 0$	$[X_i, X_j] = C^{k}_{ij} X_k$
Invariant subgroup	$\hat{C}^{\alpha}_{i\lambda} = 0$	$[X_i, X_\lambda] = C^{k}_{i\lambda} X_k$
abelian subgroup	$\hat{C}^{\mu}_{ij} = 0$	$[X_i, X_j] = 0$
Invariant abelian subgroup	$\hat{C}^{\mu}_{ij} = 0$ $\hat{C}^{\alpha}_{i\lambda} = 0$	$[X_i, X_j] = 0$ $[X_i, X_\lambda] = C^{k}_{i\lambda} X_k$

$i, j, k = 1 \cdots m$, $\alpha = m+1 \cdots r$, $\mu, \lambda = 1 \cdots r$

As mentioned in the general theory, groups which contain no invariant subgroups are called simple, while those which contain no invariant abelian subgroups are called-semi-simple. It turns out in the investigation of Lie algebras that there is a fundamental difference between Lie algebras which are simple and those which are semi-simple. (We call a Lie simple or semi-simple if the corresponding Lie group is simple or semi-simple). The difference between simple and semi-simple algebras is not so important because it can be shown without much difficulty that any semi-simple Lie algebra is the direct sum of simple Lie algebras.

From now on, we shall be concentrating our attention on the semi-simple Lie algebras and their representations. This is because, up to the present, only semi-simple Lie algebras have had any physical significance. It is convenient, therefore, to introduce at this point a criterion for the semi simplicity of a Lie algebra. The criterion is as follows: Let $C_{\mu\lambda}^{\nu}$ be the structure constants, and let us define the quantities

$$g_{\mu\lambda} = C_{\mu\sigma}^{\tau} C_{\lambda\tau}^{\sigma} \tag{9.51}$$

where a summation over σ and τ is understood. (Note that $g_{\mu\nu}$ is symmetric in μ and ν). Then the group is semi-simple if, and only if

$$\det g_{\mu\nu} \neq 0 \tag{9.52}$$

To prove that for an algebra which is not semi-simple, $\det g_{\mu\nu} = 0$ is not difficult. In fact, from the table given above we see that if the group is not semi-simple, i.e. if it contains an invariant abelian subgroup, we have

$$\begin{aligned} g_{i\lambda} &= C_{i\mu}^{\sigma} C_{\lambda\sigma}^{\mu} \\ &= C_{i\mu}^{k} C_{\lambda k}^{\mu} \\ &= C_{ij}^{k} C_{\lambda k}^{j} \\ &= 0 \end{aligned} \tag{9.53}$$

Thus the i row in the matrix $g_{\mu\nu}$ consists of zeros and since there is at least one such row, $\det g_{\mu\nu} = 0$

To prove that, conversely, if the group is semi-simple, $\det g_{\mu\nu} \neq 0$ is not so easy, and involves a number of results from the general theory of non-semi-simple groups which I do not wish to discuss in detail here. Hence I shall not give the proof of this result immediately, but for those who are interested I shall give a sketch of the proof in appendix B.

The semi-simple groups are, therefore, characterized by the fact that the matrix $g_{\mu\nu}$ is non-singular. One of the consequences of the non-singular nature of $g_{\mu\nu}$ for these group is that we can define for them a "contravariant" counterpart $g^{\mu\nu}$ to $g_{\mu\nu}$ by means of the relations

$$g_{\mu\sigma} g^{\sigma\nu} = \delta_{\mu\nu} \tag{9.54}$$

The $g_{\mu\nu}$ and $g^{\mu\nu}$ can then be used as "metric tensors" for raising and lowering indices on the $C_{\mu\lambda}^{\nu}$ i.e. we can define "completely covariant" structure constants $C_{\lambda\mu\nu}$ by the relation

$$C_{\lambda\mu\nu} = g_{\lambda\sigma} C^{\sigma}_{\mu\nu}, \tag{9.55}$$

and from these we can recover the $C^{\lambda}_{\mu\nu}$ by the relation

$$C^{\lambda}_{\mu\nu} = g^{\lambda\sigma} C_{\sigma\mu\nu}, \tag{9.56}$$

which follows from (9.54) and (9.55). Note that even for a non-semi-simple group we could define the $C_{\lambda\mu\nu}$ according to (9.55) But for such a group we could not recover the $C^{\lambda}_{\mu\nu}$ from the $C_{\lambda\mu\nu}$.

One of the advantages of introducing the $C_{\lambda\mu\nu}$ is that these quantities (unlike the $C^{\lambda}_{\mu\nu}$) are completely anti-symmetric in all the indices λ, μ, ν . .

Proof: Obviously $C_{\lambda\mu\nu}$ is anti-symmetric in μ and ν so what we have to show is that it is also anti-symmetric in λ and μ . We have

$$C_{\lambda\mu\nu} + C_{\mu\lambda\nu} = g_{\lambda\sigma} C^{\sigma}_{\mu\nu} + g_{\mu\sigma} C^{\sigma}_{\lambda\nu}$$
$$= C^{\omega}_{\lambda\tau} C^{\tau}_{\sigma\omega} C^{\sigma}_{\mu\nu} + C^{\omega}_{\mu\tau} C^{\tau}_{\sigma\omega} C^{\sigma}_{\lambda\nu} \qquad (9.54)$$

In each item on the right hand side here we have twooof the three "free" indices λ, μ, ν occurring in one C. We now use the Jacobi identity to rearrange the terms so that we have one free index for each C i.e. we have

$$C_{\lambda\mu\nu} + C_{\mu\lambda\nu} = C^{\omega}_{\lambda\tau}\left[C^{\tau}_{\mu\sigma} C^{\sigma}_{\nu\omega} + C^{\tau}_{\nu\sigma} C^{\sigma}_{\omega\mu}\right]$$
$$+ C^{\omega}_{\mu\tau}\left[C^{\tau}_{\lambda\sigma} C^{\sigma}_{\nu\omega} + C^{\tau}_{\nu\sigma} C^{\sigma}_{\omega\lambda}\right], \qquad (9.55)$$

$$= C^{\omega}_{\lambda\tau} C^{\tau}_{\mu\sigma} C^{\sigma}_{\nu\omega} - C^{\omega}_{\lambda\tau} C^{\sigma}_{\mu\omega} C^{\tau}_{\nu\sigma} \qquad (9.56)$$
$$+ C^{\tau}_{\lambda\sigma} C^{\omega}_{\mu\tau} C^{\sigma}_{\nu\omega} - C^{\sigma}_{\lambda\omega} C^{\omega}_{\mu\tau} C^{\tau}_{\nu\sigma}$$

On changing the indices in the second two terms on the right hand side of this equation, one sees that the four terms cancel in pairs, and so we get

$$C_{\lambda\mu\nu} + C_{\mu\lambda\nu} = 0 \qquad (9.57)$$

as required.

Q: Does this not indicate that there are four independent $C_{\lambda\mu\nu}$'s than $C^{\lambda}_{\mu\nu}$'s and if so, how is this compatible with the fact that we can recover the $C^{\lambda}_{\mu\nu}$'s from the $C_{\lambda\mu\nu}$'s

A: If there were no Jacobi identity, then there would be $\frac{n^2(n-1)}{2}$ independent $C^{\lambda}_{\mu\nu}$ and only $\frac{n(n-1)(n-2)}{3!}$ independent $C_{\lambda\mu\nu}$, which is, I think, what you mean. But always in the background is the Jacobi identity, and since, as you say, (9.54), (9.55) and (9.56) imply that there are the same number of independent $C_{\lambda\mu\nu}$ and $C^{\lambda}_{\mu\nu}$ one can only assume that the Jacobi identity is a stronger condition on the $C^{\lambda}_{\mu\nu}$ than on the $C_{\lambda\mu\nu}$. This is reasonable, because, in fact, the completely anti-symmetric character of the $C_{\lambda\mu\nu}$ is a <u>consequence</u> of the Jacobi identity so in assuming the anti-symmetry we have already used some of the conditions on the $C_{\lambda\mu\nu}$ implied by that identity.

Lecture 10

In the following "a" lectures, we shall start from the semi-simple Lie algebra

$$[X_\mu, X_\nu] = C^\sigma_{\mu\nu} X_\sigma \quad , \quad \mu, \nu, \sigma = 1 \ldots r \tag{10.1}$$

$$|g_{\mu\nu}| = |C^\sigma_{\mu\lambda} C^\lambda_{\nu\sigma}| \neq 0 \tag{10.2}$$

and proceed from there. There are four main things which we wish to do. First, we wish to prove Cartans theorem which reduces the Lie algebra from the general form (10.1) to the much simpler canonical form. Secondly, having got them in canonical form we wish to discuss how this form can be used to classify all the semi-simple Lie algebras by means of root diagrams. Thirdly, we wish to discuss how the representations can then be obtained by means of weight diagrams. Finally, if we have time, we wish to show how to construct all the invariants of all the semi-simple Lie algebras.

The first three lectures will be devoted to proving Cartan's theorem. We start by discussing the idea of a real and a complex Lie algebra. The Lie algebra (a) is said to be complex if in the algebra are included all linear combinations.

$$X = a_\lambda X_\lambda \tag{10.3}$$

with complex coefficients a_λ . In other words, the algebra is complex if it is defined over the field of complex numbers. For a complex Lie algebra, the structure constants may be

real, pure imaginary, or complex. The Lie algebra is said to be real if the following two conditions are fulfilled.

(1) the structure constants $C_{\lambda\mu}^{\sigma}$ are all pure imaginary.

(2) in the algebra are included only all linear combinations

$$X = a_\lambda X_\lambda \tag{10.4}$$

with real coefficients a_λ.

It is clear that a real Lie algebra can always be extended to a complex Lie algebra (simply by including as elements all linear combinations $a_\lambda X_\lambda$ with a_λ now complex). However, two or more different real Lie algebras may have the same complex extension. The question now is: given any complex Lie algebra, does there exist always a real Lie algebra of which it is the complex extension? The answer (due to Cartan) is "yes". In fact, for each complex Lie algebra there exist many real Lie algebras of which it is the complex extension. However, Contan was also able to show that out of all the real Lie algebras corresponding to a given complex one, one and only one, is the real Lie algebra of a compact Lie group. What compact means, need not concern us here, since we are dealing only with the algebra of the infinitesimal group, except for the following important result which we shall state without proof: For a compact real Lie group, the matrix $g_{\lambda\mu}$ of (10.2) is not only real and symmetric, but negative definite as well.

To illustrate the above remarks, let us consider the rotation group in three dimensions. This is a real Lie group with algebra.

$$[T_i, T_j] = i T_k, \quad i,j,k = 1,2,3 \text{ cyclically} \tag{10.5}$$

Its complex extension consists of all linear combinations

$$T = a_i T_i \tag{10.6}$$

with complex coefficients a_i. Now consider the following subset of the elements in the complex extension:

$$\begin{aligned} X_3 &= i T_3 \\ X_+ &= i(T_1 + i T_2) \\ X_- &= i(T_1 - i T_2) \end{aligned} \tag{10.7}$$

and

$$X = a_3 X_3 + a_+ X_- + a_- X_-$$

with <u>real</u> coefficients a_3, a_+, a_-. It is easily seen that this set forms a real Lie algebra with the same complex extension as that of the T_i. But it is not the same real Lie algebra of the T_i because the X's cannot be expressed as linear combinations of the T_i with real coefficients.

In this example, the X algebra is not compact, but the T-algebra is. It is easy to verify that, with (10.5),

$$g_{ij} = c_{i\lambda}^{\mu} c_{j\mu}^{\lambda} = -\delta_{ij} \tag{10.8}$$

After these preliminaries, we turn to the task of proving Cartan's theorem. We take any complex Lie algebra (10.1), <u>and we choose as our basis for it, X's, which are also a basis for the</u>

unique compact real Lie algebra of which it is the complex extension. We see we can do this from the above discussion.

We now form $g_{\lambda\mu}$. This matrix is symmetric. But because we are using a real Lie algebra it is also real, and because we are using compact real Lie algebra it is negative definite. Hence from the standard theory of matrices, it can be brought to diagonal form with -1's along the diagonal by a real orthogonal transformation i.e $A_{\sigma\nu}$ exists such that

$$A A^T = A^T A = 1 \tag{10.9}$$

$$A^T g A = A_{\sigma\nu} g_{\sigma\lambda} A_{\lambda\mu} = -\delta_{\nu\mu} \tag{10.10}$$

If we apply the transformation A to our basis X_λ we obtain a new basis $\hat{X}_\lambda$ where

$$\hat{X}_\lambda = A_{\lambda\nu} X_\nu \tag{10.11}$$

The structure constants for the new basis will be, from 10.1

$$\hat{C} = \hat{C}^{\hat{\lambda}}_{\mu\nu} = A_{\mu\sigma} A_{\nu\tau} (A^{-1})_{\lambda\omega} C^{\omega}_{\sigma\tau} = A_{\mu\sigma} A_{\nu\tau} A_{\lambda\omega} C^{\omega}_{\sigma\tau} \tag{10.12}$$
$$= A A C A^T$$

and the corresponding g will be

$$\hat{g}_{\mu\nu} = \hat{C}^{\omega}_{\mu\sigma} \hat{C}^{\sigma}_{\nu\omega} = A_{\mu\mu'} A_{\sigma\sigma'} A_{\omega\omega'} A_{\nu\mu''} A_{\omega\omega''} A_{\sigma\sigma''} \times C^{\omega'}_{\mu'\sigma'} C^{\sigma''}_{\mu''\omega''}$$

$$= A_{\mu\mu'} A_{\nu\mu''} C^{\omega}_{\mu'\sigma} C^{\sigma}_{\mu''\omega}$$

$$= A_{\mu\mu'} A_{\nu\nu'} g_{\mu'\nu'}$$

$$= -\delta_{\mu\nu} \tag{10.13}$$

from (10.10). Thus by the change of basis (10.12) we can bring g to the simple form (10.13). Note particularly that since A is real this transformation does not take us out of the space of the real compact algebra with which we started. From now on we drop the *hat* on $\hat{X}_\lambda$ and assume that we are working with a basis inthe real compact algebra such that holds.

Let us now choose any element

$$X = a_\lambda X_\lambda \tag{10.13}$$

out of the real compact algebra. From 10.1 we have,

$$[X, X_\nu] = a_\lambda [X_\lambda, X_\nu] = a_\lambda C^\sigma_{\lambda\nu} X_\sigma = L_{\nu\sigma} X_\sigma \tag{10.14}$$

where

$$L_{\nu\sigma} = a_\lambda C^\sigma_{\lambda\nu} \tag{10.15}$$

This show that the operation of commuting X with the X_ν is a <u>linear</u> operation in the space of the X_ν. <u>Thus, to each X in the Lie algebra corresponds a linear operator $L_{\nu\sigma}$ in the</u> space X_ν. Our proof of Cartan's theorem will bebased on the fact that each X can be "represented" by an L in this way.

Note that the L_μ corresponding to the X_μ in this way are just the infinitesimal elements of the <u>adjoint</u> representation of the algebra mentioned earlier. In particular they satisfy the relation

$$[L_\mu, L_\nu] = C^\lambda_{\mu\nu} L_\lambda$$

Q: To each X corresponds a unique L but is the converse true?

A: I am just coming to that. The answer is that the correspondence is unique. i.e.

$$X \leftrightarrow L \tag{10.16}$$

if , and only if, the algebra is semi-simple

Proof: Suppose

$$\begin{aligned} X_1 &\rightarrow L_1 \\ X_2 &\rightarrow L_1 \\ X_1 &\neq X_2 \end{aligned} \tag{10.17}$$

Then since the correspondence is linear

$$X_1 - X_2 \longrightarrow 0 \tag{10.18}$$

Hence

$$\left[X_1 - X_2 \, , X_\nu\right] = 0 \tag{10.19}$$

for all X_ν. Consider now the group generated by $X_1 - X_2$ can hold only if the group is not semi-simple. Conversely, it is easy to see that if it is not semi-simple we obtain relations of the kind (10.13).

We have, therefore, a one-to-one correspondence between X and L . The rule for getting from X to L and vice versa is obvious. If X commutes with a vector $b_\sigma X_\sigma, L$ multiplies it to give the same result If L multiplies $b_\sigma X_\sigma$, X commutes with it to give the same result. Of course not every matrix M acting on the r space X_λ has a corresponding X . We are considering only the subset of M which have an X.

Our next step is to choose out of all the X in the compact real algebra one definite element X_0 . Let L_0 be the corresponding L . We now diagonalize L_0 . Note that we can talk about diagonalizing L_0 because it is a matrix. To talk about diagonalizing X_0 would make no sense. This illustrates some of the advantage to be gained by introducing L .

Q: How do you know that L_0 can be completely diagonalized?

A: We have from (10.11).

$$L_{0\nu\lambda} = g_\sigma C^{\lambda}_{\sigma\nu} = a_{0\sigma} C_{\mu\sigma\nu} g^{\mu\lambda} = L'_{0\mu\nu} g^{\mu\lambda} \quad (10.15)$$

But the $a_0\sigma$ are real, the $C^{\lambda}_{\sigma\nu}$ are pure imaginary, the $g^{\mu\lambda}$ are real, and the $C_{\mu\sigma\nu}$ are pure imaginary. On the other hand we know from the general theory that the $C_{\mu\sigma\nu}$ are anti-symmetric in all indices. Thus $L'_{0\mu\nu}$ is anti-symmetric and pure imaginary. In other words, it is hermitian. Similarly since $g^{\mu\lambda}$ is real and symmetric it is hermitian. Further, it is non-singular, by the semi-simple hypothesis.

i.e. the group

$$D = 1 + i\varepsilon (X_1 - X_2) \quad (10.20)$$

where ε is a parameter. This is a one-parameter and therefore abelian, subgroup of the original group. But by (10.19) it is invariant. Thus the original group contains an invariant abelian subgroup and is not semi-simple. Thus (10.17) can hold only if the group is not semi-simple. Conversely, it is easy to see that if the group is not semi-simple, we obtain relations of the kind (10.17).

We have, therefore, a one to one correspondence between X and L. The rule for getting from X to L and vice versa is obvious. If X commutes with a vector $b_\sigma X_\sigma$, L "multiplies" it to give the same result. If L "multiplies" $b_\sigma X_\sigma$, X commutes with it to give the same result. Of course, not every matrix M acting on the r space X_λ has a corresponding X. We are considering only the subset L of M which have X

Let us now consider one definite element $X_0 = a^0_\sigma X_\sigma$ out of the compact real Lie algebra. Let L_0 be the corresponding L. We now diagonalize L_0. Note that we can talk about diagonalizing L_0 because it is a matrix. To talk about diagonalizing X_0 would be meaningless. It is for reasons such as this that X_0 is introduced.

Q: How do you know that L_0 can be diagonalized completely?

A: By definition.

$$L_0 = a^0_\sigma C^\lambda_{\sigma\mu} = - a^{(0)}_\sigma C_{\sigma\lambda\mu} \qquad (10.21)$$

on stepping down the index λ with the matrix tensor (10.13) But we know from the general theory that the $C_{\sigma\lambda\mu}$ are anti-symmetric in allindices. Hence for any $a^{(0)}_\sigma$, L is an anti-symmetric matrix. Further since the $a^{(0)}_\sigma$ are real, and the $C^\lambda_{\sigma\mu}$ pure imaginary, L_0 is pure imaginary. Hence L_0 is actually hermitian, and so can be diagonalized completely. Furthermore, its eigenvalues will be real.

This might be a good place to mention some other properties of L_0 which follow immediately from the fact that L_0 is both hermitian and anti-symmetric. First, we note that if V is an eigen-vector of L_0 with eigenvalue λ ,

$$L_0 V = \lambda V \quad , \tag{10.22}$$

then V^* is an eigenvector with eigenvalue $-\lambda$. From this it follows that

(a) The non-zero eigenvalues of L_0 occur in pairs, λ and $-\lambda$ with the same degenerach for λ and $-\lambda$

(b) The eigenvectors $V_1 X_1 + V_2 X_2 + \cdots + V_r X_r$ and $V_1^* X_1 + V_2^* X_2 + \cdots + V_r^* X_r$ belonging to the non-zero eigenvalues of L_0 have not got completely real coefficients (otherwise V and V^* could not be orthogonal) and hence do not belong to the real compace Lie algebra. Thus to diagonalize L_0 we must move out of the space of the compact Lie algebra and into its complex extension.

(c) If V is an eigenvector of L_0 with zero eigenvalue, then so is V^*. Hence we can choose a real basis in the eigenspace of L_0 so that the base X's still belong to the real compact algebra. The importance of this will be seen later.

All the properties which I have just mentioned depend very largely, of course, on the assumptions that the algebra is semi-simple and that we are using the real compact algebra and the form $-\delta_{\mu\nu}$ for $g_{\mu\nu}$ to start with.

Note that L_o has always at least one zero eigenvalue. This is because

$$L X_o = [X_o, X_o] = 0 \tag{10.24}$$

Thus in diagonal form L_o will look something like

$$L_o = \left(\begin{array}{ccccc|c} 0 & & & & & X_o^i \\ & \alpha & & & & X_\alpha^r \\ & & \beta & & & X_\beta \\ & & & \gamma & & X_\gamma^r \\ & & & & & \vdots \end{array}\right) \tag{10.25}$$

To the right of L_o we have put in the basis with respect to which L_o is diagonal. As just mentioned, the X_α^r, X_β, ... are not members of the original real compact algebra, but the X_o^i (and indeed the combinations $X_\alpha^r + X_{-\alpha}^r$, $i(X_\alpha^r - X_{-\alpha}^r)$, ...) are.

A word here about notation. We have labelled our base elements with the subscripts α, β, ... corresponding to the eigenvalues of L_o to which they belong. We have given them a superscript which is used to differentiate between the different elements belonging to the same (degenerate) eigenvalue. For reasons of convenience we shall use i, j, k, for the superscript in the zero eigenspace and r,s,t for the superscript in the non-zero ones.

We use the same letters r,s,t for each of the latter spaces, but this is not meant to imply that they will have the same range in each of these spaces. Thus, $r = 1, 2, 3$ in α space, but $r = 1, 2$ in γ-space. We hope that this will lead to no confusion We should also mention that in (10.1) we have used $\mu, \nu, \lambda = 1 \cdots r$ The μ, ν, λ should not be confused with the α, β, γ (from the beginning of the alphabet)which are used above in quite a different way.

I should like to conclude this lecture by proving the following theorem, which is known as the "step up-step down" theorem.

Theorem: The vector $\left[X_\alpha^r, X_\beta^s \right]$ is either zero, or is in the space $X_{\alpha+\beta}^t$.

Proof:
$$L_0 \left[X_\alpha^r, X_\beta^s \right] = \left[X_0, \left[X_\alpha^r, X_\beta^s \right] \right]$$
$$= \left[\left[X_0, X_\alpha^r \right] X_\beta^s \right] + \left[X_\alpha^r, \left[X_0, X_\beta^s \right] \right] \quad (10.24)$$
$$= \left[\alpha X_\alpha^r, X_\beta^s \right] + \left[X_\alpha^r, \beta X_\beta^s \right]$$
$$= (\alpha + \beta) \left[X_\alpha^r, X_\beta^s \right] \quad Q.E.D.$$

This theorem is known as the "step up-step down" theorem because if X_β^r is in the β-space and α is positive, we can step X_β^r up into the space $X_{\beta+\alpha}^t$ by operating on it with L_α^r. Similarly, by operating on it with $L_{-\alpha}^r$ we can step it down into the space $X_{\beta-\alpha}^t$. Note also the following corollary to this theorem if $\alpha + \beta$ is not an eigenvalue of L_0, $\left[F_\alpha^r, F_\beta^s \right]$ is necessarily zero.

Lecture 11

In the last lecture we saw that we could diagonalize the matrix of the adjoint representation L_o corresponding to any element X_o of a compact Lie algebra, and that L_o in that case took the form . We shall begin this lecture by showing that if X_o^i are the base elements in the zero-eigenvalue space of L_o as in (then the X_o^i have the "block" form

$$X_o^i = \begin{pmatrix} M_i & & & \\ & A_i & & \\ & & b_i & \\ & & & c_i \end{pmatrix} \begin{pmatrix} X_o^i \\ \hline X_\alpha^h \\ \hline X_\beta \\ \hline X_\gamma^h \\ \hline i \end{pmatrix} \qquad (11.1)$$

with respect to the same basis.

Proof: Since X_o^i belongs to the compact algebra, L_o^i is diagonalizable. But by the rule of correspondence (R.C.) $X \leftrightarrow L$, we have

$$\begin{aligned} [L_o, L_o^i] X_\mu &= L_o L_o^i X_\mu - L_o^i L_o X_\mu \\ &= [X_o [X_o^i X_\mu]] - [X_o^i [X_o X_\mu]] \\ &= [[X_o, X_o^i] X_\mu] \qquad (11.2) \\ &= [L_o X_o^i, X_\mu] \\ &= [0, X_\mu] \\ &= 0 \end{aligned}$$

Since this is true for all X_μ, we have

$$[L_o, L_o^i] = 0$$

from which (11.1) follows by a standard theorem on matrices.

Corollary:

$$[X_o^i, X_o^j] = L_o^i X_o^j = c_k X_o^k \qquad (12.3)$$

Up to now the L_o we have chosen has been completely aribitrary. We shall now choose it so that

(1) Of all the elements of the compact algebra L_o is one of the ones with the least number of zeros (l zeros, say).

(2) Of all the elements with l zeros, it is one of the ones with the maximum number of distinct non-zero eigenvalues.

For such as L_o we shall show that the corresponding L_o^i are not merely "blocked" as in (11.1) but are actually of the form

$$L_o^i = \left(\begin{array}{c|c|c|} 0 & & 0 \\ \hline 0 & \alpha_i I & 0 \\ \hline 0 & 0 & \beta_i I \\ \hline \end{array}\right) \qquad (11.5)$$

The proof of this (which will be the main part of this lecture) is as follows: Let

$$\hat{L}_o = L_o + \varepsilon L_o^i \qquad (11.6)$$

where ε is an arbitrarily small real number, and L_o^i is any arbitrary L_o^i. We choose ε so that

$$|\varepsilon| \ll |\alpha|, |\beta|, |\gamma| \ldots , \tag{11.7}$$

and

$$|\varepsilon| \ll |\alpha-\beta|, |\alpha-\gamma|, |\beta-\gamma|, \ldots \tag{11.8}$$

i.e. ε is much smaller than the non-zero eigenvalues of L_o, and also much smaller than their differences.

By (11.7), $\hat{L}_o$ has no zero eigenvalues except in the X_o^i space. But by hypothesis $\hat{L}_o$ has at least ℓ zero eigenvalues. Hence all the eigenvalues of $\hat{L}_o$ in X_o^i -space are zero. But L_o is zero in this space. Hence from (11.6) all the eigenvalues of L_o^i in X_o^i space are zero. Hence L_o^i is identically zero in X_o^i -space. Hence

$$X_o^i \leftrightarrow L_o^i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_i & 0 \\ 0 & 0 & C_i \end{pmatrix} \tag{11.9}$$

at any rate. Note that $\hat{L}_o$ like L_o has exactly ℓ zero eigenvalues.

Now we use (11.8). This condition tells us that $\hat{L}_o$ has at least as many distinct non-zero eigenvalues as L_o since $\hat{L}_o$ has at least one eigenvalue in the ε -neighbourhood of α, of β, of γ etc. and by (11.7)

these neighbourhoods do not overlap. But since $\hat{L}_o$ has only ℓ zero eigenvalues, by the definition (2) of L_o, it cannot have more distinct non-zero eigenvalues than L_o. Hence it must have the same number. But obviously this can happen only if $\hat{L}_o$ is a multiple of the identity in each "block". In that case, from (11.6), we see that L'_o is also a multiple of the identity in each block, as required (11.5).

Q: Why not carry the proof through with L'_o itself instead of $\hat{L}_o = L_o + \epsilon L'_o$?

A: On account of the conditions (11.7), (11.8) we are sure that $\hat{L}_o$ has no zeros outside the X'_o block. For L'_o this may not be true, and the proof will break down.

Q: Does using $\hat{L}_o$ not also avoid the difficulty that L'_o may have more than one distinct eigenvalue in a given block, but not more distinct eigenvalues than L_o altogether?

A: Yes. That is exactly right.

Corollary I: $$[X'_o, X'_o] = L'_o X'_o = 0 \qquad (11.10)$$

This (refinement of (10.22)) means that the set of ℓ operators X'_o, $i = 1 \cdots \ell$, form an abelian subalgebra of the original Lie algebra (10.1). This subalgebra of order ℓ is called the Cartan algebra.

Q: Does this not contradict the assumption that the group is semi-simple?

A: No. The subalgebra is abelian, but it is not invariant. Only invariant abelian subalgebras are forbidden.

We are now nearing our goal of writing the Lie algebra (10.1) in its canonical form. We have, namely, now obtained the result that $\{X_\lambda \cdots\} \rightarrow \{\{\cdots X_0^i \cdots\}, \{\cdots X_\alpha^\beta \cdots\}, \{X_\beta\}, \cdots\}$ where

$$[X_0^i, X_0^j] = 0$$
$$[X_0^i, X_\alpha^\beta] = \alpha_i X_\alpha^\beta$$
$$[X_\alpha^\beta, X_\beta^s] = \sum_t C_{\alpha\beta}(\beta s t) X_{\alpha+\beta}^t \quad \text{if } \alpha+\beta \text{ is a non-zero eigenvalue of } L_0$$
$$= C^i_{\alpha_\beta \beta_s} X_0^i \quad , \quad \text{if } \alpha+\beta = 0$$
$$= 0 \quad , \quad \text{otherwise} \quad , \qquad (11.11)$$

(β does not necessarily have the same range for $\alpha, \beta, \cdots$)

Note, in particular, that the α_i are <u>real</u>. This is because, as mentioned earlier, the X_0^i lie in the space of the real compact algebra. The corresponding L_0^i are therefor hermitian and since the α_i are their eigenvalues, these are real.

What remains to be shown is that the spaces X_α^β, X_β^β are all one-dimensional. (The space X_0^i is not in general 1-dimensional) This lecture will be concluded by deducing some properties of the matrix $g_{\mu\nu}$ from the results we have already obtained.

The main result we wish to obtain is that will the basis defined in this section $g_{\lambda\mu}$ will be of the form:

$$g_{\lambda\mu} = \left[\begin{array}{c|c|c|c} g_{ij} & 0 & 0 & \\ \hline 0 & \begin{matrix} 0 & g_{\beta_r \alpha_s} \\ g_{\alpha_r \alpha_s} & 0 \end{matrix} & 0 & \\ \hline 0 & 0 & \begin{matrix} 0 & g_{\beta_r - \beta_s} \\ g_{\beta_r - \beta_s} & 0 \end{matrix} & \\ \hline & & & \ddots \end{array}\right] \quad : \quad \left[\begin{array}{c} X_0^i \\ \hline X_\alpha^\beta \\ X_{-\alpha}^\beta \\ \hline X_\beta^\beta \\ X_{-\beta}^\beta \\ \hline \vdots \end{array}\right] \qquad (11.12)$$

where we have written the basis alongside for convenience.

Proof:

$$g_{\alpha_k \tau} = C^{\lambda}_{\mu\, \alpha_k}\, C^{\mu}_{\lambda \tau}$$

$$= C^{\lambda}_{i\, \alpha_k} C^{i}_{\lambda \tau} + C^{\lambda}_{\beta_s\, \alpha_k}\, C^{\beta_s}_{\lambda \tau}$$

$$= C^{\alpha_k}_{i\, \alpha_k} C^{i}_{\alpha_k \tau} + C^{(\beta+\alpha)_t}_{\beta_s\, \alpha_k}\, C^{\beta_s}_{(\beta+\alpha)_t\, \tau}, \qquad (11.13)$$

using (11.11) and the "step-up-step down" theorem. But, by the same theorem, $C^{i}_{\alpha_k \tau}$ and $C^{\beta_s}_{(\beta+\alpha)_t\, \tau}$ are zero unless $\tau = -\alpha_u$, some u. Hence the only non-zero elements in the α_k-row of $g_{\lambda\mu}$ are those which are in the $-\alpha$-"column" i.e. in some $-\alpha_s$-column. But this is just what is expressed by (11.12)

Note that we can also verify by this method our earlier result that if α is an eigenvalue of L_0, then so is $-\alpha$ since if $-\alpha$ were not an eigenvalue the g_{α_k} now would consist of zeros.

We can also obtain from (11.11) same information concerning the g_{ik}. We have, namely.

$$g_{ik} = C^{\lambda}_{\mu i}\, C^{\mu}_{\lambda k}$$

$$= C^{\lambda}_{\alpha_s\, i}\, C^{\alpha_s}_{\lambda k}$$

$$= C^{\alpha_s}_{\alpha_s\, i}\, C^{\alpha_s}_{\alpha_s\, k}$$

$$= \sum_{\alpha} m_{\alpha}^{2}\, \alpha_i \alpha_k \,, \qquad (11.14)$$

where m_α is the number of X_α^λ's in the X_α space $(\lambda = 1, \cdots m_\alpha)$

From this and the reality of the α_i it follows that g_{ik} is positive definite, in the sense that if x_i is any real vector in ℓ - space

$$g_{ik} x_i x_k = \sum_\alpha m_\alpha^2 (\alpha_i x_i)^2 \geq 0 \quad (11.15)$$

From this it follows that by a suitable transformation in X_0^i-space g_{ik} can be brought to the form

$$g_{ik} = \delta_{ik} \quad (11.16)$$

Q: If g_{ik} satisfies (11.13) it is certainly true that a transformation of the kind $A_{\lambda i} g_{ik} A^T_{ks} = \delta_{\lambda s}$ where $A^T = A$ - transpose. But how do you know that under a transformation $A_{\lambda i} X_0^i$ in X_0^i -space g_{ik} transforms in this way? Why not as $A g A^{-1}$, for example?

A: We have

$$[X_0^i, X_\alpha^\lambda] = \alpha_i X_\alpha^\lambda \quad (11.17)$$

Hence, if we let

$$X_0^{i'} = A_{ij} X_0^j \quad , \quad (11.18)$$

we obtain

$$[X_0^{i'}, X_\alpha^\lambda] = \alpha_i' X_\alpha^\lambda \quad , \quad (11.17)$$

where

$$\alpha_i' = A_{ij}\,\alpha_j \qquad (11.20)$$

In other words, the α_i transform in the same way as the X_o^i.

But from (11.14)

$$g_{ik} = \sum_\alpha m_\alpha^2\, \alpha_i' \alpha_k' = \sum_\alpha m_\alpha^2 A_{ij} A_{js}\, \alpha_r \alpha_s$$

$$= A_{ir} A_{js}\, g_{rs} = A_{ir}\, g_{rs} A^T_{sj}, \qquad (11.21)$$

as required,

Finally, we prove a result which will be of importance in the next section, namely,

$$C^{\cdot\cdot i}_{-\alpha_r\,\alpha_s} = -\,g_{-\alpha_r\,\alpha_s}\, g^{ik}\, \alpha_{k\cdot}. \qquad (11.22)$$

Proof:

$$-\,C^{\cdot\cdot i}_{-\alpha_r\,\alpha_s} = g^{i\lambda}\, C_{\lambda\,-\alpha_r\,\alpha_s}$$

$$= g^{ik}\, C_{k\,-\alpha_s\,\alpha_s}, \text{ from (11.12)}$$

$$= -g^{ik}\, C_{-\alpha_s\,k\,\alpha_s},$$

using the complete anti-symmetry of $C_{\lambda\mu\nu}$,

$$= -g^{ik}\, g_{-\alpha_s\,\lambda}\, C^{\cdot\cdot\lambda}_{k\,\alpha_s}$$

$$= -g^{ik}\, g_{-\alpha_r\,\alpha_s}\, C^{\cdot\cdot\alpha_s}_{k\,\alpha_s}, \text{ from (11.11)}$$

$$= -g^{ik}\, g_{\alpha_r\,\alpha_s\,\alpha_t} \qquad (11.23)$$

as required.

Corollary: $C^{i}_{-\alpha_r \alpha_s \alpha_i} \neq 0$ (11 24)

for at least one α_s

Proof: From (11 22)

$$C^{i}_{-\alpha_r \alpha_s \alpha_i} = -g_{-\alpha_r \alpha_s} g^{ik} \alpha_i \alpha_k$$

$$= -\underline{g}_{\alpha_r \alpha_s} \sum_{\alpha} m_\alpha^2 (\alpha_i^2)^2, \text{ from (11 14)}$$

The expression on the right hand side here cannot be zero unless $g_{-\alpha_r \alpha_s} = 0$, , and this cannot be true for every α_s, since $\det g_{\mu\lambda} \neq 0$

Lecture 12.

So for we have shown that for a semi-simple Lie algebra of order r

$$[X_\mu, X_\lambda] = c_{\mu\lambda}^{\sigma} X_\sigma \qquad \mu, \lambda, \sigma, = 1 \ldots r \tag{12.1}$$

$$|g_{\lambda\mu}| \neq 0 \tag{12.2}$$

we can set $X = X_\mu a_\mu$ in one to-one correspondence with a set of linear transformations $L = a_\mu L_\mu$, in the space of the X_α themselves and that using this correspondence we can divide the X_α into the sets $\{X_0^i\}, \{X_\alpha^r\}, \{X_\beta^s\}, \ldots$ where the "degeneracy" superscript i may not run over the same range of values for each set, such that with these as basis we have

$$X_0^r \longleftrightarrow L_0^r = \left(\begin{array}{c|c|c|c} 0 & & & \\ \hline & \alpha_r I & & \\ \hline & & \beta_r I & \\ \hline & & & \end{array}\right) \quad \vdots \quad \left(\begin{array}{c} X_0^i \\ \hline X_\alpha^r \\ \hline \vdots \end{array}\right) \tag{12.3}$$

and there is at least one particular $L_0 = \lambda_i L_0^i$ such that

$$X_0 \longleftrightarrow L_0 = \left(\begin{array}{c|c|c|c} 0 & & & \\ \hline & \alpha I & & \\ \hline & & \beta I & \\ \hline & & & \end{array}\right) \tag{12.4}$$

where $\alpha, \beta, \ldots$ are non-zero. As a result, we have

$$[X_0^i, X_0^j] = 0 \tag{12.5}$$

$$[X_0^i, X_\alpha^r] = \alpha_i X_\alpha^r \tag{12.6}$$

It is of the utmost importance for what follows, to note that in (12.6) the same X^{i}_{α} occurs on both sides of the equation. The subscript i in α_i belongs to X_0 not to X_α This is a consequence of the fact that the L^{0}_{ij} are not merely in block form in (12.3), but are multiples of the identity in each block. Hence not only do they leave the sets X^{i}_{α} invariant, but leave even the individual X^{i}_{α} all i unchanged.

We have also collected some results along the way:

(1) $\left[X^{r}_{\alpha}, X^{s}_{\beta}\right] = \sum_{k} C_{\alpha\beta}(rst) X^{t}_{\alpha+\beta}$ if $\alpha+\beta$ is an eigenvalue of L_0; $=0$ otherwise (12.6a)

(2) With the basis indicated $g_{\lambda\mu}$ has the form given in (11.15)

(3) If α is an eigenvalue of L_0 so is $-\alpha$

(4) $C^{i}_{-\alpha_k \alpha_j \alpha_i} \neq 0$, for all α_s. (12.7)

This lecture will be devoted to the task of proving that the spaces X^{i}_{α} are one-dimensional. This does not include the space X^{i}_{0}. To prove this, we focus our attention on the spectrum of L_0 and in particular to that part associated with the eigenvalue α, namely the eigenvalues $0, \alpha, 2\alpha, \ldots k_\alpha, -\alpha, -2\alpha, \ldots -k'_\alpha$ of which, of course, these are a finite number. These may be degenerate of course and we draw them schematically in Fig (12.1) allowing for this degeneracy. The $X^{i}_{0}, X^{k}_{\alpha}, X^{k}_{2\alpha}, \ldots$ are the eigenvectors corresponding to the eigenvalues. For example $X^{1}_{3\alpha}$ $X^{2}_{-\alpha}, X^{3}_{-\alpha}$ correspond to the points labelled 1,2,3 respectively.

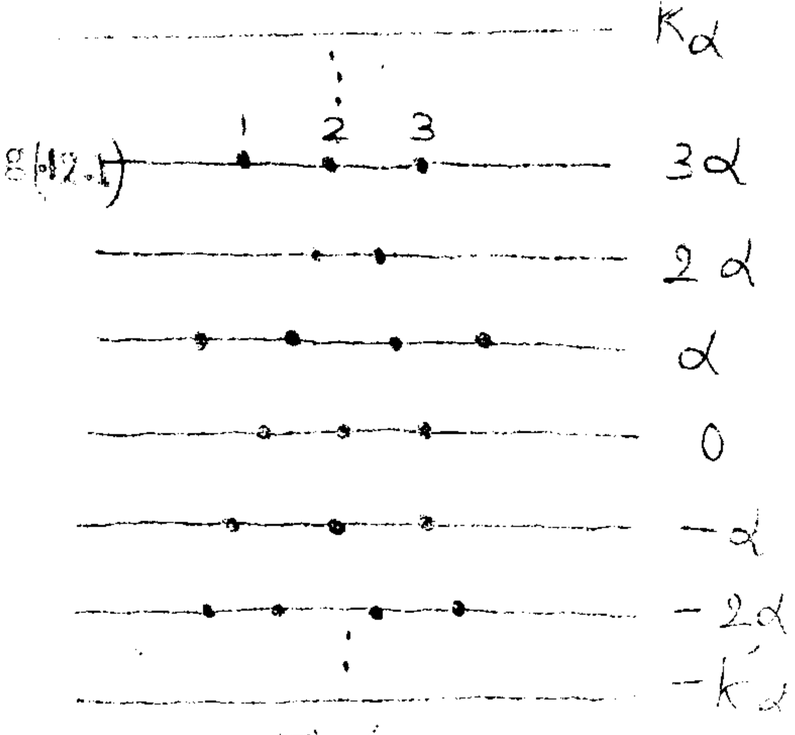


Fig (12.1)

We now consider out of this set of vectors $X^{i}_{0}, X^{\gamma_k}_{m\alpha}, m=\pm1,\pm2,\cdots$ the subset $X^{\gamma_1}_{-\alpha}, X^{i}_{0}, X^{\gamma_2}_{\alpha}, X^{\gamma_2}_{-2\alpha} \cdots$ i.e X^{i}_{0}, all the non-negative $X^{\gamma_2}_{\alpha}$, and the single negative $X^{\gamma_1}_{-\alpha}$ (which exists from (12·3) since it is a subspace of the total space X above). These vectors form a vector space $\overrightarrow{X}_{S}$ of all the X_α ;

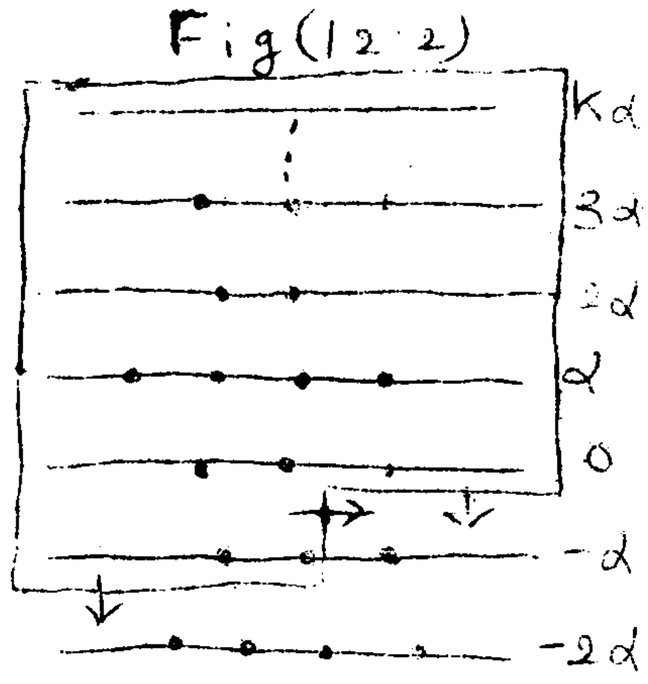
Now let us consider the corresponding L ; the space $\overrightarrow{L}_S = L^{\gamma_1}_{-\alpha}, L^{i}_{0}, L^{\gamma_2}_{\alpha}, \cdots$ We show that the space $\overrightarrow{X}_S$ is an invariant subspace for these operators

$$\text{ie} \qquad L_S \overrightarrow{X}_S = \overrightarrow{X}_S \qquad (12\cdot8)$$

Once we show that, the rest will be easy.

Proof: From Fig.(12·2) we see that $\overrightarrow{X}_S$ is the space enclosed in the box. Thus we have to show that operating with $\overrightarrow{L}_S$, on the vectors in the box, we do not go outside it.

Fig (12·2)



(a) we consider the $L^{i}_{\alpha}, L^{i}_{2\alpha}, \cdots L^{i}_{k\alpha}$. From the result (1) above we see that these are "step-up" operators i.e. they transform the vectors of the kind $X_{m\alpha}$ into vectors of the kind $X_{(m+1)\alpha}$, $X_{(m+2)\alpha}, \cdots$. Hence acting on the vectors in the box they lead only to vectors of the kind $X_{m\alpha}$, $m \geq 0$. But we have included all such vectors in the box. Thus the $L^{\gamma_1}_{\alpha}, L^{\gamma_2}_{-2\alpha}, \cdots$ do not lead out of the box.

(b) We consider the L^{i}_{0}. Acting on $X^{i}_{0}, X^{\gamma_2}_{m\alpha}$, $m>0$ they transform these into themselves, or to zero. Pictorally, we may say that the L^{i}_{0} "act horizontally".

The question is: what about $L_0^i X_{-\alpha}^{\lambda_1}$?

Is this a linear combination of $X_{-\alpha}^{\lambda_1}$ and $X_{-\alpha}^{\lambda_2}$, $\lambda_2 \neq \lambda_1$.

The answer is no, because from (12.6) and the remark following it,

$$L_0^i X_{-\alpha}^{\lambda_1} = \left[X_0^i, X_{-\alpha}^{\lambda_1}\right] = (-\alpha)_i X_{-\alpha}^{\lambda_1} \quad (12.9)$$

Thus the L_0^i do not lead out of the box.

(c) Finally we consider $L_{-\alpha}^{\lambda_1}$ itself. This is a step down operator (with α). Hence it leaves the $X_{m\alpha}^{\lambda}$ $m > 0$ in the box. But what about the X_0^i and $X_{-\alpha}^{\lambda_1}$?

As regards the X_0^i, we have, again using (12.6) and the remark following it,

$$L_{-\alpha}^{\lambda_1} X_0^i = \left[X_{-\alpha}^{\lambda_1}, X_0^i\right] = -(-\alpha)_i X_{-\alpha}^{\lambda_1} \quad (12.10)$$

Thus $L_{-\alpha}^{\lambda_1}$ transforms all of the X_0^i into $X_{-\alpha}^{\lambda_1}$ which is within the box.

Finally,

$$L_{-\alpha}^{\lambda_1} X_{-\alpha}^{\lambda_1} = \left[X_{-\alpha}^{\lambda_1}, X_{-\alpha}^{\lambda_1}\right] = 0 \quad (12.11)$$

Thus the $L_{-\alpha}^{\lambda_1}, L_0^i, L_\alpha^i, \ldots$ do not lead out of the box, although, a prior, one could imagine that they could lead out in the directions of the arrows shown in fig. (12.2) Thus

$$\overrightarrow{L_s}\,\overrightarrow{X_s} = \overrightarrow{X_s}, \quad (12.12)$$

as required. With this result in our hands, the rest is easy.

From the result (11) above, we have

$$\left[X^{\beta_1}_{-\alpha} , X^{i}_{\alpha} \right] = C^{i}_{-\alpha\beta_1,\alpha_t} X^{i}_{\sigma} \qquad (12.13)$$

from which

$$\left[L^{\beta_1}_{-\alpha} , L^{t}_{\alpha} \right] \overrightarrow{X_s} = C^{i}_{-\alpha\beta_1,\alpha_t} L^{i}_{o} \overrightarrow{X_s} \qquad (12.14)$$

Hence, taking the trace of both sides with respect to the $\overrightarrow{X_s}$ space, we have

$$\sum_{\mu} \left\langle X^{\mu}_{s} \middle| \left[L^{\beta_1}_{-\alpha} , L^{t}_{\alpha} \right] \middle| X^{\mu}_{s} \right\rangle = C^{i}_{-\alpha\beta_1,\alpha t} \sum_{\mu'} \left\langle X^{\mu}_{s} \middle| L^{i}_{o} \middle| X^{\mu'}_{s} \right\rangle$$

where $X^{1}_{s} = X^{1}_{-\alpha}$, $X^{2}_{s} = X^{1}_{o}$, $X^{3}_{s} = X^{2}_{\alpha}$, $\ldots$ $X^{n}_{s} = X^{\beta}_{\alpha}$ (12.15)

but

$$\underline{1} = \sum_{\sigma} X^{\sigma} \rangle\langle X^{\sigma} \qquad (12.16)$$

where the X^{σ} are a basis in the total space $\overrightarrow{X}$ not $\overrightarrow{X_s}$ (i.e. $\sigma = 1 \ldots \beta$). Inserting this identity into (12.15) we obtain

$$\sum_{\mu'} \sum_{\sigma} \left\langle X^{\mu}_{s} \middle| L^{\beta_1}_{-\alpha} \middle| X^{\sigma} \right\rangle \left\langle X^{\sigma} \middle| L^{t}_{\alpha} \middle| X^{\mu'}_{s} \right\rangle$$
$$= C^{i}_{-\beta_1,\alpha_t} \sum_{\mu'} \left\langle X^{\mu}_{s} \middle| L^{i}_{o} \middle| X^{\mu'}_{s} \right\rangle - \left\langle X^{\mu}_{s} \middle| L^{t}_{\alpha} \middle| X^{\sigma} \right\rangle \left\langle X^{\sigma} \middle| L^{\beta_1}_{-\alpha} \middle| X^{\mu}_{s} \right\rangle \qquad (12.17)$$

But since the $\overrightarrow{X_s}$ space is invariant for $L^{1}_{-\alpha}$ and L^{k}_{α} we have

$$\sum_{\mu} \sum_{\nu} \left\langle X^{\mu}_{s} \middle| L^{1}_{-\alpha} \middle| X^{\nu}_{s} \right\rangle \left\langle X^{\nu}_{s} \middle| L^{k}_{\alpha} \middle| X^{\mu}_{s} \right\rangle$$
$$= C^{i}_{-\alpha\beta_1,\alpha_t} \sum_{\mu'} \left\langle X^{\mu}_{s} \middle| L^{i}_{o} \middle| X^{\mu'}_{s} \right\rangle - \left\langle X^{\mu}_{s} \middle| L^{k}_{\alpha} \middle| X^{\nu}_{s} \right\rangle \left\langle X^{\nu}_{s} \middle| L^{i}_{-\alpha} \middle| X^{\mu}_{s} \right\rangle$$

In other words, although, a priori, one must use all the X^{σ} in (12.16), on account of the invariance of the subspace $\overrightarrow{X_s}$ with respect to $\overrightarrow{L_s}$, we need in fact use only the X^{ν}_{s} when (12.16) is inserted in (12.15). On exchanging the indices in the second in terms on the left hand side of (12.18) we see that it just cancels the first term on that side. Hence we have

$$0 = C^{i}_{-\alpha s_1 \alpha t} \sum_{\mu} \left\langle X^{\mu}_{s} \middle| L^{i}_{o} \middle| X^{\mu}_{s} \right\rangle$$

$$= C^{i}_{-\alpha s_1 \alpha t} \left\{ \left\langle X^{i}_{-\alpha} \middle| L^{i}_{o} \middle| X^{i}_{\alpha} \right\rangle + \sum_{j} \left\langle X^{j}_{o} \middle| L^{i}_{o} \middle| X^{j}_{o} \right\rangle + \sum_{\mu} \left\langle X^{\mu}_{2\alpha} \middle| L^{i}_{o} \middle| X^{\mu}_{\alpha} \right\rangle + \sum_{\mu} \left\langle X^{\nu}_{2\alpha} \middle| L^{i}_{o} \middle| X^{\nu}_{2\alpha} \right\rangle + \cdots \right\}$$

$$= C^{i}_{-\alpha s \alpha t} \left\{ -\alpha_i + 0 + m_1 \alpha_i + m_2 (2\alpha_i) + \cdots \right\}$$

where m_1 is the number of X^{i}_{α}, m_2 the number of $X^{i}_{2\alpha}$ etc.

$$= C^{i}_{-\alpha s_1 \alpha t} \alpha_i \left\{ -1 + m_1 + 2 m_2 + \cdots \right\} \qquad (12.19)$$

But from the result (4) listed above, the first term here is not zero for all α_t. Hence, since the m_s' are integers, we have

$$\begin{aligned} m_1 &= 1 \\ m_2 &= 0 \\ m_3 &= 0 \\ &\vdots \end{aligned} \qquad (12.20)$$

Thus the dimension of the space X_α^i is one, as required. In addition, we obtain the information that $2\alpha, 3\alpha, \cdots$ are not eigenvalues of L_0.

This completes the proof of Cartan's theorem. A summary of the results will be given in the next lecture.

Q: Could you repeat the proof that $\overrightarrow{L_s}\overrightarrow{X_s} = \overrightarrow{X_s}$?

A: Repitition of proof.

Lecture 13

In the last lecture we completed the proof of Cartans theorem by proving that the vector spaces $X_\alpha^{\mathcal{R}}$ belonging to the non-zero eigenvalues of L_0 were non-degenerate. This means that we can reduce the form of the Lie algebra given in (10.1) to the even simpler form

$$[H_i, H_j] = 0, \qquad i = 1, \cdots \ell,$$
$$[H_i, E_\alpha] = \alpha_i E_\alpha, \qquad \alpha = \ell+1, \cdots \mathcal{R}, \qquad (13.1)$$

$$[E_\alpha, E_\beta] = N_{\alpha\beta} E_{\alpha+\beta}, \quad \text{if } \alpha+\beta \text{ is a non-zero eigenvalue of } L_0,$$
$$= \mu^i_{(\alpha)} H_i, \quad \text{if } \alpha = \beta,$$
$$= 0, \quad \text{otherwise,}$$

where to follow conventional notation we have written H_i for X^i_0 and E_α for X_α. We can also sharpen some of the results obtained earlier, as follows:

First, with respect to $g_{\lambda\mu}$ we see that (11.12) now reduces to

$$g_{\lambda\mu} = \left(\begin{array}{c|cc|cc} g_{ik} & & & & \\ \hline & 0 & g_{\alpha-\alpha} & & \\ & g_{\alpha-\alpha} & 0 & & \\ \hline & & & 0 & g_{\beta-\beta} \\ & & & g_{\beta-\beta} & 0 \end{array}\right), \qquad (13.2)$$

while (11.14) reduces to

$$g_{ik} = \sum_{\alpha} \alpha_i \alpha_k \qquad (13.3)$$

note that $|g_{\lambda\mu}| \neq 0$ implies $|g_{ij}| \neq 0$, $g_{\alpha-\alpha} \neq 0$ (13.4)

It is implies in (13.2) of course, that if E_α exists so does $E_{-\alpha}$. From (13.3) we see that g_{ik} is positive definite. Hence by a suitable change of basis in H_i space, and by absorbing the non-zero $g_{\alpha-\alpha}$ into the E_α and $E_{-\alpha}$ we can produce a basis with respect to which

$$g_{\lambda\mu} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0\ 1 & & \\ & & 1\ 0 & & \\ & & & 0\ 1 & \\ & & & 1\ 0 & \\ & & & & \ddots \end{pmatrix} \qquad (13.5)$$

Secondly, with respect to the structure constants in (13.1) we can show that

(a) $(-\alpha)_i = -\alpha_i$ (13.6)

and

(b) $\mu_i = -g^{ik} g_{\alpha-\alpha} \alpha_k$, (13.7)

$= -\alpha_i$ with respect to the basis of (13.5) (13.8)

Proof: (b) follows directly from (11.23) for non-degenerate α-space. It then follows from (13.4) that $\mu_i \neq 0$ all i. But

$$0 = [H_i, \mu_j H_j] = [H_i, [E_\alpha, E_{-\alpha}]] = [[H_i, E_\alpha] E_{-\alpha}] + [E_\alpha [H_i, E_{-\alpha}]]$$
$$= (\alpha_i + (-\alpha)_i)[E_\alpha, E_{-\alpha}] = (\alpha_i + (-\alpha)_i)\mu_j H_j \qquad (13.9)$$

Hence (a) holds, as required.

We see, therefore, that the structure constants $\mu_{ij}(\alpha)$ are completely determined by the α_i . Later we shall show that the same is true of the $N_{\alpha\beta}$ (appendix B). Thus all the structure constants, and hence the Lie algebra, are determined (up to transformations of the basis X_μ) by the α_i . Since there are r^3 structure constants $C_{\mu\nu}^{\sigma}$ and only $l(r-l)$ α_i's , this represents an enormous reduction.

The study of the Lie algebras reduces therefore to a study of the α_i . To facilitate this study, it is convenient to regard the l α_i for each α as components of a vector in an l - dimensional space. The number l which is called the rank of the Lie algebra is of course equal to the number of H_i If we use g_{ik} as the metric tensor in this space, the space becomes Euclidean, since by (13·3) if $x = x_i$ is any real vector in the space ,

$$(x, x) = g_{ik} x_i x_k = \sum_{\alpha} \sum_{i} (\alpha_i x_i)^2 \geq 0 \quad (13\cdot 10)$$

with equality, if, and only if $x = 0$. If we choose the H_i so that g_{ik} is as in (13·5) then the basis in the l-space is orthonormal.

Let us now illustrate the idea of the l -space for $l = 2$ and $l = 3$. For $l = 2$

we have a 2-space as in fig.(13.1)
We suppose for definiteness that L_0 has six non-zero (as well as 2 zero) eigenvalues $\pm\alpha, \pm\beta, \pm\gamma$. The twelve basic structure constants $\pm\alpha_i, \pm\beta_i, \pm\gamma_i$ $i = 1, 2,$ can be read off directly from the diagram. Note that the relation (13.6) above is expressed diagrammatically by the fact that to each vector corresponds a vector of equal length in the opposite direction (which we have drawn dashed in in the figure

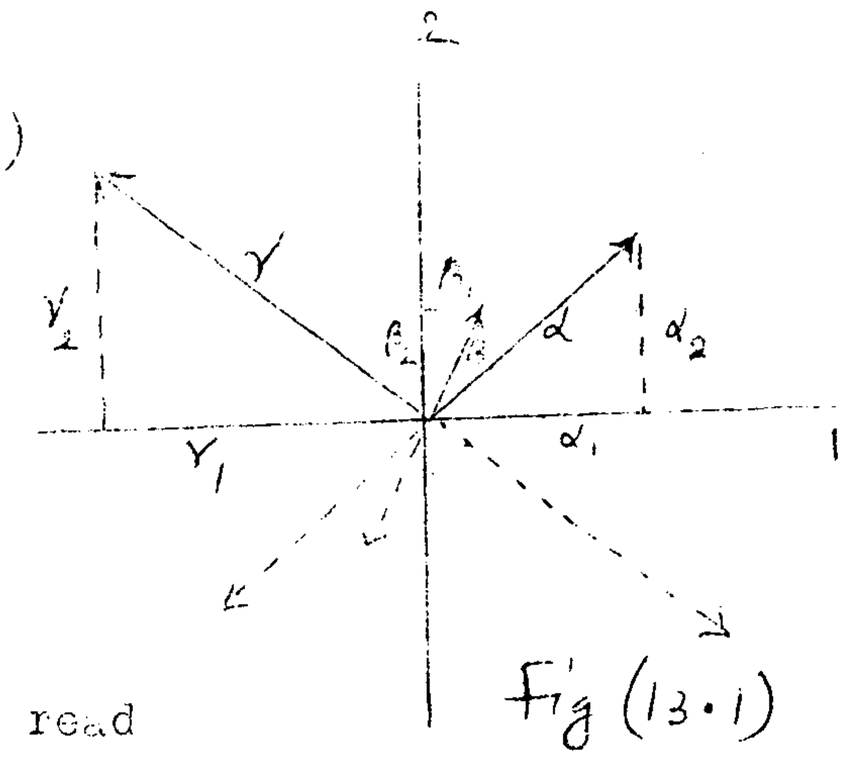


Fig (13.1)

For $l = 3$ (fig 13.2), we draw only one vector α and its opposite number $-\alpha$. In general, of course, there will be many more.

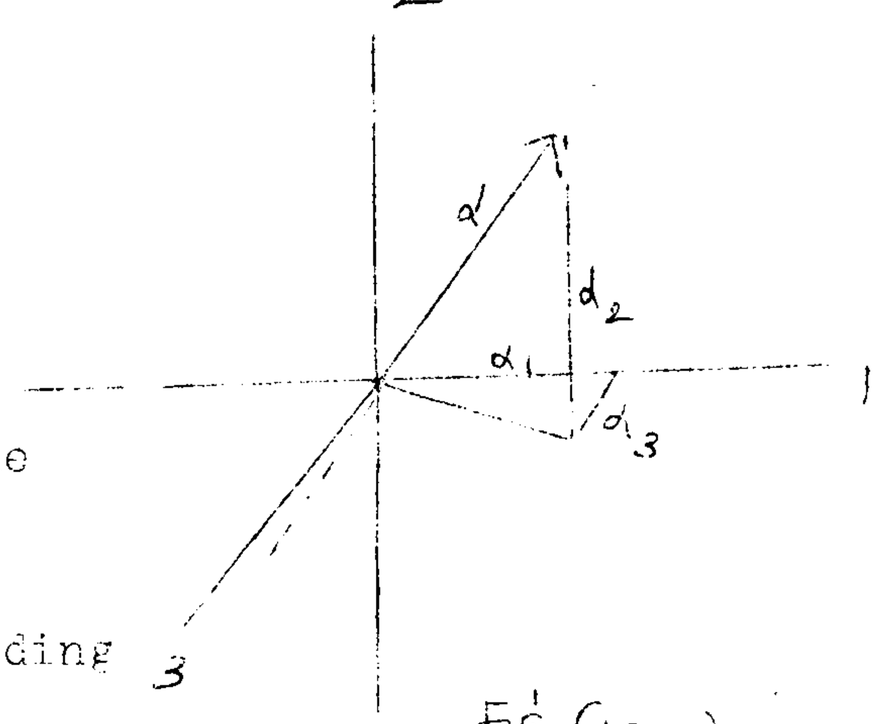


Fig (13.2)

It is usual to refer to the vectors α_i as <u>root vectors</u> or roots, and to call the corresponding diagrams the root diagrams.

The question now is: Do the vectors α, β, γ Lie around at random in the l-space or are only certain configurations allowed? The answer, as you probably know, is that the only certain discrete configurations are allowed. This result follows from the basic theorem on root diagrams which we shall now state. The proof of it will be given

in the next lecture.

Theorem: Let α and β be any two root vectors. Then $2\frac{(\alpha\cdot\beta)}{\alpha\cdot\alpha}$ is an integer, and $\beta - 2\frac{(\alpha\cdot\beta)}{(\alpha\cdot\alpha)}\alpha$ is a root vector.

Here by $(\alpha\cdot\beta)$ is meant $g^{ik}\alpha_i\beta_k$. One can see at once from this theorem that only special configurations of root vectors are allowed. Suppose, for example, that α is a fixed root vector, and β another root vector satisfying the condition

$$2\frac{(\alpha\cdot\beta)}{(\alpha\cdot\alpha)} = \text{integer} \qquad (13.10)$$

If we now vary β continuously from its position in ℓ space the vectors we obtain in the neighbourhood of β will certainly no longer satisfy (13.10). Thus no vector in the neighbourhood of a root vector can qualify as a root vector. The condition (13.10) implies actually much more than this, of course, but we shall be discussing this in more detail later.

Note finally that the vector $\beta - \frac{2(\alpha\cdot\beta)}{(\alpha\cdot\alpha)}\alpha$ has a very simple geometrical significance, namely, that it is the reflexion of β in the plane perpendicular to α. Thus we could phrase the root theorem as follows: If is a root, so is its reflexion in the plane perpendicular to any other root α.

Lecture 14

In the last lecture we saw that we could write our Lie algebra in the canonical form

$$\begin{aligned}
[H_i, H_j] &= 0 && i,j = 1 \dots \ell, \\
[H_i, E_\alpha] &= \alpha_i E_\alpha, && \alpha = \ell+1, \dots s \\
[E_\alpha, E_\beta] &= N_{\alpha\beta} E_{\alpha+\beta} && \text{if } \alpha+\beta \text{ is a non-zero root,} \\
&= \alpha^i H_i && \text{if } \alpha+\beta = 0, \\
&= 0 && \text{otherwise.}
\end{aligned} \tag{14.1}$$

As I mentioned these, the α_i determines the α^i and $N_{\alpha\beta}$ and so determine the Lie algebra, and we saw that we could represent the α_i diagramatically by considering them as vectors in a real ℓ-dimensional space. The configurations allowed for the α_i in this space are determined by the root theorem, namely.

Root Theorem: If α and β are roots, $\beta - \frac{2(\alpha\cdot\beta)}{(\alpha\cdot\alpha)}\alpha$ i.e. the reflexion of β in the plane perpendicular to α is a root, and $\frac{2(\alpha\cdot\beta)}{(\alpha\cdot\alpha)}$ is an integer. $\left[(\alpha\cdot\beta) = \alpha^i\beta_i = g^{ik}\alpha_i\alpha_k\right]$

In this lecture, we shall prove this theorem

Q: Could you repeat the proof that $\alpha^i = g^{ik}\alpha_k$?

A: $$\alpha^i = c^i{}_{\alpha,-\alpha} = g^{i\lambda} c_{\lambda,\alpha,-\alpha} = g^{ik} c_{k,\alpha,-\alpha} = g^{ik} c_{\alpha k,-\alpha}$$
$$= -g^{ik} g_{\alpha\lambda} c^\lambda{}_{k,-\alpha} = -g^{ik} g_{\alpha,-\alpha} c^{-\alpha}{}_{k,-\alpha} = g^{ik} g_{\alpha,-\alpha} g_k$$

and since we have normalized $g_{\alpha,-\alpha}$ to one, the result follows.

Proof of Theorem: The idea behind the proof is very simple that it involves some algebraic manipulations. The idea is as follows:

Let α and β be vectors as in fig.(14.1). Now consider the operation

$$[E_\alpha, E_\beta] \qquad (14.2)$$

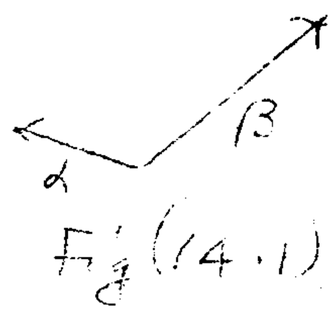


Fig (14.1)

From (14.1) this gives us either 0 or $E_{\alpha+\beta}$ i.e. the operator corresponding to the root $\beta_1 = \beta + \alpha$, if this is a root. Thus we may conceive of this operation as a "stepping up" operation from the root β to the root $\beta+\alpha$ (if this is a root). We draw this in Fig.(14.2). We now repeat the procedure with $\beta+\alpha$ instead of β to obtain $\beta+2\alpha$ if this is a root, Fig.(14.3)

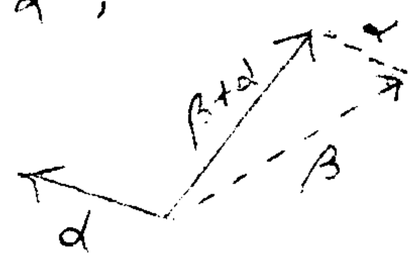


Fig (14.2)

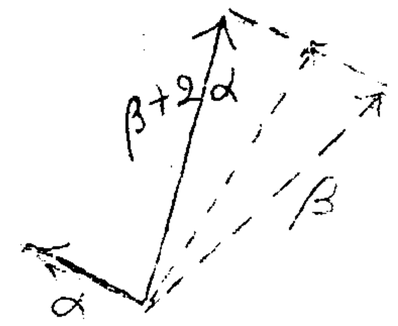


Fig (14.3)

Proceeding in this way, we eventually obtain zero, since otherwise we would obtain an infinite number of roots and this is impossible since each root corresponds to a base element of the Lie algebra Let $\beta_0 = \beta + m\alpha$ be root for which $[E_\alpha, E_{\beta_0}] = 0$ (14.3)

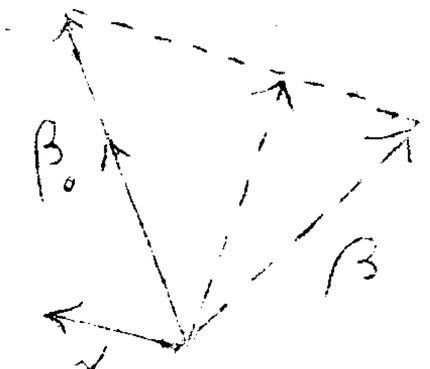


Fig (14.4)

Fig (14.4)

Let us now start with E_{β_0} and consider the operator

$$\left[E_{-\alpha}, E_{\beta_0}\right] \qquad (14.4)$$

This either zero, or yields a multiple of $E_{\beta_0 - \alpha}$ which we shall call $\hat{E}_{\beta_0 - \alpha}$. Thus

$$\left[E_{-\alpha}, E_{\beta_0}\right] = \hat{E}_{\beta_0 - \alpha} \qquad (14.5)$$

In other words $E_{-\alpha}$ is a "step-down" operator. Repeating the procedure for $\hat{E}_{\beta_0 - \alpha}$ we get either zero or a multiple of $E_{\beta_0 - 2\alpha}$ which we call $\hat{E}_{\beta_0 - 2\alpha}$,

$$\left[E_{-\alpha}, E_{\beta_0 - \alpha}\right] = \hat{E}_{\beta_0 - 2\alpha} \qquad (14.6)$$

Proceeding in this way we obtain a sequence of operators

$\hat{E}_{\beta_0}, \hat{E}_{\beta_0 - \alpha}, \hat{E}_{\beta_0 - 2\alpha}, \ldots$ according to the scheme

$$\left[E_{-\alpha}, \hat{E}_{\beta_0 - s\alpha}\right] = \hat{E}_{\beta_0 - s\alpha - \alpha} \qquad (14.7)$$

This sequence must terminate as explained above for the sequence $E_\beta, E_{\beta+\alpha}, \ldots E_{\beta_0}$. Suppose it terminates with $\hat{E}_{\beta_0 - n\alpha}$.

ie $\left[E_{-\alpha}, \hat{E}_{\beta_0 - n\alpha + \alpha}\right] = \hat{E}_{\beta_0 - n\alpha} \neq 0$ (14.8)

$$\left[E_{-\alpha}, E_{\beta_0 - n\alpha}\right] = 0 \qquad (14.9)$$

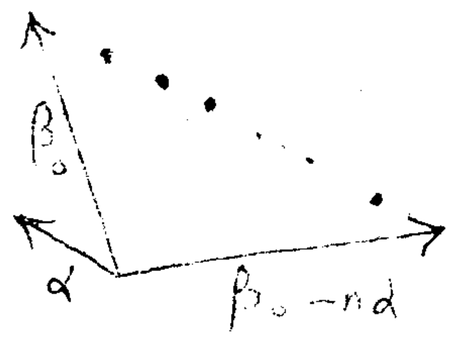


Fig (14.5)

Now we start with $\hat{E}_{\beta_0 - n\alpha}$ and use again the stepping up process, i.e. we consider the sequence

$$[E_\alpha, \hat{E}_{\beta_0 - n\alpha}] = a_{n-1} \hat{E}_{\beta_0 - n\alpha + \alpha}$$

$$[E_\alpha, E_{\beta_0 - n\alpha + \alpha}] = a_{n-2} \hat{E}_{\beta_0 - n\alpha + 2\alpha}$$

$$\vdots$$

$$[E_\alpha, \hat{E}_{\beta_0 - \alpha}] = a_0 \hat{E}_{\beta_0} \equiv a_0 E_{\beta_0}$$

$$[E_\alpha, \hat{E}_{\beta_0}] = 0 \qquad (14.10)$$

where, having defined the $\hat{E}$'s on the way down we are not free to to absorb the constants $a_0, \ldots a_{n-1}$ into them. Note also that we cannot assume, a priori, that the a's are not zero. It might conceivably happen that from a certain a_{n-p} down all the a's are zero. We shall show that this does not happen but, a priori, we cannot assume this.

Q: Does it not follow from (14.1) that the $N_{\alpha\beta}$ and hence the a's are not zero.

A: No all that we have proved is that if $\alpha + \beta$ is not a root $[E_\alpha, E_\beta] = 0$, and that if it is a root $[E_\alpha, E_\beta]$ belongs to the space $E_{\alpha+\beta}$. But it might be the zero vector in this space.

To return to our theorem, we shall now prove to our theorem, we shall now prove it as follows: Using the Jacobi identity, and (14.1) we shall obtain from the sequences (14.7) and (14.10)

(1) an expression for a_0 in terms of α and β

(2) an recurrence relation for the a_k in terms of α and β

(3) an expression for a_{n-1} in terms of α and β

From the recurrence relation and a_0 we shall obtain a second expression for a_{n-1} in terms of α and β. By equating the two expressions for a_{n-1} we shall obtain the result we require.

We obtain the required relations as follows:

(1) By definition (14.10)

$$a_0 \hat{E}_{\beta_0} = [E_\alpha, \hat{E}_{\beta_0 - \alpha}] \qquad (14.11)$$

But by the definition of $\hat{E}_{\beta_0 - \alpha}$, (14.7) we then have

$$a_0 \hat{E}_{\beta_0} = [E_\alpha, [E_{-\alpha}, \hat{E}_{\beta_0}]] \qquad (14.12)$$

Using the Jacobi identity, this becomes

$$a_0 \hat{E}_{\beta_0} = [[E_\alpha, E_{-\alpha}] \hat{E}_{\beta_0}] + [E_{-\alpha} [E_\alpha, \hat{E}_{\beta_0}]],$$

$$= [[E_\alpha, E_{-\alpha}] \hat{E}_{\beta_0}], \qquad (14.13)$$

by (14.3) which expresses the fact that β_0 is an "end-point"

Now using (14.1) for $[E_\alpha, E_{-\alpha}]$, we obtain

$$a_0 \hat{E}_{\beta_0} = [\alpha^i H_i, \hat{E}_{\beta_0}], \qquad (14.14)$$

and using (14.1) for $[H_i, \hat{E}_{\beta_0}]$, this becomes

$$a_0 \hat{E}_{\beta_0} = \alpha^i \beta_{0i} \hat{E}_{\beta_0}, \qquad (14.15)$$

from which we obtain the required expression for a_0 namely

$$a_0 = (\alpha \cdot \beta_0) \qquad (14.16)$$

(2) To obtain the recurrence relation for a_R we use practically the same trick: From the definition (14.10) we have,

$$a_R \hat{E}_{\beta_0 - R\alpha} = [E_\alpha, \hat{E}_{\beta_0 - R\alpha - \alpha}] \qquad (14.17)$$

and from (14.17) this becomes

$$= [E_\alpha, [E_{-\alpha}, \hat{E}_{\beta_0 - R\alpha}]]$$

$$= [[E_\alpha, E_{-\alpha}], \hat{E}_{\beta_0 - R\alpha}] + [E_{-\alpha}, [E_\alpha, \hat{E}_{\beta_0 - R\alpha}]] \qquad (14.18)$$

At this point there is a difference, because the second term in (14.18) is not zero. In fact, from (14.10), (14.8) becomes

$$= [[E_\alpha, E_{-\alpha}], \hat{E}_{\beta_0 - R\alpha}] + [E_{-\alpha}, a_{R-1} \hat{E}_{\beta_0 - R\alpha + \alpha}]$$

$$= [\alpha^i H_i, \hat{E}_{\beta_0 - R\alpha}] + a_{R-1} [E_{-\alpha}, \hat{E}_{\beta_0 - R\alpha + \alpha}]$$

$$= \alpha^i (\beta_0 - R\alpha)_i \hat{E}_{\beta_0 - R\alpha} + a_{R-1} \hat{E}_{\beta_0 - R\alpha} \qquad (14.19)$$

From this, we have the required recurrence relation

$$a_R = a_{R-1} + \alpha \cdot (\beta_0 - R\alpha) \qquad (14.20)$$

We can solve this relation immediately. In fact,

$$a_R = a_{R-2} + \alpha \cdot (\beta_0 - (R-1)\alpha) + \alpha \cdot (\beta_0 - R\alpha)$$

$$= a_{R-3} + \alpha \cdot (\beta_0 - (R-2)\alpha) + \alpha \cdot (\beta_0 - (R-1)\alpha) + \alpha \cdot (\beta_0 - R\alpha)$$

$$- - - - -$$

$$= a_0 + \alpha \cdot (\beta_0 - \alpha) + \alpha \cdot (\beta_0 - 2\alpha) + \cdots + \alpha \cdot (\beta_0 - R\alpha) \qquad (14.21)$$

But from (1) above $a_0 = \alpha \cdot \beta_0$ Hence

$$a_k = \alpha \cdot \beta_0 + \alpha \cdot (\beta_0 - \alpha) + \alpha \cdot (\beta_0 - 2\alpha) + \cdots + \alpha \cdots (\beta_0 - k\alpha),$$

$$= (k+1)\alpha \cdot \beta_0 - \frac{k(k+1)}{2}(\alpha \cdot \alpha) \qquad (14.22)$$

In particular

$$a_{n-1} = n(\alpha \cdot \beta_0) - \frac{n(n-1)}{2}(\alpha \cdot \alpha) \qquad (14.23)$$

To obtain the direct relation for a_{n-1} we have to change our tactics a little. Hitherto we have started with the "natural" combination $a_k \hat{E}_{\beta_0 - k\alpha}$ Now we start with

$a_{n-1} \hat{E}_{\beta_0 - n\alpha}$, not $a_{n-1} \hat{E}_{\beta_0 - n\alpha + \alpha}$ we have

$$a_{n-1} \hat{E}_{\beta_0 - n\alpha} = a_{n-1} \left[E_{-\alpha}, \hat{E}_{\beta_0 - n\alpha + \alpha} \right]$$

$$= \left[E_{-\alpha}, a_{n-1} \hat{E}_{\beta_0 - n\alpha + \alpha} \right]$$

$$= \left[E_{-\alpha}, \left[E_{\alpha}, \hat{E}_{\beta_0 - n\alpha} \right] \right]$$

$$= \left[\left[E_{-\alpha}, E_{\alpha} \right] \hat{E}_{\beta_0 - n\alpha} \right] + \left[E_{\alpha}, \left[E_{-\alpha}, \hat{E}_{\beta_0 - n\alpha} \right] \right]$$

But because $\beta_0 - n\alpha$ is an end-point, the second term vanishes, by (14.9)

Thus $a_{n-1} \hat{E}_{\beta_0 - n\alpha} = \left[\left[E_{-\alpha}, E_{\alpha} \right] \hat{E}_{\beta_0 - n\alpha} \right]$

$$= \left[-\alpha^i H_i, \hat{E}_{\beta_0 - n\alpha} \right]$$

$$= -\alpha \cdot (\beta_0 - n\alpha) \hat{E}_{\beta_0} \qquad (14.25)$$

from which $$a_{n-1} = -\alpha\cdot(\beta_0 - n\alpha) \qquad (14.26)$$

Note the minus sign which stems from the fact that we have $[E_{-\alpha}, E_\alpha]$ not $[E_\alpha, E_{-\alpha}]$ in (14.25).

If we now compare this result with our previous expression (14.23) for a_{n-1} we see that

$$a_{n-1} = -\alpha\cdot(\beta_0 - n\alpha) = n(\alpha\cdot\beta_0) - \frac{n(n-1)}{2}(\alpha\cdot\alpha) \qquad (14.27)$$

from which

$$\frac{n(n+1)}{2}(\alpha\cdot\alpha) = (n+1)(\alpha\cdot\beta_0)$$

or

$$2\frac{(\alpha\cdot\beta_0)}{(\alpha\cdot\alpha)} = n \qquad (14.28)$$

But

$$\beta_0 = \beta + m\alpha \qquad (14.29)$$

where m is an integer. Hence

$$2\left(\frac{\alpha\cdot\beta}{\alpha\cdot\alpha}\right) = n - 2m \qquad (14.30)$$

is an integer, as required. Further, since we then have

$$\beta - 2\frac{(\alpha\cdot\beta)}{(\alpha\cdot\alpha)}\alpha = \beta - (n-2m)\alpha = \beta_0 - (n-m)\alpha, \qquad (14.31)$$

this is a root, as is also required.

Thus each root β has a corresponding root which is simply its reflexion in the plane perpendicular to α

Further it is clear from the above that for any root β in the sequence $\beta_0, \beta_0-\alpha, \cdots, \beta_0-n\alpha$, the reflected root is also in the sequence. In particular $\beta_0 - n\alpha$ is the reflexion of β_0

Thus the sequence of vectors $\beta_0, \beta_0 - \alpha, \ldots \beta_0 - n\alpha$ are such that their end points lie at distances $|\alpha|$, along a line which is parallel to α, and is bisected by the plane perpendicular to α. Fig.(14.6)

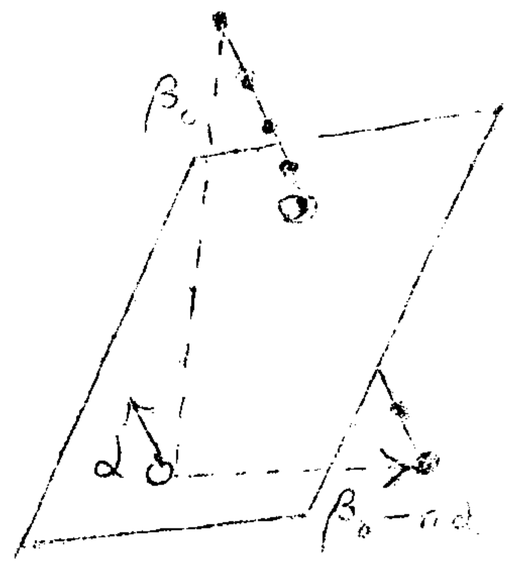


Fig (14.6)

If n is even, the root $\beta_0 - \frac{n}{2}\alpha$ will lie in the plane.] Such a line of root end-points is called a string.

Since the above considerations apply to any root β, we see that all the root vector-ends must lie on such strings. Thus for each α the end points of the roots will be as in Fig.(14.7) Note that this, in turn, is true for each root α. The conditions which this imposes on the possible root configurations will be discussed in the next lecture.

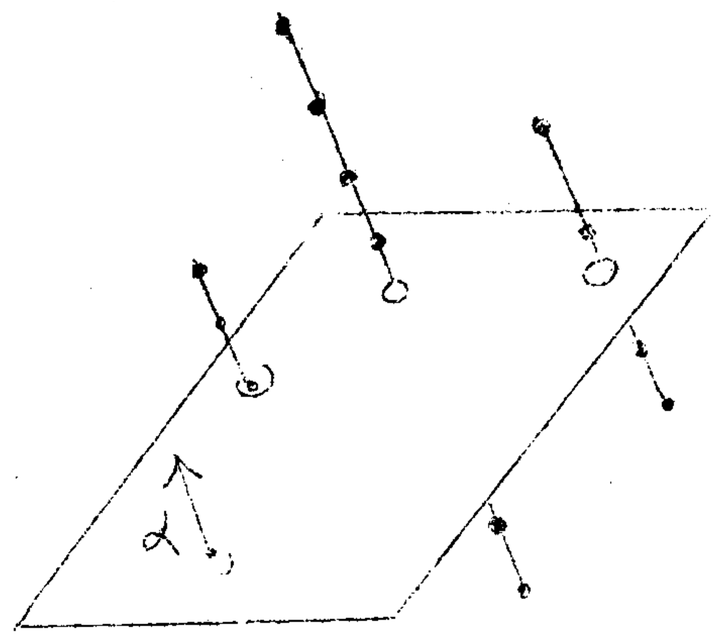


Fig (14.7)

Q: Can one of the strings parallel to α include α itself.

A: Yes the end points of $\alpha, 0, -\alpha$ lie on a string. However they are the only points lying on this string. To see this, let $\beta = k\alpha$ where k is any real number. From the root theorem $2k = \frac{2(\beta \cdot \alpha)}{\alpha \cdot \alpha} = n$ If n is even, then $\beta = \frac{n}{2}\alpha$ is an integral multiple of α and we know from previous results ($2\alpha, 3\alpha, \ldots$ are not roots), that this is possible only if $\beta = \alpha, 0, -\alpha$

If n is odd, then $\beta' = \frac{\alpha}{2}$ lies on the string and so $\alpha' = 2\beta'$ and α' would not be not a root. Thus any string through the origin has is of the form $(\alpha, 0, -\alpha)$

Q: To return again to the a_s can you now show that they are non-zero?

A: From (14.22)

$$a_s = (s+1)(\alpha\cdot\beta_0) - \frac{s(s+1)}{2}(\alpha\cdot\alpha)$$

$$= \frac{(s+1)(\alpha\cdot\alpha)}{2}\left[\frac{2(\alpha\cdot\beta_0)}{(\alpha\cdot\alpha)} - s\right]$$

$$= \frac{(s+1)(\alpha\cdot\alpha)}{2}[n-s] \qquad (14.32)$$

Since s has the range $0, 1, \ldots n-1$, we see that the a_s are not merely non-zero but are even all positive. Note by the way, that it is not necessary to establish that the $a's$ are non-zero, in order to establish that the sequence $\beta_0, \beta_0 - \alpha, \cdots, \beta_0 - n\alpha$ are roots. These are roots by the definition (14.7)

LECTURE 15.

In the last lecture we proved the root theorem which says that if α and β are roots, then

(1) $2\,\alpha\cdot\beta/\alpha\cdot\alpha$ is an integer

(2) the reflexion of β in the plane $\perp$ to α (i.e. $\beta - 2\frac{\alpha\cdot\beta}{\alpha\cdot\alpha}\alpha$ is a root.

We shall now make use of this result to clarify all the possible root diagrams. First, we have

$$2\,\frac{\alpha\cdot\beta}{\alpha\cdot\alpha} = m\,, \qquad 2\,\frac{\alpha\cdot\beta}{\beta\cdot\beta} = n\,, \tag{15.1}$$

which m and n are integers. (Note that they must have the same sign). Hence

$$4\cos^2\phi = 4\frac{(\alpha\cdot\beta)^2}{(\alpha\cdot\alpha)(\beta\cdot\beta)} = mn\,, \tag{15.2}$$

where ϕ is the "angle" between α and β. From this we see that the only possible values of mn are 0,1,2,3,4. But 4 can be discarded since if $mn = 4$, $\phi = 0$, $\beta = k\alpha$ hence from the result obtained in the last lecture, $\beta = \pm\alpha$. We then obtain the following table:

$\frac{\alpha\cdot\alpha}{\beta\cdot\beta} =$	mn	0	1	2	3
	$\cos^2\phi$	0	$\frac{1}{4}$	2/4	$\frac{3}{4}$
	$\cos\phi$	0	$+\frac{1}{2}$	$1/\sqrt{2}$	$\sqrt{3}/2$
	n, m	0,0	1,1	2,1	3,1
	$\frac{n}{m}$	$\frac{0}{0}$	1	2	3
	$\frac{\lvert\alpha\rvert}{\lvert\beta\rvert}$	$\frac{0}{0}$	1	$\sqrt{2}$	$\sqrt{3}$
	ϕ	90°	60°	45°	30°

Here, we have not considered the possibility given by $\cos\phi = -\frac{1}{2}, -\frac{1}{\sqrt{2}}, -\sqrt{3}/2$, since this amounts only to taking $-\beta$ instead of β. Also, since we shall be interested only in the ratio n/m, we do not need to include the case where both n and m are negative.

We are now in a position to start classifying the possible root diagrams. We classify them according to the increasing dimensions of ℓ, the dimension of the space.

$\ell = 1$.

If α is a root, any other root must be of the form $k\alpha$. But we have already seen that in that case $k = \pm 1$.

Hence the only diagram in a one dimensional space is in Fig. (15.1). The corresponding commutation relation are

$$\begin{aligned} [H, E_\alpha] &= \alpha E_\alpha , \\ [H, E_{-\alpha}] &= -\alpha E_\alpha , \\ [E_\alpha, E_{-\alpha}] &= \alpha H , \end{aligned} \qquad (15.3)$$

Fig. (15.1)

which, on writing $T_3 = \frac{1}{2} H$, $T_1 = \frac{1}{\alpha\sqrt{2}}(E_\alpha + E_{-\alpha})$, $T_2 = \frac{1}{i\alpha\sqrt{2}}(E_\alpha - E_{-\alpha})$, become the familar commutation relations for the rotationgroup in 3 dimensions. Let us call this group A_1. $\ell = 2$

First we consider the case of all $\phi = 90^\circ$.

The only possible vectors are then as in Fig. (15,2). The lengths of α and β are undetermined.

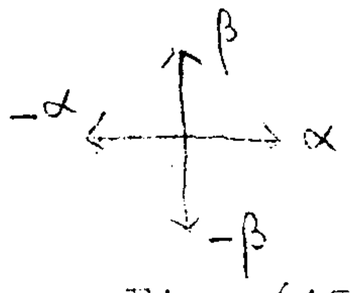


Fig. (15.2)

The corresponding commutation relations are clearly

$$\begin{aligned} [H_1, E_\alpha] &= \alpha_1 E_\alpha = |\alpha| E_\alpha , \\ [H_1, E_\beta] &= \beta_1 E_\beta = 0 \\ [H_2, E_\alpha] &= \alpha_2 E_\alpha = 0 \\ [H_2, E_\beta] &= \beta_2 E_\beta = |\beta| E_\beta , \end{aligned} \qquad (15.4)$$

and, finally, since $\pm\alpha \pm\beta$ are not roots,

$$\left[E_{\pm\alpha}, E_{\pm\beta}\right] = 0.$$

But this means that the set of six quantities $\mathcal{G} = \{E_{\pm\alpha}, E_{\pm\beta}, H, H\}$ can be divided into the two classes $\mathcal{G}_\alpha = \{E_{\pm\alpha}, H,\}$, $\mathcal{G}_\beta = \{E_{\pm\beta}, H.$ such that each member of $\mathcal{G}_\alpha$ commutes will each member of $\mathcal{G}_\beta$. In such a case $\mathcal{G}$ is said to be the direct sum of $\mathcal{G}_\alpha$ and $\mathcal{G}_\beta$. Any $\mathcal{G}$ which cannot be broken up into a direct sum of two other $\mathcal{G}$'s in this way is called a simple Lie algebra. Thus the above result may be written: $\mathcal{G}$ is not simple, but is the direct sum of two simple Lie algebras $\mathcal{G}_\alpha$ and $\mathcal{G}_\beta$. We may express this diagramatically as in Fig.(15.3)

The result $\mathcal{G} = \mathcal{G}_\alpha + \mathcal{G}_\beta$, is a special case of a general theorem which states: Any semi-smiple Lie algebra is either simple or is the direct sum of a number of simple groups.

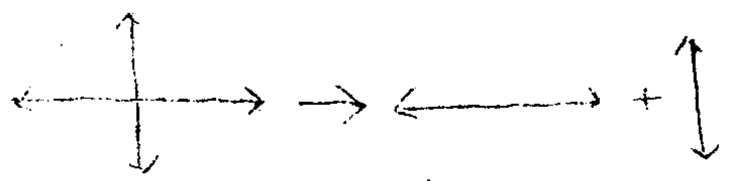


Fig.(15.3)

It is easy to verify that a simple Lie algebra as defined above is the Lie algebra of a simple Lie group, according to the usual definition of simplicity for groups (Lecture 1), and conversely, that a simple Lie group has a simple Lie algebra.

The next case we consider for $\lambda = 2$ is that of all $\phi = 60^\circ$. Let α be as in Fig.(15.4) and draw β as shown at 60° to α. From the table, we see that $|\alpha| = |\beta|$

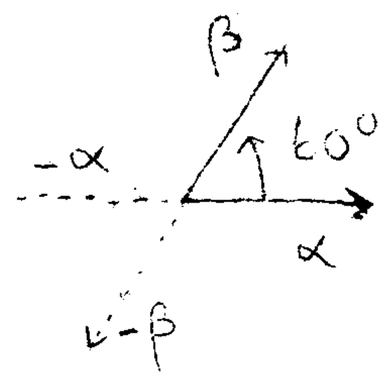


Fig. (15.4)

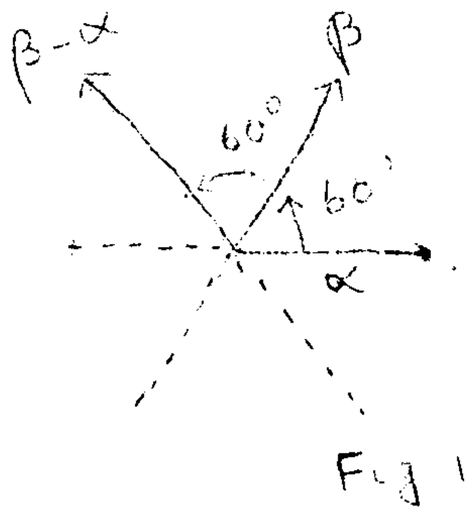


Fig 15.5

How we use the root reflexion theorem by reflecting in the plane (i.e. line) perpendicular to α. We get, as a result, the diagram of Fig.(15.5). Further, this diagram completes the scheme, for the diagram is invariant under all reflexions perpendicular to the roots and any other vector inserted in the diagram will make an angle of $\phi < 60^\circ$ with at least one of the six vectors obtain. Fig (15.5) is therefore the vector diagram for a simple Lie algebra. This Lie algebra is called SU_3. Its complex extension is called A_2, the subscript 2 referring to the value $\ell = 2$. The next case on the list for $\ell = 2$ is that for all $\phi = 45^\circ$. We let α be as shown in Fig.(15.6),

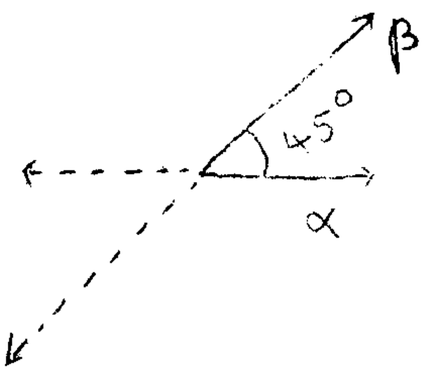
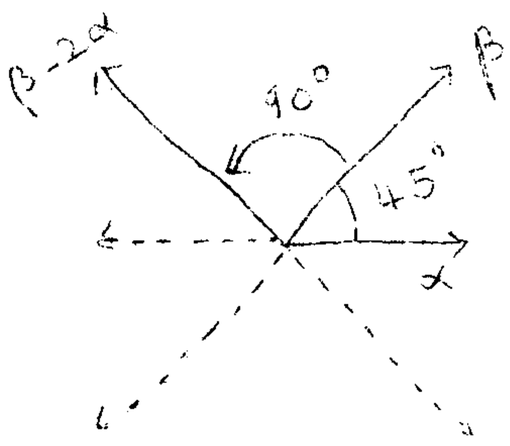
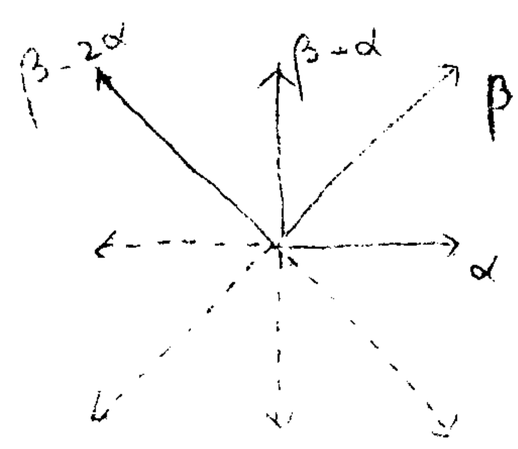


Fig 15.6

and let us draw β at 45° to this. From the table, we see that we must take $|\beta| = \sqrt{2}|\alpha|$ [or $|\alpha| = \sqrt{2}\beta$, but, this involves only a relabeling of α and β.] By reflecting β in the plane (i.e. line) perpendicular to α, we get fig.(15.7), and by reflecting α in the line perpendicular to $\beta - 2\alpha$ (i.e. β itself) we get the diagram of Fig. (15.8).



Fig(15.7)



Fig(15.8)

This diagram completes the scheme, for it remains invariant under all possible reflexions in the lines perpendicular to the roots, and we cannot insert any new vectors in the scheme, since any such vector would make an angle of less than 45° with at least one vector of Fig. (15.8). For reasons which will appear later, the compact real Lie group corresponding to Fig. (15.8) is called either B_4 or C_5, while the complex extension is called either B_2 or C_2.

We finally consider the case $\phi \geq 30^\circ$. We start as usual with α and β separated by the lest possible angle /Fig. (15.9)_/ From our table, we see that, in this case, $|\beta| = \sqrt{3}|\alpha|$ By a reflexion of β in the line perpendicular to α, we obtain the root $\beta - 3\alpha$ Fig.(15.10). By a reflexion of α in the line perpendicular to $(\beta - 3\alpha)$ we get $(\beta - 2\alpha)$, Fig. (15.11). By a reflexion of $(\beta - 2\alpha)$ in the line perpendicular to α we then obtain $(\beta - \alpha)$, Fig. (15.12), and, finally, by a reflexion of β in the line perpendicular $(\beta - 3\alpha)$ we obtain $(2\beta - 3\alpha)$ Fig. (15.13).

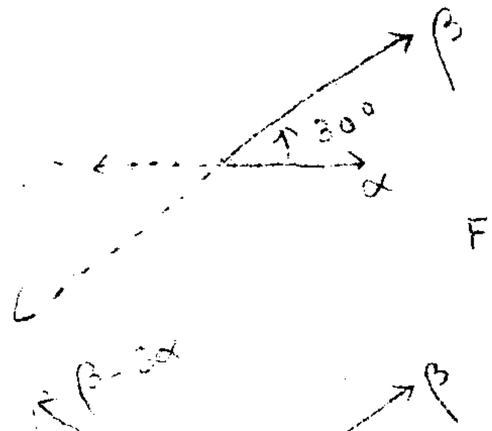


Fig 15.9

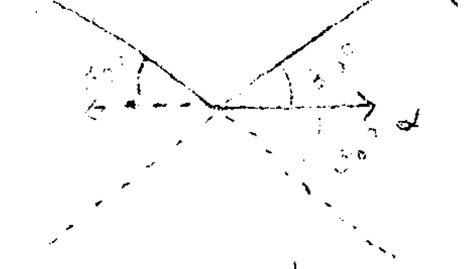


Fig (15.10)

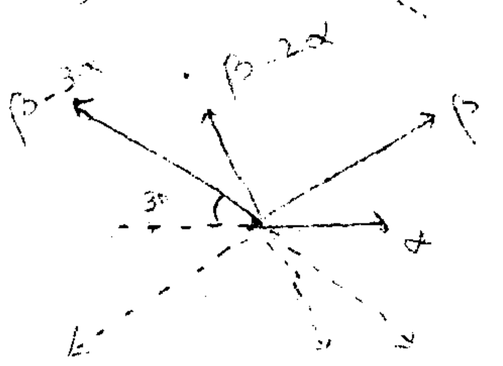


Fig (15.11)

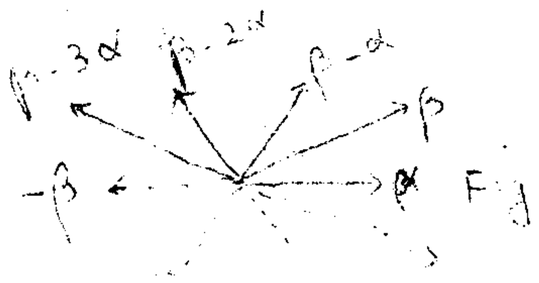


Fig (15.12)

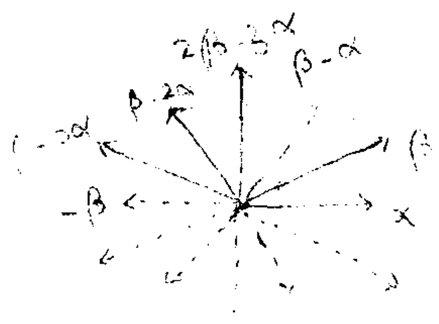


Fig (15.13)

Fig. (15.13) completes the scheme, since it is invariant under any reflexion perpendicular to the roots, and if any further vector is inserted in it, it will make an angle of less than 30° with at least one of roots already obtained.

The corresponding group is called G_2. We have now listed all the possible groups for $\ell = 2$. There is one non-simple group which is the direct sum of A_1 and A_1 and there are three simple groups A_2, B_2 and C_2 with diagrams as in Figs. (15,5), (15.8) and (15.13). (Note that for B_2 and G_2 the lines perpendicular to the roots coincide with other roots. Note also that the diagram for G_2 can be obtained by taking the diagram for A_2, and superimposing on it the same diagram related through an angle 30° and extended by a scale factor $\sqrt{3}$).

Q: Is G_2 the diagran which is sometimes called "the star of David" diagram?

A: Yes. If one joins the end-points of G_2, one obtains a Star shaped figure.

This completes the dicsaussion for $\ell = 1$ and $\ell = 2$. In the next lecture we shall discuss the cases $\ell = 3$ and $\ell > 3$ But before going on to those cases, we should like to complete the present lecture by introducing the concept of <u>primitive</u> roots:

Consider again the diagram (15.5) for A_2. If we call the vector which we have priviously called α a and the one which we have previously called $(\beta - \alpha)$ b, we see that the vectors of the diagram are a, b, a + b, and their negatives.

Consider now the diagram (15.8) for B_2. If we call the previous α, a and the previous $(\beta - 2\alpha)$. β , we see that the vectors of the diagram are a, b, b + a, b + 2a and their negatives.

Consider finally the diagram (15.13) for G_2. If we call the previous α a and the previous $\beta - 3\alpha$, β , we see that the vectors of the diagram are a, b, b + a, b + 2a, b +3a, 2b + 3a and their negatives.

Thus we see that for each of the diagrams for $\ell = 2$, we can choose two roots such that all the other roots are linear combinations of these two with coefficients which are, first of all, integers and secondly, either all positive or all negative.

Q: Is this not simply due to the fact that in an ℓ -dimensional space only ℓ vectors are linearly independent.

A: That the other roots can be expressed as linear combinations of only two is certainly due to this. But that the coeffi-cients are integers and either all positive or all negative is not. This result is due to the specific structure of the diagrams, which must satisfy the conditions of our table.

Now the point is that it can be shown that the above result holds for a diagram of any dimension ℓ, that is to say, for any root diagram of order ℓ there exist ℓ root vectors such that all the other root vectors can be expressed as linear combinations of these with coefficients which are integers and are either all positive or all negative. Such a set of ℓ root vectors are called <u>primitive roots.</u> The vectors which have positive coefficients with respect to these are called positive

LECTURE 16.

In this lecture we shall indicate how the root vector diagrams can be constructed for $\ell = 3$ and $\ell > 3$

Consider first the case when any two roots of the diagram, α and β say, Fig. (16.1) are at an angle 30° to each other. Let γ be any vector not in the same plane. If $|\gamma| < |\alpha|$ or $|\gamma| = \sqrt{2}|\alpha|$ then $|\gamma| < \frac{1}{\sqrt{3}}|\beta|$ and $|\gamma| = \sqrt{2/3}\,|\beta|$ respectively, and such ratios are not allowed by our table. Hence $|\gamma| = |\alpha| = \frac{1}{\sqrt{3}}|\beta|$ or $|\gamma| = \sqrt{3}|\alpha| = |\beta|$. In the first case, γ is at an angle $\pm 60^\circ$, $\pm 120^\circ$ to α , and $\pm 30^\circ$, $\pm 150^\circ$ to β and in the second case, it is the other way arrived. By constructing the 30°, 60° and 150° cones with 0 as origin and α and β as axes, it is easy to see that this is impossible unless γ lines in the $\alpha\beta$ -plane, contrery to hyplothesis. Thus we have the result: there are no 3-dimension diagrams in which some roots lie at an angle 30° to each other. Of course, we have left out the possibility here that there is just one γ is at 90° to α and β . But in this case, it is easy to see that the corresponding Lie algebra is just the sum $G_2 + A_1$.

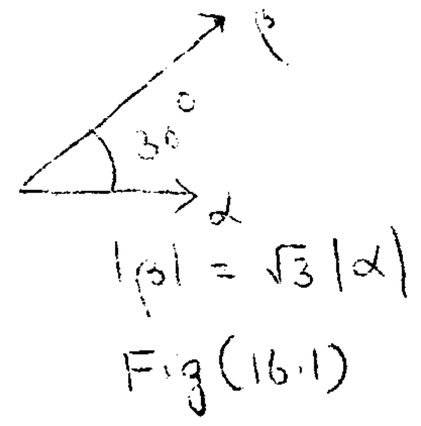


Fig (16.1)

Next we consider the case where two roots α, β are at an angle 60° Fig. (16.2). We have already seen that a third vector γ not in the same plane cannot be at an angle 30° to α or β. That leaves as possibilities $\alpha\gamma = 45^\circ$ or 60° , with

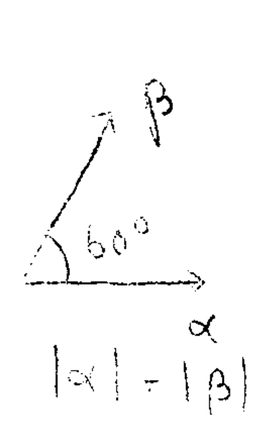
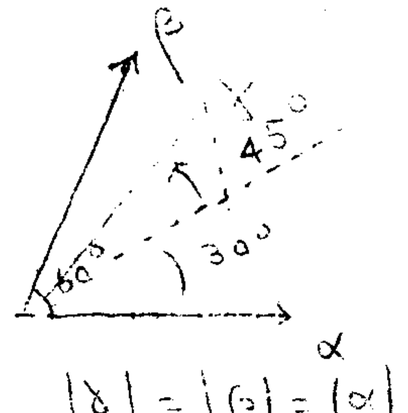


Fig (16.2)

$\hat{\alpha\beta} = 45^\circ$ or 60° respectively (since $|\alpha| = |\beta|$). We again neglect the case of a single γ perpundicular to α and β, which would give only $A_1 + A_2$.

If $\hat{\alpha\gamma} = \hat{\gamma\beta} = 60^\circ$, $|\gamma| = |\beta| = |\alpha|$ and we obtain a diagram as in Fig.(16.3) The projection of γ in the $\alpha\beta$-plane (with which γ makes an angle of 45°) bisects $\hat{\alpha\beta}$. By repeated reflexions in the planes perpendicular to the roots, as described earlier for $\ell = 2$, we obtain from this diagram of Fig. (16.4), which is closed under such relfexions and admits no further roots. The way in which this diagram is to be imagined is as follows: Six of the roots (α, β, δ and their negatives) line in a plane, forming the diagram for A_2 in that plane, The circle represents a cone whose axis is perpendicular to the plane and whose generators are at 45° to the plane. If we regard its axis as the "time" axis it is the "null-cone" for the 2-dimensional space of the paper. On this cone at a distance $= |\alpha|$ out along the generators or $\frac{1}{\sqrt{2}}|\alpha|$ above the α, β, δ plane, one draws a circle in the cone parallel to the plane. On this circle one marks the three points shown. These are the end-points of three roots. There are three corresponding roots below the plane. Thus the diagram contains 12 roots altogether. The group of which thus is the diagram is called A_3.



$|\gamma| = |\beta| = |\alpha|$
Fig (16.3)

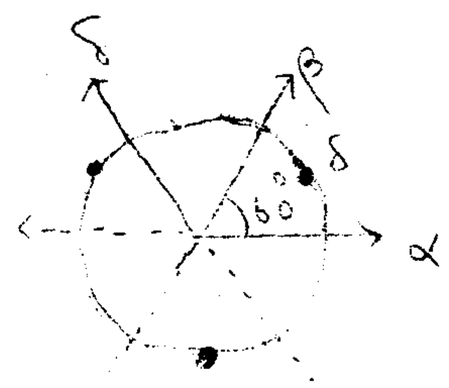


Fig (16.4)

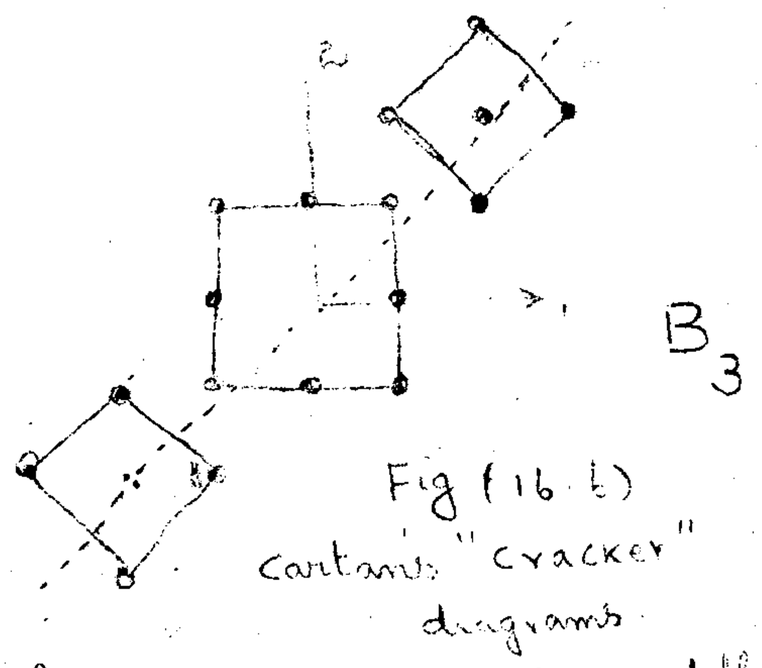
If on the other hand, $\alpha\beta = 60^\circ$, but $\gamma\alpha = \gamma\beta = 45^\circ$, there are two possibilities i.e. $|\gamma| = \sqrt{2}\,|\alpha|$ and $|\gamma| = \frac{1}{\sqrt{2}}|\alpha|$. Using each of these γ's one can in exactly the same way (i.e. by successive reflexi ns) generate two further diagrams. However, fo. reasons which will become clear later, we disregard these for the moment, and pass on to consider the cases when the original α, β are at 45° and 90°.

For $\alpha\beta$ at 45° [Fig. (16.5)] we construct a γ which does not lie in the same plane, and we find that there are two possible independent γ's from which we can generate diagrams. i.e.

$(\gamma\cdot\alpha) = 45^\circ \qquad \gamma\beta = 60^\circ$

$(\gamma\cdot\alpha) = 60^\circ \qquad \alpha\beta = 45^\circ$

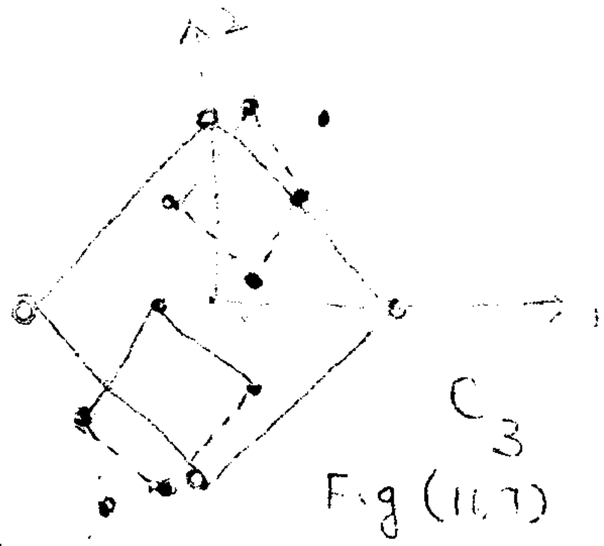
with these we can generate the diagrams shown in Fig. (16.6) and Fig. (16.7). Although each of these diagrams contains 18 root vectors whose ends are marked with a point, they are independent for no rotation of one can bring it into coincidence with the other. (This will be seen more easily from the analytic expressions for the roots which we shall give below). The groups corresponding to these diagrams are called B_3 and C_3 respectively.



B_3

Fig (16.6)
Cartan's "cracker" diagrams

Total 18 vectors → 8 for middle square + 5 each for the other two squares.



C_3

Fig (16.7)
"Toast Rack" diagram

18 vectors → 8 for the middle square + 4 each for other two squares + 1 each away from these two.

The reason for our earlier disregard of the two diagrams which can be obtained from $\alpha\beta = 60^\circ$ with $\gamma = 45^\circ$ may be mentioned here. It is that those two diagrams may be brought into coincidence with B_3 and C_3, and so yield no new group. It is convenient, however, to draw B_3 and C_3 as in Figures. (16.6) and (16.7) and not as they would be if generated by $\hat{\alpha\beta} = 60^\circ$, $\gamma = 45^\circ$. Note that whereas we can generate A_3, B_3 and C_3 from $\hat{\alpha\beta} = 60^\circ$, we can generate only B_3 and C_3 from $\hat{\alpha\beta} = 45^\circ$. This is because there is no plane in the A_3 diagram containing the $B_2 = C_2$ diagram.

We do not have to consider the case $\hat{\alpha\beta} = 90^\circ$, since if $\hat{\alpha\beta} = 90^\circ$ we can consider $\alpha\gamma$ or $\beta\gamma$ where γ is not coplanor with α and β. If $\hat{\alpha\gamma} \neq 90^\circ$ or $\hat{\gamma\beta} \neq 90^\circ$ then this is one of the cases we have already considered and if $\hat{\beta\gamma} = \hat{\alpha\gamma} = 90^\circ$, then we have $A_1 + A_1 + A_1$. However, for reasons which will be clear later it is convenient to rotate A_3 to the position shown in Fig. (16.8) in which position it it is called D_3, although, of course, $D_3 = A_3$. [It looks as if it were generated from the case $\hat{\alpha\beta} = 90^\circ$, $|\alpha| = |\beta|$]

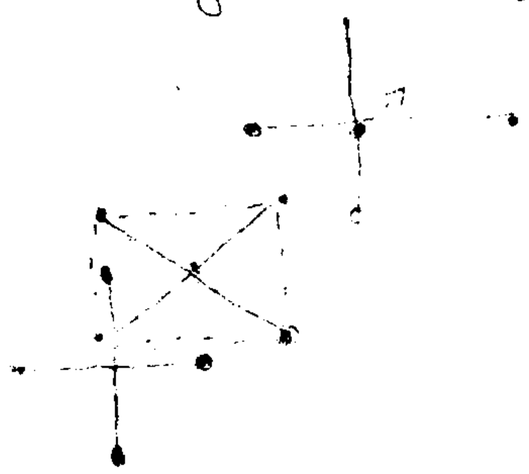


Fig (16.8)
"St. Brigids Cross" diagram

This completes our discussion of the case $\ell = 3$.

General ℓ ;

The possible diagrams for general ℓ are obtained by methods similar to these just outlined. One might expect that the number of possible diagrams increases with ℓ, but in fact, this is not the case. In fact, apart from five expectional diagrams G_2, F_4, E_6, E_7, E_8, for $\ell = 2, 4, 6, 7, 8$ respectively (note G_2 of the last lecture is one) there exist for each ℓ only 4 independent diagrams. These are called A_ℓ, B_ℓ, C_ℓ and D_ℓ, which explains the nomenclature used above for $\ell = 2$ and $\ell = 3$. Thus for each ℓ we get for Lie algebras, except for $\ell = 2, 4, 6, 7, 8$, in which cases we get five.

The diagrams for A_ℓ, B_ℓ, C_ℓ and D_ℓ, $\ell > 2$ may be regarded as generalizations of the diagrams for $\ell = 2$ and $\ell = 3$ and are classified as follows:

D_ℓ : Consider the diagram D_3, Fig. (15.8). The roots in this diagram are clearly the vectors $\pm e_i \pm e_k$, $i \neq k$ where the e_i are the base vectors as shown in the diagram $i,k = 1, 2, 3$. Similiarily, D_ℓ consists of all the vectors $\pm e_i \pm e_k$, $i \neq k$ where the e_i are base vectors in the ℓ-dimensional space. There are obviously $2\ell(\ell-1)$ such vectors.

C_ℓ : We see C_3 (Fig. (15.7)) consists of the vectors $\pm 2e_i$, $\pm e_i \pm e_k$, $i,k = 1,2,3$. For general ℓ, we have the same result with $i,k = 1,2,\ldots\ldots \ell$. There are $2\ell + 2\ell(\ell-1) = 2\ell^2$ such vectors.

B_ℓ : For B_3 (Fig. (15.6)) we have the vectors $\pm e_i$, $\pm e_i \pm e_k$, $i,k = 1,2,3$. For B_ℓ, the same holds with $i.k. = 1\ldots\ldots \ell$. There are again $2\ell^2$ vectors. B_ℓ is

distinct from C_ℓ except for $\ell = 2$ in which case the diagrams differ only by a rotation through 45° and a multiplication by a scale factor $\sqrt{2}$.

A_ℓ : To describe the generalization to A_ℓ , it is more convenient to return to A_2 rather than A_3. One then notes that the roots of A_2 can be obtained as follows:

Let e_i , $i = 1, 2, 3$ be a set of orthonormal vectors in a 3 -space and draw the tetrahedon $(e_i - e_j)$ Fig (16.9). These vectors (and their negatives) lie in the 2-space formed by the face of the tetrahedon i.e. in a plane with direction cosines $\frac{1}{\sqrt{3}}(1,1,$ But these vectors (transferred to the origin) are just the root vectors of A_2 taken in the same plane. In the same way the root vectors for A_ℓ are just the vectors $(e_i - e_j)$ of an orthonormal set e_i in $(\ell+1)$ space $i, j = 1 \ldots (\ell+1)$. They lie in the ℓ -space which has direction cosines $\frac{1}{\sqrt{\ell+1}}(1, 1 \ldots 1)$ with respect to the $(\ell+1)$ space.

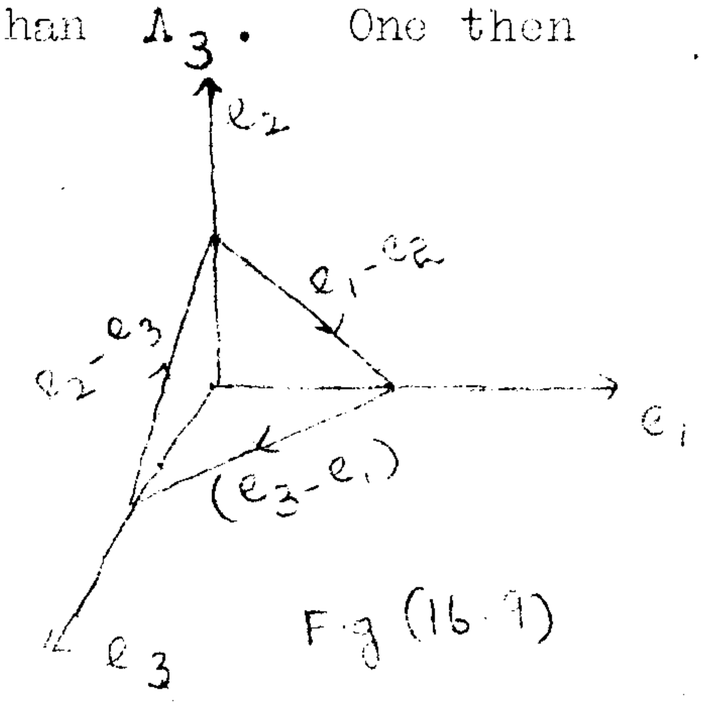


Fig (16.9)

We have noted earlier that $B_2 = C_2$ and $A_3 = D_3$. In a similar way are easily seen from the above results $A_1 = B_1 = C_1$, which D_1 does not exist. The question now arises as to whether there are any similar relations (isomorphisms) for $\ell > 3$. The answer is that there are not. Thus $A_1 = B_1 = C_1$, $B_2 = C_2$, $A_3 = D_3$ is the complete list of isomorphisms.

We shall conclude this lecture by mentioning the <u>discrete Weyl</u> group connected with any Lie algebra. From the root theorem we know that any root diagram will be invariant under the reflections in the planes perpendicular to the roots. Let α denote the roots and S_α the corresponding reflexions. Then the root diagrams will be invariant under the set of operations $S_W = \{S_\alpha, S_\alpha S_\beta, S_\alpha S_\beta S_\gamma, \cdots\}$. It is clear that this set of operations is finite $W = 1, \cdots N$, and that they form a group under the operation of applying them successively $S_{W''} = S_{W'} S_W$. The group S_W is called the discrete Weyl group.

It should be emphasized that the S_W are not all reflexions. Only those elements which are of the form

S_α, $S_\beta S_\alpha S_\beta^{-1}$, $S_\gamma S_\beta S_\alpha S_\beta^{-1} S_\gamma^{-1}$, ... are reflexions.

It can also be shown that a reflexion in any root can be generated by a product of the type S_{α_k}, $S_{\beta_k} S_{\alpha_k} S_{\beta_k}^{-1}$, $(S_{\gamma_k} S_{\beta_k} S_{\alpha_k} \cdot S_{\beta_k}^{-1} S_{\gamma_k}^{-1})$ where the $\alpha_k, \beta_k, \ldots$ are <u>primitive</u> roots. Thus the Weyl group is generated not merely by the set of reflexions perpendicular to all the roots, but even by the subset of this consisting of all reflexions perpendicular to the primitive roots.

This completes our discussion of the classification of the local Lie groups. In the next lectures we shall be considering the question of their representations.

LECTURE 17

In this, the "c" series of lectures, I should like to discuss the question of the representations of the various Lie groups which we have classified in the "b"-lectures.

For the question of representations there is a sharp distinction between the cases of compact and non-compact groups. For compact groups the representation theory is much simpler. The simplicity is closely connected with the fact that for the compact groups one can prove the following results:

(1) Any representation of a compact Lie group is equivalent to a unitary representation.

(2) Any reducible representation of a compact Lie group is fully reducible.

(3) Any irreducible representation of a compact Lie group is finite dimensional.

We shall not prove these results, but we shall see below how they are used in our investigation. The difference between a representation which is reducible and a representation which is fully reducible, is that the former need only leave a subspace R, of the representation space invariant, while the latter must leave R, and its complement $(R - R_1)$ invariant.

If we have a representation of a compact Lie group, then, as we saw in the "b" lectures, the infinitesimal generators I_λ of the representation will satisfy the commutation relations

$$\left[I_\lambda, I_\mu\right] = C_{\lambda\mu}^{\sigma} I_\sigma , \qquad (17.1)$$

of the compact Lie group. Further from the results (1),(2) (3) above, the I_λ will be finite-dimensional, either irreducible or fully reducible, and equivalent to a hermitian set I_λ^H.

Conversely, if we have such a set of I_λ satisfying (17.1) they will generate a representation of the local compact group. Thus the problem of finding the representations of the group reduces to the problem of finding all the irreducible, finite dimensional, hermitian matrices I_λ which satisfy (17.1). The advantage of attacking the problem of finding the representations of the Lie group, by changing it into the problem of finding the "representations" of the Lie algebra, is that for the algebra the matrices I_λ are simply numerical matrices, whereas, for the group, the representation matrices $D(a) = e^{i a_\rho I_\rho}$ are functions of the parameters a_ρ .

To find the representations of the Lie algebra, it is convenient to write the operator algebra, not as in (17.1), but in the Cartan canonical form derived earlier i.e.

$$
\begin{aligned}
&[H_i , H_j] = 0 , \qquad i, j = 1, \ldots, \ell \\
&[H_i , E_\alpha] = \alpha_i E_\alpha \\
&[E_\alpha , E_\beta] = N_{\alpha\beta} E_{\alpha+\beta} , \quad \text{for } (\alpha+\beta) \text{ a non-zero root} , \\
&\qquad = -\alpha \cdot H , \quad \text{for } \beta = -\alpha , \\
&\qquad = 0 \qquad \text{otherwise}
\end{aligned}
\tag{17.2}
$$

Since these are the equations which must be satisfied by the infinitesimal generators of any representation, we shall from now on take the H_i and E_α in (17.2) to be such infinitesimal generators (i.e. N x N matrices, where N is the dimension of the representation). We must only note that whereas in the compact form (17.1) the I_λ would all be hermitian, here the quantities

H_i, $i(E_\alpha + E_{-\alpha})$, $i(E_\alpha - E_{-\alpha})$ should be hermitian, since in our earlier lectures we saw that it is the abstract elements corresponding to these quantities which are elements of the compact Lie algebra.

Our problem, therefore, is to find all irreducible, finite dimensional matrices satisfying (17.2) and subject to the condition that H_i, $i(E_\alpha + E_{-\alpha})$, $i(E_\alpha - E_{-\alpha})$ be hermitian. This problem will be the subject of this series of lectures.

We start by noting that, because the H_i of any given $N \times N$ representation are hermitian, they can be diagonalized sumultaneously. Let V be a simultaneous eigenvector, i.e.

$$H_i V = m_i V \quad , \quad i = 1, \ldots \ell \qquad (17.3)$$

where m_i are the simultaneous (real) eigenvalues. The set $(m_1, m_2, \ldots m_\ell)$ of eigenvalues is called the <u>weight</u> of V. Note that the weight consists of the set m_i each m_i is not a weight. Since there are just ℓ m_i, it turns out to be very convenient to consider the weight as a vector in the ℓ-dimensional space introduced earlier for the roots. (Fig. (17.1)) Thus for every vector V which is a simultaneous eigenvector of the H_i we obtain a <u>weight</u> $\vec{m}$ in ℓ-space.

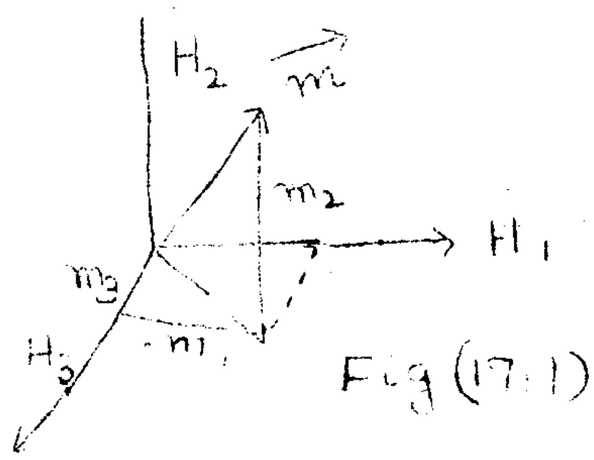


Fig (17.1)

Q: Will the simultaneous eigenvalues m_i be degenerate?

A: Yes, in general they will. In other words, if for each distinct set of m_i we take one eigenvector V, the V so obtained will not form a complete basis for the N-dimensional

representation space. Another way of saying this is to say that, in general, the H_i do not form a maximal abelian set for the algebra of the H_i, E_α and all polynomials formed out of them. There may, and, in general, there will, exist some polynamials in the H_i and E_α which commute with all the H_i but cannot be expressed as functions of the H_i alone.

Note that the V's corresponding to two distinct weights $\vec{m}$ and $\vec{m}'$ will be orthogonal. Thus there are at least as many base vectors in the ℓ-space as there are distinct weights. Alternatively we may say, that for a given finite N-dimensional representation the number of distinct weights is $\leq N$. In general, the inequality will hold. Of couse, this does not mean that the weight vector $\vec{m}$ will be orthoganal to each other. (In fact, as we shall see later, there will always be more than ℓ of them, so they could not be all orthogonal.) One must take care to differentiate between the two vector spaces with which we are now dealing namely, the ℓ-space for the vectors $\vec{m}$, and the N-space for the vectors V.

Having introduced the weights in the above way, we should now like to prove two theorems for them which are the analogues of two theorems which we proved earlier for the roots. They are:

(1) "Step-up step-down theorem" for weights, which states:
If V is an eigenvector belonging to the weight $\vec{m}$, $E_\alpha V$ (which, of course, is again a vector in N-space) is an eigenvector belonging to the weight $\vec{m} + \vec{\alpha}$.

Proof:

$$H_i(E_\alpha V) = [H_i, E_\alpha]V + E_\alpha H_i V$$

$$= \alpha_i E_\alpha V + E_\alpha m_i V$$

$$= (\alpha_i + m_i) E_\alpha V \cdot \quad Q\ E\ D \tag{17.4}$$

(2) "Reflection Theorem" for weights, which states:

If m is a weight and α a root, $\frac{2(m\cdot\alpha)}{(\alpha\cdot\alpha)}$ is an integer, and the reflexion, $m - 2\frac{(m\alpha)}{(\alpha\alpha)}\alpha$, of m in the plane perpendicular to α is a root.

Proof: The proof goes in exactly the same way as in the root case, except for one part which I shall now describe. One starts with the two vectors $\vec{m}$ and $\vec{\alpha}$ in ℓ-space, and by the preceding theorem "steps up" $\vec{m}$ to $\vec{m}+\vec{\alpha}$, $\vec{m}+2\vec{\alpha}$, ...etc until eventually we get an $\vec{m}^o = \vec{m} + k\vec{\alpha}$ such that

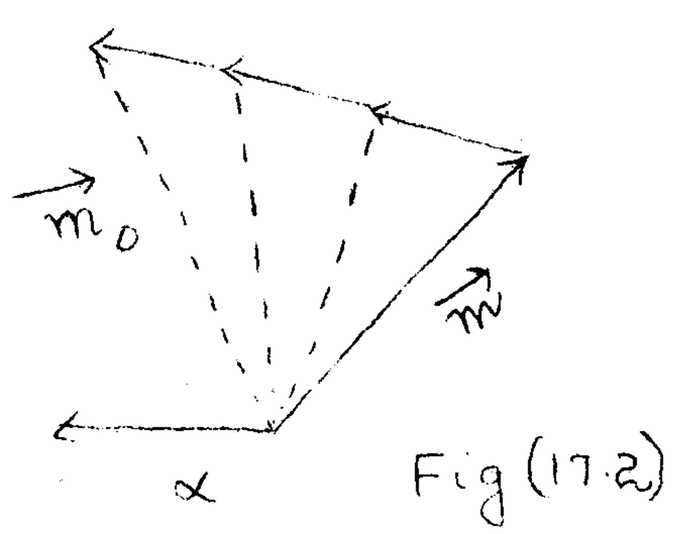


Fig (17.2)

$$E_\alpha u_{m^o} = 0 \tag{17.5}$$

(Fig. (17.2)). This must happen after a finite number of steps since, as we saw above, the number of distinct weights is bounded by N. (Note how we are using here the result that the irreducible representations of a compact group are finite dimensional). Now starting with $\vec{m}_o$, one steps it down with $E_{-\alpha}$ in exactly the same way as in the case of the roots (Fig. (14.5)), obtaining

a series of vectors $\hat{u}_{m^0-r\alpha}$,

defined by

$$E_{-\alpha} u_{m^0} = \hat{u}_{m^0-\alpha},$$

$$E_{-\alpha} \hat{u}_{m^0-\alpha} = \hat{u}_{m^0-2\alpha},$$

$$\cdots\cdots$$

$$E_{-\alpha} \hat{u}_{m^0-n\alpha+\alpha} = \hat{u}_{(m^0-n\alpha)}. \qquad (17.6)$$

$$E_{-\alpha} \hat{u}_{m^0-n\alpha} = 0$$

One now starts to step up there vectors with E_α as in the root case, but it is here that the difference comes in. In the root case, because the eigenspace of the H_i were non-degenerate, we could write

$$E_\alpha \hat{u}_{m^0-n\alpha} = a_{n-1} \hat{u}_{m^0-n\alpha+\alpha}, \qquad (17.7)$$

where a_{n-1} was a numerical coefficient. But here, on account of the degeneracy, we know only that

$$E_\alpha \hat{u}_{m^0-n\alpha} = \hat{\hat{u}}_{m^0-n\alpha+\alpha}, \qquad (17.8)$$

where $\hat{\hat{u}}_{m^0-n\alpha+\alpha}$ is some vector in the eigenspace corresponding to the weight $(m^0-n\alpha+\alpha)$. We do not know in advance that

$$\hat{\hat{u}}_{(m^0-n\alpha+\alpha)} = a'_{(n-1)} \hat{u}_{(m^0-n\alpha+\alpha)} \qquad (17.9)$$

Similiarily, we can write only

$$E_\alpha \hat{u}_{m^0-n\alpha+\alpha} = \hat{\hat{u}}_{m^0-n\alpha+2\alpha},$$

$$\cdots\cdots$$

$$E_\alpha \hat{u}_{m^0-\alpha} = \hat{u}_{m^0} \qquad (17.10)$$

where each $(m^0-r\alpha)$ is known in advance to be some vector in the $(m^0-r\alpha)$- eigenspace, but is not known to be a multiple

of $\hat{u}_{m^\circ - r\alpha}$. In fact, we do not even know in advance whether the step up process stops at $\hat{\hat{u}}_{m^\circ}$. It might stop either before or after it.

The point, however, is that one can show that in fact, the $\hat{\hat{u}}_{m^\circ}$ are (multiples) of the $\hat{u}_{m^\circ}$ (and that the process does stop at u_{m°). The proof of this is by induction: In the first place,

$$\begin{aligned} E_\alpha \hat{u}_{m^\circ - \alpha} &= E_\alpha E_{-\alpha} u_{m^\circ} \\ &= \left[E_\alpha, E_{-\alpha}\right] u_{m^\circ} + E_{-\alpha} E_\alpha u_{m^\circ} \qquad (17.11) \\ &= \alpha^i H_i u_{m^\circ} + 0 \\ &= (\alpha \cdot m_\circ) u_{m^\circ} \end{aligned}$$

so that $\hat{\hat{u}}_{m^\circ - \alpha}$ is just $(\alpha \cdot m^\circ)$ times u_{m°. Secondly, suppose that for any r

$$E_\alpha \hat{u}_{m^\circ - r\alpha} = a'_{r-1} \hat{u}_{m^\circ - r\alpha + \alpha}, \qquad (17.12)$$

Then

$$\begin{aligned} E_\alpha \hat{u}_{m^\circ - r\alpha - \alpha} &= E_\alpha E_{-\alpha} \hat{u}_{m^\circ - r\alpha} \\ &= \left[E_\alpha, E_{-\alpha}\right] \hat{u}_{m^\circ - r\alpha} + E_{-\alpha} E_\alpha \hat{u}_{m^\circ - r\alpha} \\ &= r \cdot (m^\circ - r\alpha) \hat{u}_{(m^\circ - r\alpha)} + E_{-r\alpha} a'_{r-1} \hat{u}_{(m^\circ - r\alpha + \alpha)} \\ &= \left[r \cdot (m^\circ - r\alpha) + a'_{r-1}\right] \hat{u}_{(m^\circ - r\alpha)}, \end{aligned} \qquad (17.13)$$

and so (17.12) holds also for $(r+1)$, with $a'_r = r.(m^0 - r\alpha) + a_{r-1}$ which gives a recurrence relation for the a'_r similar to that obtained earlier for the a_r in the case of the root vectors.

Thus in spite of the degeneriey of the eigenspaces corresponding to a given weight, the step up step down process does not carry us out of the space of the $\hat{\mu}(m^0 - r\alpha)$. Once this is established, it is easily seen that the proof goes through as in the case of the roots, and so I shall not bother to give the details.

Q: Is there also a result that if m and m' are weights, the reflexion of m in the plane perpendicular to m' is a weight and $2(m.m')/(m'.m')$ is an integer.

A: No. In fact it does not even follow that if m is a weight, $-m$ (the reflexion in the plane perpendicular to m itself) is a weight. An example is the "self" -representation of A_2, whose weight diagram is as in Fig. (17.4).

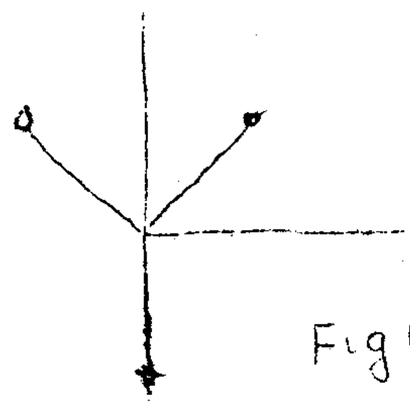


Fig (17.4)

LECTURE 18

In the last lecture the idea of weight vectors, was introduced and we proved two general results: (a) that if u is a vector in the representation space belonging to the weight m, then $(m+\alpha)$ is a root or $E_\alpha u = 0$; and (b) that if m is a weight, so is its reflexion in the plane perpendicular to any root α, and $\frac{2(m\cdot\alpha)}{(\alpha\cdot\alpha)}$ is an integer.

To proceed further, we have to introduce a new idea, namely, that of positive weights. Let $(m_1, m_2, m_3, \ldots\ldots m_\ell)$ be the components of any weight m. m is said to be positive if

$$\begin{aligned} &m_1 > 0 \\ &m_1 = 0\,,\ m_2 > 0\,, \\ &m_1 = 0 = m_2\,,\ m_3 > 0\,, \\ &\cdots\cdots \\ &m_1 = 0\,,\ m_2 = 0\,,\ \ldots\ m_{(\ell-1)} = 0\,,\ m_\ell > 0 \end{aligned} \qquad (18.1)$$

This means, of course, that positiveness is defined only relative to the choice of basis H_i. Suppose now that we have two weights m and m'. m is said to be more positive or higher than m' if the difference

$$m - m' = \{(m_1 - m_1'), (m_2 - m_2'), \ldots\ (m_\ell - m_\ell')\}$$

is positive. In this way, the idea of positiveness allows us to order the weights (according to height). The idea of one weight being higher than another becomes much more clear when we look at a diagram Fig. (18.1) Of the three weights shown a is the highest, simply because it lies farthest to the right (H_1-direction)

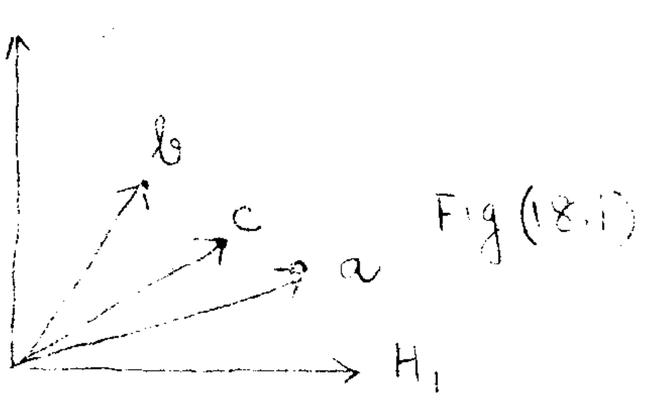


Fig (18.1)

The second highest is β , because although it lies equally far to the right as C , it lies further up (i.e in H_2 - direction) than C . As an example in 3-dimensions consider the root-diagram of D_3 Fig. (18.2). (Note that the roots are weights also (for the adjoint representation)). In this diagram there are four weights equally for out in the H_1 -direction. But one of these, a , lies further out in the H_2 -direction than the others. Hence α is the highest weight. With respect to H_1 and H_2 there are two contendes for the position of second highest weight, namely C and β , and we must go to H_3 to decide between them. Clearly, with the direction of H_3 shown, β is the one which qualifies. C is the third highest weight, d the fourth, e the fifth, and so on. Similiarly, for any number of dimensions.

Suppose, now that we have any representation of a group and that m_i^0 is the highest weight for this representation. Let u_0 be the corresponding vector in the N-dimensional representation space i.e.

$$H_i u_0 = m_i^0 u_0 . \qquad (18.2)$$

Then

$$E_\alpha u_0 = 0 . \qquad (18.3)$$

where α is any positive root. (Note that since roots are a special case of weights positivity is defined for the roots too.)

Proof: If (18.3) does not hold then $m^0_i + \alpha_i$ is a root belonging to the same representation and since $(m^0_i + \alpha_i) - m^0_i = \alpha_i$ is positive m^0_i is not the highest root contrary to hypothesis.

The importance of the result ( ) is that, using it, we can establish that there is a one-to-one correspondence between the irreducible representations of the Lie group and their highest weights. Thus the problem of finding the irreducible representations will reduce to the problem of finding the possible highest weights for each group. The one-to-one connexion between the irreducible representationss and the highest weights is obtained from the following two theorems:

Theorem 1: The vector u_0 corresponding to the highest weight of an irreducible representation is non-degenerate.

Proof: Let u'_0 and u_0 be two vectors belonging to m^0_i.

$$H_i u_0 = m^0_i u_0 \,,\; E_\alpha u_0 = 0 \,;$$
$$H_i u'_0 = m^0_i u'_0 \,,\; E_\alpha u'_0 = 0. \qquad (18.3')$$

Let

$$u'_0 = M u_0 \qquad (18.4)$$

where M is some matrix operating on the representation space. But by Burnsides' theorem since the $H_i\ E_\alpha$ are irreducible, for every matrix in the space there exists a polynomial $f(H_i, E_\alpha)$ which is equal to it. Let $f_M(H_i, E_\alpha)$ be this polynomial for M. Then

$$u'_0 = f_M(E_\alpha, H_i)\, u_0 \qquad (18.5)$$

What we wish to show is that

$$u'_0 = \lambda u_0 \qquad (18.6)$$

where λ is some constant, Clearly to show this it will be sufficient to show that for any product such as $E_\alpha E_\beta H_i E_\delta H_j E_\delta$ etc., occurring in F_M

$$E_\alpha E_\beta H_i E_\gamma H_j E_\delta u_0 = \tilde{\lambda} u_0 \qquad (18.7)$$

where $\tilde{\lambda}$ is some constant. To prove (18.7), we note first of all that the H_i H_j drop out since they are equivalent to multiplicative constants $(m_j^0 + \delta_j)$ and $(m_i^0 + \gamma_i + \delta_i)$ respectively. Next we note that since u_0' belongs to m_i^0 we must have

$$H_i E_\alpha E_\beta E_\gamma E_\delta u_0 = (\alpha_i + \beta_i + \gamma_i + \delta_i + m_i^0) \qquad (18.8)$$

from which $\times E_\alpha E_\beta E_\gamma E_\delta u_0 = m_i^0 E_\alpha E_\beta E_\gamma E_\delta u_0$

$$(\alpha_i + \beta_i + \gamma_i + \delta_i) = 0 \qquad (18.9)$$

Thus at least one of the $\alpha, \beta, \gamma, \delta$ must be positive. Suppose, for the sake of argument, that γ and δ are not positive but β is. Then

$$E_\alpha E_\beta E_\gamma E_\delta u_0 = E_\alpha \left[E_\beta, E_\gamma E_\delta \right] u_0$$

$$= N_{\beta\gamma} E_\alpha E_{\beta+\gamma} E_\delta u_0 + N_{\alpha\delta} E_\alpha E_\gamma E_{\beta+\delta} u_0 \qquad (18.10)$$

Thus we have reduced the expression $E_\alpha E_\beta E_\gamma E_\delta u_0$ with four E's to a sum of terms with only three E's each. On repeating the argument for the terms with three E's, and roots $(\alpha + (\beta + \gamma) + \delta) = 0$ etc. then repeating it for the derived terms with two E's, and so on we eventually arrive at (18.7). It is clear that similar considerations will apply no matter which roots of $\alpha, \beta, \gamma, \delta$ are positive, and that a similar argument can be used for any term in F_m. Thus the theorem is proved.

Note that the theorem holds only for an irreducible representation. Only for such a representation can we use Burnside's theorem. Note also that we have proved a little more than is

stated, because, as we noted above, the definition of highest weight is abitrary to the extent that we are always free to change our basis in $\mathcal{L}$ -space. Thus what we have proved is that all the weights which can qualify as highest weights by a suitable choice of basis are non-degenerate. In particular, the Weyl relfexion of a highest weight is non-degenerate, since on making the same Weyl reflexion (which is an orthogonal transformation) of the basis, the reflected vector becomes the highest weight with respect to the new basis.

The second theorem we have to prove is the following:

Theorem II : If two irreducible representations have the same highest weight they are equivalent.

Proof: Let $X_\lambda = H_i, E_\alpha$ be the infinitesimal generators of the first representation, and $X'_\lambda = H'_i, E'_\alpha$ these of the second, and let

$$H_i u_0 = m_i^0 u_0 ,$$

$$H'_i u'_0 = m_i^0 u'_0 . \qquad (18.11)$$

m_i^0 being the same in each equation. As we saw above from Bursides theorem, all vectors in the irreducible vector space can be expressed as linear combinations of vectors of the form

$$u = X_\alpha X_\beta \cdots X_\gamma u_0 \qquad (18.12)$$

To prove that the two representation are equivalent, what we have to show is that

$$(u, X_\alpha v) = (u', X'_\alpha v') . \qquad (\quad)$$

where u and v are arbitrary vectors in the representation space. But using (18.12), this means that what we have to prove is that

$$\left(\left[X_\beta X_\gamma \cdots X_\delta\right]u_0, X_\alpha\left[X_\varepsilon X_\phi \cdots X_\omega\right]u_0\right) = \left(\left[X'_\beta X'_\gamma \cdots X'_\delta\right]u'_0, X'_\alpha\left[X'_\varepsilon X'_\phi \cdots X'_\omega\right]u'_0\right) \tag{18.14}$$

On account of the hermiticity conditions

$$H_i^* = H_i$$
$$E_\alpha = -E_{-\alpha}^* \tag{18.15}$$

mentioned in the last lecture, however, this means that what we have to show is that

$$\left(u_0, X_\alpha X_\beta \cdots X_\gamma u_0\right) = \left(u'_0, X'_\alpha X'_\beta \cdots X'_\gamma u'_0\right) \tag{18.16}$$

for any product $(X_\alpha X_\beta \cdots X_\gamma)$. The proof of this is by induction. First of all, we have by hypothesis

$$\left(u_0, H_i u_0\right) = m_i^0 = \left(u'_0, H'_i u'_0\right),$$
$$\left(u_0, E_\alpha u_0\right) = 0 = \left(u'_0 E'_\alpha u'_0\right)$$

$$\left(u_0, E_{-\alpha} u_0\right) = -\left(E_\alpha u_0, u_0\right) = 0 = -\left(E'_\alpha u_0, u'_0\right) = \left(u'_0, E'_{-\alpha} u'_0\right) \tag{18.17}$$

so that, at any rate

$$\left(u_0, X_\alpha u_0\right) = \left(u'_0, X'_\alpha u'_0\right) \tag{18.18}$$

Now suppose that (18.16) holds for a product of n X's, and consider a product of $(n+1)$ X's. We have

$$\left(u_0, \underbrace{x_\beta \cdots x_\delta}_{\leftarrow n \text{ terms} \rightarrow} E_\alpha u_0\right) = 0 = \left(u_0' x_\beta' x \cdots x_\delta' E_\alpha' u_0'\right),$$

$$\left(u_0, x_\beta \cdots x_\delta H_i u_0\right) = m_i^0 \left(u_0, x_\beta \cdots x_\delta u_0\right) \tag{18.19}$$

$$= m_i^0 \left(u_0', x_\beta' \cdots x_\delta' u_0'\right) = \left(u_0', x_\beta' \cdots x_\delta' H_i' u_0'\right) \tag{18.20}$$

and

$$\left(u_0, x_\beta \cdots x_\delta E_{-\alpha} u_0\right) = \left(u_0, E_{-\alpha} x_\beta \cdots x_\delta u_0\right) + \left(u_0, \left[E_{-\alpha}, x_\beta \cdots x_\delta\right] u_0\right)$$

$$= -\left(E_\alpha u_0, x_\beta \cdots x_\delta u_0\right) + \underbrace{C_{-\alpha\beta}^{\gamma}\left(u_0, x_\gamma \cdots x_\delta u_0\right)}_{\leftarrow n \text{ terms} \rightarrow} + \text{Similar terms}$$

$$= 0 + \underbrace{C_{-\alpha\beta}^{\gamma}\left(u_0', x_\gamma' \cdots x_\delta' u_0'\right)}_{\leftarrow n \text{ terms} \rightarrow} + \text{Similar terms}$$

$$= 0 + \left(u_0', \left[E_{-\alpha}', x_\beta' \cdots x_\delta'\right] u_0'\right)$$

$$= 0 + \left(u_0', E_{-\alpha}' x_\beta' \cdots x_\delta' u_0'\right) + \left(u_0', x_\beta' \cdots x_\delta' E_{-\alpha}' u_0'\right)$$

$$= 0 - \left(E_\alpha' u_0', x_\beta' \cdots x_\delta' u_0'\right) + \left(u_0', x_\beta' \cdots x_\delta' E_{-\alpha}' u_0'\right)$$

$$= 0 - 0 + \left(u_0', x_\beta' \cdots x_\delta' E_{-\alpha}' u_0'\right). \tag{18.21}$$

Thus,

$$\left(u_0, \underbrace{x_\beta \cdots x_\delta}_{\leftarrow n \rightarrow} X_\chi u_0\right) = \left(u_0', x_\beta' \cdots x_\delta' X_\chi' u_0'\right) \tag{18.22}$$

and (18.16) holds for $n+1$ terms. Q.E.D.

These 2 theorems show that the irreducible representations of a Lie algebra are completely determined by their highest weights. Thus the problem of finding the irreducible representations is reduced to the problem of determining the possible highest weights, subject to the condition on the weights derived in the previous lecture, namely, that if α is a root $\frac{2\, m\cdot\alpha}{\alpha\cdot\alpha}$ must be an integer.

We shall close this lecture by mentioning a result which will be of use for construcing the irreudicble representations in the following lectures. Let $D(x)$ and $D'(x')$ be two representations and let $D(x) \times D(x')$ be the Kronecker product representation. Then the weights of $D(x) * D(x')$ are all possible seems of the weights of $D(x)$ and $D(x')$.

Proof: Let

$$D(\varepsilon) = (1 + \varepsilon_\alpha X_\alpha),$$
$$D'(\varepsilon) = (1 + \varepsilon_\alpha X'_\alpha), \qquad (18.23)$$
$$D(\varepsilon) \times D'(\varepsilon) = 1 + \varepsilon_\alpha X^K_\alpha .$$

Hence,

$$X^K_\alpha = X_\alpha \times 1 + 1 \times X'_\alpha \qquad (18.24)$$

In particular,

$$H_i^K = H_i \times 1 + 1 \times H'_i \qquad (18.25)$$

Hence the eigenvectors of H_i^K are just the products $u_m u'_{m'}$ of the eigenvectors of H_i and H'_i, and

$$H_i^K u_m u'_{m'} = H_i u_m u'_{m'} + H'_i u_m u'_{m'} \qquad (18.26)$$
$$= (m_i + m'_i)\, u_m u'_{m'} .$$

Q.E.D.

LECTURE 19.

In this and the following lectures we shall show how the irreducible representations of the groups A_ℓ, B_ℓ, C_ℓ and D_ℓ are obtained. In this lecture we shall treat C_ℓ because it is the simplest case and therefore allows one to see more clearly the ideas involved.

First of all, we recall that the root vectors for C_ℓ where $\alpha = \pm 2e_i$, $\pm e_i \pm e_k$ $(i \neq k)$, and that if m is any weight

$$2\,\frac{(m\cdot\alpha)}{\alpha\cdot\alpha} = \text{integer} \tag{19.1}$$

Inserting $\pm 2e_i$ for α we obtain from this the condition

$$m_i = \text{integer} \tag{19.2}$$

and inserting $\pm e_i \pm e_k$ yields no further condition. Thus the components of the weight vectors must be integers.

Let us now reflect $m = (m_1, m_2, \ldots m_\ell)$ in the plane perpendicular to $2e_1$. We obtain $m' = (-m_1, m_2 \ldots m_\ell)$ Similiarly for any m_i. Reflecting in the plane perpendicular to $e_1 - e_2$ we obtain $m' = (m_1' m_2' \ldots m_\ell')$, where

$$\begin{aligned} m_1' &= m_1 - 2\,\frac{(m_1 - m_2)}{2}\cdot 1 = m_2 \\ m_2' &= m_2 - 2\,\frac{(m_1 - m_2)}{2}\cdot 1 = m_1 \\ m_i' &= m_i \qquad i \neq 1,2 \end{aligned} \tag{19.3}$$

Thus under this reflexion m_1 and m_2 permute. In general under the reflexion perpendicular to $e_i - e_j$, m_i and m_j permute. Reflecting perpendicular to $(e_i + e_j)$ yields nothing new, We have thus considered all possible reflexions perpendicular to the roots and seen that they produce either a change of sign in

one of the m_i or a permutation of two m_i. Thus the Weyl group generated by taking any product of such operations consists of the group of all permutations of the m_i and all changes of sign. Its order is therefore $2^{\ell}\,\ell!$

From these results it follows that if m is the highest weight for any representation

$$m_1 \geq m_2 \geq m_3 \geq \cdots \geq m_\ell \geq 0 \qquad (19.4)$$

Proof: Suppose this is not true. First, if any m_i is negative we can simply change its sign by a reflexion and so obtain a higher weight. Thus all the m_i must be positive. Next, reading from the left let m_r be the first m_i which is followed by a greater m_i i.e.

$$m_r < m_{r+1} \qquad (19.5)$$

Reflecting this so that m_r and m_{r+1} interchange we obtain a weight m' (belonging to the same representation) such that

$$m' = (m_1, m_2, \cdots m_{r-1}, m_{r+1}, m_r, \cdots) \qquad (19.6)$$

and so

$$(m' - m) = (0, 0, 0, \cdots, (m_{r+1} - m_r), \cdots) \qquad (19.7)$$
$$> 0$$

contrary to hypothesis. Q.E.D.

Thus the highest weights of the various representations must satisfy the conditions m_i = integer and $m_i \geq m_{(i+1)}$ Oour problem now is to find the irreducible representations corresponding to these highest weights.

We start by finding the representation which corresponds to the lowest possible highest weight, namely (1 0 0 0 ... 0).

This representation is obtained as follows:

Let J be the $(2\ell \times 2\ell)$ matrix.

$$J = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}, \tag{19.8}$$

Where I_ℓ is the unit $(\ell \times \ell)$ matrix. Consider the set of all real non-singular $(2\ell \times 2\ell)$ matrices D such that

$$\tilde{D} J D = J, \tag{19.9}$$

where $\sim$ denotes transpose. Under matrix multiplication this set of matrices D forms a group, since if D_1 and D_2 are two such matrices

$$\widetilde{(D_1 D_2)} J (D_1 D_2) = \tilde{D}_2 \tilde{D}_1 J D_1 D_2 = \tilde{D}_2 J D_2 = J, \tag{19.10}$$

the unit element $(D = I_{2\ell})$ is a member of the set and for the inverse D^{-1} also, (19.9) is true. The group D is called the real sympletic group in 2ℓ dimensions. Correspondingly, if the matrices D are allowed to be complex we get the complex sympletic group in 2ℓ dimensions.

The matrices D have the property that if we form any pair of 2ℓ-vectors of the kind $X = (x_1, \ldots x_\ell, x'_1 \ldots x'_\ell)$ $Y = (y_1, \ldots y_\ell, y'_1 \ldots y'_\ell)$ the D's leave the sympletic form

$$\sum_{r=1}^{\ell} (x_r y'_r - x'_r y_r) = X J Y, \tag{19.11}$$

invariant. This follows immediately from (19.9). Let us now express (19.9) in infinitesimal form. Let

$$D = 1 + \varepsilon_\alpha \hat{X}_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \varepsilon_\alpha \begin{pmatrix} X_\alpha & Y_\alpha \\ Z_\alpha & W_\alpha \end{pmatrix}. \tag{19.12}$$

From (19.9) we obtain

$$\hat{X}_\alpha \tilde{J} = - \tilde{J} \hat{X}_\alpha \ , \qquad (19.13)$$

and substituting into this equation from (19.12), we obtain

$$\begin{aligned} X_\alpha &= -\tilde{W}_\alpha \\ Y_\alpha &= \tilde{Y}_\alpha \\ Z_\alpha &= \tilde{Z}_\alpha \end{aligned} \qquad (19.14)$$

This the generators of this group are the set of all linearly independent matrices which satisfy (19.14). There are clearly $\ell^2 + 2\,\frac{\ell(\ell+1)}{2} = (2\ell^2 + \ell)$ such matrices. Thus the group is of order $r = 2\ell^2 + \ell$.

What we now wish to show is that this group of matrices is the irreducible representation of C_ℓ with highest weight $(1,0,0,\ldots 0)$. This we shall show by proving that the group is, in fact, isomorphic to C_ℓ , and that the highest weight is $(1,0,0,\ldots 0)$. To do this we choose the following basis for the sympletic group.

$$H_1 = \begin{pmatrix} 1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & -1 & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} \qquad H_2 = \begin{pmatrix} 0 & & & & & \\ & 1 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & -1 & \\ & & & & & 0 \end{pmatrix} \quad \text{etc}$$

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ & & & & 0 & 0 & 0 & 0 \\ & & & & -1 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{etc} \qquad E_{-12} = E_{12}^{\dagger} \qquad (19.15)$$

$$E'_{11} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad E_{-11} = E_{11}^{\dagger}$$

$$E^{11}_{12} = \left[\begin{array}{c|cccc} 0 & 0 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 \\ \hline 0 & & & & 0 \end{array}\right] \quad \text{etc.} \quad E^{11}_{-12} = E^{11\dagger}_{12} \text{ etc}$$

Having obtained the self-representation or relation of C_ℓ, with highest weight $(1,0,0,\ldots..0)$, we now obtain from this ℓ representations as follows. Calling the self-representation D_1 we construct the product representation $D_1 \times D_1$. This is reducible. The next irreducible representation is obtained by taking the leading representation in the antisymmetrics part of $(D_1 \times D_1)$. What we mean by the anti-symmetric part is the following: Let $(u_1, \ldots u_\ell)$ be the base vectors as described above of the first D_1 and $(v_1, \ldots . v_\ell)$ the corresponding basis of the second D_1. The representation of $(D_1 \times D_1)_{A.S}$ is the product space $(u_i v_j - v_i u_j)$, and $(D_1 \times D_1)_{A.S}$ is simply the projection of $(D_1 \times D_1)$ into this space. Thus we obtain a second irreducible representation.

$$D_2 = \text{Leading Representation in } (D_1 \times D_1)_{A.S}$$

The question is: What is the highest weight of D_2. From the result of the last lecture that the weights of $D_1 \times D_1$, is the sum of the highest weights of D_1 and D_1 i.e. $(1\ 0\ 0 \ldots 0) + (1\ 0\ 0 \ldots 0) = (2\ 0\ 0 \ldots 0)$. But this does not belong to D_2 because the

corresponding eigenvector is $u_1 v_1$ and when we anti-symmetrize this we get zero. However, the next highest weight is obtained by taking the highest weight (1 0 0 ... 0) of D_1 together with the second highest weight (0 1 0 ... 0) of D_1. This gives a weight (1 1 0 0 ... 0) and this coertainly belongs to D_2 since its eigenvector is $u_1 v_2$ which can be anti-symmetrized to $(u_1 v_2 - u_2 v_1)$. Thus the highest weight of D_2 is (1 1 0 0 ... 0).

Continuing in the same way, we form $D_1 \times D_1 \times D_1$ and project out the completely anti-symmetric leading part in $(D_1 \times D_1 \times D_1)_{C.A.S.}$. It turns out that this is again irreducible, and it is easy to see by an argument similar to that given for D_2 that the highest weight is (1 1 1 0 ... 0).

We carry out this process ℓ times and obtain the following ℓ representations:

	Weight
D_1	$m^{(1)} = (1,0,0,\ldots\ldots 0)$
D_2 = Leading representation in $(D_1 \times D_1)_{AS}$	$m^{(2)} = (1,1,0,\ldots\ldots 0)$
D_3 = Leading irreducible representation in $(D_1 \times D_1 \times D_1)_{C.A.S}$	$m^{(3)} = (1,1,1,0,\ldots 0)$
$\ldots\ldots$	$\ldots\ldots$
D_ℓ = " in $(\underbrace{D_1 \times D_1 \times \ldots \times D_1}_{\ell})_{C.A.S}$	$m^{(\ell)} = (1,1,1,\ldots\ldots 1)$

These ℓ irreducible representations are called the ℓ <u>fundamental</u> representations of C_ℓ.

The importance of the fundamental representation lies in the fact that from them one can construct all the irreducible representations. The construction is as follows:

Let m be any highest weight. Since the m_i are then integers such that $m_1 \geq m_2 \geq \cdots \geq m_{\ell} \geq 0$, it is easy to see that m can be expressed in the form

$$m = k_1 m^{(1)} + k_2 m^{(2)} + \cdots\cdots +$$

where

$$k_1 = m_1 - m_2 ,$$

$$k_2 = m_2 - m_3 ,$$

$$\vdots$$

$$k_{\ell-1} = m_{\ell-1} - m_{\ell}$$

$k_{\ell} = m_{\ell}$, are non-negative integers.

Consider now the direct product

$$\mathcal{D} = \underbrace{D_1 \times D_1 \times \cdots\cdots \times D_1}_{\longleftarrow\ k_1\ \longrightarrow} \times \underbrace{D_2 \times \cdots\cdots \times D_2}_{\longleftarrow\ k_2\ \longrightarrow} \times \cdots\cdots \times \underbrace{D_{\ell} \times \cdots \times D_{\ell}}_{\longleftarrow\ k_{\ell}\ \longrightarrow}$$

This representation is reducible, Let D_m be the irreducible representation contained in it which has the highest weight.

Then we have

$D = D_m \dot{+}$ representations with lower highest weights.

But then the highest weight of of D_m = highest weight of D, and the latter is just the sum of the highest weights of D_1, taken k_1 times, D_2 taken k_2 times, etc. i.e. it is just

$$m = k_1 m^{(1)} + k_2 m^{(2)} + \cdots\cdots + k_{\ell} m^{(\ell)} .$$

Thus D_m is the irreducible representation with the weight m.

Thus, given any m, we can construct the corresponding irreducible representation D by taking the leading irreducible representation in the corresponding product representation D.

Q: How do you know that there are not two or more representation with the weight m in D ?

A: From the fact that to the highest weight in D corresponds only one unique vector in the representation space, namely, that formed with the product of all the vectors belonging to the highest weights in the component representations $D_1, \ldots D_1, D_2, \ldots D_2, \ldots\ldots D_\ell, \ldots\ldots D_\ell$

Q: Since, the $D_1, D_2, \ldots\ldots D_\ell$ were constructed out of direct products of D_1 with itself, why not call D_1 the one fundamental weight ?

A Certainly, D_1 is more fundamental than $D_2, \ldots\ldots D_\ell$. But the point is that for the ℓ fundamental representation; we can construct any representation by taking the leading term in the suitable direct product, whereas to construct the $D_i \left(1 \leq i \leq \ell\right)$ out of D_1 we must take the completely anti-symmetric parts of the direct products. Thus the procedure is quite different in the two cases. Furthermore, as we shall see later, whereas the result that out of the ℓ -fundamental representations we can construct all irreducible representations by the above construction, carries over to any Lie group, the result that all of the ℓ fundamental representations can be constructed out of the "self"-representation in the above way does not. For some groups, certain fundamental representations (the "spinor" representations) are constructed in an entirely different way.

LECTURE 20.

In this lecture we shall discuss the group A_ℓ. We start as usual with the root diagram which consists of all vectors $e_i - e_j$ in an $\ell+1$ dimensional space. Since the roots are most easily described in the $(\ell+1)$ space, it is convenient to consider the weights also as vectors in the $(\ell+1)$ space, but because they will not then automatically lie in the ℓ-space with normal $\frac{1}{(\ell+1)^{1/2}}(1,1,1,\ldots,1)$ we must impose on them the condition

$$\sum_{i=1}^{(\ell+1)} m_i = 0 \qquad (20.1)$$

Then they lie in the same ℓ-plane as the roots.

The condition that $2\,\frac{m\cdot\alpha}{\alpha\cdot\alpha}$ be an integer becomes in this case

$$2\,\frac{m\cdot(e_i-e_j)}{(e_i-e_j)^2} = 2\,\frac{(m_i-m_j)}{2} = (m_i-m_j) = \text{integer} \qquad (20.2)$$

Thus the differences of the components of the m's are integers.

Then from (20.1)

$$\begin{aligned} m_{\ell+1} &= m_{\ell+1}\,, \\ m_\ell &= m_{\ell+1} + \text{integer}, \\ m_{\ell-1} &= m_{\ell+1} + \text{integer}, \\ &\cdots\cdots\cdots \\ m_1 &= m_{\ell+1} + \text{integer} \end{aligned} \qquad (20.3)$$

Hence

$$0 = (\ell+1)\,m_{\ell+1} + \text{integer}$$

$$m_{\ell+1} = \frac{\text{integer}}{\ell+1} \qquad (20.4)$$

and so,

$$m_i = \frac{\text{integer}}{\ell+1} \qquad (20.5)$$

Equation (20.1), (20.2) and (20.3) are the conditions that the weights must satisfy for this group.

Under a reflection in the plane perpendicular to $(e_i - e_j)$ we have

$$m \longrightarrow m' = m - 2\,\frac{m\cdot(e_i - e_j)}{(e_i - e_j)^2}(e_i - e_j)$$

$$= m - (m_i - m_j)(e_i - e_j)$$

$$= m\left(m_i \rightleftarrows m_j\right) \qquad (20.6)$$

Thus the Weyl group S in this case consists of all the permulations of the components of m. It is therefore of order $(\ell+1)!$

Q: What about the auxiliary condition (20.1) which implies that not all the m_i are independent? Will this not reduce the possibilities for S.

A: No. The point is that if m satisfies (20.1) so does Sm automatically. Hence (20.1) implies no conditions on S. That m and Sm really lie in an ℓ-space does not mean that they will not be distinct.

An immediate consequence of this is that the highest weight for any representation must satisfy the condition.

$$m_1 \geq m_2 \geq \cdots\cdots \geq m_\ell \qquad (20.7)$$

This is similiar to the condition obtained for C_ℓ but it differ in the fact that the m_i in (20.7) need not be positive (except for m_1, since by (20.1) they cannot all be negative)

To find the representations for this group we proceed in exactly the same way as in the case of C_ℓ i.e. we first obtain the self-representation, then construct out of that

ℓ-fundamental representations, and finally use these to build up any representation.

We obtain the fundamental representation as follows: We consider the group of all $(\ell+1)\times(\ell+1)$ unimodular matrices $(SL_{\ell+1})$. It is easy to see that the infinitesimal generators of this group are simply the set of all traceless $(\ell+1)\times(\ell+1)$ matrices. If we neglect the unimodularity condition, and we shall do so for convenience, the set of infinitesimal generators becomes the set of all $(\ell+1)\times(\ell+1)$ matrices. We choose as a basis for their algebra the matrices

$$\begin{aligned} \overline{H_i} &= |i\rangle\langle i| , \\ E_\alpha &= |i\rangle\langle j| \quad , \quad |j\rangle\langle i| \\ E_{-\alpha} &= |i\rangle\langle j| \quad , \quad |j\rangle\langle i| \end{aligned} \tag{20.8}$$

where the $|i\rangle$ for a set of orthonormal vectors in $(\ell+1)$ space. The commutation relations $[\overline{H_i}, E_\alpha]$ are then

$$[\overline{H_i}, E_\alpha] = [|i\rangle\langle i| \quad , \quad |j\rangle\langle k|] = (\delta_{ij} - \delta_{ik})|j\rangle\langle k| = (\delta_{ij} - \delta_{ik})E_\alpha \tag{20.9}$$

Thus the roots are just $e_i - e_j$, and this group is isomorphic to A_ℓ. Thus the self-representation or realization of A_ℓ is the group of unimodular matrices in $(\ell+1)$-dimensions.

In the last lecture, we saw that because the sympletic matrices (19.15) were unitary when they were real, the real sympletic group was the compact group corresponding to the complex sympletic group. Here, the situation is a little different, because the ××××××××××××××× unimodular matrices are not unitary when they are real. The compact group of $SL_{\ell+1}$ is the group of unitary unimodular matrices ($SU_{(\ell+1)}$).

In (20·8) and (20·9) we have, of course, neglected the unimodular condition. We can introduce it by defining instead of the $(\ell+1)\ \overline{H}_i$, $(\ell+1)\ H_i$ by the relation

$$H_i = \overline{H_i} - \frac{1}{(\ell+1)}\ \overline{I} \qquad (20.10)$$

where $\overline{I}$ is $(\ell+1)\times(\ell+1)$ unit matrix. The matrices H_i are traceless in conformity with the unimodular condition, and further

$$\sum_{i=1}^{\ell+1} H_i = 0 \qquad (20.11)$$

So that they are not linearly independent, but satisfy the same sort of subsidiary condition as we put earlier on the roots and weights.

Q: Are the commutation relations (20.9) preserved under the change $\overline{H_i} \rightarrow H_i$? ?

A: Yes, because I in (20.10) commutes with everything. Thus the H_i take the following form:

$$\frac{1}{(\ell+1)}\begin{pmatrix} \ell & & & \\ & -1 & & \\ & & -1 & \\ & & & \ddots \\ & & & & -1 \end{pmatrix} = H_1 ; \ \ldots\ldots \ H_{\ell+1} = \frac{1}{(\ell+1)}\begin{pmatrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & \ell \end{pmatrix} \qquad (20.12)$$

Their simultaneous eigenvectors are obviously the vectors,

$$u_1 = (1,0,0,0,\ldots 0)\ ,\ u_2 = (0,1,0,\ldots 0)\ ,\ \ldots u_{\ell+1} = (0,0,0,\ldots 1)\ , \qquad (20.13)$$

and the corresponding weights are

$$\frac{1}{\ell+1}(\ell,-1,-1,\ldots -1)\ ,\quad \frac{1}{\ell+1}(-1,\ell,-1\ \ldots -1)\ ,\ \ldots\ldots\ \frac{1}{\ell+1}(-1,-1,\ldots\ell) \qquad (20.14)$$

The highest weight here is obviously $\frac{1}{\ell+1}(\ell,-1,-1,\ldots -1)$. Note that, as in the case of C_ℓ , the vectors $u_1, u_2, \ldots u_\ell$ are non-degenerate.

Q: Do you know as in the case of C_ℓ, that $\frac{1}{\ell+1}(\ell,-1,\cdots-1)$ is the lowest possible highest weight ?

A: One could probably show this at this point, but I shall not do so because it will become obvious later.

Having obtained the above self-representation, which we shall call D_1, we proceed in exactly the same way as in the case of C_ℓ to construct $(D_1\times D_1)_{AS}$, $(D_1\times D_1\times D_1)_{CAS}$, ... etc. Just as before each of these turns out to be an irreducible representation, and so we get the set of ℓ irreducible representations and their highest weights:

$$
\begin{aligned}
&D_1 \quad \text{with} && m^{(1)} = \tfrac{1}{\ell+1}(\ell,-1,-1,\cdots-1)\\
&D_2 = (D_1\times D_1)_{AS} \quad \text{with} && m^{(2)} = \tfrac{1}{\ell+1}(\ell-1,\ell-1,-2,\cdots-2)\\
&D_3 = (D_1\times D_1\times D_1)_{CAS} \quad \text{with} && m^{(3)} = \tfrac{1}{\ell+1}(\ell-2,\ell-2,\ell-2,-3,\cdots-3)\\
&\qquad\cdots\cdots\cdots && \qquad\cdots\cdots\cdots\\
&D_\ell = \underbrace{(D_1\times D_1\times\cdots\times D_1)}_{\leftarrow\ \ell\ \rightarrow}{}_{CAS} \quad \text{with} && m^{(\ell)} = \tfrac{1}{\ell+1}(1,1,1,\cdots-\ell)
\end{aligned}
\tag{20.15}
$$

The weight of D_2 is obtained by adding the highest and second highest weights of D_1, that of D_3 by adding the first, second and third highest of D_1, and so on.

It remains to show that these ℓ representations are fundamental representations ××××××××××××××××××× in the same sense as the corresponding ones for C_ℓ. To show this let m be any highest weight. It then satisfies the conditions

$$
\sum_{i=1}^{\ell+1} m_i = 0, \quad m_i - m_k = \text{integer}, \quad m_i = \frac{\text{integer}}{(\ell+1)} \tag{20.16}
$$

$$
m_i \geq m_{i+1}
$$

If we let

$$k_1 = m_1 - m_2,$$
$$k_2 = m_2 - m_3, \qquad (20.17)$$
$$\cdots\cdots$$
$$k_\ell = m_\ell - m_{\ell+1}$$

then we see by inspection that

$$m = k_1 m^{(1)} + k_2 m^{(2)} + \cdots\cdots + k_\ell m^{(\ell)} \qquad (20.18)$$

But the k_i, $i = 1 \ldots\ldots \ell$ are integers by (20.16) (note that since we have $(\ell+1)$ m_i's but only ℓ k_i's all the k's including the last, can be expressed as <u>differences</u> of m's). Further, the k's are non-negative integers. Hence, if we form the direct product

$$\mathcal{D} = \underbrace{D_1 \times D_1 \times \cdots\cdots \times D_1}_{k_1} \times \underbrace{D_2 \times D_2 \times \cdots\cdots \times D_2}_{k_2} \times \cdots\cdots \underbrace{D_\ell \times D_\ell \times \cdots \times D_\ell}_{k_\ell}$$
$$= D_m + \cdots\cdots$$

the leading irreducible representation D_m occurring in D (i.e. that with the highest highest weight) will have the highest weight in D, namely m. In this way we can construct the irreducible representation D_m with any given highest weight m out of the ℓ representations (20.15). These are therefore fundamental as required.

LECTURE 21.

In this lecture we shall find the representations of the group B_ℓ.

The roots of B_ℓ are $\pm e_i, \pm e_i \pm e_j \quad i \neq j = 1, \ldots \ell$ from lecture 16. Let m be any weight vector. Then from the weight theorem

$$2\,\frac{m\cdot\alpha}{\alpha\cdot\alpha} = 2\,\frac{m\cdot e_i}{e_i\cdot e_i} = \text{integer}, \quad m_i = \frac{\text{integer}}{2}$$

$$2\,\frac{m\cdot\alpha}{\alpha\cdot\alpha} = 2\,\frac{(m\cdot e_i - e_j)}{(e_i - e_j)^2} = m_i - m_j = \text{integer}, \tag{21.1}$$

$$2\,\frac{m\cdot\alpha}{\alpha\cdot\alpha} = 2\,\frac{(m\cdot e_i + e_j)}{(e_i + e_j)^2} = m_i + m_j = \text{integer}$$

These three relations are not independent, either of the second following from the first plus the other. Note that, in contrast to the cases C_ℓ and A_ℓ which we have already considered, the m_i can here take on $\frac{1}{2}$-integer values. The contrast with C_ℓ is particularly sharp since the latter group has the very similar roots $\pm 2e_i, \pm e_i \pm e_j$. However, the factor 2 in $\pm 2e_i$ makes a great difference, since for C_ℓ the equation corresponding to the first equation above is

$$2\,\frac{m\cdot\alpha}{\alpha\cdot\alpha} = 2\,\frac{m\cdot 2e_i}{2e_i\cdot 2e_i} = m_i = \text{integer} \tag{21.2}$$

Under a reflexion in the plane perpendicular to e_i the ith component of m changes sign. Hence the Weyl group S includes all independent changes of sign of the components of m. Under a reflexion in the plane perpendicular to $e_i - e_j$, we obtain

$$m \longrightarrow m' = m - 2\,\frac{m\cdot(e_i-e_j)}{(e_i-e_j)^2}(e_i-e_j)$$

$$= m - (m_i-m_j)(e_i-e_j) = m\left(m_i \rightleftarrows m_j\right) \qquad (21.3)$$

which means that m_i and m_j get permuted. Thus S includes also all permutations of the components of m. Reflexion in the planes perpendicular to the $e_i + e_j$ gives nothing new, so that S is the group of all possible permutations of the components of m, together with the all possible changes of sign. It follows from the permutation property that if m is a <u>highest</u> weight

$$m_1 \geq m_2 \geq \cdots\cdots \geq m_\ell, \qquad (21.4)$$

whhile it follows from the change of sign property that

$$m_1 \geq m_2 \geq \cdots\cdots \geq m_\ell \geq 0. \qquad (21.5)$$

Equations (21.1) and (21.2) are the conditions that must be satisfied by any highest weight.

Our next step is to obtain the self-representation for B_ℓ. The self-representation as we shall show below is the group of all orthogonal $(2\ell+1)\times(2\ell+1)$ matrices $O_{2\ell+1}$. The compact group corresponding to this group is the group of all <u>real</u> orthogonal $(2\ell+1)\times(2\ell+1)$ matrices, since any matrix which is real and orthogonal is unitary. The infinitesimal generators of the $O_{2\ell+1}$ are simply the anti-symmetric $(2\ell+1)\times(2\ell+1)$ matrices. By choosing a suitable basis

$$X_{mn}, \qquad m,n = 0,1,2,\ldots 2\ell$$

in the space of the generators, we obtain the well-known commutation relations for the orthogonal group, namely,

$$\left[X_{mn}, X_{rs}\right] = \delta_{nr} X_{ms} + \delta_{ms} X_{nr} - \delta_{mr} X_{ns} - \delta_{ns} X_{mr} \quad (21.5)$$

From these relations we see that a convenient, set of H_i's for this group is

$$H_1 = X_{12}\,, \quad H_2 = X_{34}\,, \quad \ldots\ldots \quad H_\ell = X_{2\ell-1\,2\ell} \quad (21.6)$$

If we now divide the remaining X's into the two classes,

E_α = certain linear combinations of the X_{mn} $m,n = 1,\ldots 2\ell$, (21.7)

E'_α = certain linear combinations of the X_{on} $n = 1\ldots 2\ell$ (21.8)

we obtain,

$$\left[H_i, E_\alpha\right] = \pm\left(e_i \pm e_j\right) E_\alpha\,, \quad (21.9)$$

$$\left[H_i, E'_\alpha\right] = \pm e_i E'_\alpha\,, \quad (21.10)$$

which estabishes the isomorphism of $O_{2\ell+1}$ and B_ℓ.

If we now diaganalize the H_i, we obtain

$$H_1 = \begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}; \quad H_2 = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 1 & & \\ & & & -1 & \\ & & & & 0 \\ & & & & & \ddots \end{pmatrix} \quad \ldots\ldots \quad H_\ell = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix} \quad (21.11)$$

The simultaneous eigenvectors of the H_i are

$$u_1 = (1\,0\,0\,0\,0\,\ldots\,0), \quad u_2 = (0\,1\,0\ldots\ldots 0) \quad \ldots\ldots \quad u_{2\ell} = (0\,0\,0\,0\ldots\ldots 0\,1) \quad (21.12)$$

These obviously span the whole representation space and so are non-degenerate. Their weights are, respectively

$$(1\,0\,0\ldots\ldots 0)\,, (-1\,0\,0\,0\,0\ldots 0)\,, (0\,1\,0\,0\,0\,0\ldots 0)\,, \ldots\ldots \quad (21.13)$$

$$\ldots\ldots (0\,0\,0\ldots\ldots 0\,-1)$$

of which the highest is clearly the first.

We now form in the by now familiar way the representations,

$D_1 = D_{self}$ with highest weight $m^{(2)} = (100\ldots0)$

$D_2 = (D_1 \times D_1)_{A.S}$ " " " " $m^{(3)} = (1100\ldots0)$

$D_3 = (D_1 \times D_1 \times D_1)_{cAS}$ " " " $m^{(4)} = (111000\ldots0)$ (2114)

. . . . . .

$D_{\ell-1} = (D_1 \times D_1 \times \ldots \times D_1)_{cAS}$ " " $m^{\ell} = (11\ldots10)$
$\overleftarrow{\quad \ell-1 \quad}\overrightarrow{\quad}$

$D_\ell = (D_1 \times D_1 \times \ldots \times D_1)_{cAS}$ " " $m^{\ell+1} = (11\ldots11)$
$\overleftarrow{\quad \ell \quad}\overrightarrow{\quad}$

These turn out to be irreducible. However, in contrast to the case of groups A_ℓ and C_ℓ, they do not contitute of fundamental set (a fact we have anticipated by writing m^{r+1} where we previously wrote m^r). The reason is that out of the D_r we can construct only "tensor" representations i.e. representations the components of whose weights are integer, whereas we have been above that weights with $\frac{1}{2}$-integer components can occur (and in fact they do occur (see below)). Thus the above set D_r consitute a fundamental set only for the tensor representations. If we wish to include also the "spinor", or $\frac{1}{2}$-integer, representations we must look for at least one $\frac{1}{2}$-integer representation to add to our fundamental set. It turns out that, in fact, on $\frac{1}{2}$-integer representation is sufficient, the representation in question being called Δ and having the highest weight $m^{(1)} = \left(\frac{1}{2}\ \frac{1}{2}\ \cdots\ \frac{1}{2}\ \frac{1}{2}\right)$

I shall not have time in these lectures to show how Δ is constructed but for a very lucid account of its construction I can refer

you to the self-contained chapter VIII of Boerner, "Representations of groups", North-Holland, Amsterdan 1963, and the papaer of Braner and Weyl, American Journal of Mathematics, 57, (1935).

We see at once that the highest weight of $\Delta \times \Delta$ is $(1\ 1\ \ldots\ 1)$. Thus $\Delta \times \Delta$ contains $\mathcal{D}_\ell$ as its leading irreducible representation. Actually the complete decomposition of $\Delta \times \Delta$ is

$$\Delta \times \Delta = \mathcal{D}_\ell \dot{+} \mathcal{D}_{\ell-1} \dot{+} \mathcal{D}_{\ell-2} \dot{+} \cdots \dot{+} \mathcal{D}_1 \dot{+} 1 \qquad (21.15)$$

Thus the $\mathcal{D}_r$ are not "elementary" but can be generated from the Δ. However, only $\mathcal{D}_\ell$ is generated in the way in which a non-fundamental irreducible representation is generated, namely, by being the leading irreducible representations in the direct product of fundamental representations. This mean that the set of ℓ irreducible representations, $\Delta, \mathcal{D}_1, \mathcal{D}_2, \ldots\ldots\ldots \mathcal{D}_{\ell-1}$ with highest weights $m^{(1)}, m^{(2)}, \ldots\ldots m^{(\ell)}$, should be the fundamental representations. It remains only to show that, in fact, they are.

To show this we need only werify that if any m satisfies (21.5) and (21.1) it can be expressed as

$$m = k_1 m^{(1)} + k_2 m^{(2)} + \cdots\cdots + k_\ell m^{(\ell)} \qquad (21.16)$$

where the k_r are non-negative integers, and it is easy to verify that this is fulfilled by taking

$$k_1 = 2m_1$$
$$k_r = m_{r-1} - m_r, \quad \ell \geq r > 1. \qquad (21.17)$$

For once this is [illegible], it is clear that D_m is the leading irreducible representation of

$$\underbrace{\Delta \times \Delta \times \cdots \times \Delta}_{k_1} \times \underbrace{D_1 \times D_1 \times \cdots \times D_1}_{k_2} \times \underbrace{D_2 \times D_2 \times \cdots \times D_2}_{k_3} \times \cdots \times \underbrace{\cdots \times D_{\ell-1} \times \cdots \times D_{\ell-1}}_{k_\ell} \qquad (21.18)$$

LECTURE 22.

This, the last lecture in the "C" series, will complete the series by discussing the representations of D_ℓ. The group D_ℓ has the roots $(\pm e_i \pm e_j)$, $i \neq j = 1, \ldots \ell$ so the weight theorem condition $2\frac{m\cdot\alpha}{\alpha\cdot\alpha}$ = integer leads to the following results:

$$2\frac{m\cdot(e_i - e_j)}{(e_i - e_j)^2} = m_i - m_j = \text{integer}, \qquad (22.1)$$

$$2\frac{m\cdot(e_i + e_j)}{(e_i + e_j)^2} = m_i + m_j = \text{integer} \qquad (22.2)$$

From this we see that the components of the weights m_i must be integer or half-integers, first as in the case of B_ℓ.

The Weyl group S for D_ℓ, however, is not the same as the corresponding group for B_ℓ. To obtain S for D_ℓ we consider the two possible types of reflexion.

$$(a) \quad m \to m' = m - 2\frac{m\cdot(e_i - e_j)}{(e_i - e_j)^2}(e_i - e_j) = m - (m_i - m_j)(e_i - e_j)$$
$$= m\,(m_i \rightleftarrows m_j) \qquad (22.3)$$

$$(b) \quad m \to m' = m - 2\frac{m\cdot(e_i + e_j)}{(e_i + e_j)^2}(e_i + e_j)$$
$$= m - (m_i + m_j)(e_i + e_j) = m\,(m_i \overrightarrow{\leftarrow} -m_j) \qquad (22.4)$$

From this one sees that the group S consists of all permutations of the m_i together with all changes of sign in pairs. It is, therefore, of order $2^{\ell-1}\,\ell!$

From the permutation part of S one sees that any highest weight must satisfy the condition

$$m_1 \geq m_2 \geq \cdots\cdots \geq m_\ell \qquad (22.5)$$

and from the changes of sign in pairs one can see that

$$m_1 \geq m_2 \geq \cdots \geq m_{\ell-1} \geq |m_\ell| \qquad (22.6)$$

One sees this as follows: if $m_{\ell-1}$ and m_ℓ are both negative a double change of sign for them leads to a higher weight of the same representation, contrary to hypothesis. Hence one of them must be positive and (22.5) implies that this must be $m_{\ell-1}$ (which, incidentally, implies that all $m_1, \ldots, m_{\ell-1}$ must be positive). If m_ℓ is positive then (22.6) follows from (22.5). Thus we need consider only the case of m_ℓ negative. Suppose $m_\ell = -|m_\ell|$. Then a double change of sign plus a permutation of m_ℓ and $m_{\ell-1}$ leads to the weight,

$$m' = m_1 \cdots m_{\ell-2}\ |m_\ell| - m_{\ell-1} \qquad (22.7)$$

if $|m_\ell| > m_{\ell-1}$ this is a higher weight than m contrary to hypothesis. Hence $m_{\ell-1} \geq |m_\ell|$, as required.

Thus the conditions on the highest weights is that their components must be integers or half-integers satisfying (22.6)

We now introduce the self-representation of D_ℓ, which is the orthogonal group in 2ℓ dimensions (in contrast to B_ℓ where it is the orthogonal group in $(2\ell+1)$ dimensions).

As for B_ℓ, the real compact subgroup is the graph of all real orthogonal matrices. The infinitesimal generators X of the group are given by the set of all anti-symmetric matrices in 2ℓ dimensions. If we choose a suitable basis

$$X = X_{mn}, \quad m,n = 1 \ldots\ldots 2\ell \tag{22.8}$$

then the satisfy the commutation relations

$$\left[X_{mn}, X_{rs}\right] = \delta_{nr} X_{ms} + \delta_{ms} X_{nr} - \delta_{mr} X_{ns} - \delta_{ns} X_{mr} \tag{22.9}$$

exactly as in the case of B_ℓ, except that here the indices m,n run from 1 to 2ℓ, not from 0 to 2ℓ. Again we choose our H_i to be $X_{12}, X_{34}, \ldots\ldots X_{2\ell-1\,2\ell}$, and our E_α to be suitable combinations of the remaining X_{mn}. However, since there are no X_{on} in this case there will be no E'_α and the root equations are simply

$$\left[H_i, E_\alpha\right] = \pm (e_i \pm e_j) E_\alpha \tag{22.10}$$

which establishes the isomorphism between this group and D_ℓ.

As before, we diagralize the H_i. In this case they become

$$H_1 = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 0 & \\ & & & \ddots \end{pmatrix}, \quad H_2 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \ldots\ldots H_\ell = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1 \\ & & & & -1 \end{pmatrix} \tag{22.11}$$

and the 2ℓ simultaneous eigenvectors are

$$u_1 = (1\,0\,0\ldots 0), \quad u_2 = (0\,1\,0\,0\ldots 0), \ldots\ldots u_{2\ell} = (0,\ldots 0\,1) \tag{22.12}$$

which are non-degenerate and have weights

$$(1\,0\,0\ldots\ldots 0), \quad (-1\,0\,0\ldots\ldots 0), \quad (0\,1\,0\ldots\ldots 0), \ldots\ldots \quad (0\,0\ldots 0\,-1) \tag{22.13}$$

respectively. The highest weight is obviously the first.

We now form the usual representations

$$D_1 = D_{Self} \text{ with highest weight } m^{(3)} = (100\ldots0)$$
$$D_2 = (D_1 \times D_1)_{AS} \quad " \quad " \quad m^{(4)} = (1100\ldots0)$$
$$\vdots \qquad\qquad\qquad\qquad (\quad)$$
$$D_{\ell-2} = \left(\underbrace{D_1 \times \cdots \times D_1}_{\ell-2}\right)_{CAS} \quad " \quad " \quad m_1^{(\ell)} = (111\ldots100)$$
$$D_{\ell-1} = \left(\underbrace{D_1 \times \cdots \times D_1}_{\ell-1}\right)_{CAS} \quad " \quad " \quad m^{\ell+1} = (111\ldots110)$$
$$D_\ell = \left(\underbrace{D_1 \times \cdots \times D_1}_{\ell}\right)_{CAS} = D_\ell^+ + D_\ell^-, \text{ with highest weights}$$
$$(1\,1\,1\ldots11) \text{ and } (1\,1\,1\ldots1\,-1),$$
respectively.

The representations $D_1, \ldots D_{\ell-1}$ are irreducible as usual, but the representation D_ℓ decomposes into D_ℓ^+ and D_ℓ^-, each of which is irreducible (for details, see Boerner loc. cit. Ch.VII).

As in the case of B_ℓ, the representations obtained here do not constitute a fundamental set, except for the tensor representations for which set $D_1, D_2, \ldots D_{\ell-2}, D_\ell^+, D_\ell^-$ are fundamental (note that $D_{\ell-1}$ is the leading irreducible representation in $D_\ell^+ \times D_\ell^-$). In order to include the spinor or half integer representations, we must look for some fundamental spinor representations. It turns out that in this case we need two fundamental spinor representations. These are called Δ^+ and Δ^- and have highest weights $m^{(1)} = \left(\frac{1}{2}, \frac{1}{2}, \ldots \frac{1}{2}, \frac{1}{2}\right)$, $m^{(2)} = \left(\frac{1}{2}, \frac{1}{2}, \ldots \frac{1}{2}, -\frac{1}{2}\right)$ respectively. It is clear from (22.6) why two such spinor representations are needed, namely, that since the last component

m_ℓ of m may be either positive or negative (and half-integer) neither Δ^+ alone together with the set (22.14) nor Δ^- alone together with the set, would be sufficient to construct all irreducible representations in the way required for fundamental representations.

We see at once that $(\Delta^+ \times \Delta^+)$, $(\Delta^- \times \Delta^-)$ and $(\Delta^+ \times \Delta^-)$ have the highest weights (11 ... 11), '11, ... 1 -1) and (11 ... 10) respectively, hence contain D^+_ℓ, D^-_ℓ and $D_{\ell-1}$ respectively as their leading irreducible representations. The complete decomposition of these products is actually

$$\begin{aligned} \Delta^+ \times \Delta^+ &= \begin{matrix} \ell \text{ even} \\ \ell \text{ odd} \end{matrix} \\ \Delta^- \times \Delta^- &= \begin{matrix} \ell \text{ even} \\ \ell \text{ odd} \end{matrix} \qquad (22.15) \\ \Delta^+ \times \Delta^- &= \begin{matrix} \ell \text{ even} \\ \ell \text{ odd} \end{matrix} \end{aligned}$$

but for details of this and the construction of Δ^+ and Δ^- together with the proof that their highest weights are $(\frac{1}{2}, \frac{1}{2} \cdots \frac{1}{2})$ and $(\frac{1}{2}\ \frac{1}{2}, \cdots\cdots\ \frac{1}{2} -\frac{1}{2})$ we must refer again to Chapter VIII of Boerner's book and the article of Braner and Weyl mentioned in the last lecture.

It now turns out, as one might expect, that the representations Δ^+, Δ^- together with the first $(\ell-2)D_r$ of (22.14) constitute the fundamental set of representations. It is, in fact, easy to show that any m satisfying the highest weight conditions (22.5) and $m_\ell = \frac{\text{integer}}{2}$ can be expressed in the form

$$m = k_1 m^{(1)} + k_2 m^{(2)} + \cdots + k_\ell m^{(\ell)} \tag{22.16}$$

where the k_i are non-negative integer, and from this it follows in the usual way that

$$\underbrace{\Delta^+ \times \cdots \times \Delta^+}_{k_1} \times \underbrace{\Delta^- \times \cdots \times \Delta^-}_{k_2} \times \underbrace{D_1 \times \cdots \times D_1}_{k_3} \times \cdots \times \underbrace{D_{\ell-2} \times \cdots \times D_{\ell-2}}_{k_\ell} \tag{22.17}$$

$$= D_m \dot{+} \text{ representations with lower highest weights.}$$

We now enumerated the representations of the groups A_ℓ, B_ℓ, C_ℓ and D_ℓ. For information concerning the group G_2 and more detailed information concerning A_2 and $B_2 = C_2$ I should like to refer you to the paper of Behrends et al. Rev. Mod. Phys. 34, (1962) 1. Here I shall simply say that the self-representation of G_2 is seven dimensional and that

$$(D_7 \times D_7)_{A.S} = D_7 \dot{+} D_{14} \tag{22.18}$$

where D_{14} is the adjoint representation. The two representations D_7 and D_{14} form the fundamental representations in this case.

Appendix A

We wish to sketch here the proof that for the Cartan canonical scheme.

$$
\begin{aligned}
[H_i, H_j] &= 0 , \\
[H_i, E_\alpha] &= \alpha_i E_\alpha , \\
[E_\alpha, E_\beta] &= N_{\alpha\beta} E_{\alpha+\beta}, \quad \text{if } \alpha+\beta \text{ is a non-zero root,} \\
&= \alpha^i H_i , \quad \text{if } \alpha+\beta = 0 , \\
&= 0 \quad \text{otherwise.}
\end{aligned} \tag{A.1}
$$

the $N_{\alpha\beta}$ are determined in terms of the α_i.

To prove this one needs the following properties of the $N_{\alpha\beta}$.

(a)
$$N_{\alpha\beta} = -N_{\beta\alpha} . \tag{A.2}$$

This property is obvious from (A.1).

(b)
$$N_{\alpha\beta} = N^*_{-\beta-\alpha} \tag{A.3}$$

This property is established as follows

$$N_{\alpha\beta} E_{\alpha+\beta} = [E_\alpha, E_\beta] = [E_\beta^\dagger, E_\alpha^\dagger]^\dagger = [E_{-\beta}, E_{-\alpha}]^\dagger$$

$$= \{N_{-\beta-\alpha} E_{-\beta-\alpha}\}^\dagger = N^*_{-\beta-\alpha} E^\dagger_{-\beta-\alpha} = N^*_{-\beta-\alpha} E_{\alpha+\beta} .$$

Note that we have used the choice

$$E_\alpha^\dagger = E_{-\alpha} ,$$

for the E_α in this proof.

(c)
$$N_{\alpha\beta} = N_{\beta\delta} , \quad \text{if} \quad \alpha+\beta+\delta = 0 . \tag{A.4}$$

This property is established in the following way: Let.

$$A = E_{\alpha} + E_{-\alpha-\beta} ,$$

$$B = E_{\beta} ,$$

where the E's are the matrices of the adjoint representation,

and let

$$C = [A, B] = N_{\alpha\beta} E_{\alpha+\beta} + N_{-\alpha-\beta,\beta} E_{-\alpha} .$$

From the identity

$$TR(AC) = TR\ A[A,B] = 0$$

we then have

$$N_{\alpha\beta} TR\ E_{\alpha} E_{\alpha+\beta} + N_{\alpha\beta} TR\ E_{-\alpha-\beta} E_{\alpha+\beta} + N_{-\alpha-\beta,\beta} TR\ E_{\alpha} E_{-\alpha}$$

or

$$+ N_{-\alpha-\beta,\beta} TR\ E_{-\alpha-\beta} E_{-\alpha} = 0 ,$$

$$N_{\alpha\beta}\, g_{\alpha\,\alpha+\beta} + N_{\alpha\beta}\, g_{-\alpha-\beta,\alpha+\beta} + N_{-\alpha-\beta,\beta}\, g_{\alpha,-\alpha} + N_{-\alpha-\beta\,\beta}\, g_{-\alpha-\beta,-\alpha} = 0$$

Using the explicit form of g given earlier for $g_{\alpha\beta}$ we obtain

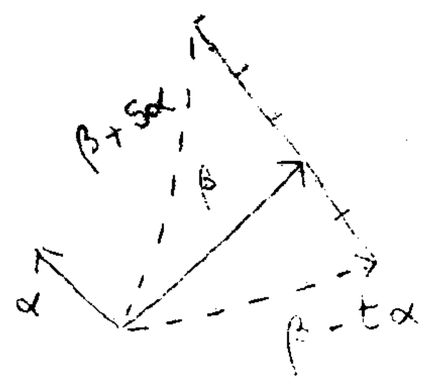
from this the relation,

$$0 + N_{\alpha\beta} 1 + N_{-\alpha-\beta,\beta} 1 + 0 = 0$$

as required. The relation (A.3) is called the <u>triangle</u> relation since the root vectors α, β, δ form a triangle.

(d) $$|N_{\alpha\beta}|^2 = \frac{\delta(l+1)}{2}(\alpha\cdot\alpha) \qquad (A.5)$$

where $\beta + S\alpha, \ldots \beta+\alpha, \beta, \beta-\alpha, \ldots \beta - t\alpha$ is a complete string. (fig. A.1).



Proof: Using β, α and $-\alpha$ one forms

the factor identity

$$\left[E_{-\alpha}, \left[E_{\alpha}, E_{\beta}\right]\right] + \left[E_{\alpha}, \left[E_{\beta}, E_{-\alpha}\right]\right] + \left[E_{\beta}, \left[E_{-\alpha}, E_{\alpha}\right]\right] = 0 ,$$

from which, we see at once that

$$N_{-\alpha,\alpha+\beta} N_{\alpha\beta} + N_{\alpha,\beta-\alpha} N_{\beta,-\alpha} + (\alpha \cdot \beta) = 0 ,$$

or

$$N_{\alpha\beta} N_{-\alpha,\alpha+\beta} = N_{\alpha,\beta-\alpha} N_{-\alpha,\beta} - (\alpha \cdot \beta) .$$

Using (A.2), (A.3) and (A.4), however, we have

$$N_{-\alpha,\alpha+\beta} = N^{*}_{\alpha\beta} ,$$

and

$$N_{-\alpha,\beta} = N^{*}_{\alpha,\beta-\alpha} .$$

Hence

$$|N_{\alpha\beta}|^2 = |N_{\alpha,\beta-\alpha}|^2 - (\alpha \cdot \beta)$$

This gives us a recurrence formula for the $|N_{\alpha\beta}|^2$ with the initial condition

$$|N_{\alpha,\beta-t\alpha}|^2 = -\alpha \cdot (\beta - t\alpha) ,$$

which is obtained by substituting $\beta - t\alpha$ for β in the above Jacobi identity. Solving these recurrence formulae, we have

$$|N_{\alpha,\beta-t\alpha+\alpha}|^2 = -\alpha \cdot (\beta - t\alpha) - \alpha \cdot (\beta - t\alpha + \alpha)$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$|N_{\alpha,\beta-\alpha}|^2 = -\alpha \cdot (\beta - t\alpha) - \cdots - \alpha \cdot (\beta - \alpha)$$

$$|N_{\alpha,\beta}|^2 = -\alpha \cdot (\beta - t\alpha) - \alpha \cdot (\beta - t\alpha + \alpha) - \cdots$$
$$- \alpha \cdot (\beta - \alpha) - \alpha \cdot \beta$$
$$= -(t+1)(\alpha \cdot \beta) + \frac{t(t+1)}{2} (\alpha \cdot \alpha)$$

But from the general theory of roots, we have the relation

$$(\alpha \cdot \beta) = \frac{(t-s)}{2}(\alpha \cdot \alpha) .$$

Hence

$$|N_{\alpha\beta}|^2 = -(t+1)\frac{(t-s)}{2}(\alpha \cdot \alpha) + \frac{t(t+1)}{2}(\alpha \cdot \alpha)$$
$$= s/2\,(t+1)(\alpha \cdot \alpha),$$

as required.

This result shows that the absolute values of the $N_{\alpha\beta}$ at any rate, are completely determined by the α's. It also shows that if $N_{\alpha\beta}$ is zero, s=o and $\alpha+\beta$ is not a root. This is the converse of the earlier result if $\alpha+\beta$ is not a root $N_{\alpha\beta}=0$, (A.1)

To show that the phases of the $N_{\alpha\beta}$ are also determined by the α_i one needs one more relation between the N's, namely

(e) $$N_{\alpha\delta}N_{\beta\gamma} + N_{\beta\delta}N_{\gamma\alpha} + N_{\gamma\delta}N_{\alpha\beta} = 0, \qquad (A.6)$$

if $$\alpha+\beta+\gamma+\delta = 0.$$

This relation is a consequence of the general facobi identity

$$\left[E_\alpha,\left[E_\beta, E_\gamma\right]\right] + \left[E_\beta,\left[E_\gamma, E_\alpha\right]\right] + \left[E_\gamma,\left[E_\alpha, E_\beta\right]\right] = 0,$$

where α, β, γ and δ are non-zero roots. From this relation, we see at once that

$$N_{\alpha,\beta+\gamma} N_{\beta\gamma} + N_{\beta,\gamma+\alpha} N_{\gamma,\alpha} + N_{\gamma,\alpha+\beta} N_{\alpha\beta} = 0$$

and on using the triangle relation this becomes

$$N_{\alpha,-(\beta+\gamma+\alpha)} N_{\beta\gamma} + N_{\beta,-(\gamma+\alpha+\beta)} N_{\gamma\alpha} + N_{\gamma,-(\alpha+\beta+\gamma)} N_{\alpha\beta} = 0$$

as required. The relation (A.6) is called the quadrangle relation since the vectors $\alpha, \beta, \gamma, \delta$ form a quadrangle. Note that (A.6) is cyclic in α, β and γ, whereas δ occurs in the same place in each term.

The point now is that using the relations a, b, c, d, and e, one can show that the phases of the $N_{\alpha\beta}$ are determined up to the following type of transformation.

$$N_{\alpha\beta} \rightarrow N'_{\alpha\beta} = N_{\alpha\beta} e^{i\delta(\alpha) + i\delta(\beta) - i\delta(\alpha+\beta)},$$

where the $\delta(\alpha)$ are arbitrary except for the condition $\delta(\alpha) = -\delta(-\alpha)$, note that there is just one phase $\delta(\alpha)$ for each root.

But from the definition

$$[E_\alpha, E_\beta] = N_{\alpha\beta} E_{\alpha+\beta}$$

of the $N_{\alpha\beta}$, it is clear that by a suitable redefinition of the E_α the phases $\delta(\alpha)$ can be absorbed, so that, in effect, the $N_{\alpha\beta}$ are completely determined.

I shall not give the proof that the relations (a) to (e) determine the $N_{\alpha\beta}$ up to the phase transformation, mentioned, as it is rather long and tedious. Anybody who is interested will

find the full proof given in Weyl's 1941 Princeton Lecture Notes

Q: Is the redefinition of the E_α compatible with the convention

$$E_\alpha^\dagger = E_{-\alpha} ,$$

which you have and earlier?

A: Yes. In fact the redefinition is simply

$$E_\alpha \to E'_\alpha = e^{-i\delta(\alpha)} E_\alpha ,$$

and then

$$E'^\dagger_\alpha = e^{i\delta(\alpha)} E^\dagger_\alpha = e^{i\delta(\alpha)} E_{-\alpha} = e^{-i\delta(-\alpha)} E_{-\alpha} = E'_{-\alpha} .$$

as base elements for this set of matrices, we choose

$$H_i = \begin{pmatrix} h_i & 0 \\ 0 & -\tilde{h}_i \end{pmatrix}, \quad E_i = \begin{pmatrix} 0 & h_i \\ 0 & 0 \end{pmatrix}, \quad E_{mn} = \begin{pmatrix} e_{mn} & 0 \\ 0 & -\tilde{e}_{mn} \end{pmatrix},$$

$$\mathcal{E}_{mn} = \mathcal{E}_{nm} = \begin{pmatrix} 0 & e_{mn}+e_{nm} \\ 0 & 0 \end{pmatrix} \quad E_i^\dagger, E_{mn}^\dagger \text{ and } \mathcal{E}_{mn}^\dagger \text{ etc} \quad (A.7)$$

where

$$i, m, n = 1, \ldots, l \qquad m < n .$$

$$h_1 = \tilde{h}_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad h_2 = \tilde{h}_2 = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \text{ etc} \quad (A.8)$$

$$i.e \ (h_i)_{rs} = (\tilde{h}_i)_{rs} = \delta_{ri}\delta_{is} ,$$

$$e_{12} = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad e_{13} = \begin{pmatrix} 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \text{ etc.} \quad i.e \ (e_{mn})_{rs} = \delta_{rm}\delta_{ns} \quad (A.9)$$

It is easy to see that for these matrices,

$$[H_i, E_j] = \begin{pmatrix} 0 & [h_i, h_j]_+ \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2\delta_{ij}h_j \\ 0 & 0 \end{pmatrix} \quad (A.10)$$

$$= 2\delta_{ij} E_j ,$$

$$[H_i, E_{mn}] = \begin{pmatrix} [h, e_{mn}] & 0 \\ 0 & -[h, e_{mn}]^{\vee} \end{pmatrix} = \begin{pmatrix} (\delta_{im} - \delta_{in}) e_{mn} & 0 \\ 0 & -(\delta_{im} - \delta_{in}) e^{\vee}_{mn} \end{pmatrix} \tag{A.11}$$

$$[H_i, \mathcal{E}_{mn}] = \begin{pmatrix} 0 & [h_i, e_{mn} + e_{nm}]_+ \\ 0 & \end{pmatrix} = \begin{pmatrix} 0 & (\delta_{im} + \delta_{in})(e_{mn} + e_{nm}) \\ 0 & 0 \end{pmatrix}$$

$$= (\delta_{im} + \delta_{in})\, \mathcal{E}_{mn}, \tag{A.12}$$

$$[H, E_\gamma] = \alpha e_\gamma E_\gamma \tag{A.13}$$

$$[H, E_{mn}] = (e_m - e_n) E_{mn}, \tag{A.14}$$

$$[H, \mathcal{E}_{mn}] = (e_m + e_n)\, \mathcal{E}_{mn} \tag{A.15}$$

Combining these commutation relations with the corresponding ones for $E^{\dagger}_{\gamma}$, $E^{\dagger}_{mn}$ and $\mathcal{E}^{\dagger}_{mn}$ we see that this group has the same roots as the group G_{χ} and is therefore isomorphic to it.

Appendix B

In this appendix we should like to sketch the proof that if $\det g_{\mu\nu} = 0$ the corresponding Lie group is not semi-simple i.e. it contains at least one invariant abelian sub-group.

The first step in the proof is to associate each element $X = a_\lambda X_\lambda$ of the Lie algebra with the corresponding element $L = a_\lambda c_\lambda^{\ \mu}$ of the adjoint representation, and then to consider for any group the set of elements $\hat{X}$ such that, for the corresponding $\hat{L}$,

$$TR\ \hat{L} L_\lambda = 0 \qquad (B.1)$$

where $L_\lambda = c_{\lambda\nu}^{\ \mu}$. This set of elements will form an <u>invariant</u> subalgebra of the Lie algebra.

Proof: Let L be any element whatsoever of the adjoint representation and let

$$\left[\hat{L}, L\right] = \tilde{L} \ . \qquad (B.2)$$

What we have to show is that $\tilde{L}$ satisfies (B.1). This is shown as follows:

$$\begin{aligned} TR\ \tilde{L} L_\lambda &= TR \left[\hat{L}, L\right] L_\lambda \\ &= TR \left(\hat{L} L L_\lambda - L \hat{L} L_\lambda\right) \\ &= TR \left(L L_\lambda \hat{L} - L \hat{L} L_\lambda\right) \\ &= TR\ L \left(L_\lambda, \hat{L}\right) \\ &= 0 \end{aligned} \qquad (B.3)$$

Q E D.

The invariant subalgebra $\hat{X}$ can be defined in this way for any Lie algebra. What characterzies the Lie algebras for which

$\det g_{\mu\nu} = 0$, however, is that for these, and only these, the invariant subalgebra so defined is non-trivial i.e. it contains at least one element other than the zero element.

Proof: (a) If $\det g_{\mu\nu} = 0$, there exists a non-zero vector b_r such that

$$g_{\mu\nu} b_\nu = 0 . \tag{B.4}$$

If we now let

$$L = b_\nu c^{\mu}_{\nu\lambda} , \tag{B.5}$$

we have

$$TR\ L L_\lambda = b_\nu c^{\mu}_{\nu\sigma} c^{\sigma}_{\lambda\mu} = g_{\nu\mu} b_\nu = 0 . \tag{B.6}$$

Thus the non-zero element $X = b_\nu X_\nu$ belongs to the invariant subalgebra.

(b) If $X = b_\nu X_\nu$ is a non-zero element belonging to the subalgebra,

$$TR\ L L_\sigma = b_\nu c^{\lambda}_{\nu\tau} c^{\tau}_{\sigma\mu} = b_\nu g_{\nu\sigma} = 0 ,$$

and so

$$\det g_{\mu\nu} = 0 \qquad \text{Q.E.D.} \tag{B.7}$$

Thus if det $g_{\mu\nu} = 0$, the Lie algebra contains a non-trivial invariant subalgebra. The question is: is this also an abelian subalgebra, since, if it is, the Lie algebra contains a non-trivial invariant abelian subalgebra and the proof that it is non-semi-simple is complete.

However, in general, the invariant subalgebra defined by (B.1) is not abelian, and to find an <u>abelian</u> invariant subalgebra we must refine the procedure. The refinement one makes is the following: One takes the invariant subalgebra defined by (B.1), and forms from it the succession of <u>derivative</u> subalgebras $\hat{X}^{(1)}, \hat{X}^{(2)}, \ldots\ldots \hat{X}^{(n)}$ as follows:

$\hat{X}^{(1)}$ consists of all elements of the form $[X_\alpha, X_\beta]$ where $X_\alpha X_\beta \in \hat{X}$,

$\hat{X}^{(2)}$ " " " " " " $[X_\alpha, X_\beta]$, $X_\alpha, X_\beta \in \hat{X}^{(1)}$,

and, in general,

$\hat{X}^{(n+1)}$ " " " " $[X_\alpha, X_\beta]$ " $X_\alpha, X_\beta \in \hat{X}^{(n)}$

It is clear that at each step the derived subalgebra is contained in the preceding one i.e. $\hat{X} \supseteq \hat{X}^{(1)} \supseteq \hat{X}^{(2)} \ldots\ldots \supseteq \hat{X}^{(n)} \supseteq \hat{X}^{(n+1)} \ldots$

Further more, it is easy to see that, by definition, each derived subalgebra is an <u>invariant</u> subalgebra.

Proof: if $X \in \hat{X}^{(n+1)}$ and $Y \in \hat{X}^{(n)}$ then, a fortiori

$X \in \hat{X}^{(n)}$ and

$$[X, Y] \in \hat{X}^{(n+1)},$$

by definition.

The point now is that one can prove that there exists, a finite integer N such that

$$\hat{X}^{(N)} \neq 0 , \quad \hat{X}^{(N+1)} = 0 . \tag{B.8}$$

But this means that $\hat{X}^{(N)}$ is abelian. Thus $\hat{X}^{(N)}$ is a non-trivial abelian invariant subalgebra of the original Lie algebra. Hence the original Lie group is not semi-simple.Q.E.D.

The difficulty is that the proof of the existence of a finite N as in (B.8) is by no means easy. Note that the existence of such an N is equivalent, to the fact that the orders of the derivative algebras $\hat{X}^{(1)}$, $\hat{X}^{(2)}$ .... form a decreasing sequence which terminates with zero i.e. instead of $\hat{X} \supseteq \hat{X}^{(1)} \supseteq \ldots \supseteq \hat{X}^{(n)} \supseteq \ldots$ we have $\hat{X} \supset \hat{X}^{(1)} \supset \hat{X}^{(2)} \supset \ldots \supset \hat{X}^{(n)} \supset \hat{X}^{n+1} = 0$

Lie algebra the orders of whose derivatives subgroups form a decreasing sequence which ends only at zero is called a <u>solvable</u> Lie algebra. Thus what one has to show to establish (B.8) is that the algebra $\hat{X}$ is solvable,

The information we are given is that if X_α are the base elements of $\hat{X}$

$$g_{\alpha\beta} = TR\, L_\alpha L_\beta = 0 \qquad (B.9)$$

Since this is a consequence of the definition (B.1) of $\hat{X}$. Thus what one has to prove is the following theorem (due to Cartan): A Lie algebra for which the metric tensor is zero, is solvable.

We shall not give the proof of this theorem here, as it would take us too far into the theory of solvable and nilpotent Lie algebras, but anybody who is interested can find the details either in Weyl's Princeton Lecture Notes or Jacobson's "Lie algebras".

PART III[+]

CONSTRUCTION OF THE INVARIANTS FOR THE SIMPLE LIE GROUPS

1. Introduction:

Recent speculations concerning the existence of 'higher symmetries' in elementary particle physics has renewed interest in the theory of representations of Lie groups. There are, in particular, three problems which are of great interest to physicists, and which have not been solved (explicitly at any rate) in the classical researches of Cartan and Weyl. They are:

(a) The problem of constructing all the independent invariants for any given Lie group.

(b) The problem of constructing a maximal abelian set for any given Lie algebra ('state labelling' problem).

(c) The problem of finding the (generalized Clebsh-Gordon) coefficients in the reduction of the direct product of any two representations of a given Lie group.

Further, one would like to be able to solve these problems in an infinitesimal way. This is because it is the infinitesimal (complex) Lie algebra, and not the Lie group itself, which is of importance for the physicist.

Let us now try the first of these three. For this problem a complete solution will be constructed in an infinitesimal way. In a sense the solution will not be new since the final result has already been given by Racah[1]. However, the present paper complements those of Racah (to which, we need hardly say, it is very much indebted) in the following ways: (1) Racah's proof that the invariants which he finds a complete set depends on a certain S-theorem (stated in 5). So far, this theorem has not been proved in an infinitesimal way, and is the only link in Racah's chain of argument which is not infinitesimal. The present paper gives an infinitesimal

---

[+] This part has been included in the lecture notes although, in fact, time did not permit it to be delivered as a set of lectures. Actually this is the content of the paper of B.Gruber and L.O'Raifeartaigh (to be published).

proof of this theorem.

(2) We show that the type of invariant considered by Racah is a very natural further generalization of the generalized Casimir operators.

(3) We give an alternative form for the Racah invariants, which allows one to write them down explicitly in a rather simple way. We also indicate the connexion between these invariants and the ones found by Okubo[2] and Biedenharn[3] for Su_3 and Su_n respectively.

(4) We show how one can understand intuitively and directly, why all of the Racah type invariants drop out except those listed in the final independent set. (In particular, we show how one can understand why the exceptional l^{th} order invariant constructed with a spinor representation of D_l, does not drop out).

(5) Finally, we include here in one paper the complete solution to the problem. Racah's original solution is contained in two papers[1),4)] both of which are relatively inaccessible, and each of which is incomplete without the other.

We shall split this ... nto two separate parts. Part B II, consisting of a preliminary section § 7 and a main section § 8, will be devoted to the infinitesimal proof of the S-theorem mentioned above under (1). In part A the programme will be as follows:

In § 2, we shall show how the Racah type invariants arise as a very natural generalization of the Casimir invariants.

In § 3, we shall prove their invariance.

In § 4, we shall list, without proof, the independent invariants for each particular kind of Lie group. We shall also derive a rather simple method for writing them down explicitly and establish the connexion between these invariants and those of Okubo[2] and Biedenhakn[3].

In § 5, we shall use the S-theorem to prove the completeness of the sets of invariants found in § 4.

In § 6, we shall give a somewhat hemristic discussion whose aim will be to make it intuitively evident why the invariants listed in § 4, should be the independent ones.

In conclusion, we should like to emphasize again the fact that the invariants listed in § 4, were originally discovered by Racah[1]. Furthermore, our use of the S-theorem in 5, to prove the completness of the set listed, is merely a more detailed exposition of the some argument used by Racah[1),4)]*.

PART I

2. General form of solution:-

We start with the Casimir operator

$$C_2 = g_{\alpha\beta} X^\alpha X^\beta = C^{\sigma 2}_{\sigma_1 \alpha} C^{\sigma 1}_{\sigma 2 \beta} X^\alpha X^\beta, \qquad (2.1)$$

which is an invariant of any Lie group. This can be generalized immediately to give a set of quantities

$$C_n = g_{\alpha.....\gamma} X^\alpha X^\beta .....X^\gamma = C^{\sigma n}_{\sigma_1 \alpha} C^{\sigma 1}_{\sigma_2 \beta} C^{\sigma n-1}_{\sigma n \gamma} X^\alpha X^\beta ....X^\gamma, \qquad (2.2)$$

which, it is easy to check, are also invariants for any Lie group. However, it is well-known that the set (2.2) does not include all of the invariants, and so must be generalized further. To generalize

it, the decisive point is to note that for the adjoint representation of the group $\overset{A}{X}_\alpha$, we have

$$\overset{A}{X}_\alpha = C^{\sigma_2}_{\sigma_1 \alpha} \quad , \tag{2.3}$$

where the σ_1 and σ_2 are to be regarded as matrix indices.

Thus (2.2) can be written in the alternative form

$$C_n = \left(Sp \ \overset{A}{X}_\alpha \overset{A}{X}_\beta \ldots\ldots \overset{A}{X}_\gamma \right) \ X^\alpha X^\beta \ldots\ldots X^\gamma \tag{2.4}$$

The generalization then consists in replacing the adjoint representation $\overset{A}{X}_\alpha$ in (2.4) by any representation $\hat{X}_\alpha$ i.e. forming the quantities

$$I_n = \left(Sp \ \hat{X}_\alpha \ \hat{X}_\beta \ldots\ldots \hat{X}_\gamma \right) \ X^\alpha X^\beta \ldots\ldots X^\gamma \tag{2.5}$$

These quantities are, in fact, the general invariants we are looking for, and we see in the above wayhow they arise as a natural generalization of the Casimir operators. But, of course, we still have to show that the I_n are invariants and that they include all the invariants (or at least all the independent invariants) of the group. Also since we have a double infinity of possibillities in (2.5) corresponding to taking any irreducible representation $\hat{X}_\alpha$, on the one hand, and the invariants of any order constructed with the given representation on the other, we must find some way of picking out the independent invariants. (The number of independent-ones turns out to be finite).

3. Proof of invariance:

In this section we shall prove the invariance of the I_n of (2.5)

Proof: Let

$$u = 1 + \varepsilon_\lambda X_\lambda \,, \tag{3.1}$$

where ε_λ, $\lambda = 1, \ldots \ldots r$, are arbitrary small numbers. Then

$$\begin{aligned} u\, X_\alpha u^{-1} &= X_\alpha + \varepsilon_\lambda [X_\lambda, X_\alpha] \\ &= X_\alpha + \varepsilon_\lambda c^\sigma_{\lambda\alpha} X_\sigma \\ &= \left(\delta^\sigma_\alpha + \varepsilon_\lambda c^\sigma_{\lambda\alpha}\right) X_\sigma \\ &= a^\sigma_\alpha(\varepsilon) X_\sigma \end{aligned} \tag{3.2}$$

Thus, under a ' u -transformation', the X_α transform as vectors, with transformation coefficients $a^\sigma_\alpha(\varepsilon)$. Ofcourse, the a^σ_α are not arbitrary but are determined by the nature of the structure constants for any given set of ε_λ.

From (2.5), we then have,

$$\begin{aligned} u\, \hat{I}_n u^{-1} &= \left(Sp\, \hat{x}_\alpha \hat{x}_\beta \ldots \hat{x}_\gamma\right) u\, (x^\alpha x^\beta \ldots x^\gamma) u^{-1} \\ &= \left(Sp\, \hat{x}_\alpha \hat{x}_\beta \ldots \hat{x}_\gamma\right) a^\alpha_\sigma a^\beta_\lambda \ldots a^\gamma_\mu\, x^\sigma x^\lambda \ldots x^\mu \\ &= Sp\, \left(a^\alpha_\sigma \hat{x}_\alpha\, a^\beta_\lambda \hat{x}_\beta \ldots a^\gamma_\mu \hat{x}_\gamma\right) x^\sigma x^\lambda \ldots x^\mu \\ &= Sp\, \left(u \hat{x}_\sigma u^{-1} . u \hat{x}_\lambda u^{-1} \ldots u x_\mu u^{-1}\right) x^\sigma x^\lambda \ldots x^\mu \\ &= Sp\, \left(\hat{x}_\sigma \hat{x}_\lambda \ldots \hat{x}_\mu\right) x^\sigma x^\lambda \ldots x^\mu \\ &= \hat{I}_n \end{aligned} \tag{3.3}$$

Thus,

$$\left[\hat{I}_n,\ u\right] = 0, \tag{3.4}$$

and since the ε_λ are arbitrary,

$$\left[\hat{I}_n\ ,\ X_\lambda\right] = 0. \qquad \text{Q.E.D.} \tag{3.5}$$

4. The independent invariance

Having established the invariance of the I_n of (2.5), our next tasks are (a) to pick out the independent ones and (b) to show that these constitute a complete set of invariants. In this section we shall simply list, without proof, the independent invariants. In the next section we shall prove that the ones listed form a complete set of independent invariants and in the following section we shall give a heuristic discussion whose aim will be to explain why these particular invariants should be picked out.

The independent invariants are:-

(1) The general groups.

For the group A_ℓ : the invariants of order 2,3,4.... 1+1, formed with the 'self-representation'

For the group B_ℓ : the invariants of order 2,4,6.... 2l, formed with the 'Self-representation'

For the group C_ℓ : the invariants of order 2,4,6.... 2l, formed with the 'self-representation'

For the group D_ℓ : the invariants of order 2,4,6.... 2l-2 formed with the 'self-representation', plus the invariants of order l, formed with the either one of the two fundamental spinor representations, Δ^+ or Δ^-

(2) The special groups.

For the group G_2; the invariants of order 2,6 formed with the 'self-representation'. For the groups F_4 etc. see appendix.

Here by the 'self-representation' we mean the lowest order faithful representation. If the group is defined as a matrix group, and not in abstracto, then this representation is just the group itself. Hence the name 'self-epresentation'.

We see therefore that with the exception of D_ℓ , the invariants can be built using the self-representation alone. Thus, the self-representation replaces the adjoint representation (which is used for the Casimir invariants) as the tool for building the invariants.

It is not yet clear, of course, why the invariants listed should be independent ones. The appearance of the representation Δ^+ or Δ^- for example, is something of a mystery as yet. Its appearance raises other questions too: 1) Why does the spinor representation of B_ℓ not appear? 2) why do we have only Δ^+ or Δ^- , but not both? 3) why is it necessary to form only <u>one</u> invariant with Δ^+ or Δ^- ?. It is to answer question(5) such as these that the hemristic section, § 6, has been included. But first, of course, one must prove that the invariants we have listed above form a complete independent set, and this we do in the next section, § 5.

Before, going onto that proof, however, we should like to complete this section by giving an alternative form to (2.5) for the invariants, and particularly to the independent invariants we have listed. The new form will be of use in § 6, but its prime purpose is, to enable us to write down the invariants for any given group in a rather simple way.

The new expressions for the invariants are obtained as follows: We have

$$\hat{I}_n = \left(Sp\ \hat{X}_\alpha \hat{X}_\beta \cdots \hat{X}_\gamma\right) x^\alpha x^\beta \cdots x^\gamma$$

$$= \hat{Sp}\left(\hat{X}_\alpha \hat{X}_\beta \cdots \hat{X}_\gamma \times x^\alpha x^\beta \cdots x^\gamma\right) \tag{4.1}$$

where $\hat{Sp}$ means that the spur is to be taken with respect to the $\hat{X}$-space only.

$$= \hat{Sp}\left(\hat{X}_\alpha \times x^\alpha\right)\left(\hat{X}_\beta \times x^\beta\right) \cdots \left(\hat{X}_\gamma \times x^\gamma\right)$$

$$= \hat{Sp}\left(\hat{A}\right)^n$$

where

$$\hat{A}_\alpha = \left(\hat{X}_\alpha \times x^\alpha\right). \tag{4.2}$$

Equations (4.1) and (4.2) give the required new expressions for the $\hat{I}_n$. To illustrate what the new form means explicitly, let us consider the group O_3 and, let $\hat{X}_\alpha$ be the self-representation, i.e.

$$\hat{X}_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \hat{X}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \hat{X}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}. \tag{4.3}$$

Then

$$\hat{A} = \begin{pmatrix} 0 & X_1 & X_2 \\ -X_1 & 0 & X_3 \\ -X_2 & -X_3 & 0 \end{pmatrix} \tag{4.4}$$

and

$$\hat{I}_2 = \hat{Sp}\left(\hat{A}\right)^2$$

$$= Sp \begin{pmatrix} -x_1^2 - x_2^2 & 0 & 0 \\ 0 & -x_1^2 - x_3^2 & 0 \\ 0 & 0 & -x_2^2 - x_3^2 \end{pmatrix} \tag{4.5}$$

$$= -2\left(x_1^2 + x_2^2 + x_3^2\right) = -2x^2.$$

Thus the idea is to let the (matrices!) X be matrix elements in the matrix $\hat{A}$. The invariants are then the spurs of the powers of $\hat{A}$.

Let us now apply this technique to the independent invariants constructed with the self-representations of A_ℓ, B_ℓ, C_ℓ and D_ℓ. In fact, as will be seen, the technique is dovetailed to suit exactly these self-representations.

A_ℓ; We ignore the unimodular restriction which is irrelevant here. Then if $|i\rangle$ is a basis in the self-representation space we can choose as our $\hat{X}_\alpha = \overset{s}{X}_\alpha$ the matrices $|i\rangle\langle j|$. Note that we thereby switch from the single index (α) to the double index (ij) notation.

Then we have

$$\hat{A} = |i\rangle\langle j| X_{ij}, \tag{4.6}$$

where for the $X_\alpha = X_{ij}$ also we switch to the double index notation, Hence

$$\hat{A} = X, \tag{4.7}$$

where X is simply the matrix with the (matrices!) X_{ij} as elements. Hence

$$\hat{I}_n = TR\, X^n = X_{ij} X_{jk} \cdots\cdots X_{ni}. \tag{4.8}$$

This gives the $\hat{I}_n$, in a very simple explicit form. This is the form used by Okubo[2] for Su_3 C A_2.

B_ℓ and D_ℓ: In an exactly similar way

$$\hat{A} = X \tag{4.9}$$

$$\hat{I}_n = TR\, X^n, \tag{4.10}$$

where X is the anti-symmetric matrix with the (matrices) X_{ij} $=-X_{ji}$ as elements. The special case of O_3 C B_l has been treated explicitly above.

C_l : In this case we can

$$\left(\begin{array}{c|c} R_{ik} & \Sigma_{ik} \\ \hline \Lambda_{ik} & -R_{ki} \end{array}\right) ; \hat{I}_n = \hat{S}\hat{p}(\hat{A})^n \tag{4.11}$$

(4.12)

where the matrices Σ_{ik} and Λ_{ik} satisfy the symmetric conditions,

$$\Sigma_{ik} = \Sigma_{ki} \ , \ \Lambda_{ik} = \Lambda_{ki}, \ i,k = 1,..... l \ . \tag{4.13}$$

these being the conditions which are satisfied by the matrices of the self-representation of the sympletic group.

These formulae can, in a sense, be simplified even further, as follows: If one chooses the variables X_α according to Cartan's canonical scheme[4], the first l being called H_i, $i = 1,..... l$, and the remainder being called $E_{\pm\alpha}$ where α are the positive roots, then for most purposes (for reasons which will be discussed at length in the next section) one is interested not so much in the invariant itself as in the highest power terms in the H_i occurring in it. Now these terms are extremely easy to calculate for the invariants listed above since for the self-representation the H_i's can be chosen to be diagonal. unImodfadt; we have:

A_l (without unimodular restriction):

$$\hat{A}' = \begin{pmatrix} X_{11} & 0 & \cdots & 0 \\ 0 & X_{22} & & 0 \\ 0 & 0 & X_{33}\cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots & X_{\ell+1\,\ell+1} \end{pmatrix} = \begin{pmatrix} H_1 & \cdots & 0 \\ 0 & H_2 & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & H_{\ell-1} \end{pmatrix} \quad (4.14)$$

$$\hat{I}'_n = \hat{S}p\,(\hat{A}')^n = \sum_i (H_i)^n \quad (4.15)$$

where the dash means that we neglect the E_α. Hence

$$\hat{I}_n = \sum_i (H_i)^n + \ldots\ldots\ldots\ldots, \quad (4.16)$$

in this case.

B_ℓ and D_ℓ : By a suitable transformation we can chose

$$\hat{A}' = \begin{pmatrix} X_{11} & & & 0 \\ & \ddots & & \\ & & -X_{\ell\ell} & \\ & & & -X_{11} \\ 0 & & & & -X_{\ell\ell} \end{pmatrix} = \begin{pmatrix} H_1 & & & 0 \\ & \ddots & & \\ & & H_\ell & \\ & & & -H_1 \\ 0 & & & & -H_\ell \end{pmatrix} \quad (4.17)$$

and so

$$\hat{I}_{2n} = \sum_i (H_i)^{2n} + \ldots\ldots$$
$$\hat{I}_{2n+1} = 0 + \ldots\ldots\ldots \quad (4.18)$$

in this case. (Later, we shall show that, in fact, $\hat{I}_{2n+1} = 0$).

C_ℓ : In a similar way

$$\hat{A}' = \begin{pmatrix} H_1 & & & 0 \\ & \ddots & & \\ & & H_\ell & \\ & & & -H_1 \\ 0 & & & & -H_\ell \end{pmatrix} \quad (4.19)$$

$$\hat{I}_{2n} = \sum_i (H_i)^{2n} + \cdots\cdots\cdots ,$$

$$\hat{I}_{2n+1} = 0 \qquad (4.20)$$

In this way we see how the invariants constructed with the self-representation ean be expressed very easily using the form (4.1), (4.2). One can do the same sort of thing for the ℓ^{th} order invariant constructed with the Δ^+ or Δ^- for D_ℓ but it is easier to write this invariant down explicitly once and for all, (or at least its leading terms in H). As os shown in the appendix A, the leading term for this invariant is just

$$H_1 H_2 \cdots\cdots H_\ell \ . \qquad (4.21)$$

Finally, we mention that we can treat the special groups G_2, F_4, etc. in a similar way. For example, for G_2 we have[5]

$$\hat{A}' = \begin{pmatrix} \frac{2}{\sqrt{3}} H_1 & & & & & & \\ & H_2 + \frac{1}{\sqrt{3}} H_1 & & & & 0 & \\ & & H_2 - \frac{1}{\sqrt{3}} H_1 & & & & \\ & & & 0 & & & \\ & & & & -H_2 + \frac{1}{\sqrt{3}} H_1 & & \\ & 0 & & & & -H_2 - \frac{1}{\sqrt{3}} H_1 & \\ & & & & & & \frac{-2}{\sqrt{3}} H_1 \end{pmatrix} \qquad (4.22)$$

so that

$$I'_{2n+1} = 0 + \ldots\ldots$$

$$\overset{S}{I}_2 = 2(H_1^2 + H_2^2) + \ldots\ldots \qquad (4.23)$$

$$\overset{S}{I}_4 = 2(H_1^2 + H_2^2)^2 + \ldots\ldots$$

$$\overset{S}{I}_6 = 2H_2^6 + 10\, H_2^4\, H_1^2 + \frac{10}{3}\, H_2^2\, H_1^4 + \frac{22}{9} H_1^6 + \ldots\ldots$$

etc.

In conclusion, we should like to mention the connexion between the above forms of the invariants and the forms found by Biedenhakn[3] for $Su_n \subset A_{n-1}$. What Biedenhakn used was the fact that for U_n (i.e. Su_n without the unimldular restriction), one has for the self-representation X_α not merely

$$\left[\overset{S}{X}_\alpha , \overset{S}{X}_\beta\right] = C_{\alpha\beta}^{\gamma}\, \overset{S}{X}_\gamma \,, \qquad (4.24)$$

but even

$$\left[\overset{S}{X}_\alpha , \overset{S}{X}_\beta\right]_+ = d_{\alpha\beta}^{\gamma}\, \overset{S}{X}_\gamma \,, \qquad (4.25)$$

or

$$\overset{S}{X}_\alpha \overset{S}{X}_\beta = d^{(1)\gamma}_{\alpha\beta}\, \overset{S}{X}_\gamma = \frac{1}{2}\left(d_{\alpha\beta}^{\gamma} - C_{\alpha\beta}^{\gamma}\right) \overset{S}{X}_\gamma \qquad (4.26)$$

This is due to the fact that U_n is so general that the $\overset{S}{X}_\alpha$ 'fill out' their representation space (see (4.6) above). Of course, the unit matrix is included on the right hand side of (4.25) and (4.26). Biedenhakn exploited this fact to obtain the invariants in the form

$$I_n^B = d^{\varepsilon}_{\alpha\beta}\, d^{\mu}_{\varepsilon\lambda} \cdots d^{\lambda'}_{\mu\gamma}\, X^{\alpha} X^{\beta} \cdots X^{\gamma} \qquad (4.27)$$

To obtain the connexion between these invariants and the ones constructed above, we write the latter in the old form

$$\hat{I}_n = (\mathrm{Sp}\ X^S_{\alpha} \cdots X^S) X^{\alpha} \cdots X^{\gamma} \qquad (4.28)$$

and use (4.26) to reduce this by successive reductions to

$$\begin{aligned} \hat{I}_n &= d^{(1)\,\varepsilon}_{\alpha\beta}\ (\mathrm{Sp}\ X^S_{\varepsilon} \cdots X^S_{\gamma})\, X^{\alpha} \cdots X^{\gamma} \\ &= d^{(1)\varepsilon}_{\alpha\beta}\, d^{(1)\mu}_{\varepsilon\lambda} \cdots d^{(1)\mu'}_{\lambda'\gamma}\, X^{\alpha} \cdots X^{\gamma} \end{aligned} \qquad (4.29)$$

But this is the same as (4.27) with d replaced by $d^{(1)}$. Since the replacement of d by $d^{(1)}$ merely changes the invariants by adding lower order invariants to it (on account of the anti-symmetry of the $C^{\gamma}_{\alpha\beta} = d^{\gamma}_{\alpha\beta} + 2\, d^{(1)\gamma}_{\alpha\beta}$) it is clear that the $\hat{I}_n$ and I_n^B are, in fact, essentially the same.

In particular, for n = 3, we have, on multiplying (4.25) by X^S_{δ}, and taking the trace of both sides,

$$d_{\alpha\beta\delta} = \mathrm{TR}\ X^S_{\alpha} X^S_{\beta} X^S_{\delta} + \mathrm{TR}\ X^S_{\beta} X^S_{\alpha} X^S_{\delta} \qquad (4.30)$$

This equation not only shows the equivalence between Biedenhakn's d's and our third order coefficients explicitly, but affords a very simple proof of the invariant tensor character of Biedenhakn's d's . In fact, using the u of §3 we have

$$T_+ \mid j > = 0, \tag{5.3}$$

and hence,

$$T^2 = \langle j \, T^2 j \rangle = j\,(j+1)- \tag{5.4}$$

If now in (5.4) we make the substitution

$$j' = (j + \tfrac{1}{2}), \tag{5.5}$$

we have

$$T^2 = j'^2 - \frac{1}{4} \tag{5.6}$$

We see that <u>T^2 is even in j' .</u> This statement is the content of the S-theorem, in this case.

The general theorem is now easy to understand. Let $F(X_\alpha)$ be any invariant and let us introduce for the X_α . Cartan's canonical H_i and $E_{\pm\alpha}$. Let us now write $F(X_\alpha)$ in normal form i.e. in such a way that in each term in $F(X_\alpha)$, the E_α are pulled to the right and the $E_{-\alpha}$ to the left. We then obtain an expression of the form

$$F\,(X_\alpha) = \Phi_0\,(H_i) + \sum_\alpha E_{-\alpha}\,\Phi_\alpha(H_i)\,E_\alpha + \sum_{\alpha,\beta,\gamma} E_{-\alpha}\Phi_{\alpha\beta}(H_i)\,E_\beta\,E_\gamma + \cdots\cdots \tag{5.7}$$

If now for any representation, $(j_1, j_2, \cdots\cdots j_\ell)$, is the highest weight (highest simultaneous eigenvalues of the H_i). and $|j_1, j_2 \cdots\cdots j_\ell\rangle$ the corresponding normalized eigenvector, we have for this representation in analogy to (5.3) and (5.4)

$$E_\alpha \mid j_1 \cdots\cdots j_\ell \rangle = 0, \text{ for all } \alpha, \tag{5.8}$$

and hence

$$F\,(X_\alpha) = \langle j_1 \cdots\cdots j_\ell \mid F\,(X_\alpha) \mid j_1 \cdots\cdots j_\ell \rangle = \Phi_0(j_i)\ . \tag{5.9}$$

Now let

$$j_i' = j_i + \frac{1}{2}\Sigma_i , \tag{5.10}$$

where the vector Σ_i is the sum of the positive roots for the α group in question, and let

$$\phi_o(j_i) = \phi_o^1(j_i'). \tag{5.11}$$

The theorem states that

$$\phi_o'(j_i') = \phi_o^1(S\, j_i'), \tag{5.12}$$

where S is the discrete Weyl group generated by the reflexions perpendicular to the roots.

We now return to the problem of proving the completeness of the sets of invariants listed in § 4. The proof will be based on the following property of the $F(X_\alpha)$ which again we first illustrate by using O_3.

For O_3, the T^2 can be expressed either in the usual form (5.1) or in the form

$$T^2 = \sum_j j(j+1)\, P_j \tag{5.13}$$

where P_j is the projection operator for the j^{th} representation. This follows from the fact that in any given irreducible representation. T^2 is a multiple of the unit matrix, with value $j(j+1)$, by (5.4). Similarly in the general case we can write $F(X_\alpha)$ in the form

$$F(X_\alpha) = \sum_{j_1,\ldots j_\ell} \phi_o(j_i)\, P_{j_1, j_2 \ldots\ldots j_\ell} \cdot , \tag{5.14}$$

where the summation is over all possible highest weights $(j_1, j_2 \ldots\ldots j_\ell)$, and $\phi_o(j_i)$ are the functions used in (5.7) and

(5.9). Thus each $F(X_\alpha)$ is specified by its $\phi_o(j_i)$, and conversely. We can therefore concentrate our attention on the $\phi_o(j_i)$. These quantities are functions of the ℓ j_i. Suppose now that we could find a set of $\phi_o(j_i)$, $\phi_o^k(j_i)$ say, such that conversely, the j_i could be expressed as functions of the ϕ_o^k . Then all other $\phi_o(j_i)$ could be expressed as functions of the ϕ_o^k since we would have the relation

$$\phi_o(j_i) = \phi_o(j_i; \phi_o^k) \qquad (5.15)$$

In that case, the set ϕ_o^k and hence their corresponding $F^{(k)}$ would be complete. Our problem is to prove that the invariants listed in § 4 correspond to such a complete set of ϕ_o^k .

First we note that since there are ℓ j's, to be determined for a group of rank ℓ , we should expect to have ℓ F's in our set $F^{(k)}$. And a quick glance at § 4 is enough to verify that for each group there are indeed just ℓ F's listed. But, of course, this does not guarantee that the corresponding $\phi_o^k(j_i)$ will determine the j_i uniquely, since when we set

$$\phi_o^k(j_i) = C^{(k)} \qquad (5.16)$$

where the $C^{(k)}$ are suitable constants, we do not know that the ℓ- equations (5.16) are independent. In the first place, we must verify that the ℓ- equations (5.16) are independent mathematical sense, i.e. that they are a set of ℓ algebraic equations for ℓ unknowns j_i which allow only a finite discrete set of solutions. For example, the two equations

$$j_1^2 + j_2^2 = c_1 , \tag{5.17a}$$

$$j_1^4 + j_2^4 = c_2 , \tag{5.17b}$$

are independent in this sense, the two equations

$$j_1^2 + j_2^2 = c_1 \tag{5.18a}$$

$$(j_1^2 + j_2^2)^2 = c_2 = c_1^2 \tag{5.18b}$$

are not. Now it is easy to verify that the ϕ_o^k belonging to the invariants of § 4 are independent in this sense. The argument is as follows: If there exists a relation of the kind (5.18) between the ϕ_o^k this relation will be independent of the values of the j_i. In particular, it will hold for very large values. Hence if we look only at the terms of highest order in j_i in the ϕ_o^k and verify that no relations of the kind (5.18) hold for the leading parts of the ϕ_o^k, that will be sufficient. But we happen to have exactly the terms we need written out explicitly in § 4, for, from the definition of $\phi_o(j)$ in (5.7), it is clear that the highest order j_i terms in the $\phi_o^k(j)$ are just the highest order H_i terms in the $F^{(k)}_{(X_\alpha)}$, with $H_i \to j_i$, and the latter are written out explicitly in (4.16), (4.18), (4.20), (4.21) and (4.23). From the latter equations it is easy to verify by inspection that the leading terms in the ℓ invariants listed as independent in § 4 are indeed independent in the above sense. It is interesting to note that for G_2 the leading term in I_4 is not independent of the leading term in I_2. This explains why I_2 and I_6 rather than I_2 and I_4

are listed in § 4 as the independent invariants for this group.

But the independence of the $\phi_0^k(j_\nu)$ in the above sense is still not sufficient to guarantee that the $\phi_0^k(j_\nu)$ will determine the j_ν uniquely. In fact, it is obvious, that they will not, since the fact that the $\phi_0^k(j_\nu)$ are polynomials is already enough to tell us that there will be a number of different discrete solutions of (5.16), the number (in general) being equal to the product of the powers of the polynomials $\phi_0^k(j_\nu)$. For example, for the set (5.17), either of the two solution of the quadratic equation (5.17a) for j_2 can be substituted into (5.16), and will lead to three solutions of the cubic equation obtained for j_1. Thus in general the system (5.3) will have six discrete solutions.

How then can the system (5.16) determine the highest weight uniquely? It is here that the S-theorem comes in. What the S-theorem gurantees is that of the finite number of discrete solutions of (5.7), only one can qualify for being a highest weight. To see this, let us regard the equations (5.16) as equations for the j'_ν defined in (5.10) rather than the j_ν. From the S-theorem we see that each equation is invariant under the transformation

$$j'_\nu \longrightarrow S\, j'_\nu \quad \text{for all } S. \tag{5.19}$$

Hence, if we have any solution j'_ν, $S j'_\nu$ is also a solution, all S. Thus the discrete solutions of (5.2) consist of sets of solutions $S j'^{(1)}_\nu$, $S j'^{(2)}_\nu$, ... , $S j'^{(k)}_\nu$. If

ν is the product of the power of the polynomials in (5.16), and h is the order of S then we must have

$$\nu = kh \tag{5.20}$$

Since both these numbers are equal to the number of solutions of (5.16). Suppose, however, that $k = 1$, or

$$\nu = h \tag{5.21}$$

Then there is only one set of solutions $S\,j'_r$. In this case, since it clear that only one vector of any single set $S\,j'_r$ can be the highest, only one vector out of the $\nu = h$ solutions can qualify for being a highest weight (or, more accurately, can qualify for determining a $j_i = j' - \frac{1}{2}\sum_i$ which is a highest weight). In other words, if $\nu = h$, (5.16) does not determine a unique j'_r, but it does determine a unique <u>set</u> $S\,j'_r$, only one of which corresponds to a highest weight and in this indirect way (5.16) determines the highest weight uniquely.

Therefore, to show that when we use in the equations (5.16) the invariants listed in § 4, they determine a highest weight uniquely and hence form a complete set, we have only to verify that for these invariants (5.21) holds, This is easily done. In fact

For A_ℓ ; $h = \nu = (\ell+1)!$,

For B_ℓ ; $h = \nu = 2^\ell\, \ell!$,

For C_ℓ ; $h = \nu = 2^\ell\, \ell!$,

For D_ℓ ; $h = \nu = 2^{(\ell-1)}\, \ell!$,

For G_2 ; $h = \nu = 12$

(for completeness, the values of $\lambda = \mu$ for F_4 are also given (appendix. C)

Our proof is now complete. In conclusion, it is interesting to note that if we left out the ℓ^{th} order spinor invariant for D_ℓ and included the self-representation invariant of order 2ℓ instead, we would obtain $\mu = 2\lambda$ and would therefore obtain from (5.16) two sets of $S j_r$ and hence two possible highest weights.

6. The independent I_n:

In this section we wish to give a somewhat hewistic explanation of how one comes to consider the independent invariants listed in § 4. The method we shall use can be made rigorous but we prefer to leave it as it is for two reasons: (a) the rigorous method is somewhat long and tends to obscure the ideas and (b) the proof of § 5 is already sufficient to show that the invariant of § 4 are, in fact, the only independent invariants.

The main ideas we shall use (at least for the next few paragraphs) is that if a representation $\hat{X}$ is "built up" out of other representation $\hat{X}_i$ by taking direct products and direct sums of the latter, than it is plausible that the invariants formed with $\hat{X}$ are functions of those formed with the $\hat{X}_i$.

This can be proved, but we shall simply accept it as a plausible assumption. Then the argument goes as follows:

(1) Since the reducible representations are built up out of the irreducible ones by taking direct sums of the latter, we should expect to need only the irreducible representations.

(2) Since any irreducible representations $\hat{D}$ can be built up out of the fuddamental representations D_{f_i} $i = 1, 2, \ldots\ldots \ell$ according to the formula

$\hat{D}$ + lower order irreducible representations =

$$D_{f_i} \times D_{f_j} \cdots \times D_{f_k} \qquad (6.1)$$

where on the right hand side any given D_{f_i} may occur more than once, we should expect to need only the ℓ fundamental representations.

We consider the fundamental representations in some more detail. Let $\overset{S}{D}$ be the self-representations of A_ℓ, B_ℓ, C_ℓ, or D_ℓ and let[6]

$\Gamma_1 = \overset{S}{D}$

$\Gamma_2 =$ anti-symmetric part of $\overset{S}{D} \times \overset{S}{D}$

$\Gamma_3 =$ totally anti-symmetric part of $\overset{S}{D} \times \overset{S}{D} \times \overset{S}{D}$

$\Gamma_\ell =$ totally anti-symmetric part of $\overset{S}{D} \times \overset{S}{D} \times \ldots\ldots \times \overset{S}{D}$ ($\leftarrow \ell$ terms $\rightarrow$)

Then[6] [7]

For A_ℓ : the Γ_ℓ are just the ℓ fundamental representations.

For C_ℓ : the Γ_ℓ are just the ℓ fundamental representations.

For B_ℓ : $\Gamma_1, \Gamma_2, \ldots, \Gamma_{\ell-1}$ are $(\ell-1)$ fundamental representations. There exists independently a fundamental (Spinor) representation $\triangle$. Γ_ℓ is irreducible, but not fundamental.

For D_ℓ : $\Gamma_1, \Gamma_2, \ldots \Gamma_{\ell-2}$ are $(\ell-2)$ fundamental representations. There exist independently two fundamental spinor representations Δ^+ and Δ^-. $\Gamma_{(\ell-1)}$ is irreducible, but not fundamental. Γ_ℓ is neither irreducible nor fundamental, but is equal to $\Gamma_\ell^+ + \Gamma_\ell^-$ where Γ_ℓ^+ and Γ_ℓ^- are irreducible but not fundamental.

Thus, except for the spinor representations, we see that the fundamental representations are built up out of the self-representations $\overset{S}{D}$. The building-up process here is more complicated than that of (6.1), since in each case from the direct product we have to extract the completely anti-symmetric part. However, it can be shown that this does not invalidate the assumption that the invariants formed with the Γ_ℓ are built up out of the invariants formed with the $\overset{S}{D}$, and hence (apart from the spinor representations) we see that we shall need only the invariants formed with the $\overset{S}{D}$.

To dispose of the spinor representations, Δ , $\Delta^\pm$ we shall have to use more mathematical and less heuristical arguments. We hope they may also make more plausible the validity of the arguments just used in (1), (2), (3).

(4) For B_ℓ , we consider Δ , and form the representation $\Delta \times \Delta$. Clearly,

$$X_\alpha(\Delta \times \Delta) = X_\alpha(\Delta) X 1 + 1 \times X_\alpha(\Delta) \qquad (6.2)$$

and so, for the $\hat{A}$ defined in (4.2) we have

$$\hat{A}(\Delta\times\Delta) = \hat{A}(\Delta)\times 1 + 1\times\hat{A}(\Delta). \tag{6.3}$$

Hence

$$\begin{aligned}
\hat{I}_n(\Delta\times\Delta) &= \hat{Sp}\left[\hat{A}(\Delta\times\Delta)\right]^n \\
&= \hat{Sp}\left[\hat{A}(\Delta)\times 1 + 1\times\hat{A}(\Delta)\right]^n \\
&= \hat{Sp}\sum_{r=1}^{n}\binom{n}{r}\hat{A}(\Delta)^r \times \hat{A}(\Delta)^{n-r} \\
&= \sum_{r=1}^{n}\binom{n}{r}\left(\hat{Sp}\,\hat{A}(\Delta)^r\right)\left(\hat{Sp}\,\hat{A}(\Delta)^{n-r}\right) \\
&= \sum_{r=1}^{n}\binom{n}{r} I_r(\Delta)\, I_{n-r}(\Delta) \\
&= 2d\, I_n(\Delta) + \text{a bilinear in lower order } \Delta \text{ invariants,}
\end{aligned} \tag{6.4}$$

where d is the dimension of Δ. On the other hand[6],

$$(\Delta\times\Delta) = 1 + \Gamma_1 + \Gamma_2 + \cdots + \Gamma_{\ell-2} + \Gamma_{\ell-1} + \Gamma_\ell \tag{6.5}$$

Hence, from (5.4)

$$2d\, I_n(\Delta) = \sum_{r=1}^{\ell} I_n(\Gamma_r) + \text{a function of lower order } \Delta \text{ invariants.} \tag{6.6}$$

Therefore, by induction in n, we see that the $I_n(\Delta)$ are functions of the $I_n(\Gamma_r)$ which, in turn, are functions of the $I_n\left(\overset{3}{D}\right)$. Thus for B_ℓ, we need not consider the spinor representation Δ.

(5) For D_ℓ, we consider Δ^+ and Δ^-, and use the relation[6]

$$\Delta^+\times\Delta^- = \Gamma_1 + \Gamma_3 + \cdots + \Gamma_{(\ell-1)} \tag{6.7}$$

to deduce in a manner similar to that just used in (4) that

$$I_n(\Delta^+) + I_n(\Delta^-) = \text{function of } I_n(\Gamma_\nu)\ \nu = 1, \ldots (\ell-1), \quad (6.8)$$
$$= \quad '' \quad\quad '' \ I_n\left(\begin{smallmatrix} S \\ D \end{smallmatrix}\right)$$

Thus the invariants formed with the Δ^+ and Δ^- are linearly related to each other, and we need consider only Δ^+ or Δ^-, but not both.

(6) On considering the relations[6]

$$\Delta^{\pm} \times \Delta^+ = 1 + \Gamma_0 + \Gamma_2 + \ldots + \Gamma_{\ell-2} + \Gamma_\ell^+, \quad \ell \text{ even},$$

$$= \Gamma_1 + \Gamma_3 + \ldots + \Gamma_{(\ell-2)} + \Gamma_\ell^-, \quad \ell \text{ odd} \quad (6.9)$$

(or the similar expressions for $\Delta^- \times \Delta^-$) we see that $\Delta^+ \times \Delta^+$ is <u>not</u> a direct sum of the Γ_ℓ (it contains $\Gamma_\ell^{\pm}$ in a form other than $\Gamma_\ell^+ + \Gamma_\ell^- = \Gamma_\ell$). Thus while we can still deduce that

$$2d\ I_n(\Delta^+) = \text{function of } I_n(\Gamma^\nu)\ \nu = 1, \ldots (\ell-2) \text{ and}$$
$$\Gamma^{\pm}_{(\ell-1)} \quad (6.10)$$

We cannot deduce from this that

$$2d\ I_n(\Delta^+) = \text{function of the } I_n\left(\begin{smallmatrix} S \\ D \end{smallmatrix}\right) \quad (6.11)$$

because we do not know that

$$I_n(\Gamma_\ell^{\pm}) = \text{function of the } I_n\left(\begin{smallmatrix} S \\ D \end{smallmatrix}\right) \quad (6.12)$$

This is because the proof that

$$I_n(\Gamma_\nu) = \text{function of the } I_n\left(\begin{smallmatrix} S \\ D \end{smallmatrix}\right)$$
$$\nu = 1, \ldots \ell. \quad (6.13)$$

does not say anything about the individual parts Γ_ℓ^+ and Γ_ℓ^- of Γ_ℓ . Hence, for D_ℓ , we cannot deduce that neither Δ^+ nor Δ^- need be considered.

(7) By a refinement of the argument, however, we can show that while we may have to retain either Δ^+ or Δ^- we shall need it to build with it at most one invariant I_ℓ. The idea is as follows: Instead of considering $\left(\Delta^+ \times \Delta^+\right)$ we consider only its anti-symmetric part, $\left(\Delta^+ \times \Delta^+\right)_{A.S}$. For this part we have

$$\left(\Delta^+ \times \Delta^+\right)_{A.S} = \begin{cases} \Gamma_2 + \Gamma_6 + \cdots + \Gamma_{\ell-2}, & \ell = 0 \pmod 4 \\ \Gamma_3 + \Gamma_7 + \cdots + \Gamma_{\ell-2}, & \ell = 1 \pmod 4 \\ 1 + \Gamma_4 + \cdots + \Gamma_{\ell-2}, & \ell = 2 \pmod 4 \\ \Gamma_1 + \Gamma_5 + \cdots + \Gamma_{\ell-2}, & \ell = 3 \pmod 4 \end{cases} \qquad (6.14)$$

and we see that the troublesome $\Gamma_\ell^\pm$ has dropped out (this is because $\Gamma_\ell^\pm$ is the highest order representation occurring in $\Delta^+ \times \Delta^+$ and therefore belongs to the symmetric part). Then, in analogy to (6.4) we have

$$I_n\left(\Delta^+ \times \Delta^+_{AS}\right) = \hat{S}p\left[\hat{A}\left(\Delta^+ \times \Delta^+_{AS}\right)\right]^n$$

$$= \hat{S}p\left[\hat{A}\left(\Delta^+\right) \times 1_{A.S} + 1 \times \hat{A}\left(\Delta^+\right)_{AS}\right]^n$$

$$= \hat{S}p \sum_{r=1}^{n} \binom{n}{r} \left[\hat{A}\left(\Delta^+\right)\right]^n \times \left[\hat{A}\left(\Delta^+\right)\right]^{n-r}_{AS} \qquad (6.15)$$

But a difference comes at this point. As is shown in the appendix

$$Sp\,(B\times C)_{AS} = \frac{1}{2}(Sp\,B)(Sp\,C) - \frac{1}{2}\,Sp(BC) \qquad (6.16)$$

Hence

$$I_n\left(\Delta^+ \times \Delta^+_{AS}\right) = \frac{1}{2}\sum_{r=1}^{n}\binom{n}{r}\left[I_r(\Delta^+)\,I_{n-r}(\Delta^+) - I_n(\Delta^+)\right]$$

$$= \frac{1}{2}\left(2d^+ - 2^n\right) I_n(\Delta^+)$$

$$+ \text{ a function of lower order invariants.} \qquad (6.17)$$

Hence, by induction in n, we have the result that

$$\left(d^+ - 2^{n-1}\right) I_n(\Delta^+) \qquad = \text{function of } I_n\left(\frac{S}{D}\right) \qquad (6.18)$$

Thus we can see that the $I_n(\Delta^+)$ are not independent, unless $d^+ = 2^{(n-1)}$. But the latter equality holds for $n = \ell$ (the dimension d^+ of Δ^+ is just $2^{(\ell-1)}$ (6) (7)). We see therefore that we can draw no conclusions concerning $I_\ell(\Delta^+)$ This is how $I_\ell(\Delta^+)$ escapes the reduction process, and why it is the only invariant to do so.

Very similar results to the above holds for the five special groups G_2 , F_4, etc. We shall not discuss these in detail, but as an example, we shall show that for G_2 the invariants formed with the second fundamental representation D_{14} are functions of those formed with the self-representation D_7 . In fact, we have

$$\left(D_7 \times D_7\right)_{AS} = D_7 + D_{14}\,. \qquad (6.19)$$

Hence

$$I_n(D_{14}) = f\left(I_n(D_7)\right) - I_n(D_7) \qquad (6.20)$$

as required.

The next question is why we need to consider only the invariants of the orders listed. An immediate explanation, of course, is that, as we saw in § 5, we need only ℓ invariants and (as we shall see below) the invariants listed in § 4, are the ℓ invariants of lowest order in each case. However, even without knowing in advance that only ℓ invariants are needed we can reduce the possibilities to those listed. The reduction goes as follows: We have seen in § 4 that the $\hat{I}_n$ can be expressed as

$$\hat{I}_n = \hat{T}_R \hat{A}^n \qquad (6.21)$$

where $\hat{A}$ is a matrix of the same dimension d as the fundamental representation $\overset{S}{D}$. But any d x d matrix satisfies its own characteristic equation

$$\hat{A}^d = \alpha_0 + \alpha_1 \hat{A} + \cdots + \alpha_{d-1} \hat{A}^{(d-1)} \qquad (6.22)$$

where the α_i's are functions of $Sp\ \hat{A}^\gamma$, $\gamma = 1, \ldots d$

On multiplying (6.22) by $\hat{A}$ and taking the spur of both sides, We have

$$\hat{I}_{d+1} = Sp\ \hat{A}^{(d+1)} = \alpha_0 Sp\ \hat{A} + \alpha_1 Sp\ \hat{A}^2 + \cdots - \alpha_d Sp\ \hat{A}^d \qquad (6.23)$$

$$= f(\hat{I}_\gamma) \quad \gamma = 1 \ldots d. \qquad (6.24)$$

Similarly for any higher order $\hat{I}_{d+b}$ $b > 0$. Thus we have for the self-representation:

For A_ℓ : $d = (\ell+1)$: Hence we need only consider the invariants of order 2, 3, ...$(\ell+1)$ which are just the ones listed in § 4.

For B_ℓ : $d = (2\ell+1)$: We apparently need the invariants of order 2, 3, ... $2\ell+1$. But because B_ℓ is the orthogonal group, $\hat{A}$ is anti-symmetric. Hence

$$I_n = \hat{Sp}\,(\hat{A})^n = \hat{Sp}\,(\hat{A}^{\sim})^n = (-1)^n\,\hat{Sp}\,(\hat{A})^n = (-1)^n I_n \qquad (6.25)$$

Thus the odd invariants drop out leaving us with the ones listed in § 4

For C_ℓ ; $d = 2\ell$: We apparently need the invariants of order 2,3, ... 2ℓ. But if we use the definition of the sympletic group as the group of $(2\ell \times 2\ell)$ matrices u satisfying

$$\tilde{u}\,J\,u = J \qquad (6.25)$$

where J is a non-singular anti-symmetric matrix, we have

$$\hat{A}^{\sim} = -J\,\hat{A}\,J^{-1} \qquad (6.27)$$

and hence

$$I_n = \hat{Sp}\,(\hat{A})^n = \hat{Sp}\,(\hat{A}^{\sim})^n = \hat{Sp}\,(-1)^n\,J\,\hat{A}^n\,J^{-1} = (-1)^n\,\hat{Sp}\,(\hat{A})^n = (-1)^n I_n \qquad (6.28)$$

Thus, once again, the odd order invariants drop out, leaving us with the set listed in § 4.

For D_ℓ : $d = 2\ell$: As for B_ℓ, the odd invariants drop out, and so from the self-representation we are left with the invariants of order 2, 4,.....2ℓ. In addition we have $I_\ell(\Delta^+)$. Thus we have the set listed in § 4, except that we still have $I_{2\ell}(\overset{S}{X})$. That this invariant is a function of the others is not so easy to see directly because it is a function of $I_\ell(\Delta^+)$ as well as the $I_{2\gamma}(\overset{S}{X})$, $\gamma = 1, \ldots (\ell-1)$ but it can be made very plausible by an inspection of equations (4.13) and (4.21), from which we see that the leading terms of $I_{2\gamma}(\overset{S}{X})$, $\gamma = 1, \ldots (\ell-1)$ and $I_\ell(\Delta^+)$. Similarly, the necessity of $I_\ell(\Delta^+)$, can be made plausible by an inspection of the leading terms of $I_{2\gamma}(\overset{S}{X})$ and $I_\ell(\Delta^+)$

the leading term in $I_{2\ell}$ can be expressed as a function of

As is shown in the appendix, Δ^+ is distinguished from Δ^- by the fact that (with a suitable basis) they have the cartan algebras

$H_1, H_2, \ldots H_\ell$ and $-H_1, H_2, \ldots H_\ell$ respectively. Thus the leading terms of $I_{2\gamma}(\overset{5}{x})$, which are even in H_1, do not distinguish between Δ^+ and Δ^-, while the leading term $H_1, H_2, \ldots H_\ell$ of $I_\ell(\Delta^+)$ does.

For G_2 : d = 7 : The odd invariants drop out, as is pleausible from (4.23) and we are left with I_2, I_4, I_6. But as we have pointed out earlier, the leading term in I_4 is just twice the square of the leading term in I_2. This makes plausible the result that I_2 and I_6, not I_2 and I_4, are the independent invariants for this group.

PART II

In this part of the paper we shall be concerned with giving an infinitesimal proof of the S-theorem stated in 5. However, for this purpose, it is first necessary to obtain same relevant properties of the roots. These properties will be deduced from Dynkens primitive root theorem in § 7. The proof proper of the S-theorem will be given in § 8

7 Relevant properties of the Roots

The properties of the roots which we shall need can all be derived from Dynkens primitive root theorem[10]. This theorem states that for every simple Lie group there exists a set of ℓ positive roots (called primitive roots) such that every

positive root can be expressed as a linear conbination of these with non-negative integer coefficients. In other wards, if α_γ, $\gamma = 1, \ldots \ell$ are the primitive roots, then any positive root α is of the form

$$\alpha = \sum_{\gamma=1}^{\ell} k_\gamma \alpha_\gamma , \tag{7.1}$$

where the k_γ we non-negative integer for all γ .

From this fundamental theorem we can deduce the following results:

Theorem I: If E_{α_γ} and E_{α_β} are the $E'\beta$ corresponding to two different primitive roots,

$$\left[E_{\alpha_\gamma}, E_{\alpha_\beta}\right] = 0 \tag{7.2}$$

Proof: Otherwise $(\alpha_\gamma - \alpha_\beta)$ and $(\alpha_\beta - \alpha_\gamma)$ would be non-zero roots, and since are of them must be positive this would contradict (7.1).

Theorem II: Let α be any positive root, and let $\alpha \to \alpha'$ under a reflection in the plane perpendicular to a primitive root . Then α' is positive, unless $\alpha = \alpha$ itself.

Proof: If $\alpha \neq \alpha$, at least are $k_\beta \neq k_\gamma$ in (7.1) is non-zero. But under the reflexion mentioned, $\alpha \longrightarrow \alpha' = \alpha - \frac{2(\alpha \alpha_\gamma)}{(\alpha_\gamma . \alpha_\gamma)} \alpha_\gamma$ so that only the coefficient of α_γ changes. Thus α' as at least are positive coefficient, namely, k_β . Thus, from (7.1) α is positive.

Theorem III: (Σ' - theorem) Let Σ' be the sum of all the positive roots excluding α . Then

$$\left(\sum' \cdot \alpha_r\right) = 0 \tag{7.3}$$

Proof: By the previous theorem all the positive roots (except α_r) remain invariant under the reflection perpendicular to α_r. Hence this reflection (since a reflection is non-singular) only permutes the positive roots (excluding α_r). Hence Σ' the sum of these, the positive roots (excluding α_r), remains invariant under the reflection. These Σ' must lie in the plane of reflection. Q.E.D.

Theorem IV: Let $k = \sum_{r=1}^{\ell} k_r$ be called the "height" of the positive root $\alpha = \sum_{r=1}^{\ell} \alpha_r k_r$. Then there exists at least one reflection in the plane perpendicular to a primitive root which reduces the 'height" of L.

Proof: If $(L \cdot \alpha_r) \leq 0$ (7.4)

all α_r, then

$$(L \cdot L) = \sum_{r=1}^{\ell} k_r (\alpha_r \cdot \alpha_r) \leq 0 \tag{7.5}$$

which is impossible. Hence there exists at least one α_r such that $(L \cdot \alpha_r) > 0$. Reflecting in the plane perpendicular to the α_r we get

$$L \rightarrow L' = L - 2 \frac{(L \cdot \alpha_r)}{(\alpha_r \cdot \alpha_r)} \alpha_r , \tag{7.6}$$

or

$$k \rightarrow k' = k - 2 \frac{(L \cdot \alpha_r)}{(\alpha_r \cdot \alpha_r)} < k \tag{7.7}$$

as required.

Corallory: For any positive root α, there exists a finite sucession of reflections in planes perpendicular to primitive roots $W_{\alpha_r}, W_{\alpha_s}, \ldots W_{\alpha_t}$ such that $W_{\alpha_r} W_{\alpha_s} \ldots W_{\alpha_t} \alpha$ is primitive.

Theorem V: (Primitive root reflection theorem) Any reflection is a plane perpendicular to a non-primitive root can be generated by a series of reflections in planes perpendicular to primitive roots (primitive reflections).

Proof: Let α be a non-primitive root, and let $W_{\alpha_r}, W_{\alpha_s}, \ldots W_{\alpha_t}$ be primitive reflections such that $\alpha' = W_{\alpha_r} W_{\alpha_s} \ldots W_{\alpha_t} \alpha$ is primitive. Consider now the operation generated by the series of primitive reflections $W_{\alpha_t} \ldots W_{\alpha_r} W_{\alpha} W_{\alpha_r} \ldots W_{\alpha_t}$. By a well-known theorem this operation is a reflection in the plane perpendicular to the vector $W_{\alpha_t} \ldots W_{\alpha_r} \alpha'$. But this vector is just α. Q.E.D.

Corollaey: The Weyl group generated by reflections in planes perpendicular to the roots, is generated by reflections in the planes perpendicular to the primitive roots alone

8. Infinitesimal Proof of the S-theorem:

The proof proceeds in three steps. The first is to prove the following result: Let $(F(X_\alpha)$ be any invariant, and let us write it in normal form i.e. as a sum of terms of the form

$$\Phi = E_{-\alpha} E_{-\beta} \ldots E_{-(\gamma+\delta)} \ldots E_{-(\lambda+\mu+\nu)} \left[\phi(\lambda,\mu)\right] E_{(\lambda'+\mu'+\nu')} \ldots E_{(\gamma'+\delta')} \ldots E_{\beta} E_{\alpha'} \quad (8.1)$$

where the α, β, ... β', α' are primitive roots. Note that

we have not only placed the destruction operators to the right and creation operators to the left, but have ordered them according to the number of primitive roots contained in this subscript. The $\Phi(H_i)$ are different for each Φ of course, Furthermore, since $F(x_\alpha)$ commutes with the H_i in particular,

$$[\alpha+\beta+\cdots-(\delta+\delta)+\cdots+(\lambda+\mu+\nu)] = [\alpha'+\beta'+\cdots-(\delta'+\delta')-\cdots+(\lambda'+\mu'+\nu')] \quad (8.2)$$

Let $|\lambda\rangle$ be the normalized eigenvector belonging to the highest weight λ in any representation, so that

$$E_\alpha |\lambda\rangle = 0 \quad (8.3)$$

for all positive roots α . The vector

$$|\lambda - n\sigma\rangle = k (E_{-\sigma})^n |\lambda\rangle \quad (8.4)$$

where

$$|k|^{-2} = \langle \lambda | (E_\sigma)^n (E_{-\sigma})^n | \lambda \rangle \quad (8.5)$$

Is the normalized eigenvector belonging to the weight $\lambda - n\sigma$

Then

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = 0 , \quad (3.6)$$

unless

$$\Phi = E_{-\alpha}^m \phi(H_i) E_\alpha^m , \quad m \leq n \quad (8.7)$$

in which case,

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = \left[n\sigma \cdot \left(\lambda - \tfrac{n-1}{2}\sigma\right)\right]\left[(n-1)\sigma \cdot \left(\lambda - \tfrac{n-2}{2}\sigma\right)\right] \quad (8.8)$$

$$\cdots\cdots \left[(n-m+1)\sigma \cdot \left(\lambda - \tfrac{n-m}{2}\sigma\right)\right]$$

$$\times \Phi\left(\lambda - (n-m)\sigma\right)$$

Proof: We have

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = k^2 \langle \lambda | E_\sigma^n \Phi E_{-\sigma}^n | \lambda \rangle \qquad (8.9)$$

But by theorem I of § 7, E_α (to the extreme right of Φ in (8.1)) commutes with $E_{-\sigma}$ unless $\alpha = \sigma$. Thus, from (8.3), $\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle$ is zero unless $\alpha' = \sigma$ In the latter case $E_{\alpha'}$ just neutralizes one of the $E_{-\sigma}$'s and we get

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = \text{const.} \langle \lambda | E_\sigma^n \Phi' E_{-\sigma}^{(n-1)} | \lambda \rangle \qquad (8.10)$$

where

$$\Phi' = E_\alpha E_\beta \cdot [\phi(H)] \cdots E_{\beta'} \qquad (8.11)$$

i.e. Φ' is equal to Φ with $E_{\alpha'}$ removed from the end. The evaluation of the constant in (8.10) will be carried out later. Now let us suppose that the number of E's belonging to primitive roots on the extreme right hand side of Φ in (8.1) is m. If, for these E's, we keep repeating the above process, there are two possibilities:

(1) $m > n$. In this case, we end up with

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = \text{const} \langle \lambda | \cdots E_{\epsilon'} | \lambda \rangle = 0 \qquad (8.12)$$

(2) $m \leq n$. In this case, we obtain

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle = \text{const} \langle \lambda | E_\sigma^n \cdots \Phi'' E_{-\sigma}^{n-m} | \lambda \rangle \qquad (8.13)$$

where $$\Phi'' = E_\alpha E_\beta \cdots [\Phi(H)] \cdots E_{(\beta'+\delta')} \qquad (8.14)$$

i.e. Φ'' is equal to Φ with all the E_α, E_β missing However, we can show that

$$|\Psi\rangle = E_{(\gamma'+\delta')} E_{-\sigma}^{n-m} |\lambda\rangle = \left[E_{\gamma'+\delta'}, E_{-\sigma}^{n-m} \right] |\lambda\rangle = 0 \qquad (8.15)$$

as follows: If $q = \gamma' + \delta' - \sigma$ is not a root, $E_{(\gamma'+\delta')}$ commutes with $E_{-\sigma}^{n-m}$ and we get zero. If q is a root, it is positive, since otherwise $\sigma = \gamma' + \delta' + |q|$ would not be primitive. Thus $|\Psi\rangle$ is a sum of terms of the form

$$|\Psi_1\rangle = E_{-\sigma}^{n-m-r-1} E_q E_{-\sigma}^{r} |\lambda\rangle = E_{-\sigma}^{n-m-r-1} \left[E_q, E_{-\sigma}^{r} \right] |\lambda\rangle \qquad (8.16)$$

If $t = q - \sigma$ is not a root, we again get zero. If t is a root, it must be positive, for if it is negative $\sigma = q + |t|$ is nonprimitive, and if it is zero $\gamma' + \delta' = q + \sigma = 2\sigma$. and is not a root. Thus each such non-zero $|\Psi_1\rangle$ is a sum of terms of the form

$$|\Psi_2\rangle = E_{-\sigma}^{n-m-s-2} \left[E_t, E_{-\sigma}^{s} \right] |\lambda\rangle \qquad (8.17)$$

Continuing in this way, and noticing that at each step the number of $E_{-\sigma}$'s present decreases by one, we easily see that eventually the $E_{-\sigma}$ disappear, and we are left with terms of the form

$$|\Psi_n\rangle = E_\omega |\lambda\rangle = 0 \,, \qquad (8.17a)$$

since ω is positive. Going back up the ladder we deduce the $\Psi_\gamma\rangle$ and $\Psi\rangle$ are zero as required.

Thus (8.13) is zero unless $\bar{\Phi}$ contains no $E_{(\delta'+\delta')}$ type terms. Similarly it can be shown that it is zero unless $\bar{\Phi}$ contains, further, no $E_{(\lambda'+\mu'+\upsilon')}$ -type terms, and so on. Hence $\langle\lambda-n\sigma|\,\bar{\Phi}\,|\lambda-n\sigma\rangle$ is zero unless $\bar{\Phi}$ is of the form

$$\bar{\Phi} = E_{\alpha-\beta}\cdots E_{(\delta+\delta)}\cdots E_{(\lambda+\mu+\upsilon)}\left[\phi(H)\right]E_{\sigma}^{m}, \quad (8.18)$$

for $m \leq n$. Applying analogous considerations to the left hand side of $\bar{\Phi}$, we find that $\langle\lambda-n\sigma|\,\bar{\Phi}\,|\lambda-n\sigma\rangle$ is zero unless $\bar{\Phi}$ is of the form

$$\bar{\Phi} = E_{-\sigma}^{m}\left[\phi(H)\right]E_{\sigma}^{m} \qquad m \leq n \qquad (8.19)$$

The power m is the same to the left and right on account of (8.2).

This proves the first part of the result stated, and it remains only to calculate:

$$\langle\lambda-n\sigma|\,\bar{\Phi}\,|\lambda-n\sigma\rangle$$

$$= \frac{\langle\lambda|\,E_{\sigma}^{n}E_{-\sigma}^{m}\left[\phi(H)\right]E_{\sigma}^{m}E_{-\sigma}^{n}\,|\lambda\rangle}{\langle\lambda|\,E_{\sigma}^{n}E_{-\sigma}^{n}\,|\lambda\rangle}$$

$$= \frac{\langle\lambda|\,E_{\sigma}^{n}E_{-\sigma}^{m}E_{\sigma}^{m}E_{-\sigma}^{n}\,|\lambda\rangle}{\langle\lambda|\,E_{-}^{n}E_{-\sigma}^{n}\,|\lambda\rangle}\,\phi\left(\lambda-(n-m)\sigma\right) \qquad (8.20)$$

We have,

$$\langle\lambda|\, E_\sigma^n E_{-\sigma}^m E_\sigma^m E_{-\sigma}^n |\lambda\rangle = \langle\lambda|\, E_\sigma^n E_{-\sigma}^m E_\sigma^{m-1}\left[E_\sigma, E_{-\sigma}^n\right]|\lambda\rangle$$

$$= \langle\lambda|\, E_\sigma^n E_{-\sigma}^m E_\sigma^{m-1} \sum_{p=0}^{n-1} E_{-\sigma}^{n-1-p}\left[E_\sigma, E_{-\sigma}\right] E_{-\sigma}^p |\lambda\rangle$$

$$= \langle\lambda| \quad " \quad " \quad " \quad \left[\gamma(\sigma).H\right] \quad " \quad |\lambda\rangle$$

$$= \langle\lambda| \quad " \quad " \quad " \quad \left[\quad .(\lambda - p\sigma')\right] \quad " \quad |\lambda\rangle$$

$$= \langle\lambda|\, E_\sigma^n E_{-\sigma}^m E_\sigma^{m-1} E_{-\sigma}^n |\lambda\rangle \sum_{p=0}^{n-1} \gamma.(\lambda - p\sigma)$$

$$= \langle\lambda| \quad " \quad " \quad |\lambda\rangle \left[n\gamma.\left(\lambda - \tfrac{n-1}{2}\sigma\right)\right]$$

$$= \langle\lambda|\, E_\sigma^{n-1} E_{-\sigma}^{m-1} E_\sigma^{m-1} E_{-\sigma}^{n-1} |\lambda\rangle \left[\quad " \quad\right]^2$$

$$= \langle\lambda|\, E_\sigma^{(n-2)} E_{-\sigma}^{m-2} E_\sigma^{m-2} E_{-\sigma}^{n-2} |\lambda\rangle \left[n\gamma.\left(\lambda - \tfrac{n-1}{2}\sigma\right)\right]^2 \times \left[(n-1)\gamma.\left(\lambda - \tfrac{n-2}{2}\sigma\right)\right]^2$$

$$= \langle\lambda|\, E_\sigma^{n-m} E_{-\sigma}^{n-m} |0\rangle . \left[n\gamma.\left(\lambda - \tfrac{n-1}{2}\right)\right]^2 \cdots \left[(n-m-1)\gamma.\left(\lambda - \tfrac{n-m}{2}\sigma\right)\right]^2 \quad (8.21)$$

and similarly,

$$\langle\lambda|\, E_\sigma^h E_{-\sigma}^h |\lambda\rangle = \langle\lambda|\, E_\sigma^{h-1}\left[E_\sigma, E_{-\sigma}^h\right]|\lambda\rangle$$

$$= \langle\lambda|\, E_\sigma^{h-1} \sum_{p=0}^{h-1} E_{-\sigma}^{h-1-p}\left[E_\sigma, E_{-\sigma}\right] E_{-\sigma}^p |\lambda\rangle$$

$$= \langle\lambda|\, E_\sigma^{h-1} E_{-\sigma}^{h-1} |\lambda\rangle \left[h\gamma.\left(\lambda - \tfrac{h-1}{2}\sigma\right)\right]$$

$$= \langle\lambda|\, E_\sigma^{h-2} E_{-\sigma}^{h-2} |\lambda\rangle \left[h\gamma.\left(\lambda - \tfrac{h-1}{2}\sigma\right)\right]\left[(h-1)\gamma.\left(\lambda - \tfrac{h-2}{2}\sigma\right)\right.$$

$$= \left[h\gamma.\left(\lambda - \tfrac{h-1}{2}\sigma\right)\right] \cdots \left[2\gamma.\left(\lambda - \tfrac{1}{2}\sigma\right)\right](\gamma.\lambda) \quad (8.22)$$

Inserting this result in (8.21) and (8.20), we obtain

$$\langle \lambda - n\sigma | \Phi | \lambda - n\sigma \rangle$$

$$= \frac{\left[(n-m)r.\left(\lambda - \frac{n-1}{2}\sigma\right)\right] \cdots\cdots (r.\lambda) \left[nr.\left(\lambda - \frac{n-1}{2}\sigma\right)\right] \cdots \left[(n-m-1)r.\left(\lambda - \frac{n-m}{2}\sigma\right)\right]}{\left[nr.\left(\lambda - \frac{n-1}{2}\sigma\right)\right] \cdots\cdots (r.\lambda)}$$

$$\times \quad \phi\left(\lambda - (n-m)\sigma\right)$$

$$= \left[nr.\left(\lambda - \frac{n-1}{2}\sigma\right)\right] \cdots\cdots \left[(n-m-1)r.\left(\lambda - \frac{n-m}{2}\sigma\right)\right]$$

$$\times \quad \phi\left(\lambda - (n-m)\sigma\right) \qquad (8.23)$$

as required. This completes the first (and longest !) step in the proof.

What we have just proved is that if $F(x_\alpha)$ is an invariant, it can be expressed in the form,

$$F(x_\alpha) = \phi_0(H) + \sum_\alpha E_{-\alpha}\phi_\alpha^{(1)}(H)E_\alpha \qquad (8.24)$$

$$+ \cdots\cdots + \sum_\alpha E_{-\alpha}^{m}\phi_\alpha^{m}(H)E_\alpha^{m} + \cdots\cdots + \chi,$$

where χ has no diagonal elements with respect to the state $|\lambda - n\sigma\rangle$. The second step in the proof consists in using this result to show that $\phi_0(\lambda)$ is even about $-\alpha/2$ under a reflection in the plane $\perp \alpha$ i.e. that

$$\phi_0\left(-\frac{\alpha}{2} + \lambda\right) = \phi_0\left(-\frac{\alpha}{2} + S_\alpha\lambda\right), \quad \text{all } \alpha, \qquad (8.25)$$

where S_α denotes the reflection $\perp \alpha$.

We show this as follows: In any irreducible representation $F(X_x)$ is a constant, f, say. Hence

$$f = \langle \lambda | F(x) | \lambda \rangle = \phi_0(\lambda), \tag{8.26a}$$

$$f = \langle \lambda - \alpha | F(x) | \lambda - \alpha \rangle = \phi_0(\lambda - \alpha) + \phi_\alpha^{(1)}(\lambda)\,(\alpha \cdot \lambda), \tag{8.26b}$$

Using (8.8). Similarly,

$$f = \langle \lambda - 2\alpha | F(x) | \lambda - 2\alpha \rangle = \phi_0(\lambda - 2\alpha) + \phi_\alpha^{(1)}(\lambda - \alpha)\left[2\alpha \cdot (\lambda - \tfrac{1}{2}\alpha)\right] + \phi_\alpha^{(2)}(\lambda)\left[2\alpha \cdot (\lambda - \tfrac{1}{2}\alpha)\right](\alpha \cdot \lambda), \tag{8.26c}$$

$$f = \langle \lambda - 3\alpha | F(x) | \lambda - 3\alpha \rangle = \phi_0(\lambda - 3\alpha) + \phi_\alpha^{(1)}(\lambda - 2\alpha)\left[3\alpha \cdot (\lambda - \tfrac{3}{2}\alpha)\right]\left[2\alpha \cdot (\lambda - \tfrac{1}{2}\alpha)\right] + \phi_\alpha^{(3)}(\lambda)\left[3\alpha \cdot (\lambda - \tfrac{3}{2}\alpha)\right]\left[2\alpha \cdot (\lambda - \tfrac{1}{2}\alpha)\right](\alpha \cdot \lambda), \tag{8.26d}$$

etc.

Equations (a) and (b) imply

$$\phi_0(\lambda - \alpha) - \phi_0(\lambda) = -\phi_\alpha^{(1)}(\lambda)(\alpha \cdot \lambda). \tag{8.27}$$

Eliminating ϕ_0 from (a) (b) and (c) we get a similar equation

$$\phi_\alpha^{(1)}(\lambda - \alpha) - \phi_\alpha^{(1)}(\lambda) = -\phi_\alpha^{(2)}(\lambda)\left[2\alpha \cdot (\lambda - \tfrac{1}{2}\alpha)\right], \tag{8.28}$$

Similarly, eliminating ϕ_0 and ϕ_1 from the first 4 equations, we obtain

$$\phi_\alpha^{(2)}(\lambda - \alpha) = \phi_\alpha^{(2)}(\lambda) - \phi_\alpha^{(3)}(\lambda)\left[3\alpha \cdot (\lambda - \tfrac{3}{2}\alpha)\right] \tag{8.29}$$

and so on. Generally,

$$\phi_\alpha^r(\lambda-\alpha) - \phi_\alpha^r(\lambda) = -\phi_\alpha^{r+1}(\lambda)\left[r\alpha\cdot\left(\lambda-\tfrac{r-1}{2}\alpha\right)\right] \quad (8.30)$$

At this stage we need the following Lemma.

Lemma: Let $\phi_\alpha^{r+1}(\lambda)$ in (8.30) satisfy

$$\phi_\alpha^{r+1}\left(\tfrac{r\alpha}{2}+\chi\right) = \phi_\alpha^{r+1}\left(\tfrac{r\alpha}{2}+S_\alpha\chi\right), \quad (8.31)$$

where χ is any vector, and S_α is a reflection in the plane $\perp\alpha$. Then

$$\phi_\alpha^r\left(\tfrac{r-1}{2}\alpha+\chi\right) = \phi_\alpha^r\left(\tfrac{r-1}{2}\alpha+S_\alpha\chi\right). \quad (8.32)$$

In other words, if ϕ_α^{r+1} is invariant under S_α about $\frac{r\alpha}{2}$, ϕ_α^r is invariant about $\frac{r-1}{2}\alpha$

Proof: Let y be any vector $\perp\alpha$ and n any integer. Then from (8.30), with $\lambda = \frac{r-1}{2}\alpha + y + (n+1)\alpha$ we have

$$\begin{aligned}\phi_\alpha^r\left(\tfrac{r-1}{2}\alpha+y+(n+1)\alpha\right) &= \phi_\alpha^r\left(\tfrac{r-1}{2}\alpha+y+n\alpha\right)\\ &+ \left[r\alpha\cdot(1+n)\alpha\right]\phi_\alpha^{r+1}\left[\tfrac{r}{2}+\left(n+\tfrac{1}{2}\right)\alpha+y\right].\end{aligned} \quad (8.33)$$

On the other hand, using (8.30) with $\lambda = \frac{r-1}{2}\alpha + y - n\alpha$, we have

$$\begin{aligned}\phi_\alpha^r\left[\tfrac{r-1}{2}\alpha+y+(n+1)\alpha\right] &= \phi_\alpha^r\left(\tfrac{r-1}{2}\alpha-y-n\alpha\right)\\ &- \left[-(n+1)\alpha\right]\phi_\alpha^{r+1}\left[\tfrac{r}{2}-\left(n+\tfrac{1}{2}\right)\alpha+y\right],\end{aligned} \quad (8.34)$$

But, by (8.31)

$$\phi_\alpha^{r+1}\left(\frac{r}{2}\alpha+(n+\tfrac{1}{2})\alpha+y\right)=\phi_\alpha^{r+1}\left(\frac{r}{2}\alpha+S_\alpha(n+\tfrac{1}{2})\alpha+S_\alpha y\right) \quad (8.35)$$

$$=\phi_\alpha^{r+1}\left(\frac{r}{2}\alpha-(n+\tfrac{1}{2})\alpha+y\right).$$

Hence, the last terms on the right hand sides of (8.35) and (8.34) are equal. Substracting, we get

$$\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y+(n+1)\alpha\right)=\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y+n\alpha\right), \quad (8.36)$$

$$-\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y-(n-1)\alpha\right)$$

$$=-\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y-n\alpha\right),$$

from which, by induction in n, we have

$$\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y+n\alpha\right)=\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+y-n\alpha\right). \quad (8.37)$$

But since the ϕ's are polynomials and this is true for all integers n we can replace n by any number, μ. Hence if x is any vector, and we denote by $x_\perp$ its projection in the plane $\perp\alpha$, we have

$$\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+x\right)=\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+x_\perp+\frac{\alpha\cdot x}{\alpha\cdot\alpha}\alpha\right), \quad (8.38)$$

$$=\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+x_\perp-\frac{\alpha\, x}{x\, x}x\right),$$

$$=\phi_\alpha^{r}\left(\frac{r-1}{2}\alpha+S_x x\right),$$

where we have used (8.37) in the central equality with $n \to \mu = \frac{\alpha \cdot \chi}{\alpha \cdot \alpha}$. This establishes the Lemma.

Applying the Lemma to the above equations, we see that ϕ_0 is invariant under S_α about $-\alpha/2$ if $\phi_\alpha^{(1)}$ is invariant about zero, which is true if $\phi_\alpha^{(2)}$ is invariant about $\frac{\alpha}{2}$, and so on. But the powers of the ϕ_α^{γ} decrease with γ. Hence we eventually arrive at one which is constant (may be zero). This will certainly have the required invariance, so continuing back up the ladder, we find that

$$\phi_0\left(-\frac{\alpha}{2}+\chi\right) = \phi_0\left(-\frac{\alpha}{2} + S_\alpha \chi\right), \tag{8.39}$$

as required.

This brings us to the final step in the proof, namely, showing that if (8.39) holds for each primitive root α, then

$$\phi_0'(\lambda') = \phi_0'(S_\alpha \lambda'), \tag{8.40}$$

where

$$\phi_0(\lambda) = \phi_0'(\lambda'), \tag{8.41}$$

and

$$\lambda' = \lambda + \tfrac{1}{2}\sum, \tag{8.42}$$

where $\sum$ is the sum of the positive roots.

Proof: Let $-\frac{\alpha}{2} + \chi = \lambda$ in (8.39). Then we have

$$\begin{aligned}\phi_0'(\lambda') = \phi_0(\lambda) &= \phi_0\left[-\frac{\alpha}{2} + S_\alpha\left(\lambda + \frac{\alpha}{2}\right)\right] \\ &= \phi_0\left[-\frac{\alpha}{2} + S_\alpha\left(\lambda + \frac{\alpha}{2}\right) + \frac{\alpha}{2}\frac{\sum}{\alpha} - \frac{\alpha}{2}\frac{\sum}{\alpha}\right],\end{aligned} \tag{8.43}$$

where Σ' is the sum of the positive roots, excluding α.
But by the Σ- theorem of §7 we have

$$S_\alpha \Sigma' = \Sigma' \tag{8.44}$$

Hence

$$\begin{aligned}\phi_0'(\lambda') &= \phi_0\left[-\frac{\Sigma}{2} + S_\alpha\left(\lambda + \frac{\Sigma}{2}\right)\right] \\ &= \phi_0\left[-\frac{\Sigma}{2} + S_\alpha(\lambda')\right] \\ &= \phi_0\left[S_\alpha(\lambda')\right],\end{aligned} \tag{8.45}$$

as required.

But by the corrolary to the primitive root reflection theorem of §4, the Weyl group S is generated by the S_α with α primitive, Hence from (8.45)

$$\phi_0'(\lambda') = \phi_0'(S\lambda'), \tag{8.46}$$

which proves the S-theorem.

APPENDIX A

The Δ^+ and Δ^- representations of D_ℓ can be obtained as follows (7). Let τ_1, τ_2 and τ_3 be the 3 Pauli-matrices, and let

$$\begin{aligned}
\gamma_1 &= \tau_1 \times 1 \times 1 \times \cdots \times 1 \quad (\ell-1 \text{ terms}), \\
\gamma_2 &= \tau_3 \times \tau_1 \times 1 \times \cdots \times 1, \\
\gamma_3 &= \tau_3 \times \tau_3 \times \tau_1 \times \cdots \times 1, \\
&\cdots\cdots \\
&\cdots\cdots \\
\gamma_{\ell-1} &= \tau_3 \times \tau_3 \times \cdots \times \tau_3 \times \tau_1, \\
\gamma_\ell &= \tau_3 \times \tau_3 \times \cdots \times \tau_3 \times \tau_3, \\
\gamma_{\ell+1} &= \tau_2 \times 1 \times \cdots \times 1, \\
\gamma_{\ell+2} &= \tau_3 \times \tau_2 \times 1 \times \cdots \times 1, \\
&\cdots\cdots \\
&\cdots\cdots \\
\gamma_{2\ell-1} &= \tau_3 \times \tau_3 \times \cdots \times \tau_3 \times \tau_2,
\end{aligned} \tag{A.1}$$

So that the γ_μ, $\mu = 1, \ldots (2\ell-1)$ are the generators of the Clifford algebra,

$$(\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu) = 2\delta_{\mu\nu},$$

in $(2\ell-1)$ dimensions.

Now let

$$\begin{aligned}
X_\alpha^{\pm} &= \{ \cdots + \gamma_\mu, \cdots [\gamma_\mu, \gamma_\nu], \cdots \}, \\
\mu, \nu &= 1, \ldots (2\ell-1), \quad \alpha = 1, \ldots \frac{2\ell(2\ell-1)}{2}.
\end{aligned} \tag{A.2}$$

Then $X_\alpha^\pm$ are the infunitesimal generators of the $\Delta^\pm$ representations of D_ℓ. A convenient Cartan algebra $H_i^\pm$ can be obtained by letting

$$\begin{aligned}
\hat{H}_1^{\pm} &= \xi_\ell = \tau_3 \times \tau_3 \times \tau_3 \times \cdots \times \tau_3 , \\
\hat{H}_2^{\pm} &= -i\left[\xi_1, \xi_{\ell+1}\right] = \tau_3 \times 1 \times \cdots \times 1 , \\
H_3 &= -i\left[\xi_2, \xi_{\ell+2}\right] = 1 \times \tau_3 \times 1 \times 1 \cdots \times 1 , \\
&\cdots\cdots \\
&\cdots\cdots\cdots \\
H_\ell &= -i\left[\xi_{\ell-1}, \xi_{2\ell-1}\right] \\
&= \left(1 \times 1 \times 1 \times \cdots \times 1 \times \tau_3\right) .
\end{aligned} \tag{A.3}$$

It is easy to verify that these H_i's commute with each other and that no other $X^\pm = a_\alpha X_\alpha^\pm$ commutes with them.

From these results we can deduce:

(1) $\Delta^+ \rightarrow \Delta^-$ by letting $\xi_\mu \rightarrow -\xi_\mu$, $\mu = 1, \cdots (2\ell-1)$,

in which case

$$\begin{aligned}
H_1 &\longrightarrow -H_1 , \\
H_i &\rightarrow H_i , \quad i \neq 1 ,
\end{aligned} \tag{A.4}$$

and (2) the leading term in $I_\ell(\Delta^+)$ will be

$$\sum TR\left(\overset{\Delta}{H_i}\overset{\Delta}{H_j}\cdots\overset{\Delta}{H_k}\right) H^i H^j \cdots H^k . \tag{A.5}$$

But from (A.3) one can see that all the terms in this summation will be zero except those for which $i, j, \ldots k$ are all different. Thus the leading term in $I_\ell(\Delta^+)$ will be

$$H, H_2 \ldots \ldots H\ell \qquad (A.6)$$

APPENDIX B

We wish to show that

$$Sp\,(A \times B)_{A\,S} = \tfrac{1}{2}(Sp\,A\;\; Sp\,B) - \tfrac{1}{2}\,Sp(AB) \qquad (B.1)$$

We have,

$$(A \times B)_{ij|mn} = A_{im}\,B_{jn}. \qquad (B.2)$$

Hence

$$(A \times B)_{A\,S} = \tfrac{1}{4}\Big[A_{im}B_{jn} - A_{jm}B_{in} - A_{in}B_{jm} + A_{jn}B_{im}\Big], \qquad (B.3)$$

and

$$\begin{aligned} Sp\,(A \times B)_{A\,S} &= \tfrac{1}{4}\Big[A_{mm}B_{nn} - A_{nm}B_{mn} - A_{mn}B_{nm} + A_{mn}B_{nm}\Big] \\ &= \tfrac{1}{2}\,(A_{mm}B_{nn} - A_{nm}B_{mn}) \\ &= \tfrac{1}{2}\,(TR\,A)(TR\,B) - \tfrac{1}{2}\,TR(AB), \end{aligned} \qquad (B.4)$$

as required.

APPENDIX C

The independent invariants of the groups G_2, F_4, E_6 E_7, E_8, are all formed with the self-representations of these groups i.e. the lowest order non-trial representations. The order of the invariants are[1]:

G_2	:	invariants of order	2, 6
F_4	:	invariants of order	2, 6, 8, 12.
E_6	:	invariants of order	2, 5, 6, 8, 9, 12.
E_7	:	invariants of order	2, 6, 8, 10, 12, 14, 18.
E_8	:	invariants of order	2, 8, 12, 14, 18, 20, 24, 30.

It can be verified that for these five groups the order of the corresponding Weyl groups are $S \simeq 2.6$, 2.6.8.12, 2.5.6.8.9.12, 2.6.8.10.12.14.18 and 2.8.12.14.18.20.24.30 respectively.

0Ooo.·.ooO0

REFERENCES to Part III

1. G.Racah, Rend. Lincei 8, (1950) 108.
2. S.Okubo, Prog. Theor. Phys. 27, (1962) 949.
3. L.C. Biedenhakn, J. Math. Phys. 4, (1963) 436.
4. G. Racah, Princeton Lecture Notes, CERN 61-8 (1961).
5. R.E. Behrends, J. Dreithein, C.Fronsdal and W.Lee, Revs. Mod. Phys. 34, (1962) 1.
6. R.Branek and H. Weyl, Amer. Jour. Math. 57, (1935) 425.
7. H. Baekner, Representations of groups, North-Holland, Amsterdam (1963).
8. A. Klein, Jour. Math. Phys. 4, (1963) 1283.
9. M. Umezawa, Nucl. Phys. 48, (1963) 111.
10. E.B. Dynkin, Amer. Math. Soc. Trans. I, 9.