

CLASSICAL GAUGE THEORY

Richard Wilhelm Janse van Rensburg

**A Dissertation Submitted to the Faculty of Science
University of the Witwatersrand, Johannesburg
in Partial Fulfilment of the Requirements of the
Degree of Master of Science**

Johannesburg 1981

Abstract

CLASSICAL GAUGE THEORY

JANSE VAN RENSBURG, Richard Wilhelm, M.Sc. University of the
Witwatersrand, 1981.

This Dissertation presents a systematic study of classical gauge field theories generated by Lie Groups. It is based on an action principle in which the Lagrangian is required to transform as invariant under the action of a Lie Group. This leads to a set of invariance identities which the functional derivatives of the Lagrangian must satisfy. These identities, together with the equations of motion, yield conservation laws and reveal the structure of the theory without explicit knowledge of the Lagrangian. Theories derived in this way are extremely general and a prescription for constructing simpler, minimally coupled Lagrangians by the use of the invariance identities is given.

The work deals separately with an arbitrary internal group and with the Poincare Group but draws many parallels between them. The Poincare Group is intimately related to gravity and allows a general theory to be formulated in which it is possible to discern clearly the roles of the curvature and torsion. It also transpires that certain constraints (or a special choice of Lagrangian) must be imposed to ensure the conservation of the gravitational stress tensors. Finally, a number of specific theories which naturally suggest themselves are analysed in terms of the general theory.

To Mama Resnik
whose guiding hand
will never be forgotten.

DECLARATION

I declare that this dissertation is my own, unaided work.
It is being submitted for the degree of Master of Science
in the University of the Witwatersrand, Johannesburg. It
has not been submitted before for any degree or examination
in any other University.



R.W. Janse van Rensburg

23rd day of October, 1981.

CONTENTS

	Page
Preface	.v.
INTRODUCTION	.1
NOTATION	.8

PART I: Preliminaries

CHAPTER 1: Group Theory

1.1	Introduction	.10
1.2	Groups: Basic Definitions and Terminology	.11
1.3	Representations	.12
1.4	Topological Groups	.14
1.5	Lie Groups	.16
1.6	Structure Constants	.17
1.7	Group Generators	.19
1.8	The Exponential Form of the Representations	.21
1.9	The Lie Algebra of a Lie Group	.22
1.10	Group Invariants	.23
1.11	Groups of Special Interest.	.24

Appendix 1: Group Transformations of the Generators	.30
---	-----

CHAPTER 2: Global Gauge Invariance and the Lagrangian Formalism

2.1	Introduction	.31
2.2	The Free Fields and the Group Action	.32

2.3	The Action Principle and the Equations of Motion . . .	.33
2.4	Symmetry Transformations	.36
2.5	Internal Symmetry.	.36
2.6	The Conservation Laws	.39
2.7	The Poincaré Group	.41

PART II : The Internal Theory

CHAPTER 3: The Local Gauge Symmetry of an Internal Group

3.1	Introduction	.48
3.2	Local Symmetry and the Gauge Connection	.49
3.3	The Invariance Identities of the Fully Dynamic Theory	.52
3.4	The Gauge Potentials	.55
3.5	The Gauge Curvature and the Gauge Field Strength	.59
3.6	The Complete Internal System	.60
3.7	The Current and its Conservation.	.62
3.8	Minimal Coupling	.63
3.9	A Particular Lagrangian	.68
3.10	The Generalized Lorentz Condition	.70

Appendix 3A:	The Construction of the Invariance Identities of an Internal Group.	.73
Appendix 3B:	The Transformation Laws of the Concomitants	.75
Appendix 3C:	The Tensorial Concomitants.	.77

CHAPTER 4: The Physical and Geometrical Properties of the Gauge Potentials

4.1	4.1 Introduction	.78
4.2	The Action of the Poincaré Group.	.78
4.3	Energy and Momentum	.79
4.4	Angular Momentum	.83
4.5	The Geometrical Interpretation	.87

PART III: Gravity as a Gauge Theory

CHAPTER 5: The Gauge Fields of the Poincaré Group

5.1	Introduction	.91
5.2	The Base Manifold	.92
5.3	Translational Invariance	.96
5.4	Local Lorentz Transformations and Representations	.105
5.5	The Gauge Fields of the Poincaré Group	.107
5.6	Connections, Curvatures and Torsions	.109

CHAPTER 6: Gravity

6.1	Introduction	.114
6.2	The Invariance Identities	.115
6.3	The Identities in Covariant Form.	.118
6.4	The Equations of Motion	.120
6.5	The Stress Tensors and their Conservation	.123
6.6	Minimal Coupling	.126
6.7	Detailed Conservation	.132
6.8	Conservation of Spin-Stress and the Symmetry of the Stress Tensors	.133
6.9	Summary	.134
Appendix 6A: The Derivation of the Invariance Identities.		.136
Appendix 6B: The Transformation Laws of the Concomitants		.142
Appendix 6C: The Tensorial Concomitants		.144

CHAPTER 7: Applications: Some Particular Theories of Gravity

7.1	Introduction	.145
7.2	The General Case	.145
7.3	Einstein's Theory	.146
7.4	The Einstein-Cartan Theory	.148

7.5	The Quadratic Theory	150
7.6	Theories which Include Torsion in a Non-trivial Way. . .	151

CONCLUSION	155
----------------------	-----

BIBLIOGRAPHY	158
------------------------	-----

Preface

Gauge theory has become a fundamental part of the modern description of interacting fields and should therefore form part of every theoretical physicist's basic knowledge.

This Dissertation evolved out of the author's need to familiarize himself with the concepts and tenets of the theory. In this the author wishes to thank his supervisors, Professor David Mason and Dr. Nigel Bishop of the Applied Mathematics Department, University of the Witwatersrand, who allowed him a free hand to explore and discover. This work is the fruit of their confidence in him.

The author also wishes to thank especially Professor Hanno Rund whose standards of excellence remain a source of inspiration.

INTRODUCTION

In the early part of this century Herman Weyl formulated a theory in which the presence of an electromagnetic field altered the scales of measurement of geometric lengths. This was an attempt to unify the only two forces known at the time, namely electromagnetism and gravity. Needless to say the attempt failed but its name lived on - Gauge Theory. Later on Weyl went on to introduce many of the concepts involved in what we now know as gauge theory (Weyl (1922), (1950)). The name itself has become somewhat misleading since the scope of the theory has broadened very considerably and no longer deals only with electrodynamics and gravity.

The generalization to arbitrary groups, of which electrodynamics is a particularly simple instance, was made by Yang and Mills (1954) in the context of isotopic spin. They showed that an interacting theory, conceptually similar to electrodynamics, could be formulated by 'gauging' the group of isotopic spin rotations and made it clear that the technique could be applied to other groups.

This was taken up in 1955 by Utiyama who generalized the work of Yang and Mills and gave a treatment of gravity as the gauge theory of the Lorentz Group although he did not consider translations. He seems also to have been one of the first to recognise that gauge theory is the 'common language' of interactions in that they could all be formulated as gauge theories.

This work was further extended and refined by Kibble (1961) and Sciama (1962). Kibble gauged the full Poincaré Group in his theory of gravity while Sciama concentrated on the gravitational role of the intrinsic spin of matter which arises naturally in a gauge theory of gravity based on the Poincaré Group. He argued that there is an analogy between spin and electric charge; to the extent that the spin gives rise to a force which would deflect a particle from the geodesic along which Einstein's General Relativity would have it move. The spin represents an internal structure of matter which is not recognised in Einsteinian gravity and, since it is known that elementary particles carry spin, it is clear that this theory of gravity would have to be broadened somewhat to encompass this phenomenon.

On the macroscopic scale the intrinsic spins of the material particles may well average out in which case Einsteinian gravity is probably adequate. However, on the microscopic scale this is not the case and a more comprehensive theory is needed. Such a theory must be able to reproduce the Einsteinian predictions under the appropriate circumstances but must also be applicable in extreme situations as may exist in certain cosmic phenomena, for example a neutron star where the neutron spins may be aligned. Furthermore, as an issue of principle, we are once more at a point where a unification of all the forces of nature may be contemplated and, as a necessary prerequisite, they must all be given a common

formulation. In this context it turns out, as we shall endeavour to show, that more general theories of gravity arise when the Poincaré Group is gauged and it is our hope that one of these will be related to the 'correct' theory.

After the historical developments we have outlined work concentrated mainly on the quantum field theories which stem from gauged groups in particular their renormalisability. In this context gravity proved to be highly non-renormalisable - a problem which is still unsolved (Deser and van Nieuwenhuizen (1974)). However, we will not concern ourselves with these issues but concentrate only on the

classical aspects of the theory. Suffice it to say that, in the author's opinion this indicates that our search for the correct theory has far to go still. We will add some further comments on this matter in the Conclusion to this work.

This work originally arose as an attempt to understand gauge theories and has emerged as a systematic analysis of the symmetries embodied by a gauge group. We have based the theory on an action principle, a group and its representations and a manifold on which these representations constitute the material fields of the theory. These fields and their derivatives are used as the arguments of the Lagrangian to which the action principle is applied to obtain the equations of motion. The theory is then made to carry the group symmetry by demanding that the Lagrangian transform as a group invariant when the material fields are subjected to the action of the group. Our analysis is based on a set of invariance identities which the functional derivatives of the Lagrangian must satisfy in order that it be a group invariant. These identities turn out to be powerful tools which we use to uncover the overall structure and content of the theory without having to know the detailed structure of the Lagrangian. We are then in a position to postpone the choice of the Lagrangian to the end and to compare theories based on different Lagrangians. This is particularly important in the case of gravity because of its inherent complexity, but also in general because we always begin with the most general case - not making any initial assumptions about things like minimal coupling. Rather we let the structure of a minimally coupled Lagrangian be dictated by the identities themselves.

This work is divided into three parts: Part I consists of Chapters 1 and 2 and deals with general preliminaries; Part II consists of Chapters 3 and 4 and concerns the so-called Internal Theorem and finally, Part III consists of Chapters 5, 6 and 7 and deals with Gravity.

In Chapter 1 we give a brief outline of standard group theory beginning with the group axioms and proceeding on to topological groups. This

chapter serves to introduce the group theoretical concepts and results we will need for the development of the later theory and ends with a discussion of the basic properties of the Poincaré Group. In Chapter 2 we define the material fields as group representation spaces attached to each point of a manifold - in this case a Minkowski manifold characterised by the Minkowski metric at each point. We then introduce the action principle by using these fields and their derivatives as the arguments of a Lagrangian from which we deduce the equations of motion. We also define the global action of the group as that in which the same group operator acts on all its representations at all points of the manifold. The invariance of the theory under such an action together with the equations of motion allows us to define certain conserved canonical quantities. We deal with two cases: that of an arbitrary internal group, in which case the conserved quantity is a generalised current (in the sense of an electrical current) and the Poincaré Group, in which case the conserved quantities are energy-momentum and angular momentum. We also identify the intrinsic spin of the matter fields by analogy with orbital angular momentum since they both appear in the same conservation law. It also transpires that the divergence of the intrinsic spin gives rise to a non-symmetric energy-momentum tensor for the material field.

Part II begins with Chapter 3 in which we concentrate on an internal group which operates only on the components of the matter field and not on the co-ordinates as the Poincaré Group does. We 'gauge' the group by allowing its operators to depend differentiably on the co-ordinates of the manifold. This means that different group operators are acting at different points of the manifold and we say that the group is acting 'locally'. Under such an action the theory loses its invariance because it depends on the derivatives of the matter fields but it turns out that we can restore the invariance by introducing a set of vector fields, one for each dimension of the group, whose transformation properties are such that the unwanted contributions of the field derivatives are cancelled exactly. We then include these fields and their derivatives, together with the

matter field, into a general Lagrangian which we use to calculate the invariance identities. We define a generalised covariant current which is found to act as a source in the new equations of motion and whose detailed structure is given by the identities. We also deduce its conservation law and find that, in the general case, it is not conserved without any further assumptions being made. However, taking advantage of the transformation properties of the functional derivatives of the Lagrangian we define a minimally coupled Lagrangian in which the current is conserved. We find also that the gauge fields carry a gauge dependent charge which is not present in electrodynamics and which represents an additional non-linearity in the theory. This theory, as a whole, gives a prescription for coupling together the matter and gauge fields and to explore this further we calculate, in Chapter 4, the canonical energy-momentum and angular momentum tensors associated with the gauge fields. We do this by taking the matter fields to be elements of a Lorentz representation on which the Poincaré Group acts globally. The conservation laws then show how energy, momentum and angular momentum are exchanged between the fields. It is found that there is no covariant intrinsic spin which can be associated with the gauge fields - as far as the conservation of angular momentum is concerned they make only an orbital contribution.

Part III concerns itself with gravity as the gauge theory of the Poincaré group and in Chapter 5 we discuss the gauge fields and the matter fields which are required for the local action of this group. At the outset the theory is formulated on a four-dimensional manifold which carries a symmetric, non-singular metric and a non-symmetric affine connection. We first present general co-ordinate transformations as the local action of the translation group and show how its gauge fields may be interpreted as a field of orthonormal vectors, known as a tetrad field, which are unique only up to Lorentz transformations and which may be used to define the manifold metric. We exploit the non-uniqueness to define local Lorentz representations which then constitute the matter fields of the theory. We are thus able to gauge the translation group and the Lorentz group simultaneously.

To gauge the Lorentz Group we introduce a set of Lorentz 'potentials' which are also manifold vectors and we next determine the manifold connection in terms of the Lorentz potentials and the tetrads by demanding that the covariant derivative of the tetrad vanishes identically. This also allows us to write the manifold torsion and curvature in terms of the gauge fields which then determine the geometric properties of the manifold completely.

In Chapter 6 we calculate the invariance identities from a general Lagrangian which contains all the fields and their derivatives as arguments. Our analysis proceeds along lines similar to that of the internal theory in Chapter 3 and we draw roughly similar conclusions. We define covariant stress tensors for both the tetrads ('stress') and the Lorentz potentials ('spin-stress') and find that they are sources in the field equations. In addition, they are not conserved a priori but can be made so by a special choice of Lagrangian or by the imposition of certain constraints. These are apparently new results, being revealed by our generalised systematic analysis, and they reduce to the usual ones in the standard theories. We then give a prescription for constructing a minimally coupled Lagrangian in which the roles of the individual fields are clearly separated. We find a stress tensor and a spin-stress tensor associated with each of the fields: matter, tetrad and Lorentz potential. The conservation laws then show clearly how each field affects and is affected by the others. The identities also give information about the symmetry properties of the stress tensors which is closely linked to the conservation of spin-stress. A feature which distinguishes this theory from the internal theory is that the tetrad is a gauge field which is a gauge tensor unlike any of the other gauge fields which are more like connections. This has far-reaching consequences, in particular we may define a covariant spin-stress for it which, in the general case, interacts with the spin-stress of matter. This is in contrast to the Lorentz potential for which no covariant spin-stress may be defined.

Finally, in Chapter 7 we consider a number of theories based on particular Lagrangians. We begin with the simplest, namely the Einstein theory with vanishing torsion, and we see, as expected, that it cannot support a material field with spin. We next include torsion to get the Einstein-Cartan theory and find that the torsion is generated locally by the spin of the matter field but does not propagate as a dynamic field. We find also that the Einstein tensor is not symmetric and that spin-stress is not conserved. We then consider a theory which is quadratic in the curvature devised as a simple means of ensuring the conservation of spin-stress. So far these are all standard results.

Many authors have noted a similarity between the roles of the curvature and the torsion and in this formulation this is made absolutely clear. We therefore consider, finally, a theory in which the usual roles of torsion and curvature are interchanged - the curvature is taken to vanish identically while the torsion is non-zero - the so-called New General Relativity. In this theory the Lorentz stress vanishes identically (the generalized equivalent of the Einstein tensor) to be replaced by the tetrad stress which has been zero in all the theories we have considered thus far. The tetrad stress is symmetric while the matter stress is not and the total stress, which behaves as a source for the propagation of torsion, is conserved. Lastly, the total spin-stress vanishes identically which leads to the spin-stress of matter being algebraically related to the tetrad spin-stress.

It must be emphasized that these are the simplest cases which naturally suggest themselves. We could consider many more exotic cases within the framework presented here - including certain classes of non-minimally coupled Lagrangians as well as hybrids in which both curvature and torsion are dynamic entities. These, however, must remain as the subject of future work.

NOTATION

The fields of the theory are denoted by:

Matter: $\psi^a; \psi^A$

Gauge: $A_a^A; e_\mu^a; \omega_\mu^{ab}$

Gravitational: $g^{\mu\rho}; \Gamma_{\mu\rho}^\nu; R_{\mu\rho}^a; S_{\mu\rho}^a$

Reduced uppercase Latin letters: $A, B, C, D = 1, \dots, M$
are group representation indices; M is the dimension of the representation.

Lower-case Latin letters from the beginning of the alphabet:

$$a, b, c, d, e, f, g, h = 0, \dots, 3$$

are Lorentz vector representation indices. The remaining lower-case letters:

$$k, m, n, p, r, s = 1, \dots, N$$

are an N -dimensional group parameter indices.

The upper-case Greek letters:

$$\Psi; T; X; \Omega$$

are used to denote the functional derivatives of the Lagrangian (the concomitants).

The specific lower-case Greek indices:

$$\alpha, \beta, \gamma, \delta, \epsilon = 1, \dots, P$$

are Lorentz Group representation indices, where P is the dimension of the representation. The remaining lower-case Greek letters

$$\mu, \nu, \rho, \sigma, \tau, \dots = 0, \dots, 3$$

are manifold co-ordinate indices.

For an arbitrary N -dimensional Lie Group the parameters are denoted by:

$$e^k \quad k = 1, \dots, N,$$

the generators in an M -dimensional representation by:

$$T_k^A \quad A, k = 1, \dots, M,$$

and the structure constants by:

$$C_{mn}^k.$$

The Translation Group parameters are:

$$\xi^\mu \quad \mu = 0, \dots, 3,$$

while for the Lorentz Group they are:

$$e^{ab} \quad a, b = 0, \dots, 3.$$

The Lorentz Group generators are

vector representation: S_{ab}^c

P-dim. representation: S_{ab}^a

and its structure constants: C_{ab}^{ef}

The symbol $\left| \right|_0$ enclosing a functional derivative indicates that it is to be evaluated when all the parameters and their derivatives have been set to zero.

The summation convention will be used throughout the dissertation

PART I: Preliminaries

CHAPTER 1

GROUP THEORY1.1 Introduction

This chapter serves to present a very brief survey of groups and their representations. It has been kept to a minimum, dealing only with that which is necessary for the later development of the general theory. For further details and a more rigorous treatment the reader is referred to the very extensive literature a selection of which is given in the bibliography.

We begin by defining an abstract group and listing some of its properties. The representation of a group is the means by which it may be realised in some way as an operation on some physical or mathematical object. In our case we will deal exclusively with linear representations in which the group elements are matrices operating on vectors in spaces of suitable dimensions. By introducing homomorphisms between groups we may define many different representations of the same group.

The theory to be developed in the sequel is based on groups which have an infinite number of elements and which admit a topological structure - the so-called Lie Groups. For these groups we may define a set of numbers called the group structure constants which determine the group uniquely. It is also found that any element of such a group in a particular representation is determined by a special set of matrices called the generators associated with that representation. The structure constants and the generators are therefore of principal concern in the theory and together they may be used to define the Lie Algebra of the group.

Finally we discuss in more detail two groups which will concern us later on: the Lorentz Group and the Translation Group. Together these constitute the Poincaré Group which, as we shall elaborate, is fundamentally related to gravity.

1.2 Groups: Basic Definitions and Terminology

A group G is defined to be a set of elements together with a binary operation $\circ$ which satisfies the following axioms:

AI: If $a, b \in G$ then $a \circ b \in G$ (closure)

AII: $a \circ (b \circ c) = (a \circ b) \circ c$ for all $a, b, c \in G$
(associativity)

AIII: There exists a unique element $e \in G$ such that

$$e \circ a = a \circ e = a \quad \text{for all } a \in G$$

e is called the *identity element* of G .

AIV: For each $a \in G$ there exists a unique element $b \in G$ such that

$$a \circ b = b \circ a = e$$

b is called the *inverse* of a and is denoted by a^{-1} .

If the composition rule is commutative i.e. if

$$a \circ b = b \circ a \quad \text{for all } a, b \in G$$

then G is said to be *Abelian*. We will see that for the more interesting groups this is not the case.

An immediate classification of G is given by the number of its elements. If they are finite in number then G is called *finite*, if they are denumerably infinite (countable) then G is called *discrete* and if they are uncountable then G is *infinite*.

A subset $H \subset G$ is called a *subgroup* of G if its elements also form a group under the same composition rule.

If H is a subgroup of G then we can define a *conjugate subgroup* $H_c(g)$ of H for each $g \in G$ by

$$H_c(g) = \{g \circ h \circ g^{-1} \mid h \in H\} = g \circ H \circ g^{-1}.$$

The subgroup H is said to be *invariant* if

$$H_c(g) = H \quad \text{for all } g \in G.$$

G is called *simple* if it has no invariant subgroups and *semi-simple* if it has no abelian invariant subgroups. In a later section we will give Cartan's criterion for establishing the semi-simplicity of an infinite group.

A group G is said to be *homomorphic* to another group G' if there exists a mapping (possibly many to one) which preserves the composition rule

If $a, b, c \in G$ have images $a', b', c' \in G'$ and their respective composition rules are $\circ$ and $\circ'$ then

$$a \circ b = c \Rightarrow a' \circ' b' = c'.$$

If this map is one-to-one then G and G' are *isomorphic*.

1.3 Representations

So far the group G is an abstract set with the property that the composition of two elements yields a third and that each element has a unique inverse defined in terms of a unique identity. Such a structure may be realised by interpreting the elements as operators:

whose composition rule has the same properties. We will work with linear operators in the form of square matrices in which case the composition rule becomes simply matrix multiplication. It is easy to see that the set of all square matrices of given dimension forms a group under multiplication. We will usually work with subgroups of this general matrix group defined by having some restriction on the matrices (for example having their determinants equal to one) which is preserved under multiplication. Thus a *representation* will be a homomorphism between the abstract group and a subgroup of the general matrix group such that:

if $a \in G$ and a has the image $D(a)$ where $D(a)$ is an $n \times n$ matrix, then we must have

$$a \circ b = c \Rightarrow D(a) \times D(b) = D(a \circ b) = D(c)$$

and $D(e) = I$ the $n \times n$ identity matrix.

The objects on which the matrix group acts are of course n - dimensional vectors occupying a vector space V_n and we will denote them by

$$\psi^A \quad A = 1, \dots, n$$

and the group action is then given by

$$\bar{\psi}^A = D^A_B \psi^B.$$

As often happens the group itself may be defined by a set of matrices and we then speak of these matrices as the *self-representation* of the group.

We may also define combinations of representations in the same way as vector spaces may be combined into direct sums and products. This leads to the concept of *reducibility* in which a compound representation may be reduced into the direct sum of a set of representations of

lower dimensionality. However, we shall not concern ourselves to any great extent with these aspects (although they are of fundamental importance) except to introduce the direct product representation of two groups.

Suppose we have two groups G and G' represented in the spaces V_n and V_m by the matrices D^a_{α} and D^b_{β} respectively, where $\alpha, \beta = 1, \dots, n$ and $a, b = 1, \dots, m$. Then the direct product representation operates in the space $V_n \times V_m$ whose vectors are denoted by the 'mixed' quantities ψ^{ab} . In addition, the representation matrices are required to commute. The order of the indices is immaterial. We have then the direct product group denoted by $G \times G'$ whose action on the elements of $V_n \times V_m$ is given by:

$$\bar{\psi}^{ab} = D^a_{\alpha} D^b_{\beta} \psi^{\alpha\beta}$$

If G' is G itself then the group so obtained is, in general, reducible.

In addition to the representation space V_n we may also introduce the dual space V_n^* inhabited by the vectors $\bar{\psi}_A$ which are such that the inner product $\bar{\psi}_A \psi^A$ is invariant under the group operations. It follows immediately that the group action on these vectors is given by:

$$\bar{\psi}_A = D^{-1B}_A \bar{\psi}_B$$

which is well defined since the inverse D^{-1B}_A always exists.

1.4 Topological Groups

We now specialise to infinite groups which are such that we may define a topological structure on their elements. This means that we are able to define a set τ of subsets of G such that each element of G is in at least one subset, the null set and G itself are in τ

and the unions and intersections of elements of τ are also elements of τ . Such a topology allows us to define the concept of *continuity* in the group. If we view the composition law

$$a \circ b = c \quad a, b, c \in G$$

as a map: $\varphi : G \times G \rightarrow G$

given by

$$c = \varphi(a, b)$$

then φ may be regarded as continuous.

Continuity also allows us to define a *connected* group as one in which every pair of elements may be joined by a continuous curve lying entirely within the group. It is *simply-connected* if every closed curve can be continuously deformed to a point. It may also occur that there are more than two distinct closed curves which cannot be deformed either into a single point or into each other (they are said to be *homotopically inequivalent*). If there are m such curves then the group is said to be *m-fold connected*.

The topological structure also allows us to speak about *neighbourhoods* of elements of the group. The simplest kind of topological group is one for which there exists an isomorphism between points of its neighbourhoods and points of open neighbourhoods in an n dimensional Euclidean space R^n . Any set which satisfies this requirement is called a *manifold* and so the space made up of the group elements becomes an n dimensional manifold.

1.5 Lie Groups

A topological group G which is an n dimensional manifold is an n dimensional *Lie Group*, provided the manifold is C^k , $k \geq 3$.

The one-to-one mapping between the group and R^n allows us to assign co-ordinates to the elements of the group. Thus each $a \in G$ is labelled by a set of numbers a^i $i = 1, \dots, n$ which we shall call its *parameters* and we suppose that the dimension n is such that the $a^1, \dots, a^n$ are *essential* in that none may be expressed in terms of the others.

The group composition law defines a third element from the composition of any two. We would thus expect that the parameters of the third element be functionally related to the parameters of the original two, i.e.

$$\begin{aligned} \text{if} \quad & c = a \circ b \\ \text{then} \quad & c^i = \varphi^i(a^j, b^k). \end{aligned} \quad (1.1)$$

It can then be shown that the set of parameter composition functions φ^i are analytic in their arguments (which condition is often used to define a Lie Group).

In the group the identity e is a special element and we will always choose our correspondence so that it is mapped onto the origin of R^n , i.e. it has all its parameters equal to zero. Note also that if U is a neighbourhood of some element $g \in G$ then $g^{-1}U$ is a neighbourhood of e and so we need only work with neighbourhoods of e when we investigate the local properties of the group. The global properties (connectedness etc.) can only be analysed by means of non-local objects such as continuous curves.

1.6 Structure Constants

The group is now a manifold whose points are continuously related to each other by the composition law (1.1) and we may ask how the group axioms restrict the composition functions φ^i .

Firstly, we have an identity (AIII) which obeys

$$e \circ e = e$$

and has all its parameters zero. This demands that

$$\varphi^i(0,0) = 0; \quad (1.2)$$

also

$$e \circ a = a \circ e = a \quad \text{for all } a \in G$$

hence

$$\varphi^i(a^j, 0) = a^i \quad (1.3)$$

and

$$\varphi^i(0, a^j) = a^i. \quad (1.4)$$

We also have an inverse (AIV):

$$a \circ a^{-1} = a^{-1} \circ a = e$$

so

$$\varphi^i(a^j, a^{-1} k) = \varphi^i(a^{-1} k, a^j) = 0 \quad (1.5)$$

where $a^{-1} k$ are the parameters of a^{-1} .

Since φ^i is analytic we may develop it in a power series:

$$\begin{aligned}\varphi^i(a,b) = & m^i + k^i_j a^j + l^i_j b^j + g^i_{jk} a^j a^k + f^i_{jk} a^j b^k + \\ & + h^i_{jkl} b^j b^k + \dots\end{aligned}\quad (1.6)$$

By (1.2) we can drop the first term. Using (1.3) and (1.4) and equating coefficients of a^i and b^j we get

$$k^i_j = l^i_j = g^i_j \quad (1.7)$$

while all coefficients of terms containing only a's or only b's vanish. (1.6) becomes therefore

$$\varphi^i(a,b) = a^i + b^i + f^i_{jk} a^j b^k + O(3) \quad (1.8)$$

where $O(3)$ denotes third and higher order terms.

The associativity property (AII):

$$a \circ (b \circ c) = (a \circ b) \circ c$$

gives

$$\varphi^m(a^i, \varphi^k(b^j, c^n)) = \varphi^m(\varphi^i(a^j, b^n), c^k). \quad (1.9)$$

By using (1.8) and defining

$$C^i_{jk} = f^i_{jk} - f^i_{kj} \quad (1.10)$$

it can be shown that the C^i_{jk} obey

$$C^k_{jm} C^i_{kn} + C^k_{mn} C^i_{jk} + C^k_{nj} C^i_{km} = 0 \quad (1.11)$$

which is called the *Jacobi Identity*.

Note that the C^i_{jk} are anti-symmetric in their lower indices.

The φ^i are not yet completely determined but it turns out that they satisfy a differential equation the integrability condition for which depends on being able to find a set of numbers C_{jk}^i which satisfy (1.11). Having found such a set we may construct the φ^i and hence a Lie Group. The C_{jk}^i therefore determine the group and they are called the *group structure constants*. Note that for an abelian group the ε_{jk}^i in (1.8) are constrained to be symmetric and hence, by (1.10), such a group will have vanishing structure constants.

For future reference we introduce the *Cartan-Killing metric*:

$$G_{jk} = C_{jm}^i C_{ik}^m \quad (1.12)$$

which can be shown to be symmetric by using the Jacobi Identity. Furthermore, if G_{jk} is non-singular then the group is semi-simple - this being Cartan's criterion of semi-simplicity.

1.7 The Group Generators

All of the above concerns only the parameters i.e. the labelling of the abstract group elements and is therefore independent of any actual representation. Consider now a representation of G in terms of $m \times m$ matrices. The matrix which corresponds to the group element having parameters a^i , $i = 1, \dots, n$ we denote by:

$$D_{\mathbf{m}}^A(a^i) \quad A, \mathbf{m} = 1, \dots, n = \dim. \text{ of the rep.}$$

since it will depend functionally on these parameters. The group identity corresponds to the matrix identity:

$$D_{\mathbf{m}}^A(0) = \delta_{\mathbf{m}}^A.$$

We can expand $D_{\mathbf{m}}^A$ in a Taylor series about the identity:

$$D_{\mathbf{m}}^A(a) = \delta_{\mathbf{m}}^A + a^i T_i^A_{\mathbf{m}} + O(a^2) \quad (1.13)$$

where

$$T_{1\mu}^A = \left(\frac{\partial D_{\mu}^A}{\partial a^1} \right)_{a=0} \quad (1.14)$$

are called the *generators* of the representation. Also, since every D has an inverse, we have

$$\begin{aligned} D^{-1A}_{\mu}(a) &= \delta_{\mu}^A - a^k T_{k\mu}^A + O(a^2) \\ &= D_{\mu}^A(-a). \end{aligned} \quad (1.15)$$

In the one dimensional scalar representation all the generators vanish.

Next we derive a commutation relation which these generators satisfy by forming the commutator of two matrix operators with arbitrary parameters:

$$\begin{aligned} D_{\mu}^A(a) D_{\nu}^B(b) - D_{\nu}^B(b) D_{\mu}^A(a) \\ = D_{\nu}^B(c_1) - D_{\mu}^A(c_2) \quad (\text{by closure}) \end{aligned} \quad (1.16)$$

where, by (1.8),

$$\begin{aligned} c_1^i &= a^i + b^i + f_{jk}^i a^j b^k + \dots \\ c_2^i &= a^i + b^i + f_{kj}^i a^j b^k + \dots \end{aligned} \quad (1.17)$$

If we now expand out the matrices in (1.16) in the manner of (1.13), collect second-order terms and use (1.10) and the arbitrariness of the parameters, we find

$$T_{\mu}^A T_{\nu}^B - T_{\nu}^B T_{\mu}^A = C_{\mu\nu}^n T_n^A \quad (1.18)$$

which we shall call the *generator structure relation*.

This relation is of fundamental importance since, given the group structure constants, any set of n square matrices $T_{k n}^A$ satisfying it will serve as generators and so define a representation of the group (we will see how to construct the representation matrices out of the generators in the next section). In particular, the structure constants themselves satisfy the Jacobi Identity which may be re-arranged to take the form (1.18) and hence may also serve as generators:

$$T_{k n}^B = C_{k n}^B \quad (1.19)$$

In so doing they define the *adjoint representation* of the group. This is an important representation, partly because it is always known.

We mention in passing that, because of the structure relation (1.18), the generators themselves transform like constant numerical tensors under the group action (see Appendix 1). Also, an arbitrary parameter transformation which preserves the identity has the form:

$$\bar{a}^i = \bar{a}^i(a^j)$$

$$\bar{a}^i(0) = 0.$$

Under such a reparameterization the generators and structure constants transform as vectors and tensors respectively so ensuring that the structure relation is preserved.

1.8 The Exponential Form of the Representations

Consider a matrix group element $D_n^A(a^k)$ and let m be an arbitrarily large integer. Then, as $m \rightarrow \infty$, the expansion

$$D_n^A(a^k/m) = \delta_n^A + \frac{a^k}{m} T_{k n}^A \quad (1.20)$$

becomes exact. The composition of two elements with parameter $\frac{a^k}{m}$

results in an element with parameter

$$\begin{aligned} c^n &= c^n\left(\frac{a^k}{n}, \frac{a^k}{n}\right) \\ &= 2 \frac{a^n}{n} + O\left(\left(\frac{a}{n}\right)^2\right) \end{aligned} \quad (1.21)$$

by (1.8). Composing n of these elements gives

$$c^n = n \frac{a^n}{n} + O\left(\left(\frac{a}{n}\right)^2\right)$$

which, to first order, is the element $D_{\mathbf{a}}^{\mathbf{A}}(a)$. Hence, in the limit as n tends to infinity,

$$\begin{aligned} D_{\mathbf{a}}^{\mathbf{A}}(a) &= \lim_{n \rightarrow \infty} [D\left(\frac{a}{n}\right)]^n \\ &= \lim_{n \rightarrow \infty} \left[\delta^{\mathbf{A}} + \frac{a^k}{n} T_k^{\mathbf{A}} \right]^n \\ &= \exp\left(a^k T_k^{\mathbf{A}}\right) \end{aligned} \quad (1.22)$$

which is the exponential form of the representation and shows that the generators in any representation are sufficient to determine the group operators in that representation.

1.9 The Lie Algebra of a Lie Group

A *Lie Algebra* is generally defined to be a set of vectors in a vector space V equipped with a binary composition $[\ , \]$ which satisfies the following axioms:

$$\text{BI: } x, y \in V \Rightarrow [x, y] \in V$$

$$\text{BII: } [x, y] = -[y, x]$$

$$\text{BIII: } [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

We have already noted that the generators are vectors under reparameterisation of the group. If we define the binary composition rule by the structure relation written in the form

$$[T_k, T_m] = C_{km}^j T_j \quad (1.23)$$

then, since the C_{km}^j are anti-symmetric in its two lower indices and satisfy the Jacobi Identity, we find that the T_k constitute the basis of a Lie Algebra. The importance of this algebra stems from a theorem which states that there is a direct correspondence between Lie Algebras and connected Lie Groups.

The generators are determined in any representation by an expansion in the neighbourhood of the identity and the fact that they form a Lie Algebra by the group axioms. Hence the Lie Algebra concerns itself only with the *local* properties of the group. Its *global* properties such as connectedness are undetermined and so two groups may have isomorphic Lie Algebras but one may be simply connected while the other is multiply connected. It then turns out that for every multiply connected group G there is a unique simply connected group $\tilde{G}$ whose Lie Algebra is isomorphic to that of G . We call $\tilde{G}$ the *covering group* of G .

1.10 Group Invariants

Schur's first lemma states that any operator which commutes with all of the operators in an irreducible representation can only be a multiple of the identity. It follows that such operators (matrices) must also commute with all of the group generators. We may then choose a complete set of mutually commuting operators and a basis in the representation space in which they are all diagonal. The set of all their eigenvalues in this basis may then be used to classify the representation (e.g. mass and spin). The maximum number of such operators which commute with all of the generators is called the *rank* of the group.

1.11 Groups of Special Interest

We review some of the properties of the Lorentz, Translation and Poincaré Groups since these will be essential to the later development of our general theory. As an illustration of the foregoing theory the Lorentz group serves as an example of a general Lie Group although its notation is more cumbersome than most.

1.11.1 The Homogeneous Lorentz Group

We will define a homogeneous Lorentz transformation as a linear transformation of the co-ordinates in a Minkowski space which preserves the Minkowski metric:

$$\{\eta_{ab}\} = \text{diag}(-1, 1, 1, 1) \quad a, b = 0, \dots, 3. \quad (1.24)$$

If the space has co-ordinates x^a then such a transformation is

$$\bar{x}^a = \Lambda^a_b x^b \quad (1.25)$$

and if this is to preserve the metric then we must have

$$\eta^{ab} = \bar{\eta}^{ab} = \Lambda^a_c \Lambda^b_d \eta^{cd} \quad (1.26)$$

where the inverse of η_{ab} is defined by

$$\eta^{ac} \eta_{cb} = \delta^a_b. \quad (1.27)$$

Using the metric to raise and lower the indices in (1.26) we have

$$\Lambda^a_d \Lambda^d_b = \delta^a_b \quad (1.28)$$

which is $\Lambda^T \Lambda = I$ or $\Lambda^{-1} = \Lambda^T$ where T denotes the transpose. This is the defining condition for matrices of the vector representation of the Lorentz Group; it is also the self-representation in our definition of the group.

A 4×4 matrix has 16 components. Equation (1.28) imposes 10 conditions on these so that Λ has only 6 independent components. We therefore infer that the Lorentz Group has 6 essential parameters. We will denote the group by L_h .

An immediate consequence of (1.28) is that

$$(\det \Lambda)^2 = 1$$

and we find

$$\det \Lambda = \pm 1, \quad (1.29)$$

We must therefore distinguish between the elements of L_h which have their determinants equal to $+1$ and -1 . Further, if we put $a = b = 0$ in (1.28), we have

$$(\Lambda^0_0)^2 = 1 + \sum_{i=1}^3 (\Lambda^i_0)^2 \geq 1 \quad (1.30)$$

so that we have two further classes of element:

$$\Lambda^0_0 \geq 1 \quad \text{or} \quad \Lambda^0_0 \leq -1.$$

We see then that the elements of the Lorentz Group each belong to one of four distinct classes. However, the identity element has

$$\det I = +1 \quad \text{and} \quad I^0_0 \geq 1$$

and so only those elements of L_h which have

$$\det \Lambda = 1 \quad \text{and} \quad \Lambda^0_0 \geq 1$$

form a subgroup. It is called the *proper orthochronous Lorentz Group* since it is a proper subgroup and it preserves time orientations. We shall confine our attention to it exclusively.

1.11.2 The Generators and Structure Constants

An infinitesimal transformation of the co-ordinates may be expressed as

$$\bar{x}^a = (\delta^a_b + e^a_b) x^b \quad (1.31)$$

where, by (1.28) the e^{ab} must satisfy

$$e^{ab} = -e^{ba} \quad (1.32)$$

i.e. the e^{ab} are 6 independent infinitesimal numbers which we shall take as the infinitesimal parameters of the group.

In infinitesimal form we write Λ as (c.f.(1.13))

$$\Lambda^a_b = \delta^a_b + e^{cd} S_{cd b}^a \quad (1.33)$$

where

$$S_{cd b}^a = -S_{dc b}^a \quad (1.34)$$

are the generators of the self-representation. Comparing (1.33) with (1.31) we find

$$e^a_b = e^{cd} S_{cd b}^a$$

from which

$$S_{cd b}^a = \frac{1}{2}(\eta_{db} \delta^a_c - \eta_{cb} \delta^a_d) \quad (1.35)$$

so that we know these generators explicitly. The commutator of a pair of generators is

$$S_{ab f}^e S_{cd g}^f - S_{cd f}^e S_{ab g}^f = C_{ab cd}^{hi} S_{hi g}^e \quad (1.36)$$

where

$$C_{ab}^{hi}{}_{cd} = \frac{1}{2}(\eta_{cb}\delta_{ad}^{hi} - \eta_{db}\delta_{ac}^{hi} + \eta_{ac}\delta_{db}^{hi} - \eta_{ad}\delta_{cb}^{hi}) \quad (1.37)$$

are the group structure constants, and

$$\delta_{ad}^{hi} \equiv \frac{1}{2}(\delta_a^h \delta_d^i - \delta_d^h \delta_a^i). \quad (1.38)$$

Although we have used the self-representation of the group to find the structure constants, we may use the structure relation (1.18) to construct any other representation we find necessary. We will, however, not delve too deeply into this complex subject and merely remark that the Lorentz is doubly connected which is the fundamental reason for the existence of two kinds of representations: *spinors* which rotate into themselves under a spatial rotation of 4π radians, and *tensors* which do so under a rotation of 2π radians. We will make no special distinction between these two but instead denote an arbitrary spin-tensor by ψ^α where the index α will range over the dimension of the representation space. A group transformation is then

$$\bar{\psi}^\alpha = D^\alpha_\beta(\epsilon^{ab}) \psi^\beta \quad (1.39)$$

where the transformation matrix has the form

$$D^\alpha_\beta(\epsilon) = \exp[\epsilon^{ab} S_{ab}{}^\alpha_\beta] \quad (1.40)$$

and the generators satisfy

$$S_{ab}{}^\alpha_\gamma S_{cd}{}^\gamma_\beta - S_{cd}{}^\alpha_\gamma S_{ab}{}^\gamma_\beta = C_{ab}{}^{ef}{}_{cd} S_{ef}{}^\alpha_\beta. \quad (1.41)$$

Note that the parameters ϵ^{ab} are no longer necessarily infinitesimal but they must still satisfy (1.32). We also consider the dual transformation given by:

$$\bar{\psi}_\alpha = D^{-1\beta}_\alpha \psi_\beta \quad (1.42)$$

where

$$C_{ab}^{hi}{}_{cd} = \frac{1}{2}(\eta_{cb}\delta_{ad}^{hi} - \eta_{db}\delta_{ac}^{hi} + \eta_{ac}\delta_{db}^{hi} - \eta_{ad}\delta_{cb}^{hi}) \quad (1.37)$$

are the group structure constants, and

$$\delta_{ad}^{hi} = \frac{1}{2}(\delta_a^h \delta_d^i - \delta_d^h \delta_a^i). \quad (1.38)$$

Although we have used the self-representation of the group to find the structure constants, we may use the structure relation (1.18) to construct any other representation we find necessary. We will, however, not delve too deeply into this complex subject and merely remark that the Lorentz is doubly connected which is the fundamental reason for the existence of two kinds of representations: *spinors* which rotate into themselves under a spatial rotation of 4π radians, and *tensors* which do so under a rotation of 2π radians. We will make no special distinction between these two but instead denote an arbitrary spin-tensor by ψ^a where the index a will range over the dimension of the representation space. A group transformation is then

$$\bar{\psi}^a = D^a_b(\epsilon^{ab}) \psi^b \quad (1.39)$$

where the transformation matrix has the form

$$D^a_b(\epsilon) = \exp[\epsilon^{ab} S_{ab}^a{}_b] \quad (1.40)$$

and the generators satisfy

$$S_{ab}^a{}_c S_{cd}^c{}_b - S_{cd}^c{}_a S_{ab}^a{}_b = C_{ab}^{ef}{}_{cd} S_{ef}^a{}_b. \quad (1.41)$$

Note that the parameters ϵ^{ab} are no longer necessarily infinitesimal but they must still satisfy (1.32). We also consider the dual transformation given by:

$$\bar{\psi}_a = D^{-1a}{}_b \psi_b \quad (1.42)$$

where

$$D^{-1\beta}_{\alpha}(\epsilon^{ab}) = D^{\beta}_{\alpha}(-\epsilon^{ab}) \quad (1.43)$$

is the inverse transformation. This ensures that

$$\psi_{\alpha} \psi^{\alpha} \quad (1.44)$$

is an invariant scalar. We also have the adjoint representation whose vectors carry double indices:

$$\bar{\psi}^{ab} = D^{ab}_{cd}(\epsilon^{ef}) \psi^{cd} \quad (1.45)$$

where

$$D^{ab}_{cd}(\epsilon^{ef}) = \exp[\epsilon^{ef} C_{ef}^{ab}] \quad (1.46)$$

and, finally, we have the scalar representation whose generators all vanish. The group action is therefore, since these objects carry no indices,

$$\bar{\psi} = \psi. \quad (1.47)$$

1.11.3 The Poincaré or Inhomogeneous Lorentz Group

We combine co-ordinate translations with rotations to define the Poincaré Group by its action on the co-ordinates of a Minkowski space:

$$\bar{x}^a = \Lambda^a_b(\epsilon^{cd}) x^b + \xi^a \quad (1.48)$$

which is a 10 parameter transformation. If we make the convention that Lorentz transformations are to be performed before translations then we can denote an element of the group by

$$(\Lambda, \xi)$$

where the identity element is $(I, 0)$. The product of two elements is then found to be

$$(\Lambda_2, \xi_2)(\Lambda_1, \xi_1) = (\Lambda_2 \Lambda_1, \Lambda_2 \xi_1 + \xi_2). \quad (1.49)$$

It is immediately apparent that the commutator of a pure translation and a Lorentz transformation does not vanish:

$$(I, \xi)(\Lambda, 0) - (\Lambda, 0)(I, \xi) = (0, [I - \Lambda]\xi) \quad (1.50)$$

so that these two group operations do not commute. This means that we cannot write the Poincaré Group as a direct product of the Translation and Lorentz groups. We may, however, use the composition law (1.49) to define a *semi-direct product*

$$P = L \ltimes T. \quad (1.51)$$

We will implement the group in the following way: at each point of the Minkowski space we attach a Lorentz representation space. Suppose that ψ_p^α is a representation vector at the point P which has co-ordinates x_p^a . Under a Poincaré transformation the components of ψ_p^α become $\bar{\psi}_p^\beta$ and the co-ordinates of P become $\bar{x}_p^b$, where

$$\bar{\psi}_p^\beta = D^\beta_\alpha(\epsilon^{cd}) \psi_p^\alpha \quad (1.52)$$

and

$$\bar{x}_p^b = \Lambda^b_a(\epsilon^{cd}) x_p^a + \xi^b. \quad (1.53)$$

Note that, under pure translations ($\epsilon^{ab}=0$) the components of ψ^α transform as scalars. We will sometimes write $\psi^\alpha(x^a)$ to indicate the values of the components situated at the point with co-ordinates x^a . Finally, by the chain rule and (1.48) we have, for the transformation of the derivatives

$$\frac{\partial \bar{\psi}_p^\alpha}{\partial \bar{x}_p^a} = \Lambda^{-1}{}^b_a D^\alpha_\beta \frac{\partial \psi_p^\beta}{\partial x_p^b}. \quad (1.54)$$

where the identity element is $(I, 0)$. The product of two elements is then found to be

$$(\Lambda_2, \xi_2)(\Lambda_1, \xi_1) = (\Lambda_2 \Lambda_1, \Lambda_2 \xi_1 + \xi_2). \quad (1.49)$$

It is immediately apparent that the commutator of a pure translation and a Lorentz transformation does not vanish:

$$(I, \xi)(\Lambda, 0) - (\Lambda, 0)(I, \xi) = (0, [I - \Lambda]\xi) \quad (1.50)$$

so that these two group operations do not commute. This means that we cannot write the Poincaré Group as a direct product of the Translation and Lorentz groups. We may, however, use the composition law (1.49) to define a *semi-direct product*

$$P = L \ltimes T. \quad (1.51)$$

We will implement the group in the following way: at each point of the Minkowski space we attach a Lorentz representation space. Suppose that ψ_p^a is a representation vector at the point P which has co-ordinates x_p^a . Under a Poincaré transformation the components of ψ_p^a become $\bar{\psi}_p^b$ and the co-ordinates of P become $\bar{x}_p^b$, where

$$\bar{\psi}_p^b = D^b_a(\epsilon^{cd}) \psi_p^a \quad (1.52)$$

and

$$\bar{x}_p^b = \Lambda^b_a(\epsilon^{cd}) x_p^a + \xi^b. \quad (1.53)$$

Note that, under pure translations ($\epsilon^{ab}=0$) the components of ψ^a transform as scalars. We will sometimes write $\psi^a(x^a)$ to indicate the values of the components situated at the point with co-ordinates x^a . Finally, by the chain rule and (1.48) we have, for the transformation of the derivatives

$$\frac{\partial \bar{\psi}_p^a}{\partial \bar{x}_p^a} = \Lambda^{-1}{}^a_b D^a_c \frac{\partial \psi_p^c}{\partial x_p^b}. \quad (1.54)$$

APPENDIX 1

Group Transformations of the Generators

We may define the transformation laws of 'mixed' group tensors. For example suppose that $M_{k\ n}^A$ transforms in the adjoint representation on its index k and in an arbitrary representation on its indices A, n then

$$M_{k\ n}^A = D_c^A D^{-1D}_n D^{-1m}_k M_{m\ D}^c.$$

In particular, if $M_{k\ n}^A = T_{k\ n}^A$ is a generator, then, since we need only work in a neighbourhood of the identity, we have, to first order

$$\begin{aligned} T_{r\ n}^A &= (\delta_c^A + \epsilon^k T_{k\ c}^A)(\delta_n^D - \epsilon^n T_{n\ D}^D)(\delta_r^m - \epsilon^p C_{p\ r}^m) T_{m\ D}^c \\ &= T_{r\ n}^A + \epsilon^k [T_{k\ c}^A T_{r\ n}^c - T_{r\ c}^A T_{k\ n}^c - C_{k\ r}^m T_{m\ n}^A] \\ &= T_{r\ n}^A \end{aligned}$$

by the structure relation (1.18). The generators are therefore considered to transform as constant numerical tensors. We also treat the structure constants themselves as generators and so it follows that they too may be considered to transform as constant numerical tensors.

It also follows that

$$\overline{(\psi_A T_{k\ n}^A \psi^B)} = D^{-1m}_k (\psi_A T_{m\ n}^A \psi^B)$$

i.e. these quantities are adjoint vectors.

CHAPTER 2 Global Gauge Invariance and the Lagrangian Formalism

2.1 Introduction

In the last chapter we introduce groups and their representations and we will now use these as the fundamental objects on which to base a theory of fields. We first develop a free-field theory in which there are no sources and the fields propagate freely through space-time. We are then in a position to define various kinematically conserved quantities which may be interpreted physically as properties which the fields transport. These properties stem directly from the group invariance by virtue of the Noether Theorem and are readily obtained for any given group.

The equations of motion for Lorentz fields may be deduced from their Poincaré transformation covariance and we will assume that these are known for the representations of interest without any further comment (we will assume that every field belongs to some representation of the Lorentz Group without always explicitly saying so). Such equations of motion may be derived from a variational principle based on a Lagrangian or, conversely, we may always construct a Lagrangian which will reproduce a known equation of motion.

From a gauge point of view the Lagrangian approach has two major advantages: firstly we may ensure that the theory as a whole is gauge covariant by requiring that the action be gauge invariant i.e. it must transform as a group scalar, and secondly, as we will see in later chapters, we may enlarge the action of the group to generate theories of interacting fields.

2.2 The Free Fields and the Group Actions

We will define a field as a group representation vector attached to each point of a base manifold. This manifold is to represent Minkowski space-time and as such it is endowed with the Minkowski metric introduced in the last chapter. We will denote its co-ordinates by x^a , $a = 0, \dots, 3$.

The action of an n parameter Lie Group G on the components of this field is given by:

$$\bar{\psi}^A(x^a) = D^A_B(e^k) \psi^B(x^a) \quad (2.1)$$

where $A, B = 1, \dots, m$ = dimension of the representation. Note that we define the components of ψ^A to be individually Lorentz scalars.

The effect of this transformation is to 'rotate' the representation vectors everywhere by the same amount depending on the values of the parameters. These parameters are constants - independent of the co-ordinates - and we call this the *global* action of the group. The group transformation does not affect the co-ordinates themselves and for this reason it is often referred to as an *internal* group.

On the other hand, the Poincaré Group affects both the co-ordinates and the field components which are now generalized Lorentz vectors. For this reason it is called an *external* group. Its action is given by:

$$\bar{\psi}^a(\bar{x}^a) = D^a_b(e^{cd}) \psi^b(x^b) \quad (2.2)$$

$$\bar{x}^a = \Lambda^a_b(e^{cd}) x^b + \xi^a \quad (2.3)$$

where a, b range over the dimension of the Lorentz representation to which the field belongs. The 10 parameters e^{cd}, ξ^a are constants and this constitutes the *global* action of the Poincaré Group.

Finally we will impose the requirement that the representation vectors be continuous, differentiable functions of the manifold co-ordinates in all regions of interest. We will denote their partial derivatives with respect to a particular co-ordinate system by a comma:

$$\psi_{,a}^{\dots} \equiv \frac{\partial \psi^{\dots}}{\partial x^a} \quad (2.4)$$

In the sequel we will often refer to the $\psi^{\dots}$ collectively as the matter fields.

2.3 The Action Principle and the Equations of Motion

We give a brief account of the derivation of the field equations from the Action Principle, presenting only a simplified version. For a more detailed and rigorous treatment the reader is referred to the literature (see, for example, D. Lovelock and H. Rund, 1975).

Define the action:

$$I \equiv \int_R L d^4x \quad (2.5)$$

where R is some region of space-time and d^4x denotes the elementary four-volume. L is the Lagrangian which we shall suppose has the functional dependence:

$$L = L(\psi^A : \psi_{,a}^A) \quad (2.6)$$

where now ψ^A stands for any collection of fields carrying whichever indices we require. Anticipating certain results we restrict L to contain only the first derivatives of the fields so that the equations of motion be of at most second order, and also to have no explicit dependence on the co-ordinates themselves.

We next subject the field components to infinitesimal variations independently of each other but all constrained to vanish on the boundary ∂R of the region R :

$$\bar{\psi}^A = \psi^A + \delta\psi^A \quad (2.7)$$

where

$$\delta\psi^A = 0 \text{ on } \partial R .$$

We define the functional variation of the action as:

$$\delta I = \int_R \delta L d^4x \quad (2.8)$$

where

$$\delta L = L(\bar{\psi}^A; \bar{\psi}^A_{,a}) - L(\psi^A; \psi^A_{,a}) \quad (2.9)$$

and the functional dependence of L on its arguments is not to be altered. Note also that the co-ordinates are not subject to variation.

The Action Principle now demands that δI vanishes for every choice of the infinitesimal variations $\delta\psi^A$, i.e.

$$\int_R \delta L d^4x = 0 \quad (2.10)$$

To first order we have

$$\delta L = \frac{\partial L}{\partial \psi^A} \delta\psi^A + \frac{\partial L}{\partial \psi^A_{,b}} \delta(\psi^A_{,b}) \quad (2.11)$$

and also, since the derivative commutes with the variation,

$$\begin{aligned} & \frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \delta\psi^A \right) \\ &= \frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \right) \delta\psi^A + \frac{\partial L}{\partial \psi^A_{,b}} \delta\psi^A_{,b} . \end{aligned} \quad (2.12)$$

The variation of the action is therefore:

$$\int_R \delta L d^4x = \int_R \left[\left(\frac{\partial L}{\partial \psi^A} - \frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \right) \right) \delta \psi^A + \frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \delta \psi^A \right) \right] d^4x \quad (2.13)$$

The last term, being a divergence, may be written as an integral over the boundary of R and hence, in view of (2.7) contributes nothing to the variation. We are left with

$$\int_R \left(\frac{\partial L}{\partial \psi^A} - \frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \right) \right) \delta \psi^A d^4x = 0 \quad (2.14)$$

which can hold for arbitrary, independent $\delta \psi^A$ only if

$$\frac{\partial}{\partial x^b} \left(\frac{\partial L}{\partial \psi^A_{,b}} \right) - \frac{\partial L}{\partial \psi^A} = 0 \quad (2.15)$$

These are called the *Euler-Lagrange Equations* which we shall take as the equations of motion. They specify a configuration of the fields (with suitable boundary conditions) which is such that the action is unaltered when they are varied slightly. Note that, since L depends only on the first derivatives of the fields, they can be, at most, of second order in the field derivatives.

We have already stated that our approach is to assume that the field equations are known for the Lorentz representations of interest (except perhaps, those for fields of high spin; see, for example, A. Aurilia and H. Umezawa, 1954). The reader will find discussions of the construction of these equations in Roman (1960), Corson (1953) and many others. It is then a simple task to construct a Lagrangian which will reproduce them in the form of the Euler-Lagrange equations and in the sequel we shall always use such a Lagrangian as the starting point of our discussions. We will, however, only use the equations of motion in the form (2.15).

2.4 Symmetry Transformations

In addition to the field configuration specified by the equations of motion we will demand that the action be invariant under transformations of the field functions induced by a group. This is our fundamental invariance requirement and ensures that the resulting theory embodies the symmetry defined by the group. Physically we may interpret these transformations as connecting pairs of observers who are using fields which differ from each other by generalized 'phase' rotations (c.f. the exponential form of the representations) and, since they are describing the same physical situation, we demand that the predictions they make concerning the outcomes of experiments be independent of these detailed phase differences.

We will see this invariance leads directly to the definitions of generalized conserved currents and charges which characterise the fields. This is called *gauge invariance of the first kind* or *kinematical gauge invariance* and is generated by the global action of the group. It is therefore also known as the Global Theory.

The detailed structure of the Lagrangian remains unspecified but we deduce a set of invariance identities which the functional derivatives of the Lagrangian must satisfy in order for the theory to exhibit the required group invariance. As we will see, these identities are extremely useful in our analysis.

2.5 Internal Symmetry

2.5.1 The Invariance Identity

Under the action of an internal group the Action transform as a scalar

$$T = I \quad (2.16)$$

when the fields transform as:

$$\bar{\psi}^A = D^A_{\ B}(\epsilon^k) \psi^B \quad (2.17)^a$$

$$\bar{\psi}^A_{,a} = D^A_{\ B}(\epsilon^k) \psi^B_{,a} \quad (2.17)^b$$

where the ϵ^k are constant parameters. In detail the transformed action is

$$I = \int_R L(\bar{\psi}^A; \bar{\psi}^A_{,a}) d^4x \quad (2.18)$$

which leads to the Lagrangian transforming under the group as a scalar:

$$L(\bar{\psi}^A; \bar{\psi}^A_{,a}) = L(\psi^B; \psi^B_{,a}) \quad (2.19)$$

We now differentiate both sides of this equation with respect to ϵ^k and then set these parameters to zero. Since the r.h.s. is independent of ϵ^k we have

$$\left(\frac{\partial L}{\partial \bar{\psi}^A} \right)_0 \left(\frac{\partial \bar{\psi}^A}{\partial \epsilon^k} \right)_0 + \left(\frac{\partial L}{\partial \bar{\psi}^A_{,a}} \right)_0 \left(\frac{\partial \bar{\psi}^A_{,a}}{\partial \epsilon^k} \right)_0 = 0 \quad (2.20)$$

(The subscript "0" indicates that the parameters have been set to zero.).

$\epsilon^k = 0$ corresponds to the identity transformation, hence:

$$\left(\frac{\partial L}{\partial \bar{\psi}^A} \right)_0 = \frac{\partial L}{\partial \psi^A} \quad (2.21)$$

$$\left(\frac{\partial L}{\partial \bar{\psi}^A_{,a}} \right)_0 = \frac{\partial L}{\partial \psi^A_{,a}} \quad (2.22)$$

Define the quantities

$$\psi_A \equiv \frac{\partial L}{\partial \dot{\psi}^A} \quad (2.23)$$

and

$$\psi_A^{a} \equiv \frac{\partial L}{\partial \dot{\psi}_{,a}^A} \quad (2.24)$$

which are called the *concomitants* associated with the Lagrangian.

By (2.17)^a

$$\begin{aligned} \left(\frac{\partial \bar{\psi}^A}{\partial \epsilon^k} \right)_0 &= \left(\frac{\partial D^A_B}{\partial \epsilon^k} \right)_0 \psi^B \\ &= T_{k}^A \psi^B \end{aligned} \quad (2.25)$$

by definition of the group generators. Similarly, by (2.17)^b,

$$\left(\frac{\partial \bar{\psi}_{,a}^A}{\partial \epsilon^k} \right)_0 = T_{k}^A \psi_{,a}^B. \quad (2.26)$$

Putting all of this into (2.20) we have

$$\psi_A T_{k}^A \psi^B + \psi_A^{a} T_{k}^A \psi_{,a}^B = 0 \quad (2.27)$$

which is the invariance identity associated with internal global group transformations. By construction any Lagrangian whose derivatives satisfy this identity embodies the group symmetry. Note that it is linear in the Lagrangian.

The concomitants themselves also transform under the group action and their transformation rule may be found by differentiating (2.19) with respect to the field and its derivative.

Thus, differentiating with respect to ψ^B :

$$\frac{\partial \bar{L}}{\partial \bar{\psi}^A} \frac{\partial \bar{\psi}^A}{\partial \psi^B} = \frac{\partial L}{\partial \psi^B}$$

and using (2.17)^a and (2.23) we get

$$\bar{\psi}_A D^A_B = \psi_B$$

or

$$\bar{\psi}_A = D^{-1B}_A \psi_B. \quad (2.28)$$

Hence $\bar{\psi}_A$ is a group vector transforming in the dual representation. Similarly, differentiating (2.19) with respect to $\psi^B_{,b}$ and using (2.17)^b we find:

$$\bar{\psi}_A^b = D^{-1B}_A \psi_A^b. \quad (2.29)$$

As far as the group is concerned the manifold index b is merely a labelling index. We will see later when we deal with Poincaré transformations that it is in fact a Lorentz vector index so that the concomitant transforms as

$$\bar{\psi}_A^b = \Lambda^b_a \psi_A^a \quad (2.30)$$

under a Lorentz transformation.

In view of these transformation properties we see that the identity itself is an adjoint tensor equation and is therefore gauge covariant.

2.6 The Conservation laws

Using (2.23) and (2.24) the equations of motion may be written:

$$\psi_{A,b}^b - \psi_A = 0 \quad (2.31)$$

which is clearly gauge covariant. We will suppose that the field configuration is such that it is satisfied.

We have:

$$\begin{aligned}
 (\psi_A^a T_k^A \psi^a)_{,a} &= \psi_{A,a}^a T_k^A \psi^a + \psi_A^a T_k^A \psi^a_{,a} \\
 &= \psi_A^a T_k^A \psi^a + \psi_A^a T_k^A \psi^a_{,a} \quad (\text{by (2.31)}) \\
 &= 0
 \end{aligned}$$

by the invariance identity. We therefore define a strictly conserved current:

$$J_k^a = \psi_A^a T_k^A \psi^a \quad (2.32)$$

whose divergence vanishes:

$$J_{k,a}^a = 0 \quad (2.33)$$

Note that J_k^a is simultaneously a set of n Lorentz vectors and a set of four adjoint vectors.

If we integrate (2.33) and suppose that the fields go to zero sufficiently fast in the approach to spatial infinity then we find that

$$Q_k = \int_{\sigma} J_k^a d\sigma_a \quad (2.34)$$

where σ is an infinite space-like hypersurface, are a set of n conserved charges. Their number being exactly equal to the dimension of the group. We see therefore that the invariance identity together with the equation of motion gives rise to a set of conserved currents and charges. Because they entail the use of the equations of motion these conservation laws are called *weak conservation laws*. Finally, note that the scalar representation of the group does not carry charge.

2.7 The Poincaré Group

The group now affects both the co-ordinates and the field components so that the calculations are a little more complex than for the internal group. The transformations are:

$$\bar{\psi}^{\alpha} = D^{\alpha}_{\beta} (\epsilon^{cd}) \psi^{\beta} \quad (2.35)$$

$$\bar{\psi}^{\alpha}_{,b} = \Lambda^{-1a}_{\ b} D^{\alpha}_{\beta} \psi^{\beta}_{,a} \quad (2.36)$$

$$\bar{x}^a = \Lambda^a_{\ b} (\epsilon^{cd}) x^b + \xi^a. \quad (2.37)$$

Since the action is required to be a Lorentz scalar and the group affects the co-ordinates the Lagrangian must be a *scalar density* $\mathcal{L}$, transforming as:

$$\bar{\mathcal{L}} = J \mathcal{L} \quad (2.38)$$

where J is the *Jacobian* of the transformation given by:

$$J = \det \left(\frac{\partial \bar{x}^a}{\partial x^b} \right). \quad (2.39)$$

However, for a Poincaré transformation (2.37),

$$J = \det (\Lambda^a_{\ b}) = +1$$

since we are dealing only with proper Lorentz transformations. We may factor a constant scalar density e out of the Lagrangian which we then treat simply as an ordinary scalar:

$$\mathcal{L} = eL \quad (2.40)$$

and

$$\bar{\theta} = \theta = \text{constant} . \quad (2.41)$$

2.7.1 The Invariance Identities of the Global Poincaré Group

(a) Translations

Let us suppose first that the Lagrangian explicitly contains the co-ordinates as arguments. Its transformation is

$$L(\bar{x}^a : \bar{\psi}^a : \bar{\psi}_{,a}^a) = L(x^b : \psi^b : \psi_{,b}^b) . \quad (2.42)$$

Differentiate both sides of this equation with respect to ξ^a and set all the Poincaré parameters to zero:

$$L_b \delta^b_a + \psi_a \left(\frac{\partial \psi^a}{\partial \xi^a} \right)_0 + \psi_a^b \left(\frac{\partial \psi_{,b}^a}{\partial \xi^a} \right)_0 = 0 \quad (2.43)$$

where now L_b denotes the *explicit* derivative of L with respect to x^b . The concomitants in this equation are defined as

$$\psi_a = \frac{\partial L}{\partial \psi^a} \quad (2.44)$$

$$\psi_a^b = \frac{\partial L}{\partial \psi_{,b}^a} . \quad (2.45)$$

Now, by (2.35),

$$\frac{\partial \bar{\psi}^a}{\partial \xi^a} = D^a_b \frac{\partial \psi^b}{\partial \xi^a} = 0$$

end, similarly, by (2.36),

$$\frac{\partial \bar{\psi}^a_c}{\partial \xi^a} \equiv \frac{\partial}{\partial \xi^a} (D^a_b \Lambda^{-1b}_c \psi^b_{,b}) = 0.$$

Hence (2.43) reduces to

$$L_a = 0 \quad (2.46)$$

which may be satisfied identically by omitting the co-ordinates from the argument list of the Lagrangian.

We will use $L_{,a} = \frac{\partial L}{\partial x^a}$ to denote the *total derivative* of the Lagrangian with respect to the co-ordinates, i.e.

$$\frac{\partial L}{\partial x^a} \equiv L_a + \frac{\partial L}{\partial \psi^a_{,a}} \psi^a_{,a} + \frac{\partial L}{\partial \psi^a_{,b}} \psi^a_{,ba}. \quad (2.47)$$

Using (2.46) and the definitions of the concomitants, we may write this as

$$L_{,a} - \psi^a_{,a} \psi^a_{,a} - \psi^b_{,a} \psi^a_{,ba} = 0 \quad (2.48)$$

which we will use as the *translational invariance identity*.

(b) Lorentz Transformations

From now on we will omit the co-ordinates from the Lagrangian the transformation law of which is:

$$L(\bar{\psi}^a : \bar{\psi}^a_{,a}) = L(\psi^b : \psi^b_{,b}). \quad (2.49)$$

Differentiating (2.49) with respect to the parameter ϵ^{cd} and setting all the parameters equal to zero we find:

$$\psi_a \left(\frac{\partial \bar{\psi}^a}{\partial \epsilon^{cd}} \right)_0 + \psi_a^b \left(\frac{\partial \bar{\psi}^a_b}{\partial \epsilon^{cd}} \right)_0 = 0. \quad (2.50)$$

By (2.35) and the definition of the Lorentz generator we have

$$\left(\frac{\partial \bar{\psi}^a}{\partial \epsilon^{cd}} \right)_0 = S_{cd}^a \psi^b. \quad (2.51)$$

Also, by (2.36),

$$\begin{aligned} & \left(\frac{\partial \bar{\psi}^a_b}{\partial \epsilon^{cd}} \right)_0 \\ &= \delta^a_b \left(\frac{\partial D^a_b}{\partial \epsilon^{cd}} \right)_0 \psi^b_{,a} + \delta^a_b \left(\frac{\partial \Lambda^{-1a}_b}{\partial \epsilon^{cd}} \right)_0 \psi^b_{,a} \\ &= S_{cd}^a \psi^b_{,b} - S_{cd}^a \psi^a_{,a} \end{aligned} \quad (2.52)$$

since $(\Lambda^{-1a}_b)_0 = \delta^a_b$ and $(D^a_b)_0 = \delta^a_b$. The equation (2.50) becomes:

$$\psi_a S_{cd}^a \psi^b + \psi_a^b (S_{cd}^a \psi^b_{,b} - S_{cd}^a \psi^a_{,a}) = 0, \quad (2.53)$$

which is the *Lorentz Invariance Identity*. Observe that the generators corresponding to two different representations are present in this identity. This is clearly due to the presence of derivatives in the Lagrangian.

By means similar to that used in the internal theory (last section) it can be shown that both concomitants are tensors and hence both identities, (2.48) and (2.53), are gauge covariant.

2.7.2 The Conservation Laws

(a) Energy and Momentum

The translational invariance identity allows us to define a conserved canonical energy-momentum tensor.

The equation of motion for this case reads:

$$\psi_{\alpha,b}^b - \psi_{\alpha} = 0. \quad (2.54)$$

Also, we have

$$(\psi_{\alpha}^b \psi_{,a}^{\alpha})_{,b} = \psi_{\alpha} \psi_{,a}^{\alpha} + \psi_{\alpha}^b \psi_{,ab}^{\alpha}$$

by the equation of motion. Commuting the derivatives in the last term and using the identity (2.48) we find

$$\begin{aligned} L_{,a} - (\psi_{\alpha}^b \psi_{,a}^{\alpha})_{,b} &= 0 \\ \text{or} \quad \partial_b (L \delta_a^b - \psi_{\alpha}^b \psi_{,a}^{\alpha}) &= 0. \end{aligned} \quad (2.55)$$

Hence define

$$T_a^b = L \delta_a^b - \psi_{\alpha}^b \psi_{,a}^{\alpha}; \quad (2.56)$$

then

$$T_{a,b}^b = 0 \quad (2.57)$$

and T_a^b is strictly conserved. It is called the *canonical energy-momentum tensor*. Note that, as it is defined, it is not symmetric and we will return to this point later after we have obtained the conservation law of angular momentum.

(b) Angular Momentum

Using the equation of motion the first and second terms of the Lorentz identity may be written together so that it becomes:

$$(\psi_a^b S_{cd}^a \psi^b)_{,b} = \psi_a^b S_{cd}^a \psi^a_{,b} . \quad (2.58)$$

Since S_{cd}^a is the generator of the self-representation of the Lorentz group, we know it explicitly:

$$S_{cd}^a = \frac{1}{2}(\eta_{bd} \delta_c^a - \eta_{bc} \delta_d^a)$$

and hence (2.58) becomes

$$\psi_a^b S_{cd}^a \psi^a_{,b} = \frac{1}{2} \psi_a^b \psi^a_{,d} \eta_{bc} - (c \leftrightarrow d) .$$

But, by definition of the energy-momentum tensor, we have

$$\psi_a^b \psi^a_{,d} \eta_{bc} = L \eta_{cd} - T_{cd}$$

and (2.58) is finally:

$$(\psi_a^b S_{cd}^a \psi^b)_{,b} = \frac{1}{2}(T_{cd} - T_{dc}) . \quad (2.59)$$

To obtain a conservation law out of this we observe that

$$\begin{aligned} & (x_c T_d^b - x_d T_c^b)_{,b} \\ &= (\eta_{ca} x^a T_d^b - \eta_{da} x^a T_c^b)_{,b} \\ &= T_{cd} - T_{dc} \quad \text{by (2.57).} \end{aligned}$$

(2.59) can be written:

$$\partial_b M_{cd}^b = 0 . \quad (2.60)$$

where

$$M_{cd}^b = \psi_a^b S_{cd}^a \psi^c - \frac{1}{2}(x_c T_d^b - x_d T_c^b). \quad (2.61)$$

The last term contains the co-ordinates explicitly and represents the *orbital angular momentum* stemming from the spatial rotation of linear momentum (for the spatial components at least) thus we interpret M_{cd}^b as the angular momentum carried by the field. The term containing the generator is then interpreted as an *intrinsic angular momentum* carried by the representation. Note that in the scalar representation it vanishes. The total angular momentum is seen to be conserved and we will call M_{cd}^b the *canonical angular momentum tensor*. Observe that it is antisymmetric in its lower indices.

From (2.59) we see that the antisymmetric part of the energy-momentum tensor is given by the divergence of the intrinsic angular momentum. This, together with the fact that the energy-momentum tensor is not uniquely determined by having a vanishing divergence, may be used to define a symmetric energy-momentum tensor which is also conserved (see Jauch and Rohrlich (1955) p 90 for details). However, it appears more natural to conclude that the presence of spinning matter necessitates the use of a non-symmetric energy-momentum tensor (Sciama (1962); Papapetrou (1951)). We shall again consider both the angular momentum and the linear momentum when we have introduced the dynamic theory (gravity) in Chapter 6.

PART II: The Internal Theory

CHAPTER 3 The Local Gauge Symmetry of an Internal Group

3.1 Introduction

In the previous chapter we considered global gauge invariance in which the theory is required to be covariant when the same group element is applied to all its representations at all points of the space-time manifold. We shall now extend this invariance and demand that the theory be covariant when the group parameters themselves are allowed to be functions of the manifold co-ordinates. In the case different group elements act at separated points on the manifold but we will demand that continuity be preserved, that is, that neighbouring group elements act at neighbouring points. This is only possible for topological groups whose operators are differentiable functions of their parameters which, in turn, are now differentiable functions of the co-ordinates. Differentiability with class C^k , $k \geq 2$ is required.

At first sight this has nothing to do with the interactions between fields but, as we will see, the local theory is in fact a prescription for coupling fields together in a unique manner. The simplest method of coupling has always been to add together Lagrangians representing the free fields which are to be coupled and then to add on an interaction Lagrangian which contains a mixture of the fields. This was generally an ad hoc procedure since there was no general principle which could be used to determine the structure of the interacting Lagrangian uniquely, except possibly the requirement of Lorentz invariance. A promising approach was found in the principle of minimal coupling in electrodynamics in which the derivatives

in the free electron field Lagrangian are modified by the electromagnetic potentials. The electron-electron interaction then occurs through the intermediary of these potentials. The addition of a free potential field Lagrangian completes the system which then constitutes a closed electron-potential dynamic interaction.

The question naturally arises as to whether this successful procedure may be generalised to all interacting systems. At present the answer appears to be in the affirmative, at least in principle. In the following we develop an interacting theory based on an arbitrary internal group which turns out to be a generalisation of the electrodynamic case. The implication of this is that the forces of nature are generated by the action of different groups - in the electrodynamic case, one of the simplest, it is the one dimensional group of rotations of the complex plane.

3.2 Local Symmetry and the Gauge Connection

We begin with the original Lagrangian

$$L = L(\psi^A; \psi^A_{,a}) \quad (3.1)$$

which we assume to be invariant under the global action of the internal group. We wish to construct a Lagrangian out of (3.1) which is invariant under the *local* action

$$\bar{\psi}^A = D^A_{B}(e^k(x^a)) \psi^B \quad (3.2)$$

where the parameters are differentiable functions of the co-ordinates. This causes the transformation law of the field derivative to become non-tensorial (in direct contrast with the global case) :

$$\bar{\psi}^A_{,a} = D^A_{B} \psi^B_{,a} + D^A_{B,a} \psi^B. \quad (3.3)$$

The Lagrangian itself must still transform as a group scalar:

$$L(\bar{\psi}^A; \bar{\psi}_{,a}^A) = L(\psi^A; \psi_{,a}^A) \quad (3.4)$$

If we differentiate both sides of this equation with respect to e^k and note that

$$D_{a,a}^A = e_{,a}^k T_{k,a}^A + O(e^2)$$

so that

$$(D_{a,a}^A)_0 = 0 \quad (3.5)$$

where the subscript '0' now indicates that both the parameter and its derivative have been set to zero, we find again the global invariance identity:

$$\psi_A T_{k,a}^A \psi^a + \psi_A^a T_{k,a}^A \psi_{,a}^a = 0 \quad (3.6)$$

We also have an additional identity which arises by differentiating (3.4) with respect to the derivatives of the parameters: $e_{,b}^k$, and then setting the parameter and its derivative to zero:

$$\left(\frac{\partial L}{\partial \bar{\psi}^A} \right)_0 \left(\frac{\partial \bar{\psi}^A}{\partial e_{,b}^k} \right)_0 + \left(\frac{\partial L}{\partial \bar{\psi}_{,a}^A} \right)_0 \left(\frac{\partial \bar{\psi}_{,a}^A}{\partial e_{,b}^k} \right)_0 = 0 \quad (3.7)$$

The r.h.s of (3.2) is independent of $e_{,b}^k$ so we have

$$\left(\frac{\partial \bar{\psi}^A}{\partial e_{,b}^k} \right)_0 = 0 \quad (3.8)$$

On the other hand, the r.h.s. of (3.3) gives:

$$\left(\frac{\partial \bar{\psi}^A}{\partial e_{,b}^k} \right)_0 = \delta_a^b T_{k,a}^A \psi^a \quad (3.9)$$

Hence (3.7) becomes:

$$\psi_A^b T_{k}^A \psi^k = 0 \quad (3.10)$$

By definition of the current ((2.32) of previous chapter) this is simply

$$J_k^b = 0 \quad (3.11)$$

In other words the field must be *chargeless*. This is true of the scalar representation, in which the generators themselves vanish, and it also holds when the Lagrangian is independent of the field derivatives, in which case the concomitant ψ_A^b vanishes and the equation of motion (2.31) reduces to an algebraic equation. These instances represent trivial solutions to the problem of finding a locally invariant theory.

It is clearly the derivative which lies at the root of the problem. By (3.3) it is not a tensor and we are led to construct a covariant equivalent. To do this we introduce a set of vector fields which are also elements of the Lie Algebra of the group. This means that they are group operators and may be used to define a new derivative

$$\psi_{;a}^A \equiv \psi_{,a}^A + A_{a}^A \psi^a \quad (3.12)$$

The group transformation law of these vector fields may be determined by demanding that the new derivative be covariant under group transformations:

$$\bar{\psi}_{;a}^A = D_{a}^A \psi_{;a}^B \quad (3.13)$$

We call $\psi_{;a}^B$ the *gauge covariant derivative*.

Since (3.12) and (3.13) must hold for an arbitrary matter vector ψ^A in the representation, the required transformation law is:

$$\bar{A}_{a}^A = D_{a}^A A_{a}^B D^{-1c}_{B} - D_{c,a}^A D^{-1c}_{B} \quad (3.14)$$

We observe that, because of the inhomogeneous term, this is the transformation law of a gauge tensor under global transformations only and not under local ones. We will call the A_{μ}^A the *gauge connections*.

Of course, the connections must be introduced into the Lagrangian since we intend to replace the ordinary derivative by the covariant one. Hence we take the Lagrangian to have the field dependence:

$$L = (\psi^A : \psi_{,a}^A : A_{\mu}^A) \quad (3.15)$$

But, to ensure the invariance we must restrict the inclusion of the connection so that it appears only as part of the derivative:

$$L'(\psi^A : \psi_{,a}^A) = L(\psi^A : \psi_{,a}^A : A_{\mu}^A) \quad (3.16)$$

Bearing in mind the transformation law (3.14) of the connection and the fact that the covariant derivative depends on the matter field, its derivative and the connection, we can deduce a new pair of invariance identities from the transformation law of the Lagrangian on the r.h.s of (3.16). The Lagrangian on the l.h.s. will then satisfy these provided only that it is obtained from the original globally invariant Lagrangian by replacing the ordinary derivatives by covariant ones (see next section). In addition, a new term appears in the identity (3.7) thus avoiding the conclusion that the matter field is chargeless. We have, therefore, succeeded in constructing a locally invariant theory out of the global one.

3.3 The Invariance Identities of the Fully Dynamic Theory

So far we have included the connection in the Lagrangian only because it transforms in such a way as to preserve the invariance. Its equation of motion as determined from the Lagrangian (3.15) is purely algebraic and to remedy this we must also include its derivative in the Lagrangian. In so doing we obtain a fully dynamic theory in which the connection and the matter interact with each other. We note that the covariant derivative is in fact a prescription for coupling the fields.

In the following analysis we will start from the beginning by taking the complete set of transformations to be:

$$L(\bar{\psi}^A; \bar{\psi}_{,a}^A; \bar{\lambda}_{a,b}^A; \bar{\lambda}_{a,b,b}^A) = L \text{ (unbarred)} \quad (3.17)$$

$$\bar{\psi}^A = D^A_{a} \psi^a \quad (3.18)^a$$

$$\bar{\psi}_{,a}^A = D^A_{b} \psi^b_{,a} + D^A_{c,a} \psi^c \quad " \quad b$$

$$\bar{\lambda}_{a,b}^A = D^A_{c} A^c_{a,b} D^{-1b}_{b} - D^A_{c,a} D^{-1c}_{b} \quad " \quad c$$

$$\bar{\lambda}_{a,b,b}^A = [D^A_{c} A^c_{a,b} D^{-1b}_{b} - D^A_{c,a} D^{-1c}_{b}]_{,b} \quad " \quad d$$

Examination of the equations (3.18) shows that the parameter and its first and second derivatives occur on the r.h.s. Hence if we differentiate both sides of (3.17) with respect to these quantities and set them all to zero, we will get three sets of identities which are (see Appendix 3A) :

$$\begin{aligned} & \psi_A T^A_{kb} \psi^b + \psi_A^2 T^A_{kb} \psi^b_{,a} \\ & + \psi_A^2 (T^A_{kc} A^c_{a,b} - A^A_{a,c} T^c_{kb}) \\ & + \psi_A^{ab} (T^A_{kc} A^c_{a,b} - A^A_{a,c,b} T^c_{kb}) = 0 \end{aligned} \quad (3.19)$$

$$\begin{aligned} & \psi_A^2 T^A_{kb} \psi^b - \psi_A^2 T^A_{kb} \\ & + \psi_A^{ba} (T^A_{kc} A^c_{b} - A^A_{b,c} T^c_{kb}) = 0 \end{aligned} \quad (3.20)$$

In the following analysis we will start from the beginning by taking the complete set of transformations to be:

$$L(\bar{\psi}^A; \bar{\psi}_{,a}^A; \bar{\lambda}_{ab}^A; \bar{\lambda}_{ab,b}^A) = L \text{ (unbarred)} \quad (3.17)$$

$$\bar{\psi}^A = D^A_{a} \psi^a \quad (3.18)^a$$

$$\bar{\psi}_{,a}^A = D^A_{b} \psi^b_{,a} + D^A_{c,a} \psi^c \quad " \quad b$$

$$\bar{\lambda}_{ab}^A = D^A_{c} \lambda_{ab}^c D^{-1b}_{b} - D^A_{c,a} D^{-1c}_{b} \quad " \quad c$$

$$\bar{\lambda}_{ab,b}^A = [D^A_{c} \lambda_{ab}^c D^{-1b}_{b} - D^A_{c,a} D^{-1c}_{b}]_{,b} \quad " \quad d$$

Examination of the equations (3.18) shows that the parameter and its first and second derivatives occur on the r.h.s. Hence if we differentiate both sides of (3.17) with respect to these quantities and set them all to zero, we will get three sets of identities which are (see Appendix 3A) :

$$\begin{aligned} & \psi_A T^A_{kb} \psi^b + \psi_A T^A_{kb} \psi^b_{,a} \\ & + \psi_A^{ab} (T^A_{kc} \lambda_{ab}^c - \lambda_{ac}^A T^c_{kb}) \\ & + \psi_A^{ab} (T^A_{kc} \lambda_{ab,b}^c - \lambda_{ac,b}^A T^c_{kb}) = 0 \end{aligned} \quad (3.19)$$

$$\begin{aligned} & \psi_A T^A_{kb} \psi^b - \psi_A^{ab} T^A_{kb} \\ & + \psi_A^{ab} (T^A_{kc} \lambda_{ab}^c - \lambda_{bc}^A T^c_{kb}) = 0 \end{aligned} \quad (3.20)$$

and

$$(\Phi_{A}^{cd} + \Phi_{A}^{dc}) T_{kA} = 0. \quad (3.21)$$

The concomitants are defined as

$$\Phi_{A}^{a} = \frac{\partial L}{\partial A_{aA}} \quad (3.22)$$

$$\Phi_{A}^{ab} = \frac{\partial L}{\partial A_{a,b}^{A}} \quad (3.23)$$

The transformation laws of the concomitants may be found by differentiating both sides of (3.17) with respect to the fields. They are (see Appendix 3B):

$$\bar{\Psi}_B = D^{-1c}_B \Psi_c + D^{-1c}_{B,b} \Psi_c^b \quad (3.24)^a$$

$$\bar{\Psi}_B^b = D^{-1c}_B \Psi_c^b \quad " \quad b$$

$$\bar{\Phi}_B^{ac} = D^{-1A}_B D^c_D \Phi_A^{aD} + (D^{-1A}_B D^c_D)_{,c} \Phi_A^{acD} \quad " \quad c$$

$$\bar{\Phi}_B^{abc} = D^{-1A}_B D^c_D \Phi_A^{abD} \quad " \quad d$$

It is important to note that the concomitants associated with the derivatives are tensors while those associated with the fields themselves are not. By using the transformation law of the connection it is possible to construct tensorial equivalents for the non-tensorial concomitants (see Appendix 3C) these are:

$$\Pi_B = \Psi_B - \Psi_A^b A_{bB}^A \quad (3.25)$$

and

$$\Pi_B^{AA} = \Phi_B^{AA} - A_{bB}^c \Phi_c^{abA} + A_{bB}^A \Phi_c^{ab c} \quad (3.26)$$

which transform as

$$\bar{\Pi}_B = D^{-1A}_B \Pi_A \quad (3.27)$$

and

$$\bar{\Pi}^{AB}_B = D^A_C D^{-1D}_B \Pi^{AC}_D \quad (3.28)$$

The set (3.19) - (3.21) constitutes the invariance identities in *integrated* form. By this is meant that the transformation laws of the various quantities are given for non-infinitesimal parameters. It is, however, more convenient to work with the gauge potentials (next section) whose transformation laws are found most easily in infinitesimal form.

3.4 The Gauge Potentials

We have introduced the gauge connections as elements of the Lie Algebra of the group and, as such, they may be expanded in terms of the group generators which form a basis of the Lie Algebra:

$$A^A_{ab} = A^k_a T^A_k \quad (3.29)$$

where the set of n vector fields A^k_a are called the *gauge potentials*. The transformation law of the potentials is easily found by expanding the connection transformation law to first order and using the structure relation (1.18) which introduces the structure constants:

$$\bar{A}^m_a = A^m_a + \epsilon^k C^m_{kn} A^n_a - \epsilon^m_{,a} + O(2) \quad (3.30)$$

This law shows that the potentials are adjoint tensors under global transformations.

Since the generators are constants it is clear that the potentials constitute the essential dynamic gauge fields in the Lagrangian. Using (3.29) we may define concomitants corresponding to these potentials and their derivatives:

$$\tau_k^a = \frac{\partial L}{\partial A_a^k} = \frac{\partial L}{\partial A_a^A} \frac{\partial A_b^A}{\partial A_a^k} \quad (3.32)$$

$$= \phi_A^{ab} T_{kb}^A \quad (3.33)$$

Similarly,

$$\tau_k^{ab} = \frac{\partial L}{\partial A_{a,b}^k} \quad (3.34)$$

$$= \phi_A^{ab} T_{kb}^A \quad (3.35)$$

Using the transformation laws of the connection concomitants and the structure relation we find the transformation laws of the potential concomitants:

$$\bar{\tau}_k^a = \tau_k^a - e^n C_{nk}^m \tau_m^a - e_{,c}^n C_{nk}^m \tau_m^{ac} + O(2) \quad (3.36)$$

$$\bar{\tau}_k^{ab} = \tau_k^{ab} - e^n C_{nk}^m \tau_m^{ab} + O(2) \quad (3.37)$$

We know, however that ϕ_A^{ab} is a group tensor so it follows that τ_k^{ab} is also a tensor. (3.37) shows that in fact it is an adjoint tensor. On the other hand τ_k^a is not a tensor but if we contract (3.26) with the generators and use the structure relation we find a tensorial equivalent for it:

$$\Pi_k^a = \Pi_A^{ab} T_{kb}^A \quad (3.38)$$

$$= \tau_k^a - \tau_n^{ac} C_{mk}^n A_c^m \quad (3.39)$$

Introducing the potential and its concomitants into the identity (3.19) and using the structure relation we get

$$\begin{aligned} \psi_A T_k^A \psi^B + \psi_A T_k^A \psi^B_{,a} \\ + T_n^C C_{km}^n A_a^m + T_n^{ab} C_{km}^n A_{a,b}^m = 0. \end{aligned} \quad (3.40)$$

Similarly the other two identities become:

$$\psi_A T_k^A \psi^B - T_k^A + T_n^{ca} C_{km}^n A_c^m = 0 \quad (3.41)$$

$$(3.42)$$

$$T_k^{cd} + T_k^{dc} = 0.$$

The identity (3.42) is in a manifestly covariant form and we next bring the other two also into this form. Using (3.39), (3.41) becomes covariant:

$$\psi_A T_k^A \psi^B - \Pi_k^A = 0. \quad (3.43)$$

Using (3.25) and (3.39) to introduce the tensors Π_A and Π_k^A into (3.40) we get

$$\begin{aligned} 0 = \psi_A T_k^A \psi^B_{,a} + \psi_A T_k^A \psi^B + \psi_c A_a^m T_m^c T_k^A \psi^B \\ + \Pi_A T_k^A \psi^B + \Pi_n^A C_{km}^n A_a^m \\ + T_n^{ac} C_{km}^n A_{a,c}^m + T_n^{ac} C_{pn}^r C_{km}^n A_a^m A_c^p. \end{aligned} \quad (3.44)$$

The structure relation may be used to reverse the order of the generators in the 2nd term of (3.44) and, on introducing the covariant derivative of ψ^B , the first two terms may be written:

$$\begin{aligned} \psi_A T_k^A \psi^B_{,a} + A_a^m C_{mk}^n (\psi_A T_n^A \psi^B) \\ = \psi_A T_k^A \psi^B_{,a} - \Pi_n^A C_{km}^n A_a^m \end{aligned} \quad (3.45)$$

where we have used the fact that the structure constants are anti-symmetric in their lower indices and the identity (3.43). By (3.42) γ^{ac}_k is antisymmetric in its co-ordinate indices so that the last term in (3.44) may be written as:

$$\begin{aligned} & \gamma^{ac}_r C^r_{pn} C^n_{km} A^m_a A^p_c \\ &= \frac{1}{2} \gamma^{ac}_k \{C^r_{pn} C^n_{km} - C^r_{mn} C^n_{kp}\} A^m_a A^p_c \\ &= -\frac{1}{2} \gamma^{ac}_r C^r_{kn} C^n_{mp} A^m_a A^p_c \end{aligned}$$

by the Jacobi identity. Hence the last two terms together are

$$\gamma^{ac}_r C^r_{kn} \{A^n_{a,c} - \frac{1}{2} C^n_{mp} A^m_a A^p_c\} \quad (3.46)$$

or, again using (3.42) and defining

$$F^{ac}_n \equiv A^n_{a,c} - A^n_{c,a} - C^n_{mp} A^m_a A^p_c \quad (3.47)$$

(3.46) is then

$$\frac{1}{2} \gamma^{ac}_r C^r_{kn} F^{ac}_n.$$

Putting all of this together, (3.44) finally becomes

$$\Pi_A T^A_{kn} \psi^B + \psi^A_A T^A_{kn} \psi^B_{;a} + \frac{1}{2} \gamma^{ac}_r C^r_{kn} F^{ac}_n = 0 \quad (3.48)$$

To see that this is covariant we still need to know the transformation law for F^{ac}_n which we will find in the next section. It is, however, interesting to note that we had to use the second pair of identities (3.42) and (3.43) to bring (3.40) into covariant form. (3.40) is, of course, the global identity and we would expect that we could not write it in a form which is covariant under local transformations without the assumption of local symmetry which is embodied in the identities (3.42) and (3.43).

3.5 The Gauge Curvature and the Gauge Field Strength

Define the *gauge curvature*:

$$F_{ab}^A = F_{ab}^n T_n^A \quad (3.49)$$

then, by (3.47) using (3.29) and the structure relation we find

$$F_{ab}^A = A_{a,b}^A - A_{b,a}^A - [A_a^C A_b^C - A_b^C A_a^C]. \quad (3.50)$$

We will call F_{ab}^A the *gauge field strength* to distinguish it from the curvature. The transformation law of the connection shows that the curvature is a second rank gauge tensor:

$$F_{ab}^A = D_c^A D^{-1D} F_{ab}^C. \quad (3.51)$$

It also follows from this that the field strength is an adjoint tensor:

$$F_{ab}^n = D_n^A F_{ab}^A. \quad (3.52)$$

The first important property of these quantities is simply that they are tensors. The potentials themselves are not tensorial and in order to construct gauge invariants out of them it is important to combine them into a tensorial object. The field strength is such an object and we may use it to characterise the gauge field in a covariant manner. From a physical point of view the components of the field strength can be taken as the observables of the theory.

Secondly, the curvature measures the non-commutation of the gauge covariant derivative:

$$\psi_{;ab}^A - \psi_{;ba}^A = F_{ab}^A \psi^A \quad (3.53)$$

and so provides a measure of the distortion which the gauge field has introduced into the matter distribution. We will call this the *Ricci Identity*.

Author Janse van Rensburg R W
Name of thesis Classical Gauge Theory 1981

PUBLISHER:

University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.