How Symmetry Became Super

Extending the mathematical foundations of symmetry to supersymmetry, with an appreciation of the underlying geometry.

Sam Richards

May 2021

Supervisor: Dr Stefano Cremonesi

Durham University

This piece of work is a result of my own work except where it forms an assessment based on group project work. In the case of a group project, the work has been prepared in collaboration with other members of the group. Material from the work of others not involved in the project has been acknowledged and quotations and paraphrases suitably indicated.

Abstract

In this report we develop the mathematical techniques necessary for understanding the presence and impact of symmetry in physical theories. The first half of the report deals with standard Lie theory and its elegance in describing the continuous symmetries of Minkowski spacetime. This culminates in a detailed discussion of the method of induced representations and the remarkable Wigner's classification. The second half of the report is dedicated to the extension of these ideas to deal with the supersymmetric case; introducing notions of grading, the Grassmann algebra, superalgebras and extending the method of induced representations to the superPoincaré algebra. We then develop the field of homogeneous space and show how it's application can provide a well defined notion of superspace on which we can build the wonderfully simplistic superfield formalism. We conclude with a brief discussion on the role of fibre bundles.

Acknowledgments

A special thank you to my supervisor Stefano Cremonesi whose advice, guidance and continued patience throughout this very difficult year was instrumental in enabling me to complete this project, I am extremely grateful. Also thank you to my parents and younger sister for all their patience and support during my unexpectedly long stay at home this year.

Contents

Chapter 1

Introduction

Symmetry is possibly the single most fundamental thing in all of reality. The majority of our understanding of physics stems from symmetry in some form; from the application of principles as fundamental as conservation of momentum to the conserved quantum numbers which provide our only real way of understanding the mysterious quantum world. But what is symmetry? To many people, the notion of symmetry is that of lines of symmetry on the regular polygons they encountered in school, or even just 'something to do with mirrors'. These symmetries are called **discrete**; the lines of symmetry on a polygon constitute a *finite, countable* set. Mathematically, the finite, countable transformations between these lines form an object called a **group**. For polygons, these groups are known as *dihedral groups*. Physically speaking, symmetry takes the following definition:

Definition 1.1. A **symmetry** of a physical system is a transformation of a set of dynamical variables under which the measurable physics of the system is invariant[8].

In physics, we are often more concerned with **continuous** symmetry. The simplest examples of such symmetries are *very* familiar; things like *translation i.e.* walking somewhere, and *rotation*. The measurable physics they preserve in reference to definition (1.1) is that of distances between points in space *i.e.* they preserve the *Euclidean inner product*. Such transformations must now be labelled, or rather parametrized, by continuous variables, such as vectors for translations or angles for rotations. As we'll see in chapter two, these ideas lead to a very special group structure defining an object called a **Lie group**. As we progress through this report, we'll explore how Lie groups describe symmetry in physical theories and the culmination of this in chapter 3 will be the quite remarkable Wigner classification; how fundamental particles themselves are the consequence of symmetry. In chapter 4, we'll introduce the weird and wonderful world of supersymmetry and discuss how we extend many of the mathematical ideas of chapters 2 and 3 to the supersymmetric case. In chapter 5, we'll extend notions of space itself to include strange anti-commuting parameters and accommodate for an elegant formalism of supersymmetric field theory. As an epilogue to all this discussion, we'll briefly investigate a breathtaking link between the ideas we develop throughout the report.

Chapter 2

Introduction to Lie Theory

Let's begin by formally introducing some group theory. More specifically, some Lie theory. Indeed, in the introduction we mentioned a need to deal with *continuous* transformations and this means dealing with groups that exhibit continuous transformations. The necessity for such a group leads us to the concept of *Lie groups* and in this chapter we will develop much of the core mathematics required to understand and apply them in the study of symmetry.

2.1 Lie Groups

2.1.1 Some Group Theory

The material in this subsection is loosely based on [8].

Definition 2.1. A **group** is a tuple $(G, \cdot)$, (or sometimes $(G, \cdot, e)$), containing a set of elements $G = \{g_1, g_2, ...\}$, and a *binary operation*,

$$\cdot : G \times G \to G, \quad (g_1, g_2) \mapsto g_1 \cdot g_2$$

that satisfies the following axioms:

1. (**Closure**): $g_1, g_2 \in G \implies g_1 \cdot g_2 \in G$
2. (**Associativity**): $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3 \quad \forall\, g_1, g_2, g_3 \in G$
3. (**Identity**): $\exists\, e \in G : \forall g \in G \;\; g \cdot e = e \cdot g = g$
4. (**Inverse**): $\forall g \in G \;\; \exists\, g^{-1} \in G : g \cdot g^{-1} = g^{-1} \cdot g = e$

A group is **abelian** if we also have $g_1 \cdot g_2 = g_2 \cdot g_1 \;\forall\, g_1, g_2 \in G$. A **subgroup** $H \subset G$ is a subset H containing the identity e for which the above axioms hold independently.

This definition is completely general, covering both discrete groups, for which the set G is discrete (group elements may be indexed by some $i = 1, 2, 3, ...$), and continuous groups for which the set G is, funnily enough, continuous (group elements are parametrized by some continuous variable e.g. $t \in \mathbb{R}$).

Example 2.2. The **Symmetric Group** of $X = \{1, 2, 3\}$ is $(\mathrm{Sym}(X), \circ)$ where

$$\mathrm{Sym}(X) = \{(123), (312), (231), (132), (213), (321)\}$$

and $\circ$ denotes composition of the permutation functions. Clearly this is discrete since $\mathrm{Sym}(X)$ is discrete. This is also denoted S_n for $X =: \{1, 2, ..., n\}$.

Important Example 2.3 (The Special Unitary Group: $SU(2)$)**.**

$$SU(2) := \{U \in U(2) | \det(U) = 1\}$$

where $U(2)$ denotes the **Unitary Group** $U(2) := \{U \in M_2(\mathbb{C}) | U^\dagger U = \mathbb{1}\}$. It's very simply to show that $SU(2)$ is a group:

1. Closure: Let $A, B \in SU(2)$. Then $(AB)^\dagger(AB) = B^\dagger A^\dagger AB = B^\dagger \mathbb{1} B = \mathbb{1}$.
2. Associativity: This is inherent since matrix multiplication is associative.
3. Identity: Trivially $\mathbb{1}^\dagger \mathbb{1} = \mathbb{1}\mathbb{1} = \mathbb{1}$.
4. Inverse: If $A \in SU(2)$ then $A^\dagger A = \mathbb{1} \implies A^{-1} = A^\dagger$. Now $(A^{-1})^\dagger(A^{-1}) = (A^\dagger)^\dagger(A^\dagger) = AA^\dagger = (A^\dagger A)^\dagger = \mathbb{1}^\dagger = \mathbb{1}$.

Really a group is a very abstract object; a collection of elements that can be thought of in some sense as abstractions of rotations. Rotations aren't, in essence, relations between points on a shape but their action, on the set that constitutes the shape, that relates two points is how we understand them. We really understand groups best by their *action* on a set:

Definition 2.4. An **action** of a group G on a set X, $G \curvearrowright X$, is a function $(\cdot, \cdot) : G \times X \to X$ for which the following holds:

1. **Identity:** $(e, x) = (x, e) = x \;\forall\; x \in X$.
2. **Associativity:** $(g \cdot h, x) = (g, (h, x)) \;\forall\; g, h \in G,\; x \in X$

In particular, the **left (regular) action** of a group G on a set X is a map

$$\mathcal{L} : G \times X \to X, \;\; (g, x) \mapsto g \cdot x =: \mathcal{L}_g(x)$$

such that $g_1 \cdot (g_2 \cdot x) = (g_1 \cdot g_2) \cdot x$ and $e \cdot x = x$, $g_1, g_2 \in G$, $x \in X$. The **right action** is defined analogously by $\mathcal{R} : X \times G \to X$, $\mathcal{R}_g(x) := x \cdot g$.

Definition 2.5. For a group G acting on a set X we have:

1. The **orbit** of an element $x \in X$ under the left action of G is the set

   $$G \cdot x := \{g \cdot x | g \in G\} \subset X$$

   If the action is the left regular action we call the orbit a **right coset**. We denote the **set of right cosets** by $G/X =: \{G \cdot x\}_{x \in X}$. Left cosets are defined analogously via the right regular action.

2. The **isotropy subgroup**, or **stabiliser**, of $x \in X$ is defined to be:

$$G_x := \{g \in G | g \cdot x = x\}$$

Another group action we will make some use of later is the following:

Definition 2.6. The **conjugate,** or **adjoint, action** of a group G is defined to be

$$G \times G \to G, \quad (g, h) \mapsto g \cdot h \cdot g^{-1}$$

Now we want to consider continuous group elements as required for the continuous symmetry we referenced in chapter 1. These are labelled by continuous parameters which we vary continuously to move from one group element to another. This suggests the group has a smooth underlying structure. To define such a group we must define the underlying structure.

2.1.2 Smooth Manifolds

The material in this subsection follows [9].

Definition 2.7. A **smooth manifold** M is a topological, Hausdorff space for which there exists an integer $n > 0$ and a collection $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ of open sets $U_\alpha \subset M$ and bijective homomorphisms $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$, for some indexing set I, such that $\bigcup\limits_\alpha U_\alpha = M$ and $\forall \alpha, \beta \in I$ s.t $U_\alpha \cap U_\beta \neq \varnothing$ the *transition map* $\phi_\beta \circ \phi_\alpha^{-1}$ is differentiable.

We say the manifold is n-dimensional. This definition looks a little terrifying but it's just formal. Essentially it's defining a well behaved space which, for all points, looks locally like $\mathbb{R}^n$ - the kind of space we occupy every day. $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ is called an **atlas** and the ϕ_α are **coordinate charts** $\phi(p \in M) = x \in \mathbb{R}^n$.

Definition 2.8. Consider a curve $\gamma : (-\epsilon, \epsilon) \to M$, $\epsilon > 0$ small, $\gamma(0) = p$. A **tangent vector** at $p \in M$ is defined to be the tangent vector of a curve at p i.e. $v := \gamma'(0) \equiv \frac{d}{dt}\gamma(t)|_{t=0}$. The **tangent space** at $p \in M$ is the vector space $T_pM \cong \mathbb{R}^n$ spanned by the vectors tangent to all curves through p.

These tangent vectors have a canonical interpretation as differential operators via the directional derivative. *i.e.* if we have a smooth function $f : M \to \mathbb{R}$ and pick coordinate chart $\phi : p \mapsto (x_1, ..., x_n)$ then it's derivative at p in the direction of tangent vector $V := \gamma'(0)$ is

$$V(f) = \frac{d}{dt}(f \circ \gamma)(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma(0)) \cdot \frac{\partial x_i}{\partial t}(0) = [\sum_{i=1}^{n} x_i'(0) \frac{\partial}{\partial x_i}] f \implies V \equiv \sum_{i=1}^{n} x_i'(0) \frac{\partial}{\partial x_i}$$

So the basis T_pM in coordinates ϕ is $\{\frac{\partial}{\partial x_i}\}$.

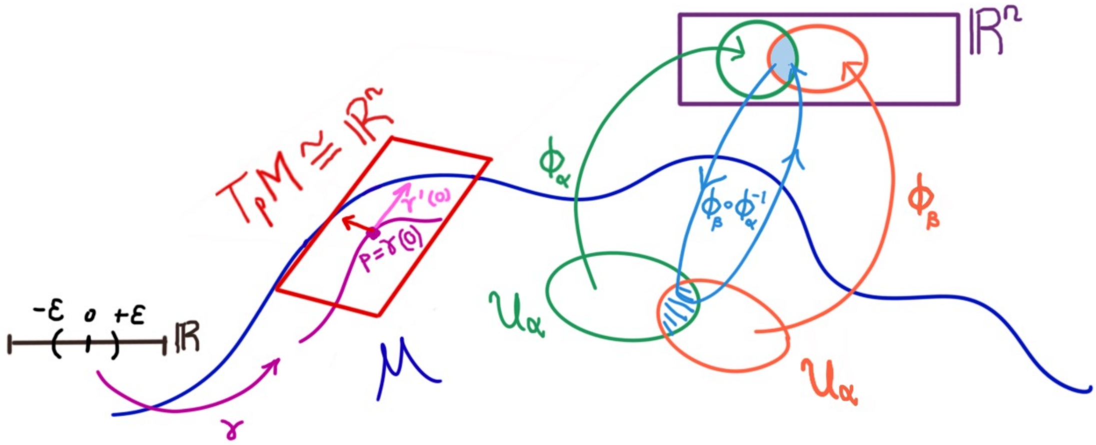

Figure 2.1: A smooth manifold M with coordinate charts $(U_\alpha, \phi_\alpha), (U_\beta, \phi_\beta)$ and tangent space T_pM.

Definition 2.9. Using definition (2.8) we can define a **vector field** which assigns to each point $p \in M$ a tangent vector $X(p) \in T_pM$. Sometimes we refer to vector fields along curves $X(t) \in T_{\gamma(t)}M$.

Definition 2.10. If M, N are smooth manifolds, $f : M \to N$ a smooth function and $\gamma : (\epsilon, \epsilon) \to M$ a smooth curve on M, with $\gamma(0) = p$, $\gamma'(0) = v \in T_pM$, then we can define the **linearization**, or **differential**, of f at $p \in M$ along γ by

$$df_p : T_pM \to T_{f(p)}N, \quad df_p(v) = (f \circ \gamma)'(0)$$

The concept of a smooth manifold gives definition to the structure we're looking for in our special group of transformation.

2.1.3 Lie Groups

The material in this subsection is based on [8], while drawing on parts of [2].

Definition 2.11. A **Lie group** is a group which is also a smooth manifold. i.e. the group operations of inverse and $\cdot : G \times G \to G$ are smooth. A subgroup $H \subset G$ is a **Lie subgroup** if H is a smooth submanifold of G.

Through the following examples we introduce an extremely important class of Lie groups that will be imperative to the rest of the report.

Example 2.12 (**Matrix Groups**). Matrix groups are groups whose set consists of matrices and whose group structure arises under matrix multiplication. For this reason they are always denoted simply by their defining set.

1. **The Orthogonal Group:** The n-dimensional orthogonal group is defined by

$$O(n) := \{A \in M_n(\mathbb{F}) \mid A^T A = \mathbb{1}\}$$

where $M_n(\mathbb{F})$ denotes the set of all $n \times n$ matrices over some algebraically closed field $\mathbb{F}$. For our purposes $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

2. **An Interesting Example:** Consider the following set of matrices:

$$\mathcal{S} := \Big\{ \begin{pmatrix} a & 0 \\ b & 1/a \end{pmatrix} \mid a, b \in \mathbb{R},\ a \neq 0 \Big\} \subset M_2(\mathbb{R})$$

We will first show it satisfies the group axioms. Note that associativity is assumed from the associativity of matrix multiplication. Let $a_i, b_i \in \mathbb{R}$, $i = 1, 2$:

- Closure: $\begin{pmatrix} a_1 & 0 \\ b_1 & 1/a_1 \end{pmatrix}\begin{pmatrix} a_2 & 0 \\ b_2 & 1/a_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 & 0 \\ b_1a_2+b_2/a_1 & 1/a_1a_2 \end{pmatrix}$
- Identity: set $a = 1$, $b = 0 \implies \begin{pmatrix} a & 0 \\ b & 1/a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
- Inverse: set $a_2 = 1/a_1$ and $b_2 = -b_1$ then $\begin{pmatrix} a_1a_2 & 0 \\ b_1a_2+b_2/a_1 & 1/a_1a_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (b_1+b_2)a_2 & 1 \end{pmatrix}$ This suggests that $\begin{pmatrix} a_1 & 0 \\ b_1 & 1/a_1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/a_1 & 0 \\ -b_1 & a_1 \end{pmatrix}$ so we need only check this inverse is commutative: $\begin{pmatrix} 1/a_1 & 0 \\ -b_1 & a_1 \end{pmatrix}\begin{pmatrix} a_1 & 0 \\ b_1 & 1/a_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Thus this is indeed a group. To argue that it's a manifold we use a cheap argument. If roughly speaking a manifold is a space which looks locally like $\mathbb{R}^n$ then we simply note that $\mathcal{S}$ is spanned by the numbers $a \neq 0, b \in \mathbb{R}$ and thus it is at the very least homeomorphic to the upper half plane $\mathbb{R}^2_{y>0}$ which is a smooth manifold automatically. Since matrix multiplication is a polynomial process it's smooth and so $\mathcal{S}$ is a Lie group!

Important Example 2.13 (The Special Unitary Group: $SU(2)$)**.**

$$\text{Let } U = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \in SU(2). \text{ Then } U^\dagger U = \mathbb{1} \text{ gives:}$$

$\begin{pmatrix} \bar{z}_1 & \bar{z}_3 \\ \bar{z}_2 & \bar{z}_4 \end{pmatrix}\begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} = \begin{pmatrix} |z_1|^2 + |z_3|^2 & \bar{z}_1 z_2 + \bar{z}_3 z_4 \\ \bar{z}_2 z_1 + \bar{z}_4 z_3 & |z_2|^2 + |z_4|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Thus we obtain:

$$|z_1|^2 + |z_3|^2 = |z_2|^2 + |z_4|^2 = 1, \quad \bar{z}_1 z_2 + \bar{z}_3 z_4 = \bar{z}_2 z_1 + \bar{z}_4 z_3 = 0.$$

The second two relations requires one of two possibilities: either $z_4 = \bar{z}_1$ and $z_3 = -\bar{z}_2$, or $z_4 = \bar{z}_2$ and $z_3 = -\bar{z}_1$. We now impose the second condition $\det(U) = +1$ which implies $z_1z_4 - z_2z_3 = 1$ thus we must have $z_4 = \bar{z}_1$ and $z_3 = -\bar{z}_2$ which along with $|z_1|^2 + |z_3|^2 = 1$ assures $\det(U) = z_1z_4 - z_2z_3 = |z_1|^2 + |z_2|^2 = 1$. Thus, setting $z_1 = a + ib$ and $z_2 = c + id$, with $a, b, c, d \in \mathbb{R}$, we obtain the general form for $U \in SU(2)$:

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix} = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

We tend to write this in the form

$$U = a_0 \mathbb{1} + i(\boldsymbol{a} \cdot \boldsymbol{\sigma}), \quad \boldsymbol{a} = (a_1, a_2, a_3) \equiv (d, c, b) \tag{2.1}$$

where the $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the **Pauli matrices**

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.2}$$

Crucially, here we note that $|z_1|^2 + |z_2|^2 = 1 \implies a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1$ *i.e.* the continuous parameters that span $SU(2)$ as the coefficients of $\boldsymbol{\sigma}$ exist in the set

$$\{(a_0, a_1, a_2, a_3) \in \mathbb{R}^4 | a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1\}$$

This is nothing but the hypersphere S^3! S^3 is a smooth manifold (there are multiple possible choices of coordinate chart including *stereographic projection*) thus we have confirmed $SU(2)$ is indeed a Lie group.

Now we introduce a couple of concepts which will assist us in understanding the Lorentz group later on.

Definition 2.14. A **connected component** of a Lie group G is a maximal subset $H \subset G$ whose elements are generated by the variation of the continuous parameter(s) of a single element.

Example 2.15. For example, in Example 2.1, $O(n)$ has two connected components, $\{O(n) | \det(A) = +1\}$ and $\{O(n) | \det(A) = -1\}$, which are related by a discrete **parity** transformation. Recall that the determinant of a matrix $A \in M_n(\mathbb{F})$ is given by

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1(1)\sigma} A_{2(2)\sigma} \cdots A_{n(n)\sigma}$$

where σ are permutations and $\operatorname{sgn}(\sigma) = +1(-1)$ if the permutation is even(odd). Now suppose we vary an element $A \in O(n)$ by acting on it with the infinitesimal transformation $B \approx (\mathbb{1} + \epsilon C) \in O(n)$, $\epsilon > 0$ small and $C \in M_n(\mathbb{F})$. Then

$$\begin{aligned} \det((\mathbb{1} + \epsilon C)A) = \det(\mathbb{1} + \epsilon C)\det(A) &= \det(A) \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(\mathbb{1} + \epsilon C)_{1(1)\sigma} \cdots (\mathbb{1} + \epsilon C)_{n(n)\sigma} \\ &= \det(A)(1 + \epsilon(C_{11} + ... + C_{nn}) + \mathcal{O}(\epsilon^2)) \\ &= \det(A)(1 + \epsilon \operatorname{Tr}(C) + \mathcal{O}(\epsilon^2)) \xrightarrow[\epsilon \to 0]{} \det(A) \end{aligned}$$

In other words, if $\det(A) = \pm 1$ then the action of an infinitesimal transformation on A will preserve the sign of it's determinant and since any transformation is the composition of infinitesimal ones, the sign of $\det(A)$ specifies connected components in $O(n)$.

Theorem 2.16. The connected component of a Lie group G that contains the identity forms an invariant Lie subgroup of G.

Remark. We won't give an explicit proof but a useful example is the case of $O(n)$ where the component $\{O(n)|\det(A) = +1\}$ forms a subgroup called the **Special Orthogonal Group** denoted $SO(n)$.

Definition 2.17. A Lie groups is **compact** if the range for it's parameters is a compact set.

Definition 2.18. The action of a Lie group G on a manifold M is **transitive** if $\forall p, q \in M$ $\exists\; g \in G$ s.t. $p = g \cdot q$

2.1.4 The Structure of a Lie Group

Lie Groups seem like very complicated objects but let's delve a little deeper into their structure. This discussion follows 2.2 in [17]. Suppose we have an n-dimensional Lie group $G = \{g(t_i) \,|\, t_i \in \mathbb{R}, i = 1, ..., n\}$. The group product must be smooth; we can impose this in full generality with the following:

$$g(t_i) \cdot g(s_i) = g(f(t_i, s_i)) \equiv g(f_i(t, s)) \tag{2.3}$$

where f is a smooth function taking the vectors of continuous parameters t, s to some new vector of continuous parameters. Now the group axioms automatically place some restrictions on f: If $t = 0$ gives the identity element $e \equiv g(0)$ then we must have

$$g(t_i) = g(t_i) \cdot e = g(t_i) \cdot g(0) = g(f(t_i, 0)) = g(f_i(t, 0)) \implies f_i(t, 0) = f_i(0, t) = t_i$$

Now since it's a Lie group and has smooth structure, for some sufficiently small subset around a given element, we may perform an expansion for a given element:

$$g(t) = g(0) + t^i X_i + \frac{1}{2} t^j t^k X_{jk} + ..., \;\; X_i = \frac{\partial}{\partial t^i} g(t)|_{t=0}, \; X_{jk} = \frac{\partial^2}{\partial t^j \partial t^k} g(t)|_{t=0} \tag{2.4}$$

Note that we're using the Einstein summation convention. Also recall that $g(0) = e$, the identity, so we will simply write that from now on. Now f is a smooth function, so it too has an expansion which must take the form:

$$f^i(t, s) = t^i + s^i + f^i_{jk} t^j s^k + ...$$

which is such that $f^i(t, 0) = t^i$, $f^i(0, s) = s^i$. (Note: if we had terms of the form $(t^i)^2$ then $f^i(t, 0) = t^i + a(t^i)^2 + ...$ This is forbidden by the group structure). Now we can apply these expansions to the general product (2.3):

$$\begin{aligned}
\text{LHS: } g(t_i)\cdot g(s_i) &= [e + t^iX_i + \frac{1}{2}t^jt^kX_{jk} + ...]\cdot[e + s^iX_i + \frac{1}{2}s^js^kX_{jk} + ...] \\
&= e + (t^i + s^i)X_i + t^js^kX_jX_k + \frac{1}{2}(t^jt^k + s^js^k)X_{jk} + ... \\
&= [e + (t^i + s^i)X_i + \frac{1}{2}(t^jt^k + s^js^k)X_{jk}] + t^js^kX_jX_k + ... \\
\text{RHS: } g(f(t_i, s_i)) &\equiv g(f_i(t, s)) = g(t^i + s^i + f^i_{jk}t^js^k + ...) \\
&= e + (t^i + s^i + f^i_{jk}t^js^k + ...)X_i + \frac{1}{2}(t^j + s^j + ...)(t^k + s^k + ...)X_{jk} + ... \\
&= e + (t^i + s^i)X_i + f^i_{jk}t^js^kX_i + \frac{1}{2}(t^jt^k + s^js^k + t^js^k + s^jt^k + ...)X_{jk} + ... \\
&= [e + (t^i + s^i)X_i + \frac{1}{2}(t^jt^k + s^js^k)X_{jk}] + f^i_{jk}t^js^kX_i + \frac{1}{2}(t^js^k + s^jt^k)X_{jk} + ...
\end{aligned}$$

When we equate these two the parts in the square brackets cancel and we obtain:

$$\begin{aligned}
\frac{1}{2}(t^js^k + s^jt^k)X_{jk} &= t^js^k(X_jX_k - f^i_{jk}X_i) \\
\implies t^js^kX_{jk} &= t^js^k(X_jX_k - f^i_{jk}X_i) \\
\implies X_{jk} &= X_jX_k - f^i_{jk}X_i
\end{aligned} \tag{2.5}$$

Where we've used that $X_{jk} = \frac{\partial^2}{\partial t^j \partial t^k}g(t)|_{t=0} \implies X_{jk} = X_{kj}$ meaning $((t^js^k + s^jt^k)X_{jk} = 2t^js^kX_{jk}$ since both indices are contracted. But in turn $X_{jk} = X_{kj}$ asserts the following:

$$\begin{aligned}
X_jX_k - f^i_{jk}X_i = X_kX_j - f^i_{kj}X_i &\implies X_jX_k - X_kX_j = (f^i_{jk} - f^i_{kj})X_i \\
\implies [X_j, X_k] = C^i_{jk}X_i,\ \ C^i_{jk} &\equiv (f^i_{jk} - f^i_{kj})
\end{aligned} \tag{2.6}$$

Where $[X_i, X_j] := X_iX_j - X_jX_i$ is called a **commutator** and the constants $C^i_{jk} \equiv (f^i_{jk} - f^i_{kj})$ are called the **structure constants**. Such a relation exists for all $i, j, k = 1, ..., n$. This is quite remarkable. We began with a complicated object, a Lie group, and using the assertions of its definition we've produced a set of unique algebraic relations defining the structure of the group near the identity e. This magically appearing algebra will be the subject of the next section, where we will give it a proper definition.

2.2 Lie Algebras

2.2.1 Formal Definition

This material loosely follows chapters 1 and 2 of [8].

Definition 2.19. A **Lie algebra** is a vector space $\mathfrak{g}$ over a field $\mathbb{F}$, equipped with a binary operation $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called a **Lie bracket**, that satisfies the following:

1. **Anti-Symmetry:** $[X, Y] = -[Y, X]$

2. **Linearity:** $[aX + bY, Z] = a[X, Z] + b[Y, Z]$
3. **Jacobi Identity:** $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$

$\forall\, X, Y, Z \in \mathfrak{g}$ and $a, b \in \mathbb{F}$. Again, for our purposes $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Remark. What is the vector space $\mathfrak{g}$? Recall that, for a Lie group G, when we expanded the element $g(t) \in G$ in (2.4) we had that $X_i = \frac{\partial}{\partial t^i} g(t)|_{t=0}$ i.e. the X_i are the tangent vectors at the identity $g(0) = e$. This gives us the answer: $\mathfrak{g}$ is the tangent space to G at the identity $e \in G$ aka $T_e G$ from Def.(2.8)!

The Jacobi identity arises naturally from the definition of the structure constants $C^i_{jk} \equiv (f^i_{jk} - f^i_{kj})$, more specifically from the antisymmetry in the interchange of j and k, and the symmetry of the interchange of i and j or k. Indeed, it takes the form:

$$C^l_{jk} C^m_{il} + C^l_{ij} C^m_{kl} + C^l_{ki} C^m_{jl} = 0$$

Important Example 2.20 (The Special Unitary Group: $SU(2)$)**.** It's reasonable to suggest that the form (2.1) implies the Pauli matrices (2.2) may be the generators of $SU(2)$. This is almost correct but there's subtlety here. Let's explore this geometrically by finding the tangent space $T_{\mathbb{1}} SU(2)$. Consider the curve $A(t) \in SU(2)$ *i.e.* $A : (-\epsilon, \epsilon) \to SU(2)$, $A(0) = \mathbb{1}$. We want to know what properties $A(t)^\dagger A(t) = \mathbb{1}$ and $\det(A(t)) = 1$ impose on tangent vectors $A'(0) \in T_{\mathbb{1}} SU(2)$:

$$\begin{aligned}
0 = \frac{d}{dt}[A(t)^\dagger A(t)]_{t=0} &= A'(0)^\dagger A(0) + A(0)^\dagger A'(0) \implies A'(0)^\dagger = -A'(0) \\
0 = \frac{d}{dt}[\det(A(t))]_{t=0} &= \frac{d}{dt}[A_{11}(t)A_{22}(t) - A_{12}(t)A_{21}(t)]_{t=0} \\
&= A'_{11}(0)A_{22}(0) + A_{11}(0)A'_{22}(0) - A'_{12}(0)A_{21}(0) - A_{12}(0)A'_{21}(0) \\
A(0) = \mathbb{1} \implies &= A'_{11}(0) + A'_{22}(0) = \mathrm{Tr}(A'(0)) \implies \mathrm{Tr}(A'(0)) = 0
\end{aligned}$$

In other words, $T_{\mathbb{1}} SU(2) = \{A \in GL_2(\mathbb{C}) | A^\dagger = -A,\ \mathrm{Tr}(A) = 0\}$. Now for the subtlety: the above result would give us the infinitesimal expansion at the identity to be:

$$U = \mathbb{1} + a_i T_i, \quad T_i^\dagger = -T_i$$

i.e. the generators T_i are anti-hermitian. But the Pauli matrices are clearly hermitian! It turns out that the space of traceless, anti-hermitian 2×2 matrices is spanned by $\{i\sigma_1, i\sigma_2, i\sigma_3\}$, with σ_i as in (2.2). As physicists we tend to choose to factor out the i in the expansion to give the form (2.1). Our Lie algebra is therefore $\mathfrak{su}(2) \equiv T_{\mathbb{1}} SU(2) = \mathrm{span}\{\sigma_1, \sigma_2, \sigma_3\}$. We can compute commutation relations, for example:

$$[\sigma_1, \sigma_2] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} = 2i\sigma_3$$

Similarly, we obtain $[\sigma_3, \sigma_1] = 2i\sigma_2$ and $[\sigma_2, \sigma_3] = 2i\sigma_1$. Taking ϵ_{ijk} to be the **Levi-Civita tensor density** we can write the general algebraic relation:

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.7}$$

For this reason we tend to choose the normalized basis $\mathfrak{su}(2) = \mathrm{span}\{T_i = \frac{\sigma_i}{2}\}_{i=1,2,3}$ giving the familiar commutation relations $[T_i, T_j] = i\epsilon_{ijk}T_k$.

2.2.2 Subalgebras

Definition 2.21. Let $\mathfrak{g}$ be a Lie algebra. A **Lie subalgebra** is a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ that satisfies Def.2.9 1-3. An **ideal** is a subalgebra $\mathfrak{h}$ s.t $[X, Y] \in \mathfrak{h} \, \forall \, X \in \mathfrak{g}, Y \in \mathfrak{h}$. The **derived algebra** is the ideal $\mathfrak{i}(\mathfrak{g}) = \{[X, Y] | X, Y \in \mathfrak{g}\}$. The **centre** is the ideal $J(\mathfrak{g}) = \{X \in \mathfrak{g} | [X, Y] = 0 \, \forall \, Y \in \mathfrak{g}\}$.

Definition 2.22. A Lie algebra $\mathfrak{g}$ is **abelian** if $J(\mathfrak{g}) = \mathfrak{g}, \mathfrak{i}(\mathfrak{g}) = \{0\}$ i.e. it's bracket is always 0 (completely degenerate). It is **simple** if $J(\mathfrak{g}) = \{0\}, \mathfrak{i}(\mathfrak{g}) = \mathfrak{g}$ i.e. it's non-abelian and has no non-trivial ideals. Finally, it is **semi-simple** if it has no non-trivial abelian ideals i.e. all its ideals are non-abelian.

Definition 2.23. The **complexification** of a real Lie algebra, with basis $\mathfrak{g} = \text{span}_{\mathbb{R}}\{X^a\}$ is the Lie algebra given by $\mathfrak{g}_{\mathbb{C}} := \text{span}_{\mathbb{C}}\{X^a\} \equiv \text{span}_{\mathbb{R}}\{X^a, iX^a\}$.

We sometimes write $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \oplus i\mathfrak{g}$.

Important Example 2.24 (The Special Unitary Group: $SU(2)$)**.** Clearly from it's commutator relations we see that $\mathfrak{su}(2)$ is non-abelian has no non-trivial ideals and is therefore simple. The basis for the complexification is $\mathfrak{su}(2)_{\mathbb{C}} = \text{span}_{\mathbb{R}}\{T_1, T_2, T_3, iT_1, iT_2, iT_3\}$. Using (2.7) we have the following two relations: $[iT_i, iT_j] = -[T_i, T_j] = -i\epsilon_{ijk}T_k$, and $[T_i, iT_j] = i[T_i, T_j] = i^2\epsilon_{ijk}T_k = i\epsilon_{ijk}(iT_k)$. Thus $\mathfrak{su}(2)_{\mathbb{C}}$ has commutation relations

$$[T_i, T_j] = i\epsilon_{ijk}T_k, \quad [iT_i, iT_j] = -i\epsilon_{ijk}T_k, \quad [T_i, iT_j] = i\epsilon_{ijk}(iT_k) \tag{2.8}$$

The generators are no longer all hermitian; for example $(iT_1)^\dagger = -iT_1$. However, they are all still traceless since $\text{Tr}(cA) = c\text{Tr}(A)$, $c \in \mathbb{C}$. This means the matrices in the corresponding Lie group have constant determinant; in fact for it to satisfy the group axiom of closure it must be $\det(A) = 1$, $A \in M_2(\mathbb{C})$. This is nothing but the **special linear group** $SL(2, \mathbb{C})$. Thus we've deduced that $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C})$. This will be instrumental in our understanding of the Lorentz group in chapter 3.

2.2.3 Why is the Lie algebra so special?

Earlier we remarked that the Lie algebra provides key structural information about its corresponding Lie group, but this came with a potentially disappointing assumption: in section 2.1.4 our expansions were done close to the identity. How is the tangent space at the identity of a vast Lie group going to tell us anything global about the group? We consider the following lemma as in chapter 2 of [8]

Lemma 2.25. The left action of a Lie group G on itself,

$$\mathcal{L}_h : G \to G, \;\; g \mapsto h \cdot g$$

is a smooth, bijective map. (i.e. it is a **diffeomorphism**).

Proof. First we prove surjectivity i.e. that for every element in the image there exists at least one element in the pre-image that maps to it. We have that

$$\mathcal{L}_h(h^{-1} \cdot g) = h \cdot (h^{-1} \cdot g) = (h \cdot h^{-1}) \cdot g = e \cdot g = g, \ \forall g \in G$$

Now we prove injectivity i.e. that for every element in the image, there exists exactly one element in the pre-image that maps to it. Suppose that $g_1, g_2 \in G$ map to the same element, $\mathcal{L}_h(g_1) = \mathcal{L}_h(g_2)$. Then

$$\begin{aligned}\mathcal{L}_h(g_1) = \mathcal{L}_h(g_2) &\implies h \cdot g_1 = h \cdot g_2 \implies h^{-1} \cdot (h \cdot g_1) = h^{-1} \cdot (h \cdot g_2) \\ &\implies (h^{-1} \cdot h) \cdot g_1 = (h^{-1} \cdot h) \cdot g_2 \implies g_1 = g_2\end{aligned}$$

So we have both surjectivity and injectivity which means $\mathcal{L}_h$ is bijective. Finally, since G is a Lie group it's group operation is smooth and as the action $\mathcal{L}_h$ is defined by the group operation it is a smooth map. Thus it is a smooth bijection. □

The following material is no longer that of [8]. We will now use this map to show something quite special. Since the map $\mathcal{L}_h$ is a smooth function on the manifold G we can take it's linearization at a point $g = \gamma(0)$, $\gamma(t)$ a curve on G, (2.10):

$$d[\mathcal{L}_h]_g(v) = (\mathcal{L}_h \circ \gamma)'(0), \ v := \gamma'(0), \ d[\mathcal{L}_h]_g : T_gG \to T_{hg}G$$

Now we say that a vector field X is **left invariant** if $\forall\ h \in G$ we have that

$$d[\mathcal{L}_h]_g(X(g)) = X(h \cdot g), \ \text{i.e. } d[\mathcal{L}_h]X = X$$

Now suppose X is a non-zero tangent vector at the identity e, $X \in T_eG$. Then we can *define* a left invariant vector field $X(g)$ by

$$X(g) := d[\mathcal{L}_g]_e(X), \ X(g) \in T_gM, \ X \in T_eM$$

Note that this is automatically left invariant since

$$d[\mathcal{L}_h]_g(X(g)) = d[\mathcal{L}_h]_g(d[\mathcal{L}_g]_e(X)) = d[\mathcal{L}_h \circ \mathcal{L}_g]_e(X) = d[\mathcal{L}_{hg}]_e(X) =: X(h \cdot g)$$

This is an immense achievement. From any tangent vector at the identity, i.e any element of the Lie algebra, we can construct a vector field which will take us to *any* other tangent space on the Lie group! However, it remains to be seen whether our wonderful structure at the identity is the same in other tangent spaces and we will explore this now. Recall that tangent vectors act on functions as differential operators. Let $f \in C^\infty(G)$ i.e. f is a smooth function on G, and let $X, Y \in \mathfrak{g} \equiv T_eG$. Then we have that:

$$\begin{aligned}[X,Y]f(g) &= (XY - YX)f(g) = X(Y(f(g)) - Y(X(f(g)) \\ &= X(Y(g)f) - Y(X(g)f) = [X(d[\mathcal{L}_g]_eY(e)) - Y(d[\mathcal{L}_g]_eX(e))]f \\ &= XY(f \circ \mathcal{L}_g) - YX(f \circ \mathcal{L}_g) = [X,Y](f \circ \mathcal{L}_g)(e) = d[L_g]_e[X,Y]f(e)\end{aligned}$$

Where in to get to the second line we've used that the tangent vector acting on the function at g is the vector field at g acting on the function i.e. $X(f(p))|_{p=g} \equiv X(p)(f)|_{p=g}$.

And then for the rest of it we've use that since $X, Y \in T_eG$ we can treat them as the origin of vector fields $X(g) := d[\mathcal{L}_g]_e(X)$.

So the information about the structure of a Lie group enclosed in it's Lie algebra is applicable across the entire group! We need only work with the Lie algebra. Should we then need to return to working with the Lie group elements there is a special map we can use on the Lie algebra to get back to the group.

2.2.4 From Lie Algebra to Lie Group

This material draws on chapter 2 of [8]. Firstly, we consider the following definition:

Definition 2.26. Let G be a Lie group. A **1-parameter subgroup** of the G is a curve $g : I \subset \mathbb{R} \to G$ which forms a subgroup i.e.

$$g(t_1) \cdot g(t_2) = g(t_1 + t_2)$$

Now suppose we have the element of the Lie algebra $X = \frac{d}{dt}g(t)|_{t=0}$, given by the expansion of the 1-parameter subgroup about the identity $g(0) = e$, and we want to re-obtain the group element $g(t)$. For sufficiently small neighbourhood about $g(0)$, G looks like Euclidean space, since it's a manifold. So to move to a group element $g(t)$ in the neighbourhood we simply act on X with $g(t)$. Thus we want to solve the system $\frac{d}{dt}g(t) = g(t)X$. This is solved by a very familiar looking function:

Definition 2.27. The **exponential** of a matrix $M \in M_n(\mathbb{F})$ is given by the Taylor expansion:

$$\exp(M) = \sum_{i=0}^{\infty} \frac{M^i}{i!} \in M_n(\mathbb{F})$$

Note that as per our discussion in example (2.20), physicist tend to take the complex exponential $\exp(iM)$.

Remark. This solves $\frac{d}{dt}g(t) = g(t)X$:

$$g(t) = \exp(tX) \implies \frac{d}{dt}g(t) = \sum_{i=0}^{\infty} \frac{t^{i-1}}{(i-1)!}X^i = X\sum_{i=0}^{\infty} \frac{t^{i-1}}{(i-1)!}X^{i-1} = g(t)X$$

and $g(0) = \exp(0) = e$ so indeed it does.

Now consider the matrices $\exp(X), \exp(Y) \in M_n(\mathbb{F})$ with $X, Y \in M_n(\mathbb{F})$. In general, for the exponential of matrices we don't have $\exp(X)\exp(Y) = \exp(X+Y)$. In fact this only holds if X and Y commute. The machinery here is captured in the following:

Definition 2.28. There exists a closed formula called the **Baker-Campbell-Hausdorff** formula, given by:

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] - [Y,[X,Y]) + ...) \tag{2.9}$$

Lemma 2.29. Let G be a Lie group. $S_X := \{g(t) =: \exp tX | t \in J \subset \mathbb{R}, X \in \mathfrak{g}\} \subset G$, for a suitable choice of J, is an abelian 1-parameter subgroup.

Proof. Consider two elements of S_X, $\exp(t_1X)$ and $\exp(t_2X)$. Obviously t_1X, t_2X commute since X commutes with itself and the t's are real numbers. Thus we have $\exp(t_1X)\exp(t_2X) = \exp((t_1+t_2)X) = \exp((t_2+t_1)X) = \exp(t_2X)\exp(t_1X)$ which gives both closure, since $t_1, t_2 \in \mathbb{R} \implies t_1 + t_2 \in \mathbb{R}$, and proves S_X is abelian. Setting $t = 0$ gives $\exp(\mathbf{0}) = \mathbb{1}$, and thus we have identity. Finally, it is simple to see that $\exp(t_1X)\exp(-t_1X) = \exp(0) = \mathbb{1}$ and thus we have inverse. ☐

2.3 Some Representation Theory

The material in this section follows chapters 3 and 4 of [8], and 4 and 5 of [11]. Earlier, I alluded earlier to the fact that we can't actually grasp groups in their essence and rely on their actions on certain sets to give us intuition. When we act with a group on a set (it could even be itself!) we do so using a *representation* of some form. Representations of groups are something students will almost certainly have used in some form long before they're introduced to them rigorously. We begin with some definitions:

2.3.1 Some Definitions

Broadly speaking, a **representation** of an algebraic object A is a pair (ρ, V), where V is a vector space called the **representation space** whose dimension gives the dimension of the representation, and $\rho : A \to \mathrm{End}(V)$ is a homomorphism. Across the literature representations are frequently referred to by one of these two things without the other.

Definition 2.30. We will refer predominantly to the following two definitions:

1. For any group G a **representation** is a linear map $D : G \to GL_n(\mathbb{F})$ such that

$$D(g_1)D(g_2) = D(g_1 \cdot g_2), \;\; \forall \, g_1, g_2 \in G$$

   Here the representation space is $\mathbb{F}^n$.

2. For any Lie algebra $\mathfrak{g}$ a **representation** is a linear map $R : \mathfrak{g} \to M_n(\mathbb{F})$ such that:
   (a) $[R(X), R(Y)] = R([X, Y]) \; \forall \, X, Y \in \mathfrak{g}$
   (b) $R(aX + bY) = aR(X) + bR(Y) \; \forall \, X, Y \in \mathfrak{g}$, $a, b \in \mathbb{F}$

Remark. Note that in the case of groups $\mathrm{End}(V) \equiv GL_n(\mathbb{F})$. This is because the representation preserves the algebraic structure and groups have inverses so it maps to a space where all the matrices have inverses. There is no such constraint on Lie algebras.

We will be exclusively considering matrix groups and this matrix Lie algebras in this report, for which we require the following two important examples of representations:

Example 2.31. The **fundamental representation** is defined $R_f(X) = X$, $\forall\ X \in \mathfrak{g}$. The **adjoint representation** is defined $R_{Ad}(X) = \mathrm{ad}_X$ where $\mathrm{ad}_X(Y) = [X, Y]$, $\forall\ Y \in \mathfrak{g}$. E.g. the funamental representation of $\mathfrak{su}(2)$ is the pauli matrices (2.2). We will come to the adjoint representation soon.

Lemma 2.32. Let $\gamma(t)$ be a curve on Lie group G, with $\gamma(0) = e$, and $D : G \to GL_n(\mathbb{F})$ a representation of G. Then

$$R(X) = \frac{d}{dt} D(\gamma(0))$$

is a representation of the Lie algebra of G. In addition, if $d : \mathfrak{g} \to M_n(\mathbb{F})$ is a representation of the Lie algebra of G then

$$D(\exp X) \equiv \exp\left(R(X)\right)$$

is a representation of G.

Definition 2.33. Two representations R_1 and R_2 of a Lie algebra are **equivalent** or **isomorphic** if there exists a non-singular matrix S such that

$$R_2(X) = S R_1(X) S^{-1},\ \ \forall\ X \in \mathfrak{g}$$

S is simply a change of basis in the representation space.

Definition 2.34. A representation R of a Lie algebra $\mathfrak{g}$ with representation space V has an **invariant subspace** $U \subset V$ if $R(X)u \in U\ \forall\ X \in \mathfrak{g},\ u \in U$. Furthermore, a representation of $\mathfrak{g}$ with no non-trivial invariant subspaces is called an **irreducible representation** or **irrep**. i.e. the only invariant subspaces of an irrep are $\{0\}$ and V.

2.3.2 The Representation Theory of $SU(2)$

In this section we will develop the representation theory of $SU(2)$ by analysing the structure of it's Lie algebra. The material in this section draws on chapter 1 of [4] in addition to elements of [8] and [11].

To understanding the structure of a Lie algebra we consider it's action on itself through the adjoint representation. Since the Lie algebra is a vector space we can study it's eigenspace decomposition under this action through the characteristic polynomial. Let's begin by considering the adjoint representation of $SU(2)$:

$$\mathrm{ad}_{T_3}(T_i) = [T_3, T_i] = \begin{cases} iT_2 & i = 1 \\ -iT_1 & i = 2 \\ 0 & i = 3 \end{cases} \implies \mathrm{ad}_{T_3} = i \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The characteristic polynomial of this is:

$$\det(\mathrm{ad}_{T_3} - \alpha \mathbb{1}) = (-\alpha)^3 - (i)(-i)(-\alpha) = -\alpha(\alpha^2 + 1) = \alpha(1 + \alpha)(1 - \alpha) = 0 \quad (2.10)$$

This has solutions: $\alpha = 0, 1, -1$. Let's compute the eigenvectors of the non-zero eigenvalues $\alpha = \pm 1$. Working in the basis $\{T_1, T_2, T_3\}$ still, we have:

$$0 = (\mathrm{ad}_{T_3} - i\mathbb{1})\boldsymbol{v} = \begin{pmatrix} -1 & i & 0 \\ -i & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -v_1 + iv_2 \\ -iv_1 + v_2 \\ -v_3 \end{pmatrix} \implies \boldsymbol{v} = v_1 \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$$

Thus the eigenvector corresponding to $\alpha = 1$ is $T_+ := T_1 + iT_2$. Carrying out the same procedure for $\alpha = -1$ we obtain $T_- := T_1 - iT_2$. We can compute the new commutation relations using (2.7):

$$\begin{gathered} T_z := T_3, \quad T_+ := T_1 + iT_2, \quad T_- := T_1 - iT_2 \\ [T_z, T_+] = T_+, \quad [T_z, T_-] = -T_-, \quad [T_+, T_-] = 2T_z \end{gathered} \tag{2.11}$$

With this we introduce some terminology: the non-zero eigenvalues of the characteristic polynomical are called **roots** and $T_\pm$ are called **root vectors**. The elements with vanishing eigenvalues, in our case just T_z, span the **Cartan subalgebra** (CSA) of $\mathfrak{su}(2)$. Our choice to consider the adjoint representation of T_3 here was trivial, however this is not the case for more complicated groups such as $SU(3)$, and we require it to be en element of the CSA. Together, the elements of the CSA and the root vectors form the **Cartan-Weyl Basis** of the Lie algebra. For a more detailed discussion of this, see Appendix A. Note that the root vectors exist in the complexification $\mathfrak{su}(2)_\mathbb{C}$. This is to be expected since the solutions to (2.10) may be necessarily complex.

Let us now consider a representation $R : \mathfrak{su}(2) \to \mathrm{end}(V)$, where V is the representation space. Taking V to be finite dimensional, and assuming $R(T_z)$ is diagonalizable, V is spanned by the eigenvectors of $R(T_z)$, v_λ s.t. $R(T_z)v_\lambda = \lambda v_\lambda$, $\lambda \in \mathbb{C}$. The eigenvalues of representations of the CSA are called **weights**[1]. Now, observe that:

$$\begin{aligned} R(T_z)R(T_\pm)v_\lambda &= (R(T_\pm)R(T_z) + [R(T_z), R(T_\pm)])v_\lambda = (R(T_\pm)R(T_z) + R([T_z, T_\pm]))v_\lambda \\ &= (R(T_\pm)R(T_z) \pm R(T_\pm))v_\lambda = (\lambda \pm 1)R(T_\pm)v_\lambda \end{aligned}$$

In other words, $R(T_\pm)v_\lambda$ is an eigenvector of $R(T_z)$ with eigenvalue $(\lambda \pm 1)$. Because of this, the $T_\pm$ are often referred to as **ladder operators**. Recall def. (2.34): if R is an irrep then V's only invariant subspaces are $\{0\}$ and itself, thus V will be spanned by the vectors $v_{\Lambda-k} = R(T_-)^k v_\Lambda$, $k \in \mathbb{N}$, where Λ is the **maximum weight**, which must exist since V is finite. Taking Λ to be the max. weight we define the actions:

$$T_z v_\Lambda = \Lambda v_\Lambda, \quad T_+ v_\Lambda = 0 \tag{2.12}$$

We know that we have $R(T_-)v_\Lambda = v_{\Lambda-1}$ and we can thus define all our eigenvectors sequentially: $v_{k-1} \equiv R(T_-)v_k$. Just as we have a maximum weight Λ, this sequence must terminate for some value q, $R(T_-)v_q = 0$. To determine the eigenvalues recursively, we must consider $R(T_+)v_{k-1}$:

$$\begin{aligned} &R(T_+)v_{k-1} = R(T_+)R(T_-)v_k = (2R(T_z) + R(T_-)R(T_+))v_k = (2k + R(T_-)R(T_+))v_k \\ \implies &R(T_+)v_{k-1} - R(T_-)R(T_+)v_k = 2kv_k \implies R(T_+)v_{k-1} \propto v_k \implies R(T_+)v_{k-1} = r_{k-1}v_k \end{aligned}$$

[1] The weights are also linear functionals on the CSA, given by the composition $\lambda \circ R : \mathfrak{h} \to \mathbb{C}$, and can be shown to be given by linear combinations of the roots.

We can determine r_{k-1} by applying the first line above with $R(T_+)v_{k-1} = r_{k-1}v_k$:

$$r_{k-1}v_k = (2R(T_z) + R(T_-)R(T_+))v_k = (2k + r_k)v_k \implies r_k = 2(k+1) + r_{k+1}$$

We can solve this by imposing $r_\Lambda = 0$. For the first couple of n: $r_{\Lambda-1} = 2\Lambda + r_\Lambda = 2\Lambda$, then $r_{\Lambda-2} = 2(\Lambda-1) + r_{\Lambda-1} = 2(\Lambda-1) + 2\Lambda$ and so on. We find:

$$\begin{aligned}
&r_{\Lambda-n} = 2(\Lambda - n + 1) + r_{\Lambda-n+1} = 2n\Lambda - 2\textstyle\sum_{i=1}^{n-1} i = 2n\Lambda - (n-1)n \\
&\implies r_{\Lambda-n} = n\Lambda - \frac{1}{2}n^2 + \frac{1}{2}n \implies r_k = 2(\Lambda-k)\Lambda - (\Lambda-k)^2 + (\Lambda-k) \\
&\implies r_k = \Lambda(\Lambda+1) - k(k+1)
\end{aligned}$$

Now we can use this and $R(T_-)v_q = 0$ to give:

$$\begin{aligned}
&0 = R(T_+)R(T_-)v_q = (R(T_-)R(T_+) + 2R(T_z))v_q = (\Lambda(\Lambda+1) - q(q+1) + 2q) \\
&\implies \Lambda(\Lambda+1) - q(q+1) + 2q = 0 \implies q^2 - q - \Lambda(\Lambda+1) = 0
\end{aligned}$$

This quadratic equation is solved very easily by completing the square to give $q = -\Lambda$ or $q = \Lambda + 1$. Since Λ is the highest weight $q = \Lambda + 1$ isn't possible so $q = -\Lambda$. Thus our eigenvalues on an irrep of highest weight Λ are $\{-\Lambda, -\Lambda+1, ..., \Lambda-1, \Lambda\}$. This is a total of $2\Lambda + 1$ eigenvalues meaning the irrep is $(2\Lambda+1)$-dimensional.

We can then construct a **Casimir operator**:

$$\boldsymbol{T}^2 := [R(T_z)^2 + 1/2(R(T_+)R(T_-) + R(T_-)R(T_+))] \tag{2.13}$$

This is an operator that exists in the **centre** of the **universal enveloping algebra** - the algebra containing all representations of a given Lie algebra. This means it commutes with all other operators. To see why this is useful consider the following:

$$\begin{aligned}
\boldsymbol{T}^2 v_k &= [R(T_z)^2 + 1/2(R(T_+)R(T_-) + R(T_-)R(T_+))]v_k \\
&= [k^2 + 1/2(r_{k-1} + r_k)]v_k \\
&= [k^2 + 1/2(2\Lambda(\Lambda+1) - k(k-1) - k(k+1))]v_k \\
&= [k^2 + \Lambda(\Lambda+1) - k^2]v_k = \Lambda(\Lambda+1)v_k
\end{aligned}$$

$\boldsymbol{T}^2$ has the same eigenvalue regardless of the eigenvector in a given irrep! This allows us to categorise our irreps using their highest weight Λ in the eigenvalue of the Casimir operator. In general, we refer to a **spin-Λ** representation. For example, $\Lambda = 0$ gives a **scalar**, spin-0, representation. This is one-dimensional and transforms trivially under $SU(2)$. $\Lambda = \frac{1}{2}$ corresponds to the 2d fundamental representation.

We round off our discussion of $SU(2)$ with an exploration of how, given it's irreps, we can produce new representations using **direct products/sums**.

Definition 2.35. Let R_1, R_2 be representations of a Lie algebra $\mathfrak{g}$ with representation spaces V_1, V_2, $\dim(V_1) = d_1$, $\dim(V_2) = d_2$ respectively.

1. The **direct sum** $R_1 \oplus R_2$ is a $(d_1 + d_2)$-dimensional representation of $\mathfrak{g}$ that acts on $V_1 \oplus V_2 = \{v_1 \oplus v_2 | v_i \in V_i, i = 1, 2\}$ in the following way:

$$(R_1 \oplus R_2)(X)(v_1 \oplus v_2) = (R_1(X)v_1) \oplus (R_2(X)v_2) \quad \forall \;\; X \in \mathfrak{g}$$

2. The **tensor product** $R_1 \otimes R_2$ is a $d_1 \cdot d_2$-dimensional representation of $\mathfrak{g}$ that acts on $V_1 \otimes V_2 = \{v_1 \otimes v_2 | v_i \in V_i, i = 1, 2\}$ in the following way:

$$(R_1 \otimes R_2)(X)(v_1 \otimes v_2) = (R_1(X)v_1) \otimes v_2 + v_1 \otimes (R_2(X)v_2)$$

Consider two irreps $(R_1, V_1), (R_2, V_2)$ with highest weights and dimensions Λ_1, Λ_2 and d_1, d_2 respectively. Direct sum representations are only relevant when $\Lambda_1 = \Lambda_2$ in which case we obtain a reducible rep. with the same spin as the irreps. Direct products are of greater interest. This gives a $\dim(V_1 \otimes V_2) = d_1 d_2$ representation with highest weight $(\Lambda_1 + \Lambda_2)$:

$$(R_1(T)v_{\Lambda_1}) \otimes u + v \otimes (R_2(T)u_{\Lambda_2}) = \Lambda_1 v_{\Lambda_1} \otimes u_{\Lambda_2} + \Lambda_2 v_{\Lambda_1} \otimes u_{\Lambda_2} = (\Lambda_1 + \Lambda_2) v_{\Lambda_1} \otimes u_{\Lambda_2}$$

We can therefore deduce the irreps of a product representation by acting on $v_{\Lambda_1} \otimes u_{\Lambda_2}$ with $(R_1 \otimes R_2)(T_-)$: $(R_1 \otimes R_2)(T_-)(v_{\Lambda_1} \otimes u_{\Lambda_2}) = v_{\Lambda_1 - 1} \otimes u_{\Lambda_2} + v_{\Lambda_1} \otimes u_{\Lambda_2 - 1}$. Acting repeatedly we obtain the formula:

$$(R_1 \otimes R_2)(T_-)^n (v_{\Lambda_1} \otimes u_{\Lambda_2}) = \sum_{k=0}^{n} \binom{n}{k} v_{\Lambda_1 - n + k} \otimes u_{\Lambda_2 - k}$$

Suppose $\Lambda_1 > \Lambda_2$. Then at $n = 2\Lambda_2$ the sum will contain $2\Lambda + 1$ terms and $u_{\Lambda_2 - n} = u_{-\Lambda_2}$. Thus for $n > 2\Lambda_2$, since $R_2(T_-)u_{-\Lambda_2} = 0$, the final term will always be reduced to zero so the sum will stay at $2\Lambda_2 + 1$ terms. That is until $n = 2\Lambda_1$ at which point for each increment of n we lose a term due to $R_1(T_-)v_{-\Lambda_1} = 0$. Since at $n = 2\Lambda_1$ there were still exactly $2\Lambda_2 + 1$ terms in the sum, we hit zero terms at $n = 2\Lambda_1 + 2\Lambda_2 + 1$ thus confirming the irrep does indeed have dimension 2(highest weight) + 1. Now the dimension of the total product rep. is $\dim(V_1 \otimes V_2) = d_1 d_2 = (2\Lambda_1 + 1)(2\Lambda_2 + 1) = 4\Lambda_1\Lambda_2 + 2\Lambda_1 + 2\Lambda_2 + 1$ which is clearly greater than $2\Lambda_1 + 2\Lambda_2 + 1$. This suggests the product rep. decomposes into a direct sum of irreps! We can find the rest of these irreps by selecting the next highest weight, $\Lambda_1 + \Lambda_2 - 1$, and acting with $(R_1 \otimes R_2)(T_-)$. This process will terminate at the weight $|\Lambda_1 - \Lambda_2|$, since $(2\Lambda_1 + 1)(2\Lambda_2 + 1) = (2(\Lambda_1 + \Lambda_2) + 1) + (2(\Lambda_1 + \Lambda_2 - 1) + 1) + ... + (2|\Lambda_1 - \Lambda_2| + 1)$. We obtain

$$R_{\Lambda_1} \otimes R_{\Lambda_2} = R_{\Lambda_1 + \Lambda_2} \oplus R_{\Lambda_1 + \Lambda_2 - 1} \oplus ... \oplus R_{|\Lambda_1 - \Lambda_2| + 1} \oplus R_{|\Lambda_1 - \Lambda_2|}$$

This is particularly nice when $\Lambda_1 = \Lambda_2$ since $V_1 \otimes V_2$ is decomposable into the symmetric and antisymmetric parts of the product, $v_i \otimes u_j \pm v_j \otimes u_i$.

Example 2.36. Let's compute the direct product of the $R_{1/2}$ representation with itself. The representation space is spanned by $\{v_{1/2}, v_{-1/2}\}$ thus $R_{1/2} \otimes R_{1/2}$ acts on representation space $V \otimes V = \text{span}\{v_{1/2} \otimes v_{1/2}, v_{1/2} \otimes v_{-1/2}, v_{-1/2} \otimes v_{1/2}, v_{-1/2} \otimes v_{-1/2}\}$. We can decompose this into symmetric and anti-symmetric pieces:

$$\{v_{1/2} \otimes v_{1/2}, v_{1/2} \otimes v_{-1/2} + v_{-1/2} \otimes v_{1/2}, v_{-1/2} \otimes v_{-1/2}\}_S \cup \{v_{1/2} \otimes v_{-1/2} - v_{-1/2} \otimes v_{1/2}\}_A$$

Now we act on our highest weight element $v_{1/2} \otimes v_{1/2}$:

$$(R_{1/2} \otimes R_{1/2})(T_-)(v_{1/2} \otimes v_{1/2}) = v_{-1/2} \otimes v_{1/2} + v_{1/2} \otimes v_{-1/2}$$
$$(R_{1/2} \otimes R_{1/2})(T_-)^2(v_{1/2} \otimes v_{1/2}) = 2v_{-1/2} \otimes v_{-1/2}$$
$$(R_{1/2} \otimes R_{1/2})(T_-)^3(v_{1/2} \otimes v_{1/2}) = 0$$

So we have 3-dimensional representation spanned by the symmetric part. Using $(R_{1/2} \otimes R_{1/2})(T_z)$ we see that they have weights $\{1, 0, -1\}$. For the anti-symmetric part: $(R_{1/2} \otimes R_{1/2})(T_-)(v_{1/2} \otimes v_{-1/2} - v_{-1/2} \otimes v_{1/2}) = 0$. Thus we've shown that the symmetric part gives a spin-1 irrep and the anti-symmetric part a spin-0 irrep.

$$R_{1/2} \otimes R_{1/2} = R_1 \oplus R_0$$

For $\Lambda_1 = \Lambda_2 = 1/2$ this is in complete agreement with our formula.

$SU(2)$ will be of central importance in the following chapter, where we will begin to work with groups derived from the symmetries of spacetime. From this we can begin to look at how invariance is encoded into our physical theories and, ultimately, how these notions of invariances can be extended in a fascinating way that forces the evolution of the ideas we've discussed in this chapter.

Chapter 3

Not-So-Super Symmetry

Supersymmetry would not be 'super' if there wasn't some predetermined notion of 'ordinary' symmetry from which it was established as an extension. In this chapter we introduce the symmetry of nature which preceded supersymmetry.

3.1 The Homogeneous and Inhomogeneous Lorentz Groups

Before the 20th century, the accepted theory of relativity was **Galilean relativity** whose symmetry group preserved the standard inner product on Euclidean space, consisting of translations, $SO(3)$ rotations and Galilean boosts $x \to x + vt$. At the dawn of the 20th century, Einstien threw a spanner in the works. Time was no longer absolute, Minkowski spacetime replaced Euclidean space, and the new isometries corresponded to preservation of the **Minkowski metric** $\eta_{\mu\nu}$. The new group of isometries was the **Poincaré group**. We will now introduce this, starting with it's homogeneous part, the Lorentz group. The material of 3.1.1 and 3.1.2 loosely refers to [15], [17], [2].

3.1.1 The Lorentz Group

In direct analogy to the defining relation of the orthogonal group, $R(\theta)^T \mathbb{1} R(\theta) = \mathbb{1}$, the Lorentz group (a **pseudo-orthogonal group**) is defined as follows.

Definition 3.1. The **homogeneous Lorentz group** or simply **Lorentz group** is:

$$O(1,3) := \{\Lambda \mid \Lambda^{\mu}{}_{\rho}\eta_{\mu\nu}\Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma}\}$$

Remark. Note that $\Lambda^{\mu}{}_{\rho}\eta_{\mu\nu}\Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma} \implies \Lambda^{\mu}{}_{\rho}\Lambda_{\mu\sigma} = \eta_{\rho\sigma}$. Multiplying through by $\eta^{\nu\sigma}$ gives: $\Lambda^{\mu}{}_{\rho}\Lambda_{\mu}{}^{\nu} = \delta_{\rho}{}^{\nu}$ thus $(\Lambda^{\mu}{}_{\rho})^{-1} = \Lambda_{\mu}{}^{\nu}$.

To explore the structure of the group, consider $\det(\Lambda^{\mu}{}_{\rho}\eta_{\mu\nu}\Lambda^{\nu}{}_{\sigma})$:

$$\det(\Lambda)\det(\eta)\det(\Lambda) = \det(\eta) \implies -\det(\Lambda)^2 = -1 \implies \det(\Lambda) = \pm 1$$

Via the same procedure as in example (2.15), we know that the sign of $\det(\Lambda)$ specifies two connected components, with $\det(\Lambda) = 1$ forming the subgroup denoted $SO(1,3)$.

These components are related by a **spatial** parity transform. In the $O(1,3)$ case there is another discrete transformation we can consider: **time-reversal**. We have:

$$\eta_{00} = \Lambda^\mu{}_0 \eta_{\mu\nu} \Lambda^\nu{}_0 \implies 1 = (\Lambda^0{}_0)^2 - (\Lambda^i{}_0)^2 \implies 1 + (\Lambda^i{}_0)^2 = (\Lambda^0{}_0)^2$$

so we have that $|\Lambda^0{}_0| \geq 1$. Thus we may have that the transformation reverses time, $\Lambda^0{}_0 \leq 1$, or that it does not, $\Lambda^0{}_0 \geq 1$. It's easy to show this specifies a connected component:

$$[(\mathbb{1} + \omega)\Lambda]^0{}_0 = [\Lambda + \omega\Lambda]^0{}_0 = \Lambda^0{}_0 + (\omega\Lambda)^0{}_0 \xrightarrow{\omega \to 0} \Lambda^0{}_0 \tag{3.1}$$

so: $\Lambda^0{}_0 \geq 1 \implies [(\mathbb{1} + \omega)\Lambda]^0{}_0 \geq 0$ for ω sufficiently small. Thus in total, the Lorentz group has four connected components which we summarise below:

1. **Proper, Orthochronous** transformations: $\det(\Lambda) = +1$, $\Lambda^0{}_0 \geq 1$.
2. **Proper, Non-Orthochronous** transformations: $\det(\Lambda) = +1$, $\Lambda^0{}_0 \leq 1$.
3. **Improper, Orthochronous** transformations: $\det(\Lambda) = -1$, $\Lambda^0{}_0 \geq 1$.
4. **Improper, Non-Orthochronous** transformations: $\det(\Lambda) = -1$, $\Lambda^0{}_0 \leq 1$.

Since the proper, orthochronous transformations contain the identity, they form a Lie subgroup which we denote $SO^+(1,3)$. Each of the four connected components can be obtained by acting on $SO^+(1,3)$ by one of the discrete transformations $\{\mathbb{1}, \mathcal{P}, \mathcal{T}, \mathcal{PT}\}$.

3.1.2 The Poincaré Group

To complete our isometries of the metric we introduce the group of 4d spacetime translations: $\mathcal{R}^{1,3}$. This is abelian since the addition of vectors inherits commutativity from the real numbers. The complete Poincaré group will have a general element of the form (Λ, a). The group multiplication operation is simple to find:

$$\begin{aligned} x' &= (\Lambda_1, a_1)x = \Lambda_1 x + a_1 \\ \implies x'' &= (\Lambda_2, a_2)x' = \Lambda_2 x' + a_2 = \Lambda_2(\Lambda_1 x + a_1) + a_2 = \Lambda_2\Lambda_1 x + \Lambda_2 a_1 + a_2 \end{aligned}$$

Thus we obtain the group product operation $(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2)$. From this procedure we can deduce that the identity is $(\mathbb{1}, 0)$ and thus the inverse of (Λ, a) is $(\Lambda^{-1}, -\Lambda^{-1}a)$. This is not a direct product group of $O(1,3)$ and $\mathcal{R}^{1,3}$. For this to be the case, both $O(1,3)$ and $\mathcal{R}^{1,3}$ would need to be **normal subgroups**, i.e. invariant under the conjugation action. While $\mathcal{R}^{1,3}$ is a normal subgroup;

$$\begin{aligned} &(\Lambda, a)(\mathbb{1}, b)(\Lambda, a)^{-1} = (\Lambda, \Lambda b + a)(\Lambda^{-1}, -\Lambda^{-1}a) = (\mathbb{1}, b), \\ &\text{i.e. } (\mathbb{1}, b) \in \mathcal{R}^{1,3} \implies (\Lambda, a)(\mathbb{1}, b)(\Lambda, a)^{-1} \in \mathcal{R}^{1,3} \end{aligned}$$

$O(1,3)$ is not: $(\Lambda, a)(L, 0)(\Lambda^{-1}, -\Lambda^{-1}a) = (\Lambda L \Lambda^{-1}, a - \Lambda L \Lambda^{-1}a) \notin O(1,3)$. When it is the case that only one of the subgroups is a normal subgroup we have what is called the **semi-direct product group**. Thus we can finally define our full group of isometries.

Definition 3.2. The **inhomogenous Lorentz group**, or simply **Poincaré group**, is defined to be the semi-direct product:

$$ISO(1,3) := \mathcal{R}^{1,3} \rtimes O(1,3)$$

Our task is now to find the Lie algebra of this group.

3.2 The Poincaré Algebra

The Poincaré group is indeed a Lie group and thus we can expand a given element close to the identity. We first examine the Lorentz group individually. The material in this section draws on [15],[17]. Consider the expansion:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} \approx (\delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}) x^{\nu} \tag{3.2}$$

Substituting this into the defining group relation for $SO(1,3)$ gives:

$$\Lambda^{\mu}{}_{\rho} \eta_{\mu\nu} \Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma} \implies (\delta^{\mu}{}_{\rho} + \omega^{\mu}{}_{\rho}) \eta_{\mu\nu} (\delta^{\nu}{}_{\sigma} + \omega^{\nu}{}_{\sigma}) = \eta_{\sigma\rho} + \omega_{\rho\sigma} + \omega_{\sigma\rho} + \mathcal{O}(\omega^2)$$
$$\implies \eta_{\rho\sigma} = \eta_{\sigma\rho} + \omega_{\rho\sigma} + \omega_{\sigma\rho}$$

Now since $\eta_{\rho\sigma} = \eta_{\sigma\rho}$ we must have $\omega_{\rho\sigma} + \omega_{\sigma\rho} = 0 \implies \omega_{\rho\sigma} = -\omega_{\sigma\rho}$. A 4-dimensional antisymmetric tensor (indeed, a 4-dimensional antisymmetric matrix) has exactly 6 degrees of freedom, thus $SO(1,3)$ contains exactly 6 independent Lorentz transformations. Adding the infinitesimal form of the translations into the mix as well, we have:

$$\Lambda^{\mu}{}_{\nu} \approx \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}, \quad a^{\mu} \approx \epsilon^{\mu} \tag{3.3}$$

Note the important subtlety: $\omega_{\rho\nu} \equiv \eta_{\rho\mu}\omega^{\mu}{}_{\nu}$ is antisymmetric, *not* $\omega^{\mu}{}_{\nu}$. $\omega^{\mu}{}_{\nu}$ and ϵ^{μ} generate the infinitesimal transform. $\omega^{\mu}{}_{\nu}$'s 6 degrees of freedom gives the general form:

$$\omega^{\mu}{}_{\nu} \equiv \omega_{01}(J^{01})^{\mu}{}_{\nu} + \omega_{02}(J^{02})^{\mu}{}_{\nu} + \omega_{03}(J^{03})^{\mu}{}_{\nu} + \omega_{12}(J^{12})^{\mu}{}_{\nu} + \omega_{13}(J^{13})^{\mu}{}_{\nu} + \omega_{23}(J^{23})^{\mu}{}_{\nu} \tag{3.4}$$

The $(J^{\rho\sigma})^{\mu}{}_{\nu}$ are called the **generators** of the Lorentz transformations, and the $\omega_{\rho\sigma}$ are the continuous parameters of the transform. By direct comparison to the expansion we considered in section 2.1.4, it's clear these are the elements of the Lie algebra. We divide them into two categories:

$$\text{Boosts: } \{K_1 \equiv J^{01}, K_2 \equiv J^{02}, K_3 \equiv J^{03}\}, \text{ Rotations: } \{L_1 \equiv J^{23}, L_2 \equiv J^{31}, L_3 \equiv J^{12}\}$$

where the $\omega_{\rho\sigma}$ are boost parameters (the **rapidity** of the boost) for the K's and angles for the L's. Now $\omega_{\mu\nu} = -\omega_{\nu\mu}$ obviously implies $(J^{\rho\sigma})_{\mu\nu} = -(J^{\rho\sigma})_{\nu\mu}$ but we also have anti-symmetry in the $\rho\sigma$ indices $(J^{\rho\sigma})_{\mu\nu} = -(J^{\sigma\rho})_{\mu\nu}$. This is intuitive; the $\rho\sigma$ indices essentially label the axes involved in the transformation, and this inverse transformation inverts the axes. This is what happens when we switch $\mu\nu$ so it should also happen if we switch $\rho\sigma$. With this, the $(J^{\rho\sigma})_{\mu\nu}$ may be written in the general form:

$$(J^{\rho\sigma})_{\mu\nu} = \eta^{\rho}{}_{\mu}\eta^{\sigma}{}_{\nu} - \eta^{\rho}{}_{\nu}\eta^{\sigma}{}_{\mu} \equiv \delta^{\rho}{}_{\mu}\delta^{\sigma}{}_{\nu} - \delta^{\rho}{}_{\nu}\delta^{\sigma}{}_{\mu} \tag{3.5}$$

Now using this form we have that:

$$(\omega_{\rho\sigma}J^{\rho\sigma})_{\mu\nu} = \omega_{\rho\sigma}(\delta^{\rho}{}_{\mu}\delta^{\sigma}{}_{\nu} - \delta^{\rho}{}_{\nu}\delta^{\sigma}{}_{\mu}) = 2\omega_{\mu\nu} \implies \omega^{\mu}{}_{\nu} \equiv \frac{i}{2}(\omega_{\rho\sigma}J^{\rho\sigma})^{\mu}{}_{\nu} \tag{3.6}$$

where the factor of i is convention. It assures that when we consider representations on a Hilbert space of quantum states, the corresponding operator will be unitary. For translations, there are 4 independent generators, one per coordinate axis: $\{P^0, P^1, P^2, P^3\}$. Note that P^0 is the *time-translation*, or more familiarly, *time-evolution* operator, *i.e.* it is the **Hamiltonian**, $P^0 = H$. We thus have the following expression:

$$\epsilon \equiv -i\epsilon^{\rho}P_{\rho} \tag{3.7}$$

A general infinitesimal transformation acting on the point x^μ thus has the form:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} + a^{\mu} \approx (\delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu})x^{\nu} + \epsilon^{\mu} = x^{\mu} + \frac{i}{2}(\omega_{\rho\sigma}J^{\rho\sigma})^{\mu}{}_{\nu}x^{\nu} - i\epsilon^{\rho}P_{\rho}x^{\mu} \tag{3.8}$$

These are the generators of the Poincaré group. They satisfy the **Poincaré Algebra**:

$$\begin{aligned} i[J^{\mu\nu}, J^{\rho\sigma}] &= \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\sigma\mu}J^{\rho\nu} + \eta^{\sigma\nu}J^{\rho\mu} \\ i[P^{\mu}, J^{\rho\sigma}] &= \eta^{\mu\rho}P^{\sigma} - \eta^{\mu\sigma}P^{\rho} \\ [P^{\mu}, P^{\nu}] &= 0 \end{aligned} \tag{3.9}$$

A full derivation of this algebra is given in appendix B. The key is to use that the conjugation action (2.6) naturally produces the Lie algebra of generators by substitution of the expansion. Suppose $H \subset G$ is a Lie subgroup of Lie group G and g and h have infinitesimal forms $g \approx (e + \omega)$, $h \approx (e + \epsilon)$. The for $h_1 \neq h_2$ we have

$$gh_1g^{-1} = h_2 \implies (e+\omega)(e+\epsilon_1)(e-\omega) = (e+\epsilon_2) \implies \epsilon_1 + [\omega, \epsilon_1] = \epsilon_2 + \mathcal{O}(\omega^2)$$

Discarding $\mathcal{O}(\omega^2)$, clearly $h_1 = h_2 \implies \epsilon_1 = \epsilon_2 \implies [\omega, \epsilon_1] = 0$ as we would expect. For completeness, we add that the discrete symmetries of parity $\boldsymbol{P}$ and time-reversal $\boldsymbol{T}$ have the following action on the rotations and boosts:

$$\begin{aligned} \boldsymbol{P}L_i\boldsymbol{P}^{-1} &= +L_i, \quad \boldsymbol{P}K_i\boldsymbol{P}^{-1} = -K_i, \quad \boldsymbol{P}P_i\boldsymbol{P}^{-1} = -P_i \\ \boldsymbol{T}L_i\boldsymbol{T}^{-1} &= -L_i, \quad \boldsymbol{T}K_i\boldsymbol{T}^{-1} = +K_i, \quad \boldsymbol{T}P_i\boldsymbol{T}^{-1} = -P_i \end{aligned} \tag{3.10}$$

We are now ready to consider building representations of this algebra.

3.3 Representations on Fields

The canonical form of physical theories that are invariant under special relativity is as a field theory. In field theory, the dynamics of a system is encoded into the **action**,

$$\mathcal{S}[\phi_a] = \int \mathcal{L}(\phi_a, \partial\phi_a)\, d^4x \tag{3.11}$$

where $\mathcal{L}$ is the **Lagrangian density** $\mathcal{L} = \mathcal{T} - \mathcal{V}$. $\mathcal{T}$ and $\mathcal{V}$ are the kinetic and potential energy densities respectively and it is these objects which are built from the fields $\phi_a(x)$ themselves. The equations of motion are given by applying the **principle of stationary action** which mathematically equates to taking the first variation of the action and asserting $\delta\mathcal{S} = 0$. In the general case, this produces the **Euler-Lagrange equations**; a system of differential equations from which equations of motion can be found:

$$\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\Big) = 0 \tag{3.12}$$

Thus by building our lagrangians out of fields which carry a representations of the Poincaré algebra, Poincaré symmetry is encoded into the theory automatically.

3.3.1 Representations of $SO(1,3)$

We draw on [15] and [3]. To develop the representation theory of $SO(1,3)$ we begin by examining the commutator relations of the rotation and boost generators, $L_i = \epsilon_{ijk}J^{jk}$ and $K_i = J^{0i}$, individually. Using (3.9) we obtain:

$$[L_i, L_j] = i\epsilon_{ijk}L_k, \quad [K_i, K_j] = -i\epsilon_{ijk}L_k, \quad [L_i, K_j] = i\epsilon_{ijk}K_k \tag{3.13}$$

These are nothing but the commutation relations we saw in (2.8) for $\mathfrak{su}(2)_\mathbb{C} \simeq \mathfrak{sl}(2,\mathbb{C})$! Furthermore, we can decouple these relations by defining the operator $N_i^\pm := \frac{1}{2}(L_i \pm iK_i)$ for which we obtain, using (3.13), the relations:

$$[N_i^+, N_j^-] = 0, \quad [N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+, \quad [N_i^-, N_j^-] = i\epsilon_{ijk}N_k^- \tag{3.14}$$

which quite clearly indicates the two of $\mathfrak{su}(2)$-like algebras. Using (3.10) we can clearly see these two algebras are related by parity. From our knowledge of $\mathfrak{su}(2)$ we can also conclude that $(N^+)^2 \equiv N_i^+N_i^+$ and $(N^-)^2 \equiv N_i^-N_i^-$ are Casimir operators for their respective copies of $\mathfrak{su}(2)$. Suppose $(N^+)^2$ has value $n_1(n_1+1)$ on a given representation, and $(N^-)^2$ has value $n_2(n_2+1)$. This means the two representations have maximum weight n_1, n_2 respectively. Now since $N_i^+ + N_i^- = L_i$ this gives the total spin of the representation of $\mathfrak{so}(1,3)$ as $n_1 + n_2$. Thus representation of $\mathfrak{so}(1,3)$ correspond to direct products of irreps of $\mathfrak{su}(2)$! We denote them (n_1, n_2) which is short hand for $R_{n_1} \otimes R_{n_2} = R_{n_1} \otimes \mathbb{1}_{n_2} + \mathbb{1}_{n_1} \otimes R_{n_2}$. This provides us with a beautifully simple way in which to discuss the irreps of $\mathfrak{so}(1,3)$ using our knowledge of $\mathfrak{su}(2)$.

As it turns out this structure of the Lorentz algebra is due to the fact that $SL_2(\mathbb{C})$ is a **double cover** of $SO(1,3)$:

$$SO(1,3) \cong SL_2(\mathbb{C})/\mathbb{Z}_2 \tag{3.15}$$

3.3.2 The Infinitesimal Field Operator and Scalar Fields

At this point, you may be wondering why, having referred to representations of the Poincaré algebra on states and fields, we've just spent a page or so talking about representations of just the Lorentz algebra. We will clarify this now with the help of the Poincaré generators and a new tool:

For an arbitrary infinitesimal transformation $x \mapsto x' = x + \delta x$ acting on any field $f_i : \mathbb{R}^{1,3} \to \mathbb{R}$ can be written as

$$\begin{aligned}\delta f_i &= f_i'(x') - f_i(x) \\ &= f_i'(x + \delta x) - f_i(x) \\ &= f_i'(x) + \delta x^\mu \partial_\mu f_i'(x)_i - f_i(x) + \mathcal{O}(\delta x^2)\end{aligned} \tag{3.16}$$

This is a very standard application of Taylor's theorem. In addition, provided the δx is sufficiently small, which it must be in order to apply Taylor's theorem, we can safely assume that $\partial_\mu f_i'(x) = \partial_\mu f_i(x)$ since $f_i'(x)$ will reside in a sufficiently small neighbourhood of $f_i(x)$ by continuity. Thus we have the following definition:

Definition 3.3. The infinitesimal change in a field $f_i : \mathbb{R}^{1,3} \to \mathbb{R}$ under infinitesimal transformation $x \mapsto x' = x + \delta x$ is given by the action of the operator $\delta = \delta_0 + \delta^\mu \partial_\mu$:

$$\delta f_i = \delta_0 f_i + \delta x^\mu \partial_\mu f_i \tag{3.17}$$

where $\delta_0 f_i \equiv f_i'(x) - f(x)$ is the **functional change** and $\delta x^\mu \partial_\mu f_i$ is known as the **transport term**.

With this we are very immediately able to see why the we need only consider the Lorentz algebra for fields. The coordinate derivatives ∂_μ, as linear operators on space, furnish a representation (satisfy (3.9)) of the translation operators, $P_\mu = -i\partial_\mu$. Indeed, the ∂_μ are tangent vectors on $\mathbb{R}^{1,3}$. Thus translations act on f_i as derivatives:

$$P_\mu : f_i \mapsto f_i' = f_i - i\epsilon^\mu P_\mu f_i, \quad i\epsilon^\mu P_\mu = \epsilon^\mu \partial_\mu$$

thus $\delta_0 f_i \equiv f_i'(x) - f(x) = -i\epsilon^\mu P_\mu f_i = -\epsilon^\mu \partial_\mu f_i$ so $\delta f = 0$. This essentially means fields automatically carry a representation of translations and we can simply concern ourselves with finding the field representations of the part of the Poincaré algebra that doesn't concern translations: the Lorentz algebra.

We are now ready to explore the field representations. We begin with the very simple singlet or **scalar** representation $(0, 0)$. It's corrsponding field is a scalar field $\phi(x)$ which transforms trivially under the Lorentz group: $\phi'(x') = \phi(x)$. We can manipulate this to reveal how the Lorentz generators act on ϕ. $\phi'(x') = \phi(x)$ implies:

$$\begin{aligned}\phi'(x) &= \phi(\Lambda^{-1}x) \approx \phi(x^\mu - \omega^\mu{}_\nu x^\nu) \approx \phi(x) - \omega^\mu{}_\nu x^\nu \partial_\mu \phi(x) \\ &= \phi(x) - \frac{i}{2}(\omega_{\rho\sigma} J^{\rho\sigma})^\mu{}_\nu x^\nu \partial_\mu \phi(x) = \phi(x) - \frac{i}{2}\omega_{\rho\sigma}(\eta^{\rho\mu}\eta^\sigma{}_\nu - \eta^\rho{}_\nu \eta^{\sigma\mu}) x^\nu \partial_\mu \phi(x) \\ &= \phi(x) - \frac{i}{2}\omega_{\rho\sigma}(x^\sigma \partial^\rho - x^\rho \partial^\sigma)\phi(x) \\ \therefore \delta\phi = 0 &\implies \delta_0 \phi = -\frac{i}{2}\omega_{\rho\sigma}(x^\sigma \partial^\rho - x^\rho \partial^\sigma)\phi\end{aligned} \tag{3.18}$$

This action defines the behaviour of a **scalar field** of the Lorentz group.

3.3.3 Spinor Fields

We follow the discussion in 1.4 of [15], and draw on [12]. Next we have the spin-$\frac{1}{2}$ half representations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, related by hermitian conjugation. These are 2-dimensional: our corresponding fields will transform like 2d vectors in the fundamental representation of $SL(2,\mathbb{C})$. We refer to these as the **left-handed spinor** and **right-handed spinor** representations respectively, and the corresponding left and right fields, $\psi_L \equiv (\frac{1}{2}, 0)$ and $\psi_R \equiv (0, \frac{1}{2})$, are called **spinor fields**.

To build Lorentz invariant forms from ψ_L and ψ_R we need to know how they transform under the Lorentz group. These transformations are matrices in $SL(2,\mathbb{C})$, obtained by taking the exponential map of generators in it's Lie algebra. Using (3.13) with (2.8) we see that in the spinor representation we can choose $L_i = \frac{1}{2}\sigma^i$ and $K_i = \frac{i}{2}\sigma^i$. We have:

$$\mathcal{M}_L := \exp\left[\frac{i}{2}\boldsymbol{\sigma}(\boldsymbol{\theta} - i\boldsymbol{\omega})\right], \quad \mathcal{M}_R := \exp\left[\frac{i}{2}\boldsymbol{\sigma}(\boldsymbol{\theta} + i\boldsymbol{\omega})\right] \tag{3.19}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices, $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ are the angles corresponding to rotations, and $\boldsymbol{\omega} = (\omega_{01}, \omega_{02}, \omega_{03})$ are the boost parameters.

Remark. These are not equivalent reps. *i.e.* $\nexists\, C \in M_2(\mathbb{C})$ s.t. $C\mathcal{M}_L C^{-1} = \mathcal{M}_R$.

Despite not being equivalent, the parity relating the matrices leads to a more helpful expression relating $\mathcal{M}_L$ and $\mathcal{M}_R$:

$$\mathcal{M}_R^\dagger = \exp\left[\frac{i}{2}\boldsymbol{\sigma}(\boldsymbol{\theta}+i\boldsymbol{\omega})\right]^\dagger = \exp\left[-\frac{i}{2}\boldsymbol{\sigma}^\dagger(\boldsymbol{\theta}-i\boldsymbol{\omega})\right] = \exp\left[-\frac{i}{2}\boldsymbol{\sigma}(\boldsymbol{\theta}-i\boldsymbol{\omega})\right] = \mathcal{M}_L^{-1} \tag{3.20}$$

Now we have our Lorentz transformations, we want to build Lorentz scalars out of them. Recall that in Chapter 2 we established $R_{1/2} \otimes R_{1/2} = R_1 \oplus R_0$, thus we have that $(\frac{1}{2}, 0) \otimes (\frac{1}{2}, 0) = (1, 0) \oplus (0, 0)$. The scalar representation arises in the anti-symmetric product so let's try defining an anti-symmetric product of spinors and see where we get. We define the anti-symmetric bilinear form $\epsilon := i\sigma^2$. To define our scalar product, we prove the following:

Proposition 3.4. Let σ^i, $i = 1, 2, 3$ denote the Pauli matrices and $\mathcal{M}_L, \mathcal{M}_R$ be as in (3.19). Then we have the following:

$$(1) : \epsilon\sigma^i\epsilon^{-1} = -(\sigma^i)^*, \quad (2) : \epsilon\mathcal{M}_L\epsilon^{-1} = \mathcal{M}_R^*, \quad (3) : \mathcal{M}_L^T\epsilon\mathcal{M}_L = \epsilon$$

Proof. To prove (1) note that $\epsilon^{-1} = -\epsilon$:

$$\epsilon\sigma^1\epsilon^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -\sigma^1 = -(\sigma^1)^*$$
$$\epsilon\sigma^2\epsilon^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma^2 = -(\sigma^2)^*$$
$$\epsilon\sigma^3\epsilon^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -\sigma^3 = -(\sigma^3)^*$$

To prove (2), observe that for $A, B \in M_n(\mathbb{C})$ we have $(BAB^{-1})^n = B(A)^nB^{-1}$ and thus $B\exp(A)B^{-1} = B(\sum_k \frac{1}{k!}(A)^k)B^{-1} = \sum_k \frac{1}{k!}(BAB^{-1})^k = \exp(BAB^{-1})$:

$$\epsilon\mathcal{M}_L\epsilon^{-1} = \epsilon e^{\frac{i}{2}\boldsymbol{\sigma}(\boldsymbol{\theta}-i\boldsymbol{\omega})}\epsilon^{-1} = e^{\frac{i}{2}\epsilon\boldsymbol{\sigma}\epsilon^{-1}(\boldsymbol{\theta}-i\boldsymbol{\omega})} \overset{(1)}{=} e^{-\frac{i}{2}\boldsymbol{\sigma}^*(\boldsymbol{\theta}-i\boldsymbol{\omega})} = \mathcal{M}_R^*$$

And finally, by (3.20) we have $\mathcal{M}_R^\dagger = \mathcal{M}_L^{-1} \implies \mathcal{M}_R^* = (\mathcal{M}_L^{-1})^T$ thus (2) gives $\epsilon\mathcal{M}_L = (\mathcal{M}_L^{-1})^T\epsilon$ which in turn, using $(A^{-1})^T = (A^T)^{-1}$, gives $\mathcal{M}_L^T\epsilon\mathcal{M}_L = \epsilon$. This proves (3) and we are done. □

The same relations exist with R and L reversed and the proof is completely analogous. Now consider two spinors χ_L, ψ_L. We define their scalar product by $(\chi_L, \psi_L) := \chi_L^T\epsilon\psi_L$. Clearly this is Lorentz invariant:

$$\chi_L'^T\epsilon\psi_L' = (\mathcal{M}_L\chi_L)^T\epsilon\mathcal{M}_L\psi_L = \chi_L^T\mathcal{M}_L^T\epsilon\mathcal{M}_L\psi_L \overset{(3)}{=} \chi_L^T\epsilon\psi_L$$

However, there's a problem. Given a spinor field ψ_L it's associated Lorentz scalar is:

$$(\psi_L, \psi_L) = \psi_L^T\epsilon\psi_L = \begin{pmatrix}\psi_{L1} & \psi_{L2}\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\psi_{L1}\\ \psi_{L2}\end{pmatrix} = \psi_{L1}\psi_{L2} - \psi_{L2}\psi_{L1} \tag{3.21}$$

So clearly $\psi_L^T\epsilon\psi_L = 0$ for $\psi_{Li} \in \mathbb{C}, i = 1, 2$. In fact we've discovered a very important point: the numbers ψ_{Li} are not ordinary complex numbers but in fact complex **Grassmann numbers**[1] which are anti-commuting; $\{\psi_{L1}, \psi_{L2}\} = \psi_{L1}\psi_{L2} + \psi_{L2}\psi_{L1} = 0$.

Van der Waerden Notation

It worth noting the similarity between the role of ϵ, for our spinor fields, and $\eta^{\mu\nu}$ for Lorentz transformations. This is more than just coincidence, and is a link best appreciated in a tensor-like conception of spinor analysis developed first by B.L. Van der Waerden in 1929. Suppose that $\mathcal{M}_L = \mathcal{M}_\alpha{}^\beta$, $\alpha, \beta = 1, 2$. Then defining ψ_L to be ψ_α our transformation takes the form $\psi_\alpha' = \mathcal{M}_\alpha{}^\beta\psi_\beta$. This set up is nice as we can make use of already familiar tensor conventions: for $\mathcal{M}_\alpha{}^\beta$ we have $(\mathcal{M}^{-1})_\beta{}^\alpha$ and $(\mathcal{M}^T)^\beta{}_\alpha$. To find the form of $\mathcal{M}_R$ we employ (3.20) and these conventions along with an additional one unique to this analysis: $\mathcal{M}_L^* = (\mathcal{M}^*)_{\dot\alpha}{}^{\dot\beta}$. We obtain:

$$\mathcal{M}_R = \mathcal{M}_L^{-1\dagger} = ((\mathcal{M}^\dagger)^{\dot\beta}{}_{\dot\alpha})^{-1} = (\mathcal{M}^{-1\dagger})^{\dot\alpha}{}_{\dot\beta} \tag{3.22}$$

Thus we can define ψ_R to be $\bar\psi^{\dot\alpha}$ and transform as $\bar\psi'^{\dot\alpha} = (\mathcal{M}^{-1\dagger})^{\dot\alpha}{}_{\dot\beta}\bar\psi^{\dot\beta}$. This notation greatly simplifies our scalar product. The ϵ tensor is:

$$\epsilon^{\alpha\beta} = \epsilon^{\dot\alpha\dot\beta} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \epsilon_{\alpha\beta} = \epsilon_{\dot\alpha\dot\beta} = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \text{ s.t } \epsilon^\alpha{}_\beta = \epsilon^{\dot\alpha}{}_{\dot\beta} = \epsilon^{\alpha\delta}\epsilon_{\delta\beta} = \mathbb{1}$$

The ϵ's raise and lower their respective (dotted/undotted) indices, and our scalar product takes the more elegant form: $\psi^\alpha\psi_\alpha = \epsilon^{\alpha\beta}\psi_\beta\psi_\alpha$, $\bar\psi_{\dot\alpha}\bar\psi^{\dot\alpha} = \epsilon_{\dot\alpha\dot\beta}\bar\psi^{\dot\alpha}\bar\psi^{\dot\beta}$. It is therefore useful to think of ψ_α, $\bar\psi^{\dot\alpha}$ as column vectors and ψ^α, $\bar\psi_{\dot\alpha}$ as row vectors. The ψ^α, $\bar\psi_{\dot\alpha}$ transform in equivalent representations to their respective column counterparts. Indeed,

[1]Strictly speaking Grassmann numbers exist in the **exterior algebra** of the complex numbers but this is mathematical rigour we need not discuss here. They are named after Hermann Grassmann.

$\psi'^{\alpha} = (\mathcal{M}^{-1T})^{\alpha}{}_{\beta}\psi^{\beta}$, where $(\mathcal{M}^{-1T})^{\alpha}{}_{\beta} = \epsilon^{\alpha\gamma}\mathcal{M}_{\gamma}{}^{\delta}\epsilon_{\delta\beta}$. This is exactly $\mathcal{M}_R^* = \epsilon\mathcal{M}_L\epsilon^{-1}$ which we saw in (3.4). Similarly, $\mathcal{M}_L^* = \epsilon\mathcal{M}_R\epsilon^{-1}$ corresponds to $\bar{\psi}_{\dot{\alpha}}$. It's worth noting that, using our conventions, $(\psi'_{\alpha})^{\dagger} = (\mathcal{M}_{\alpha}{}^{\beta}\psi_{\beta})^{\dagger}$ becomes $(\psi'_{\alpha})^{\dagger} = (\psi_{\beta})^{\dagger}(\mathcal{M}^{\dagger})^{\dot{\beta}}{}_{\dot{\alpha}}$ which implies $(\psi_{\alpha})^{\dagger} = \bar{\psi}_{\dot{\alpha}}$. Thus we may write $\psi_L^{\dagger}\psi_R \equiv (\psi_{\alpha})^{\dagger}\bar{\psi}^{\dot{\alpha}} = \psi_{\dot{\alpha}}\psi^{\dot{\alpha}}$ and similarly $\psi_R^{\dagger}\psi_L = \psi^{\alpha}\psi_{\alpha}$. This equivalent formulation of the scalar product proves useful in the next discussion.

Contracting all spinor indices is not the only way we can build Lorentz scalars from spinor fields. Consider a product of the form $(\psi_{\alpha})^{\dagger}\psi_{\alpha} = \bar{\psi}_{\dot{\alpha}}\psi_{\alpha}$. This transforms as $(\frac{1}{2}, 0) \otimes (0, \frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})$ *i.e.* as a 4-vector! Contracting all spinor and vector indices will give us a new invariant. However to contract the spinor indices we require a single object with both dotted and undotted indices. To explore this let's simply consider how $(\psi_L)^{\dagger}\psi_L$ transforms:

$$\begin{aligned}(\psi_L)^{\dagger}\psi_L \rightarrow &(e^{\frac{i}{2}\sigma^i(\theta_i - i\omega_i)}\psi_L)^{\dagger}(e^{\frac{i}{2}\sigma^j(\theta_j - i\omega_j)}\psi_L) = \psi_L^{\dagger}(e^{\frac{i}{2}\sigma^i(\theta_i - i\omega_i)})^{\dagger}e^{\frac{i}{2}\sigma^j(\theta_j - i\omega_j)}\psi_L \\ &= \psi_L^{\dagger}e^{-\frac{i}{2}\sigma^i(\theta_i + i\omega_i)}e^{\frac{i}{2}\sigma^j(\theta_j - i\omega_j)}\psi_L \approx \psi_L^{\dagger}[1 - \frac{i}{2}\sigma^i(\theta_i + i\omega_i)][1 + \frac{i}{2}\sigma^j(\theta_j - i\omega_j)]\psi_L \\ &\approx \psi_L^{\dagger}[1 - \frac{i}{2}\sigma^i(\theta_i + i\omega_i) + \frac{i}{2}\sigma^i(\theta_i - i\omega_i)]\psi_L = \psi_L^{\dagger}\psi_L + \omega_i\psi_L^{\dagger}\sigma^i\psi_L\end{aligned}$$

i.e. $\delta\psi_L^{\dagger}\psi_L = \omega_i\psi_L^{\dagger}\sigma^i\psi_L$. By the same process, and using $\{\sigma^i, \sigma^j\} = 2\mathbb{1}$ if $i = j$ and 0 otherwise, we find that $\delta\psi_L^{\dagger}\sigma^i\psi_L = \omega_i\psi_L^{\dagger}\psi_L$. These transformation properties are exactly those of a 4-vector $(\psi_L^{\dagger}\psi_L, \psi_L^{\dagger}\sigma^i\psi_L)$! This solves the problem of our mysterious object: the Pauli matrices carry one dotted and one undotted spinor index! We have:

$$\bar{\psi}_{\dot{\alpha}}(\sigma^{\mu})^{\dot{\alpha}\alpha}\psi_{\alpha},\ \sigma^{\mu} := (\sigma^0 \equiv \mathbb{1}, \sigma^i) \quad \psi^{\alpha}(\bar{\sigma}^{\mu})_{\alpha\dot{\alpha}}\bar{\psi}^{\dot{\alpha}},\ \bar{\sigma}^{\mu} := (\sigma^0 \equiv \mathbb{1}, -\sigma^i) \tag{3.23}$$

Before we move on from spinors we will consider the reducible $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation. To do this we introduce a very special and powerful representation.

3.3.4 The Clifford Algebra

This subsection is based on chapter 4 of [16]. The **Clifford Algebra** is given by the following relation:

$$\{\gamma^{\mu}, \gamma^{\nu}\} := \gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2\eta^{\mu\nu} \tag{3.24}$$

The Clifford algebra was discovered long before it's relevance in physics was known. In the 1920's, the algebra arose naturally out of Paul Dirac's attempt to produce a relativistic wave equation describing spin-$\frac{1}{2}$ particles. Taking:

$$S^{\mu\nu} := \frac{1}{4}[\gamma^{\mu}, \gamma^{\nu}] = \begin{cases} 0, & \mu = \nu \\ \frac{1}{2}\gamma^{\mu}\gamma^{\nu}, & \mu \neq \nu \end{cases} = \frac{1}{2}(\gamma^{\mu}\gamma^{\nu} - \eta^{\mu\nu}) \tag{3.25}$$

it's not difficult to show using (3.24) that $[S^{\mu\nu}, S^{\rho\sigma}]$ satisfies (3.9). But what are the γ^{μ}'s? Using (3.24) it's easy to see that $(\gamma^0)^2 = 1$, $(\gamma^i)^2 = -1$ and $\gamma^{\mu}\gamma^{\nu} = -\gamma^{\nu}\gamma^{\mu}$ $\mu \neq \nu$. Clearly the γ^μ must be matrices of some sort to satisfy this. Now for we have $\mu \neq \nu$ $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu \implies \det(\gamma^\mu)\det(\gamma^\nu) = \det(\gamma^\mu\gamma^\nu) = \det(-\gamma^\nu\gamma^\mu) = (-1)^n\det(\gamma^\nu)\det(\gamma^\mu)$. Thus if n is odd $\det(\gamma^\mu)\det(\gamma^\nu) = 0 \implies \det(\gamma^\mu) = 0$, or $\det(\gamma^\nu) = 0$. Either way, $(\gamma^0)^2 = 1 \implies \det(\gamma^0) = \pm 1$, and $(\gamma^i)^2 = -1 \implies \det(\gamma^i) = (-1)^{n/2}$ which are both non-zero. This means we need only consider $2n \times 2n$ matrices. The 2×2 case is simply the Pauli matrices with $\sigma^0 = \mathbb{1}$ however these don't satisfy (3.24) (e.g. $\{\sigma^0, \sigma^2\} = 2\sigma^2 \neq 0$) thus we consider the 4×4 case. It turns out there is a unique irrep up to equivalence for 4×4 given by the **Chiral/Weyl basis**:

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{3.26}$$

It is now natural to introduce the **Dirac spinor** $\Psi_{\hat{\alpha}} \equiv (\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$:

$$\Psi_{\hat{\alpha}} := \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \equiv \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}, \quad \hat{\alpha} = (\alpha, \dot{\alpha}) = (1, 2, \dot{1}, \dot{2})$$

This transforms under $\mathcal{M} = \mathcal{M}_L \oplus \mathcal{M}_R$ as $\Psi'(x') = \mathcal{M}\Psi(x)$. The scalar product for the Dirac spinor is thus:

$$\psi_R^\dagger\psi_L + \psi_L^\dagger\psi_R = \left(\begin{smallmatrix} \psi_R^\dagger & \psi_L^\dagger \end{smallmatrix} \right)\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \left(\begin{smallmatrix} \psi_L^\dagger & \psi_R^\dagger \end{smallmatrix} \right)\left(\begin{smallmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{smallmatrix} \right)\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \Psi^\dagger\gamma^0\Psi =: \bar{\Psi}\Psi \tag{3.27}$$

Our vector forms (3.23) become $\bar{\Psi}\gamma^\mu\Psi$. There is plenty more we could say about the Clifford algebra. It is a fruitful approach to spinor representations of the Lorentz group for many reasons: it's easy to generalize to higher dimensions which is necessary in string theory and supergravity (see [10]), it streamlines the approach to building chiral theories such as electroweak theory, and it enables us to encode properties involving parity and charge symmetry into our theories via a mere change of basis. We'll see in 3.3 how it's representations correspond to half integer spin particles, and in chapters 4 and 5 it will appear briefly as we discuss supersymmetric transformations.

3.3.5 Higher Spins

Next we have spin-1 representations: $(1, 0)$, $(0, 1)$ and $(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$. The most important spin-1 representation to mention is the **vector field** representation $(\frac{1}{2}, \frac{1}{2})$. Indeed, we've just seen how $\Psi^\dagger\gamma^\mu\Psi$ transforms as a vector. We can compute explicitly how a vector field $A_\mu(x)$ transforms:

$$\begin{aligned} A'_\mu(x') = \Lambda_\mu{}^\nu A_\nu(x) &\implies A'_\mu(x') \approx (\delta_\mu{}^\nu + \omega_\mu{}^\nu)A_\nu(x) \implies \delta A_\mu = \omega_\mu{}^\nu A_\nu(x) \\ \text{Now } A'_\mu(x + \delta x) \approx (\delta_\mu{}^\nu + \omega_\mu{}^\nu)A_\nu(x) &\implies A'_\mu(x) + \delta x^\nu\partial_\nu A_\mu(x) = A_\mu(x) + \omega_\mu{}^\nu A_\nu(x) \\ &\implies \delta_0 A_\mu = \omega_\mu{}^\nu A_\nu(x) - \delta x^\nu\partial_\nu A_\mu(x) = \omega_\mu{}^\nu A_\nu(x) - \delta x^\nu\partial_\nu A_\mu(x) \\ &\implies \delta_0 A_\mu = \frac{i}{2}(\omega_{\rho\sigma}J^{\rho\sigma})_\mu{}^\nu A_\nu(x) - \frac{i}{2}\omega_{\rho\sigma}(x^\rho\partial^\sigma - x^\sigma\partial^\rho)A_\mu(x) \end{aligned}$$

Contrary to the scalar field case, the first term in $\delta_0 A_\mu$ shows the Lorentz transform acts directly on the fields indices, not merely acting as a derivative on a function. This reflects the fact that the vector field actually has non-zero spin. It's Lorentz invariants are found in the usual way for vectors; by contracting the Lorentz indices. E.g.

$$A_\mu(x)A^\mu(x),\ \ \partial^\mu A_\mu(x),\ \ \partial_\nu A_\mu(x)\partial^\nu A^\mu(x), ...$$

Now the $(1,0)$ and $(0,1)$ are not vectors. $(1,0)$, for example, corresponds to the symmetric tensor product part of $(\frac{1}{2},0)\otimes(\frac{1}{2},0)$. Recall that this is spanned by $\{v_{1/2}\otimes v_{1/2}, (v_{1/2}\otimes v_{-1/2}+v_{-1/2}\otimes v_{1/2}), v_{-1/2}\otimes v_{-1/2}\}$, *i.e.* it corresponds to 3 independent Lorentz components that transform among themselves under the action of the Lorentz group. What object does this correspond to? Well, an anti-symmetric tensor in 4 dimensions $B_{\mu\nu}=-B_{\nu\mu}$ has $\frac{1}{2}(4)(4-1)=6$ components. Thus a **self-dual** ($B_{\mu\nu}=\frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}B_{\rho\sigma}$) anti-symmetric tensor has half the number of independent components again, *i.e.* 3. Thus the reps $(1,0)$, $(0,1)$ correspond to self-dual and anti self-dual anti-symmetric tensor fields respectively.

3.4 Application: Noether's Theorem

The material of this section is based on chapters 1 and 4 of [16]. Having built various types of field irreps, we will now give an example of their physical application, and how this links to observable quantities. There is a very profound consequence of the presence of a continuous symmetry in a physical theory that was first formalised by the brilliant Emmy Noether in the early 20th century. Noether's theorem states that every continuous symmetry of a physicall system gives rise to a **conserved current** $j^\mu(x)$ such that the equations of motion imply $\partial_\mu j^\mu(x)=\frac{\partial j^0}{\partial t}+\boldsymbol{\nabla}\cdot\boldsymbol{j}=0$. In addition, this implies the presence of a **conserved charge**:

$$Q:=\int_{\mathbb{R}^3} j^0\, d^3x \implies \frac{dQ}{dt}=\int_{\mathbb{R}^3}\frac{\partial j^0}{\partial t}\, d^3x=-\int_{\mathbb{R}^3}\boldsymbol{\nabla}\cdot\boldsymbol{j}\, d^3x=0 \tag{3.28}$$

Lets consider a general field $\phi_a(x)$, a some index. This transforms as $\phi_a\to\phi_a+\delta\phi_a$. The Lagrangian then becomes:

$$\mathcal{L}(\phi_a+\delta\phi_a,\partial_\mu\phi_a+\partial_\mu\delta\phi_a)=\mathcal{L}(\phi_a,\partial_\mu\phi_a)+\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a+\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a)$$
$$\implies \delta\mathcal{L}=\Big[\frac{\partial\mathcal{L}}{\partial\phi_a}-\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\Big)\Big]\delta\phi_a+\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\Big)=\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\Big)$$

where we've used the product rule and the Euler-Lagrange equations. Now we can similarly take the transformation to act directly on the Lagrangian: $\mathcal{L}\to\mathcal{L}+\partial_\mu F^\mu$ $\implies \delta\mathcal{L}=\partial_\mu F^\mu$. Equating the two forms for $\delta\mathcal{L}$ provides an expression for our current:

$$\partial_\mu\Big[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a-F^\mu(\phi_a)\Big]=0 \implies j^\mu=\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a-F^\mu(\phi_a)$$

For example, if we have a translation $\delta\phi_a=\epsilon^\nu\partial_\nu\phi_a$ then $\delta\mathcal{L}=\epsilon^\nu\partial_\nu\mathcal{L}$ which implies $F^\mu=\epsilon^\nu\delta^\mu{}_\nu\mathcal{L}$. Thus we obtain the current:

$$(j^\mu)_\nu=\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\nu\phi_a-\delta^\mu{}_\nu\mathcal{L}=:T^\mu{}_\nu \tag{3.29}$$

$T^{\mu\nu}$ is the **Energy-Momentum tensor**. E.g. For the Dirac Lagrangian, composed from our invariants in the following way: $\mathcal{L} = i\bar{\Psi}\gamma^\mu\partial_\mu\Psi - m\bar{\Psi}\Psi$, we have $T^{\mu\nu} = i\bar{\Psi}\gamma^\mu\partial^\nu\Psi$. Let's look at thus example a little further, recall that the Dirac spinor transformed under $SO(1,3)$ as $\Psi'(x') = \mathcal{M}\Psi$, $\mathcal{M} = \mathcal{M}_L \oplus \mathcal{M}_R$. It's not too hard to see, using the properties of exponentials of diagonal matrices, that the generators of $\mathcal{M}$ are in fact the $S^{\mu\nu}$ in (3.25) and thus we obtain:

$$\Psi'(x+\delta x) \approx (\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}S^{\mu\nu})\Psi(x) \implies \delta_0\Psi(x) = \frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}\Psi(x) - \omega^\mu{}_\nu x^\nu\partial_\mu\Psi(x)$$

Now in the second term in $\delta_0\Psi(x)$ the action is that of the Lorentz group on scalar field thus we employ (3.18) to obtain:

$$\begin{aligned}\delta_0\Psi(x) &= \frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}\Psi(x) - \frac{i}{2}\omega_{\rho\sigma}(\eta^{\rho\mu}\eta^\sigma{}_\nu - \eta^\rho{}_\nu\eta^{\sigma\mu})x^\nu\partial_\mu\Psi(x) \\ &= \omega_{\mu\nu}\frac{i}{2}[S^{\mu\nu}\Psi(x) - (x^\nu\partial^\mu - x^\mu\partial^\nu)\Psi(x)]\end{aligned}$$

The conserved current we obtain using (3.29), by substituting $\mathcal{L}$ and $\delta_0\Psi$ in is:

$$(J^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} - i\bar{\Psi}\gamma^\mu S^{\rho\sigma}\Psi \tag{3.30}$$

The conserved current here is that of total angular momentum: both external and intrinsic *i.e.* spin. After quantization, $(J^\mu)^{\rho\sigma}$ becomes an operator responsible for providing single particle states (we'll see these next) with spin - the Dirac spinor gives rise to spin-$\frac{1}{2}$ particles after quantization!.

3.5 Representations on States

In this section we consider possibly the most remarkable consequence of Poincaré symmetry in all of physics: Wigner's classification of one-particle states. Certainly it makes intuitive sense that the fundamental particles in nature should be Lorentz invariant otherwise their properties, specified by their quantum numbers, would change depending on the Lorentz reference frame. To develop Wigners classification, we first need to introduce the method of **induced representations**.

3.5.1 Induced Representations

We draw directly on material from [14]. Let G be a Lie group, $H \subset G$ a Lie subgroup, and $\rho : H \to \mathrm{End}(\mathcal{H}(\rho))$ a unitary representation of $H \subset G$. Consider the set of functions $f \in \mathcal{F} : G \to \mathcal{H}(\rho)$ (*i.e.* from the group G to the representation (Hilbert) space of ρ) that satisfy the following:

$$f(h \cdot g) = \rho(h)f(g)$$

Now for each such f we have that $\forall h \in H$:

$$\langle f(hg), f(hg)\rangle = \langle \rho(h)f(g), \rho(h)f(g)\rangle = \rho(h)^\dagger\rho(h)\langle f(g), f(g)\rangle = \langle f(g), f(g)\rangle \tag{3.31}$$

i.e. since ρ is a unitary representation of $H \subset G$ it preserves the complex inner product on $\mathcal{H}(\rho)$. This means $f(g)$ depends only on the *right coset* $H \cdot g$ w.r.t the inner product. We can then define the following vector space:

$$\mathcal{H}(U^\rho) := \{f \in \mathcal{F} \mid \|f\|^2 := |\langle f(g), f(g)\rangle|^2 := \sum_{G/H} \langle f(g), f(g)\rangle < \infty\}$$

It can be shown that this is a Hilbert space. If we define $U^\rho_{g_1}(f(g_2)) := f(g_2 \cdot g_1)$, $g_1, g_2 \in G$, then $g \mapsto U^\rho_g$ is a unitary representation of G called the representation **induced by H**. We have:

$$U^\rho_{g_1 \cdot g_2}(f)(g_3) = f(g_3 \cdot (g_1 \cdot g_2)) = f((g_3 \cdot g_1) \cdot g_2) = U^\rho_{g_2}(f)(g_3 \cdot g_1) = (U^\rho_{g_2} \circ U^\rho_{g_1})(f)(g_3)$$

So group structure is preserved. (Note: the group action is from the right but the representations act from the left so the order is correct). We will now consider the case $G \equiv ISO(1,3)$.

3.5.2 Wigner's Classification

The task now is to find the unitary irreps of $ISO(1,3)$ on the Hilbert space of eigenstates $\mathcal{H} := \text{span}\{\psi_{p,\sigma}\}_{p,\sigma}$. Note that we only consider unitary irreps since symmetries of a quantum system must manifest as unitary operators otherwise they would not preserve the inner product as in (3.31), and thus would not be symmetries. The states are labelled such that $P^\mu \psi_{p,\sigma} = p^\mu \psi_{p,\sigma}$ and σ is some other quantum number. The p^μ are called the **spectral values** of P^μ.

This procedure is adapted from a similar discussion in chapter 17 of [6]. To begin, recall that $ISO(1,3) = \mathcal{R}^{1,3} \rtimes SO(1,3)$ where $\mathcal{R}^{1,3}$ is an abelian normal subgroup. Thus the unitary irreps of $ISO(1,3)$ on $\mathcal{H}$ will be the product of the unitary irreps of $\mathcal{R}^{1,3}$ on $\mathcal{H}$ and those of $SO(1,3)$ on $\mathcal{H}$. $\mathcal{R}^{1,3}$ is the direct product of 4 1-parameter abelian subgroups and thus has the 1-parameter unitary irrep:

$$U^p_{\mathcal{R}}[(\mathbb{1}, a)] = \exp(ia_\mu P^\mu) \equiv \exp(ia_\mu p^\mu) \tag{3.32}$$

since $\exp(ia^\mu P_\mu)\psi_{p,\sigma} = \sum_n \frac{(ia^\mu P_\mu)^n}{n!}\psi_{p,\sigma} = \sum_n \frac{(ia^\mu p_\mu)^n}{n!}\psi_{p,\sigma} = \exp(ia^\mu p_\mu)\psi_{p,\sigma}$. This is the easy part. We now need to find the Lie subgroups H from which we induce our irreps. To do this, the first step is to define the **little group** of a 4-vector k^μ:

$$L(k) := \{(\Lambda, 0) \in SO(1,3) | \Lambda^\mu{}_\nu k^\nu = k^\mu\} \subset SO(1,3) \tag{3.33}$$

i.e. it is the **stabiliser/isotropy** subgroup of k^μ. Now using (3.32) we have that

$$\begin{aligned} U^k_{\mathcal{R}}[(\Lambda,0)(\mathbb{1},a)(\Lambda,0)^{-1}] &= U^k_{\mathcal{R}}[(\mathbb{1},\Lambda a)] = \exp(i(\Lambda a)_\mu k^\mu) \\ &= \exp(ia_\mu(\Lambda^{-1}k)^\mu) = U^{(\Lambda^{-1}k)}_{\mathcal{R}}[(\mathbb{1},a)] \end{aligned} \tag{3.34}$$

where we've used that $\Lambda^\mu{}_\sigma \eta_{\mu\nu} \Lambda^\nu{}_\rho = \eta_{\sigma\rho}$ implies $\eta_{\mu\nu}\Lambda^\nu{}_\rho = \eta_{\sigma\rho}(\Lambda^{-1})^\sigma{}_\mu$ to show that $(\Lambda a)_\mu k^\mu = \eta_{\mu\nu}\Lambda^\nu{}_\rho a^\rho k^\mu = \eta_{\sigma\rho}(\Lambda^{-1})^\sigma{}_\mu a^\rho k^\mu = a_\sigma(\Lambda^{-1}k)^\sigma$, which we arbitrarily relabel.

This tells us that $U^k_{\mathcal{R}}[(\mathbb{1}, \Lambda a)] = U^k_{\mathcal{R}}[(\mathbb{1}, a)]$ iff $\Lambda \in L(k)$. *i.e.* We have 1 unitary irrep of $\mathcal{R}^{1,3}$ per little group since translations produced by the action of any one little group have equivalent irreps. Now we want to identify the analogues of of the subgroup H in 3.5.1 for $ISO(1,3)$.

For step two, recall explicitly that p^μ is in the orbit of k^μ if $\exists$ $(\Lambda, 0) \in SO(1,3)$ such that $p = \Lambda k$. So clearly any member of the *left coset* $(\Lambda, 0)L(k)$ generates an element p^μ in the orbit of k^μ. Thus there is a one-to-one correspondence between the orbit of k and the set of left cosets of $SO(1,3)$ w.r.t it's little group $L(k)$. We denote the coset representative corresponding to p in the orbit of k by $(\Lambda(p,k), 0)$ *i.e.* $(\Lambda(p,k), 0)k = p$ $\implies$ $(\Lambda(p,k)^{-1}, 0)p = k$. Now if $(\Lambda_1, 0), (\Lambda_2, 0)$ belong to the same left coset then $\exists$ $(\Lambda_0, 0) \in L(k)$ such that $(\Lambda_1, 0) = (\Lambda_2, 0)(\Lambda_0, 0)$ which implies $(\Lambda_2^{-1}, 0) = (\Lambda_0, 0)(\Lambda_1^{-1}, 0)$ *i.e.* $(\Lambda_1^{-1}, 0), (\Lambda_2^{-1}, 0)$ are in the same *right coset*. Thus we can take the $(\Lambda(p,k)^{-1}, 0)$ to be our coset representatives for the decomposition of $SO(1,3)$ into the set of right cosets $L(k)\backslash SO(1,3) := \{L(k)(\Lambda(p_1,k)^{-1}, 0),\ L(k)(\Lambda(p_2,k)^{-1}, 0), ...\}$.

Proposition 3.5. *$L(p) \cong L(k)$ $\forall$ p s.t. $\exists$ $\Lambda \in SO(1,3)$ with $p = \Lambda k$, i.e. all members of the orbit of k have isomorphic little groups.*

Proof. Let $\phi_p : ISO(1,3) \to ISO(1,3)$ be the map $\phi_p(X) = (\Lambda(p,k), 0)X(\Lambda(p,k)^{-1}, 0)$. The proof is entirely analogous to (2.25): by the same method as in (2.25) we can show that the right action is a diffeomorphism and thus we know that the conjugate action, as a composition of the left and right action, must itself be a diffeomorphism. Now let $(\Lambda, 0) \in L(k)$ $(\Lambda k = k)$. Then $\phi_p((\Lambda, 0)) \in L(p)$:

$$\phi_p((\Lambda,0))p = (\Lambda(p,k)\Lambda\Lambda(p,k)^{-1}, 0)p = (\Lambda(p,k)\Lambda, 0)k = (\Lambda(p,k), 0)k = p$$

So ϕ_p maps $L(k)$ onto $L(p)$. Since the little groups are Lie subgroups (2.25) applies and thus provided they're in the same orbit they must be diffeomorphic and thus isomorphic. $\square$

This means we need only find the discrete set of orbits that partition the translations, choose a standard generating vector k and then determine it's little group. Irreps of other vectors in the orbit will be equivalent:

$$U^p_{\mathcal{R}}[\phi_p((\mathbb{1}, a))] \overset{(3.32)}{=} U^k_{\mathcal{R}}[(\mathbb{1}, a)] \tag{3.35}$$

Thus our subgroup(s) that will induce our representations are $H_k := \mathcal{R}^{1,3} \rtimes L(k)$ for each standard vector k. Note that since the right cosets $L(k)\backslash SO(1,3)$ partitioned $SO(1,3)$, the right cosets $H_k\backslash ISO(1,3)$ partition $ISO(1,3)$. We define our irreps of H_k: $(\rho_k, \mathcal{H}(\rho_k))$, $\rho_k : H_k \to \mathrm{End}(\mathcal{H}(\rho_k))$ given by:

$$\rho_k[(\mathbb{1}, a)(\Lambda, 0)] := U^k_{\mathcal{R}} \cdot \Gamma[(\Lambda, 0)] \tag{3.36}$$

where Γ is an irrep of $L(k)$. If $\{\phi_i\}_i$ is a basis for $\mathcal{H}(\rho_k)$, then the ρ_k act as:

$$\rho_k[(\Lambda, a)]\phi_n = \sum_{m=1}^{d} U^k_{\mathcal{R}}[(\mathbb{1}, a)]\Gamma[(\Lambda, 0)]_{mn}\phi_m, \quad d := \dim(\mathcal{H}(\rho_k)) \tag{3.37}$$

To complete the construction we have one final step: to define the functions $\mathcal{F}$ and write down the full irrep of $ISO(1,3)$. Let $\mathcal{F}$ be the following:

$$\mathcal{F}_k := \{\psi : ISO(1,3) \to \mathcal{H}(\rho_k) \mid \psi[(\Lambda,0)(\Lambda',a)] = \rho_k[(\Lambda,0)]\psi[(\Lambda',a)] \ \forall\ (\Lambda,0) \in L(k)]\}$$

Clearly, as mentioned in 3.5.1, these only depend on the right cosets due to their equivariant property, thus we can take their argument to simply be the vector p defining the coset representative $(\Lambda(p,k),0)^{-1}$. Now we define the full irrep, as we did in 3.5.1 using operators acting on $\mathcal{F}_k$, to be:

$$U^{\rho_k}[(\Lambda,a)]\psi(p) := \psi[(\Lambda(p,k),0)^{-1}(\Lambda,a)] \tag{3.38}$$

But since the $H_k(\Lambda(p,k),0)^{-1}$ partition $ISO(1,3)$ there must be a p' such that:

$$\begin{aligned}
&(\Lambda(p,k),0)^{-1}(\Lambda,a) \in H_k(\Lambda(p',k),0)^{-1} \implies \Lambda(p,k)^{-1}\Lambda p' = k \implies \Lambda p' = p\\
&\therefore U^{\rho_k}[(\Lambda,a)]\psi(p) := \psi[(\Lambda(p,k),0)^{-1}(\Lambda,a)(\Lambda(p',k),0)(\Lambda(p',k),0)^{-1}]\\
&\qquad = \rho_k[(\Lambda(p,k),0)^{-1}(\Lambda,a)(\Lambda(p',k),0)]\psi(p')
\end{aligned}$$

We can now define a basis $\psi_{p,n}(p') = \begin{cases}\phi_n & p = p'\\ 0 & p \neq p'\end{cases}$ such that

$$U^{\rho_k}[(\Lambda,a)]\psi_{p,n}(p') = \rho_k[(\Lambda(p',k),0)^{-1}(\Lambda,a)(\Lambda(\Lambda^{-1}p,k),0)]\psi_{p,n}(\Lambda^{-1}p')$$

which is only non-trivial if $p = \Lambda^{-1}p'$ *i.e.* $(\Lambda,0) \in L(p)$. Therefore, by substitution and a little moving things around, we have $\forall\ (\Lambda,a) \in ISO(1,3)$:

$$\begin{aligned}
&U^{\rho_k}[(\Lambda,a)]\psi_{p,n}(p') = \rho_k[(\Lambda(\Lambda p,k),0)^{-1}(\Lambda,a)(\Lambda(p,k),0)]\psi_{\Lambda p,n}(p')\\
&\overset{(3.37)}{\implies} U^{\rho_k}[(\Lambda,a)]\psi_{p,n} = U_{\mathcal{R}}^{\Lambda p}[(\mathbb{1},a)]\sum_{m=1}^{d}\Gamma[(\Lambda(\Lambda p,k)^{-1}\Lambda\Lambda(p,k),0)]_{mn}\psi_{\Lambda p,m}
\end{aligned} \tag{3.39}$$

where we've also used (3.34) to show $U_{\mathcal{R}}^{k}[(\mathbb{1},\Lambda(\Lambda p,k)^{-1}a)] = U_{\mathcal{R}}^{\Lambda(\Lambda p,k)k}[(\mathbb{1},a)] = U_{\mathcal{R}}^{\Lambda p}[(\mathbb{1},a)]$.

Casimir Operators

The remainder of 3.5.2 loosely follows 10.2 of [2], with elements of [17], [6]. To classify our little groups, motivated by our examination of the connected components of $SO(1,3)$, we use the values of $p^2 := \eta_{\mu\nu}p^\mu p^\nu$ and the sign of p^0. These are summarised in Table 3.1 along with the simplest choice of the 'standard' vector k. P^2 is clearly a Casimir operator since it takes the same eigenvalue, p^2, on all the momentum vectors in the orbit of k. In fact, the equation

$$p^2 \equiv (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = M^2 \tag{3.40}$$

describes various forms of hyperboloid depending on the value and sign of M^2, summarised in (3.1). If $p^2 = -M^2$, we call the corresponding hyperboloid the **mass shell**.

Orbit	k^μ	$L(k)$
(a) $p^2 = -M^2 < 0, p^0 > 0$	$(M, 0, 0, 0)$	$SO(3)$
(b) $p^2 = -M^2 < 0, p^0 < 0$	$(-M, 0, 0, 0)$	$SO(3)$
(c) $p^2 = 0, p^0 > 0$	$(\kappa, 0, 0, \kappa)$	$ISO(2)$
(d) $p^2 = 0, p^0 < 0$	$(-\kappa, 0, 0, \kappa)$	$ISO(2)$
(e) $p^2 = N^2 > 0$	$(0, N, 0, 0)$	$SO(1, 2)$
(f) $p^2 = 0, p^\mu = 0$	$(0, 0, 0, 0)$	$SO(1, 3)$

Table 3.1: The Wigner Little Groups and their standard k vectors [17]

To be *on-shell* is to satisfy the relativistic energy equation (3.40). Representations on states are called *on-shell representations*, whereas representations on fields are called *off-shell* since they're not automatically constrained in this manner. They become on-shell when constrained by an equation of motion.

To fully classify our representations we need a second Casimir operator whose values correspond to the irreps of the little groups themselves. To see what this is consider the infinitesimal form of $\Lambda^\mu{}_\nu k^\nu = k^\mu$:

$$k^\mu = \Lambda^\mu{}_\nu k^\nu \approx (\delta^\mu{}_\nu + \omega^\mu{}_\nu)k^\nu = k^\mu + \omega^\mu{}_\nu k^\nu \implies \omega^\mu{}_\nu k^\nu = 0 \tag{3.41}$$

To satisfy this condition we can choose $\omega_{\mu\nu} = -\epsilon_{\mu\nu\rho\sigma}s^\rho k^\sigma$ with s^ρ some spacelike 4-vector. This gives $\omega_{\mu\nu}k^\nu = -\epsilon_{\mu\nu\rho\sigma}s^\rho k^\sigma k^\nu$ which is clearly a totally antisymmetric part $-\epsilon_{\mu\nu\rho\sigma}s^\rho$ contracted with a totally symmetric part $k^\sigma k^\nu$: $-\epsilon_{\mu\nu\rho\sigma} = \epsilon_{\mu\sigma\rho\nu} \implies -\epsilon_{\mu\nu\rho\sigma}s^\rho k^\sigma k^\nu = \epsilon_{\mu\sigma\rho\nu}s^\rho k^\sigma k^\nu = \epsilon_{\mu\sigma\rho\nu}s^\rho k^\nu k^\sigma = \epsilon_{\mu\nu\rho\sigma}s^\rho k^\sigma k^\nu \implies \epsilon_{\mu\nu\rho\sigma}s^\rho k^\sigma k^\nu = 0$. Now using (3.6) we can define the **Pauli-Lubanski four vector**:

$$W^\mu := -\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}P_\nu J_{\rho\sigma} \tag{3.42}$$

These are the generators of the little groups. W^2 is our second Casimir operator. It takes eigenvalues $-p^2 j(j+1)$ where j is the spin of the irrep of the little group. The $j(j+1)$ part may look a little out of nowhere, but if we set $p = k$ in (3.39) then $U^{\rho_k}[(\Lambda, a)]\psi_{k,n}$ is only non-zero if $\Lambda k = k$ *i.e.* $\Lambda \in L(k)$. Thus, setting $a = 0$, we obtain:

$$U^{\rho_k}[(\Lambda, 0)]\psi_{k,n} = \sum_{m=1}^{d} \Gamma[(\Lambda, 0)]_{mn}\psi_{k,m} \implies J_{\mu\nu}\psi_{k,n} = \sum_{m=-j}^{j} \Gamma^j[J_{\mu\nu}]_{mn}\psi_{k,m}$$

where Γ^j is the $(2j+1)$-dimensional irrep of the $J_{\mu\nu}$ that generate the little group. For example, in case (a) the little group is $SO(3)$, whose irreps are simply those of it's double cover $SU(2)$ (in the same way $SL(2, \mathbb{C})$ is the double cover of $SO(3)$). Thus they're given by $j = 0, \frac{1}{2}, 1, ...$ etc. Since the orbits are infinite, all unitary representations of $ISO(1, 3)$ are infinite dimensional apart from $p^\mu = 0$, whose orbit contains only $p^\mu = 0$.

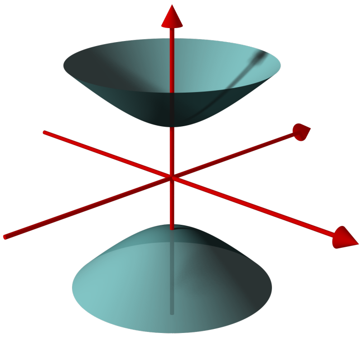

Figure 3.1: Two Sheet Hyperboloid: The $p^2 = -M^2$ Mass Shells $p^0 > 0$ and $p^0 < 0$

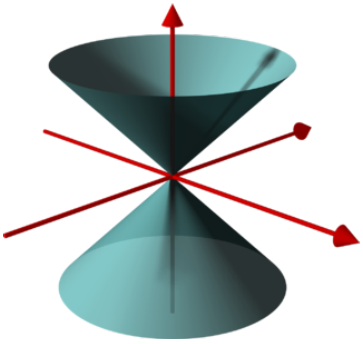

Figure 3.2: Double Cone: $p^2 = 0$ Mass Shells $p^0 > 0$, $p^0 < 0$ and the origin $p^\mu = 0$

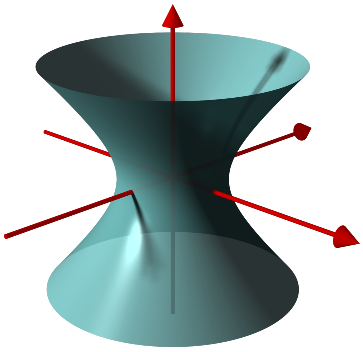

Figure 3.3: One Sheet Hyperboloid: The $p^2 = N^2$ Mass Shell

Example

To show how all this works, let's consider case (c) from 3.1. To find the little group it's actually easier to work in $SL(2,\mathbb{C})$ using the map $\rho : \mathbb{R}^{1,3} \to M_2(\mathbb{C})$, $\rho(x^\mu) = x^\mu \sigma_\mu$.

$$K := k^\mu \sigma_\mu = \kappa\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) + \kappa\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2\kappa & 0 \\ 0 & 0 \end{smallmatrix}\right)$$

Now let $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{C})$ $(ad - bc = 1)$ be such that $K = AKA^\dagger$. Then:

$$\left(\begin{smallmatrix} 2\kappa & 0 \\ 0 & 0 \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} 2\kappa & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a^* & c^* \\ b^* & d^* \end{smallmatrix}\right) = \begin{pmatrix} 2|a|^2\kappa & 2ac^*\kappa \\ 2ca^*\kappa & 2|c|^2\kappa \end{pmatrix} \implies c = 0, |a|^2 = 1$$

Now $\det(A) = 1 \implies ad = 1$ thus $d = a^{-1}$. So our little group is:

$$\tilde{L}(k) = \{A \in SL(2,\mathbb{C}) | A = \begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix}, |a|^2 = 1\} \tag{3.43}$$

This is indeed a Lie group, in fact this is exactly what we showed in Example 2.10 way back at the start! But what group actually is this? To see this, let $a = e^{-\frac{i}{2}\theta}$, $b = (\alpha + i\beta)e^{\frac{i}{2}\theta}$, $\alpha, \beta \in \mathbb{R}$ w.l.o.g and let $A(\alpha_1, \beta_1; \theta_1), A(\alpha_2, \beta_2; \theta_2) \in \tilde{L}(k)$. Then:

$$A(\alpha_1,\beta_1;\theta_1)A(\alpha_2,\beta_2;\theta_2) = \begin{pmatrix} e^{-\frac{i}{2}\theta_1} & (\alpha_1+i\beta_1)e^{\frac{i}{2}\theta_1} \\ 0 & e^{\frac{i}{2}\theta_1} \end{pmatrix}\begin{pmatrix} e^{-\frac{i}{2}\theta_2} & (\alpha_2+i\beta_2)e^{\frac{i}{2}\theta_2} \\ 0 & e^{\frac{i}{2}\theta_2} \end{pmatrix}$$
$$= \begin{pmatrix} e^{-\frac{i}{2}(\theta_1+\theta_2)} & [\alpha_2\cos\theta_1+\beta_2\sin\theta_1+\alpha_1+i(\beta_2\cos\theta_1-\alpha_2\sin\theta_1+\beta_1)]e^{\frac{i}{2}(\theta_1+\theta_2)} \\ 0 & e^{\frac{i}{2}(\theta_1+\theta_2)} \end{pmatrix}$$

i.e. $A(\alpha_1,\beta_1;\theta_1)A(\alpha_2,\beta_2;\theta_2) = A(\alpha_2\cos\theta_1 + \beta_2\sin\theta_1, \beta_2\cos\theta_1 - \alpha_2\sin\theta_1 + \beta_1; \theta_1 + \theta_2)$ This is a rotation by θ_1 and translation by (α_1, β_1) *i.e.* it's the standard Euclidean group $ISO(2) = \mathcal{R}^2 \rtimes SO(2)$! It has Lie algebra:

$$[L_3, A] = iB, \ \ [L_3, B] = -iA, \ \ [A, B] = 0 \tag{3.44}$$

where $A := L_2 + K_1$, $B := -L_1 + K_2$ act on phase space states as

$$A\psi^{\theta}_{k,\alpha,\beta} = (\alpha\cos\theta - \beta\sin\theta)\psi^{\theta}_{k,\alpha,\beta}, \quad B\psi^{\theta}_{k,\alpha,\beta} = (\alpha\sin\theta + \beta\cos\theta)\psi^{\theta}_{k,\alpha,\beta}$$

Like for $ISO(1,3)$, the unitary representations of this group are all infinite dimensional (since $\alpha, \beta \in \mathbb{R}$) apart from when the translation vector (α, β) is zero and the group reduces to $SO(2)$. Massless particles in nature exhibit no such infinite continuous degree of freedom that would correspond to such representations, so we only consider the irreps of $SO(2)$. The covering group of $SO(2)$ is $U(1) = \{e^{i\theta}, \theta \in [0, 2\pi)\}$, the irreps of which are all 1-dimensional given by $e^{\frac{i}{2}j\theta}$, $j = 0, \pm\frac{1}{2}, \pm 1, ...$ This quantity j isn't actualy spin as such, but rather **helicity**; the normalized spin in the direction of the 3-momentum. Our representation is therefore:

$$U^{\rho(\kappa,0,0,\kappa)}[(\Lambda, x)] = \exp(ix_\mu k^\mu)\sum_{j=0}^{n}\exp(\frac{i}{2}j\theta(\Lambda)) \tag{3.45}$$

3.5.3 The Spin-Statistics Theorem

To close out this chapter and motivate the next couple of chapters we will discuss the **Spin-Statistics Theorem**. A few years before Wigner released his classification of one-particle states, various problems in statistical mechanics were troubling physicists such as Satyendra Nath Bose, Albert Einstein, Enrico Fermi and Paul Dirac. Bose was working on the statistical properties of photons due to their indistinguishability leading him an Einstein to develop **Bose-Einstein Statistics**. These can be summarised by considering the wavefunction of a system containing two identical particles 1 and 2, $\Psi(1,2)$. Bose-Einstein statistics correspond to the relation $\Psi(1,2) = \Psi(2,1)$, and particles possessing this symmetry when placed in a system with an identical particle are called **Bosons**. At the same time, concerned with the statistical mechanics of electrons in a metal Fermi and Dirac published **Fermi-Dirac Statistics** which can be summarised by $\Psi(1,2) = -\Psi(2,1)$ in the same 2-identical-particle set up as for bosons. Particles exhibiting this behaviour are called **Fermions**. The remarkable conclusion of the spin-statistics theorem directs relates the 'spin' of Wigners representations to the physical statistics of a particle; *all integer spin particles are bosons and all half integer spin particles are fermions.*

Beyond the elegant and profound nature of this result, why is this relevant to us? Well, all the representations we considered so far are either bosonic or fermionic according to this theorem! By the very maths itself, our representations cannot be both spin $\frac{\mathbb{Z}}{2}$ and spin $\mathbb{Z}$ since the ladder operators of 2.3.2 raise/lower indices by 1. Thus Wigners classification together with the spin-statistics theorem forbids transformations between bosonic and fermionic states; there is no symmetry here! There is no irrep containing both! Such a transformation, should it exist, would extend our conventional notions of symmetry; it would be a... super-symmetry.

Chapter 4

The Supersymmetry Algebra

The culmination of our discussion in the previous chapter was the remarkable Wigner classification of fundamental particles according to Poincaré symmetry. The spin-statistics theorem then kept these particles firmly in their place; separating the world into the integer spin bosons and half integer spin fermions. The consequence of the mathematics was clear: no boson-fermion transformations since they cannot exist in the same irreps! This final statement is very rigid, and indeed there seems no particular need to broaden it. However, during the explosion of particle theory in the mid-20th century, theoretical physicists found themselves investigating increasingly exotic symmetries to explain the increasing array of particle phenomena, and a desire arose to build an extended symmetry group, fusing both Poincaré symmetry and some additional internal symmetry.

4.1 Extending the Poincaré Algebra

The exciting proposal was a theory of $SU(6)$, which suggested the possibility of a relativistic symmetry group which was not just a direct product containing $ISO(1,3)$. Unfortunately, all attempts to realise such a theory failed miserably and gradually a series of **no-go** theorems were produced, placing restrictions on the symmetry group a successful Quantum Field Theory (QFT) could respect. The last and most powerful of these was the **Coleman-Mandula Theorem**.

4.1.1 The Coleman-Mandula Theorem

We begin with a summary of the introduction to [5]. In QFT, all the information describing a particle interaction is contained in the **S-Matrix** (S for scattering) therefore symmetry in a QFT manifests itself as symmetries of the S-matrix. The statement of the Coleman-Mandula Theorem is as follows: Let G be a connected symmetry group of the S-matrix and let the following assumptions hold:

1. **Lorentz Invariance**: G contains a subgroup locally isomorphic to the Lorentz group.

2. **Particle Finiteness**: All particle types correspond to a positive energy representation of $ISO(1,3)$, For any finite M there are only a finite no. of particles with $p^2 = -m^2$, $m < M$.

3. **Weak Elastic Analyticity**: Elastic scattering amplitudes are analytic functions of of the centre of mass energy and the conserved momentum transfer.

4. **Non-trivial Scattering**: Scattering occurs non-trivially.

5. **The "Ugly" Assumption**: There is a neighbourhood $U \subset G$ about the identity $e \in G$ such that $\forall\, g \in G$ there exists a one parameter subgroup on which g lies.

Then G is locally isomorphic to the direct product of an internal symmetry group and $ISO(1,3)$.

The assumptions in this theorem are very important. (1) is necessary since in QM there are many spin-independent Galilean theories where the theorem doesn't hold. (2) again removes classes of theories that break the theorem but had no physical relevance at the time. (3) is in there since it appears to be a property of reality *i.e.* theories in $3+1$ dimensions will have this. Symmetries are known in $1+1$ dimensions, where only forwards and backwards scattering is allowed, which is obviously not analytic and breaks the theorem. (5) is in there since, as physicists, we like to work with infinitesimal generators and this allows us to do so via our discussion in section 2.2.5. Coleman and Mandula refer to this assumption as ugly multiple times in the paper - presumably since it removes some of the shine from the generality of the theorem.

This theorem heavily restricts us in our quest to extend the Poincaré symmetry: it's formal conclusion is that the Lie algebra of the most general QFT symmetry group G contains only the generators $P_\mu, J_{\mu\nu}$ and B_l where the B_l generate some internal symmetry. These correspond to transformations with commutative parameters, whose generators obey commutator relations leading to strictly boson-to-boson and fermion-to-fermion symmetry (since our irreps, the multiplets, contain only either bosons or fermions). Suppose that we tried to transform a scalar field into a spinor field. This is clearly a boson-fermion transformation. What would this look like? Based on our discussion in chapter 3 we know this transformation would have to involve a Grassmann parameter otherwise the spinor field obtained would have zero scalar product. But such parameters anticommute, what would this look like algebraically? Before we push this idea any further, it's worth taking a closer look at the behaviour of Grassmann numbers.

4.1.2 Grassmann Numbers

We follow the material from Appendix A in [1], with some elements from chapter 20 of [7]. As we've mentioned, Grassmann numbers (G-numbers) are anti-commuting. If θ_i, θ_j are two G-numbers, then:

$$\{\theta_i, \theta_j\} = 0 \tag{4.1}$$

Notice immediately that in the case $i = j$ this implies $\theta_i^2 = 0$. G-numbers and arbitrary products thereof generate a **Grassmann algebra**. Suppose we have n G-numbers $\{\theta_1, ..., \theta_n\}$ and an identity 1. For a set of indices $\{i_1, i_2, ..., i_k\} \subset \{1, 2, ..., n\}$ with $k \leq n$ the arbitrary product of G-numbers is defined:

$$\Theta_k := \theta_{i_1}\theta_{i_2}\cdots\theta_{i_k} = \epsilon_{i_1 i_2 \dots i_k}\theta_1\theta_2\cdots\theta_k, \quad 1\cdot\theta_{i_1}\theta_{i_2}\cdots\theta_{i_k} = \theta_{i_1}\theta_{i_2}\cdots\theta_{i_k}\cdot 1 \tag{4.2}$$

The **degree** of a product such as the one above is $\deg(\Theta_k) = (-1)^k$. An arbitrary element of the algebra then takes the form:

$$\mathcal{G} = \sum_{k=0}^{n}\sum_{i_1<\dots<i_k} g_{i_1 i_2 \dots i_k}\theta_{i_1}\theta_{i_2}\cdots\theta_{i_k}, \quad g_{i_1 i_2 \dots i_k} \in \mathbb{R} \text{ or } \mathbb{C}$$

The real(complex) Grassmann algebra generated by these n G-numbers is denoted $B\mathbb{R}_n(B\mathbb{C}_n)$ and has real(complex) dimension 2^n.

Example 4.1. The algebra $B\mathbb{C}_3$ is spanned by $\{1, \theta_1, \theta_2, \theta_3, \theta_1\theta_2, \theta_1\theta_3, \theta_2\theta_3, \theta_1\theta_2\theta_3\}$ with complex coefficients. This does indeed have dimension $2^3 = 8$.

Functions of G-numbers may be defined using Taylor expansion, given that for any given G-number $\theta^2 = 0$ so the series truncates at finite order. For example, the exponential functions is $e^\theta = 1 + \theta$. Or a field of one Grassmann and one normal parameter becomes $\phi(x, \theta) = \phi_1(x) + \theta\phi_2(x)$. We can also differentiate w.r.t a Grassmann variable:

$$\frac{\partial}{\partial\theta_j}\theta_i = \delta_{ij} \implies \frac{\partial}{\partial\theta_i}(\theta_j\theta_k) = \frac{\partial}{\partial\theta_i}(\theta_j)\theta_k - \frac{\partial}{\partial\theta_i}(\theta_k)\theta_j = \delta_{ij}\theta_k - \delta_{ik}\theta_j \tag{4.3}$$

Using the above formula to compute $\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j}(\theta_k\theta_l)$ and $\frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_i}(\theta_k\theta_l)$ it's easy to see that:

$$\{\frac{\partial}{\partial\theta_i}, \frac{\partial}{\partial\theta_j}\} = 0 \tag{4.4}$$

which in turn gives $\frac{\partial^2}{\partial\theta_i^2} = 0$. Now integration is a little different. In the context of G-numbers forming an **exterior algebra** this makes sense, and this is easily looked up. Essentially integration coincides with differentiation:

$$\int d\theta_1\, d\theta_2 ...\, d\theta_n f(\theta_1, \theta_2, ..., \theta_n) = \frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2}\cdots\frac{\partial}{\partial\theta_n} f(\theta_1, \theta_2, ..., \theta_n) \tag{4.5}$$

In particular: $\int d\theta = 0$ and $\int d\theta\theta = 1$. Defining the Dirac delta function as normal we see it takes a very simple form in the context of G-numbers:

$$\int d\theta\delta(\theta - \tilde{\theta})f(\theta) = f(\tilde{\theta}) \implies \delta(\theta - \tilde{\theta}) = (\theta - \tilde{\theta}) \tag{4.6}$$

We can now use this newfound knowledge to consider our proposal: what if a transformation also contained a Grassmann parameter? Think back to section 2.1.4 and consider how to commutators arose. First of all in (2.4) we had that $X_{jk} = \frac{\partial^2}{\partial t^j \partial t^k} g(t)$ and thus $X_{jk} = X_{kj}$ however if t is Grassmann valued we must use (4.4) and thus $X_{jk} = -X_{kj}$.

Now assuming both t and s are Grassmann numbers we can follow the same procedure through until we reach our equivalent of (2.6) which is:

$$\begin{aligned} X_{jk} = -X_{kj} \implies X_j X_k - f^i_{jk} X_i = f^i_{kj} X_i - X_k X_j \\ \{X_j, X_k\} = (f^i_{kj} + f^i_{jk}) X_i \end{aligned} \tag{4.7}$$

where $\{X_j, X_k\} := X_j X_k + X_k X_j$ is the **anti-commutator**. Note that Grassmann parameters behave normally with normal numbers *i.e* they commute with normal numbers. This means a 'Lie group' with both Grassmann and normal parameters will have a 'Lie algebra' containing commutators AND anti-commutators! What are the consequences of such a strange algebra? Is this well defined? Is this even a Lie algebra? This brings us to the next section.

4.1.3 Graded Lie Algebras

We follow chapter 20 of [7]. The notion of **grading** is one already familiar to pretty much everyone; mathematician or not. Consider the natural numbers $\mathbb{N}$. They're all either *even* or *odd* and under multiplication they obey the structure *even·even=even*, *even·odd=even*, *odd·odd=odd*. This is a **grade-1** or $\mathbb{Z}/2$-**graded** structure. This notion can be extended to **grade-n** or $\mathbb{Z}/(n+1)$-**graded** structure but we need not discuss this here.

Definition 4.2. A **grade-1 vector space** is an $(m+n)$-dimensional vector space V, $m, n \in \mathbb{Z}_{>0}$. Let $\{v_1, ..., v_m, v_{m+1}, ..., v_{m+n}\}$ be a basis for V. This is graded if we decree that elements $\{v_1, ..., v_m\}$ are even and elements $\{v_{m+1}, ..., v_{n+m}\}$ are odd. Elements that are solely even or odd *i.e.* $v_E = \sum_{i=1}^{m} a_i v_i$ or $v_O = \sum_{i=m+1}^{m+n} a_i v_i$ are called **homogeneous**. We differ between the two by defining the **degree** of an element to be $\deg(v) = 0$ or 1 if v is even or odd respectively. V is therefore decomposed into an even and odd subspace $V = V_0 \oplus V_1$.

Graded vector spaces are really only meaningful when endowed with some sort of product between elements. If the product is associative it becomes an **associative superalgebra**. An example of this would be the Grassmann algebra we discussed in the previous section. In the $n = 3$ case $B\mathbb{C}_3 = B\mathbb{C}_{3,0} \oplus B\mathbb{C}_{3,1}$ with $B\mathbb{C}_{3,0} = \{1, \theta_1\theta_2, \theta_1\theta_3, \theta_2\theta_3\}$ and $B\mathbb{C}_{3,1} = \{\theta_1, \theta_2, \theta_3, \theta_1\theta_2\theta_3\}$. There is another example we will discuss briefly, as it introduces an object which will be of use later on.

Definition 4.3. A **supermatrix** is a $(p+q) \times (r+s)$ matrix of the form:

$$M = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \tag{4.8}$$

where $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\boldsymbol{D}$ are $(p \times r), (p \times s), (q \times r)$ and $(q \times s)$ submatrices belonging to different parts of the Grassmann algebra $B\mathbb{C}_n$: M is a $(p+q) \times (r+s)$ **even supermatrix** if $\boldsymbol{A}, \boldsymbol{D} \in B\mathbb{C}_{n0}$ and $\boldsymbol{B}, \boldsymbol{C} \in B\mathbb{C}_{n1}$. M is a $(p+q) \times (r+s)$ **odd supermatrix** if $\boldsymbol{A}, \boldsymbol{D} \in B\mathbb{C}_{n1}$ and $\boldsymbol{B}, \boldsymbol{C} \in B\mathbb{C}_{n0}$. The degree is then defined as before.

Clearly if M, M' are $(p+q)\times(r+s)$ and $(p'+q')\times(r'+s')$ supermatrices respectively, matrix multiplication is only defined for $r = p'$, $s = q'$. The matrix superalgebra $M(p|q;\mathbb{C})$ with $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\boldsymbol{D}$ being $(p\times p), (p\times q), (q\times p)$ and $(q\times q)$ submatrices respectively is an associative superalgebra. Since Lie algebra's are vector spaces, we may now extend their definition to suit our needs:

Definition 4.4. A **Lie superalgebra** is a grade-1 vector space $L = L_0 \oplus L_1$, with $\dim(L_0) = m$ and $\dim(L_1) = n$, endowed with a *supercommutator* $[A, B\} \ \forall\ A, B \in L$. This retains the properties of closure and linearity of the usual commutator on a normal Lie algebra. In addition, we have $\forall\ A, B, C \in L,\ a, b, c := \deg(A), \deg(B), \deg(C)$:

1. $[A, B\} = -(-1)^{ab}[A, B\}$
2. $[A, B\} \in L_0$ iff $A, B \in L_0$, $A \in L_0$ and $B \in L_1$ or $B \in L_0$ and $A \in L_1$. $[A, B\} \in L_1$ iff $A, B \in L_1$. This is the grading structure we expect from Def. (4.2).
3. $(-1)^{ac}[A, [B, C\}\} + (-1)^{ba}[B, [C, A\}\} + (-1)^{bc}[C, [A, B\}\} = 0$

Remark. Relation (2) is called the **generalized Jacobi identity**. Some simple algebraic manipulation shows that if $A, B \in L_0$, $A \in L_0$ and $B \in L_1$, or $A \in L_1$ and $B \in L_0$ we have that $[A, B\} \equiv [A, B]$ where $[A, B]$ is the normal commutator. If $A, B \in L_1$ then $[A, B\} \equiv \{A, B\}$. This will be key in our construction in section 4.2. To generalise this to a grade-n structure we use that $[L_i, L_j\} \in L_{(i+j) \bmod (n+1)}$ and $[L_i, L_j\} = -(-1)^{(ij) \bmod (n+1)}[L_j, L_i\}$.

The Lie superalgebra is a **non-associative superalgebra**. With this definition, we have all the tools to extend the Poincaré algebra.

4.2 The Full $N = 1$ Supersymmetry Algebra

This discussion is based on chapter 2 of [3]. Let's consider introducing a fermionic generator $Q_{\alpha_1\ldots\alpha_a,\dot{\alpha}_1\ldots\dot{\alpha}_b}$. This transforms in the irrep $(\frac{a}{2}, \frac{b}{2})$, with a, b both odd, and we take it to be complex Grassmann valued thus it's conjugate $\bar{Q}_{\dot{\alpha}_1\ldots\dot{\alpha}_a,\alpha_1\ldots\alpha_b}$ is also in the algebra and transforms as $(\frac{b}{2}, \frac{a}{2})$. The anticommutator $\{Q_{\alpha_1\ldots\alpha_a,\dot{\alpha}_1\ldots\dot{\alpha}_b}, \bar{Q}_{\dot{\beta}_1\ldots\dot{\beta}_a,\beta_1\ldots\beta_b}\}$ will transform as $(\frac{a}{2}, \frac{b}{2}) \otimes (\frac{b}{2}, \frac{a}{2})$ and which has spin $(a + b)$. Now by the previous remark, if these are to be elements of the odd part of a grade-1 Lie algebra, then the anticommutator occupies the even part. The restriction placed on the generators by the Coleman-Mandula theorem means this can either be the B_l which are scalars *i.e.* $(0, 0)$, the translation generators P_μ which transform in $(\frac{1}{2}, \frac{1}{2})$ or the Lorentz generators $J_{\mu\nu}$ which transform in the reducible $(1, 0) \oplus (0, 1)$. Since a, b are positive integers, the only possibility is $a = 1, b = 0$ or arbitrarily vice versa. Thus we introduce the algebra $L = L_0 \oplus L_1$, where $L_0 = \mathfrak{iso}(1, 3)$ and $L_1 = (Q^I_\alpha, \bar{Q}^I_{\dot{\alpha}})$, $I = 1, ..., N$ with N the complex dimension of L_1. We can now construct the algebra:

1. $\{Q^I_\alpha, \bar{Q}^J_{\dot{\beta}}\}$: from our discussion above we know $\{Q^I_\alpha, \bar{Q}^J_{\dot{\beta}}\} \sim P_\mu$. In fact we have $\{Q^I_\alpha, \bar{Q}^J_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu \delta^{IJ}$ where the δ^{IJ} indicates the choice of a diagonal basis for L_1.

2. $[P_\mu, Q^I_\alpha]$: This transforms as $(\frac{1}{2},\frac{1}{2})\otimes(\frac{1}{2},0) = (1,\frac{1}{2})\oplus(0,\frac{1}{2})$. Terms of $(1,\frac{1}{2})$ are forbidden by Coleman-Mandula thus we expect something of the following form $[P_\mu, Q^I_\alpha] = C^I_J(\sigma_\mu)_{\alpha\dot{\beta}}\bar{Q}^{J\dot{\beta}}$, and analogously $[P_\mu, \bar{Q}^I_{\dot{\alpha}}] = (C^I_J)^*(\bar{\sigma}_\mu)_{\dot{\alpha}\beta}Q^{J\beta}$, where C^I_J is some as yet unknown matrix. Here we can employ the usual Jacobi identity:
$$\begin{aligned}0 &= [[Q^I_\alpha, P_\mu], P_\nu] + [[P_\mu, P_\nu], Q^I_\alpha] + [[P_\nu, Q^I_\alpha], P_\mu]\\ &= -C^I_J(\sigma_\mu)_{\alpha\dot{\beta}}[\bar{Q}^{J\dot{\beta}}, P_\nu] + C^I_J(\sigma_\nu)_{\alpha\dot{\beta}}[\bar{Q}^{J\dot{\beta}}, P_\mu],\ \text{since } [P_\mu, P_\nu] = 0\\ &= C^I_J(C^J_K)^*(\sigma_\mu)_{\alpha\dot{\beta}}(\bar{\sigma}_\nu)^{\dot{\beta}}_\gamma Q^{K\gamma} - C^I_J(C^J_K)^*(\sigma_\nu)_{\alpha\dot{\beta}}(\bar{\sigma}_\mu)^{\dot{\beta}}_\gamma Q^{K\gamma} = 4(CC^*)^I{}_K(\sigma_{\mu\nu})_{\alpha\gamma}Q^{K\gamma}\end{aligned}$$
where we've introduced $(\sigma^{\mu\nu})_\alpha{}^\beta := \frac{1}{4}(\sigma^\mu_{\alpha\dot{\gamma}}(\bar{\sigma}^\nu)^{\dot{\gamma}\beta} - \sigma^\nu_{\alpha\dot{\gamma}}(\bar{\sigma}^\mu)^{\dot{\gamma}\beta})$ (this and the corresponding $(\bar{\sigma}^{\mu\nu})$ give the left/right components of the generators $S^{\mu\nu}$ in (3.25) *i.e.* an alternative form of the generators of $\mathcal{M}_L$,$\mathcal{M}_R$). This implies $CC^* = 0$. It will turn out that $C \equiv 0$ but for this we must obtain an additional constrain which involves the next anticommutator.

3. $\{Q^I_\alpha, Q^J_\beta\}$: This must transform as $(\frac{1}{2},0)\otimes(\frac{1}{2},0) = (0,0)\oplus(1,0)$ thus it will have the form $\{Q^I_\alpha, Q^J_\beta\} = \epsilon_{\alpha\beta}Z^{IJ} + \epsilon_{\beta\gamma}(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}$. The Z^{IJ} are Lorentz scalars and given $\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha}$ and $\{Q^I_\alpha, Q^J_\beta\} = \{Q^J_\beta, Q^I_\alpha\}$ we must have $Z^{IJ} = -Z^{JI}$. Similarly, since clearly $\epsilon_{\beta\gamma}(\sigma^{\mu\nu})_\alpha{}^\gamma = \epsilon_{\alpha\gamma}(\sigma^{\mu\nu})_\beta{}^\gamma$, we must have $Y^{IJ} = Y^{JI}$. Now
$$\begin{aligned}&[\{Q_{I\alpha}, Q_{J\beta}\}, P_\mu] = \{Q_{I\alpha}, [Q_{J\beta}, P_\mu]\} + \{Q_{J\beta}, [Q_{I\alpha}\}, P_\mu]\\ \implies &[\{Q_{I\alpha}, Q_{J\beta}\}, P_\mu] = [\epsilon_{\alpha\beta}Z^{IJ}, P_\mu] + [\epsilon_{\beta\gamma}(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}, P_\mu] = [\epsilon_{\beta\gamma}(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}, P_\mu]\end{aligned}$$
where the $[\epsilon_{\alpha\beta}Z^{IJ}, P_\mu] = 0$ since the P_μ commute with scalars (since derivatives do). Now $\epsilon^{\alpha\beta}(\sigma^{\mu\nu})_\alpha{}^\gamma = 0$ (since the Pauli matrices are traceless and they're closed under commutator), thus $\epsilon^{\alpha\beta}[\epsilon_{\beta\gamma}(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}, P_\mu] = 0$. From this we obtain:
$$\begin{aligned}0 =&\{Q_{I\alpha}, [Q_{J\beta}, P_\mu]\} + \{Q_{J\beta}, [Q_{I\alpha}, P_\mu]\}\\ &= \epsilon^{\alpha\beta}C_I{}^K(\sigma_\mu)_{\beta\dot{\beta}}\{Q_{J\alpha}, \bar{Q}^{\dot{\beta}}_K\} - \epsilon^{\alpha\beta}C_J{}^K(\sigma_\mu)_{\beta\dot{\beta}}\{Q_{I\alpha}, \bar{Q}^{\dot{\beta}}_K\} = 2(C_{IJ} - C_{JI})P_\mu\end{aligned}$$
which yields $C_{IJ} = C_{JI}$. Thus along with (2) we obtain $CC^\dagger = 0$ which implies $C = 0$. To complete this part, we substitute $[Q_{J\beta}, P_\mu] = [Q_{I\alpha}, P_\mu] = 0$ to obtain $[\{Q_{I\alpha}, Q_{J\beta}\}, P_\mu] = 0 \implies [Z^{IJ}, P_\mu] + [(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}, P_\mu] = 0 \implies [(\sigma^{\mu\nu})_\alpha{}^\gamma J_{\mu\nu}Y^{IJ}, P_\mu] = 0 \implies Y^{IJ} = 0$ since $[J_{\mu\nu}, P_\rho] \neq 0$.

4. In general, the supercharges Q will carry a representation of the internal symmetry group *i.e.* $[Q^I_\alpha, B_l] = (b_l)^I{}_J Q^J_\alpha$ and $[\bar{Q}^I_{\dot{\alpha}}, B_l] = -\bar{Q}^J_{\dot{\alpha}}(b_l)^I{}_J$. The largest possible internal symmetry group that can act non-trivially on the Q's is $U(N)$. This is called the **R-symmetry group**.

By further use of the Jacobi identities we can show that the Z^{IJ} actually commute with all generators of the supersymmetry algebra. For this reason, they are called **central charges**. The full algebra is then (3.9) along with:

$$[P_\mu, Q^I_\alpha] = [P_\mu, \bar{Q}^I_{\dot{\alpha}}] = 0 \quad [J_{\mu\nu}, Q^I_\alpha] = i(\sigma_{\mu\nu})_\alpha{}^\beta Q^I_\beta, \quad [J_{\mu\nu}, \bar{Q}^{I\dot{\alpha}}] = i(\bar{\sigma}_{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}} \bar{Q}^{I\dot{\beta}} \tag{4.9}$$

$$\{Q^I_\alpha, \bar{Q}^J_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu \delta^{IJ}, \quad \{Q^I_\alpha, Q^J_\beta\} = \epsilon_{\alpha\beta} Z^{IJ} \quad \{\bar{Q}^I_{\dot{\alpha}}, \bar{Q}^J_{\dot{\beta}}\} = \epsilon_{\dot{\alpha}\dot{\beta}} (Z^{IJ})^* \tag{4.10}$$

$$[B_l, B_m] = i f_{lm}^{\;\;n} B_n, \;\; [P_\mu, B_l] = [J_{\mu\nu}, B_l] = 0 \tag{4.11}$$

$$[Q^I_\alpha, B_l] = (b_l)^I{}_J Q_\alpha{}^J, \quad [\bar{Q}_{I\dot{\alpha}}, B_l] = -\bar{Q}_{J\dot{\alpha}} (b_l)^J{}_I \tag{4.12}$$

This construction is very general. We will predominantly consider the $N = 1, D = 4$ superPoincaré algebra in which the L_1 has only one complex dimension. In this case, $I = J = 1$ and thus the central charges must vanish.

Remark. In the $N = 1$ case, we have two supersymmetry generators, or **supercharges**, corresponding to one Majorana spinor $Q^I = \begin{pmatrix} Q^I_\alpha \\ \bar{Q}^{I\dot{\alpha}} \end{pmatrix}$. By relationship (1) we see that in local supersymmetry where $Q^I = Q^I(x^\mu)$ the translation generators P_μ MUST also depend on spacetime coordinate x^μ. This is a theory invariant under general coordinate transforms *i.e.* gravity arises naturally in theories of local supersymmetry! Such theories are called **supergravity** theories.

4.2.1 The SuperPoincaré group $\overline{Osp(4|1)}$

Before we move on to finding the irreps of the unextended superPoincaré algebra on the quantum state space, we will give a brief description of the corresponding *supergroup*. We draw on chapter 1 of [7]. First of all, we give the following example of a supermatrix associative superalgebra. Let $M \in M(2p|N; \mathbb{C})$ be the $(2p+N) \times (2p+N)$ supermatrix given below, with degree $m := \deg(M)$. We can define it's **supertranspose** as follows:

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \implies M^{sT} = \begin{pmatrix} A^T & (-1)^m C^T \\ -(-1)^m B^T & D^T \end{pmatrix} \tag{4.13}$$

Now consider the set: $\mathfrak{osp}(2p|N; \mathbb{C}) := \{M | M^{sT} K + (-1)^m K M = 0\}$ where:

$$K := \begin{pmatrix} J_p & 0 \\ 0 & \mathbb{1}_N \end{pmatrix}, \quad J_p := \begin{pmatrix} 0 & \mathbb{1}_{(2p)} \\ -\mathbb{1}_{(2p)} & 0 \end{pmatrix} \tag{4.14}$$

Now substituting $M_0 = \left(\begin{smallmatrix} A & 0 \\ 0 & D \end{smallmatrix}\right) \in L_0$ and $M_1 = \left(\begin{smallmatrix} 0 & B \\ C & 0 \end{smallmatrix}\right) \in L_1$ into $M^{sT}K + (-1)^m KM = 0$ individually, and using (4.13) and (4.14) we obtain the relations:

$$(1)\;\; A^T J_p + J_p A = 0, \;\; (2)\;\; D^T \mathbb{1}_N + \mathbb{1}_N D = 0, \;\; (3)\;\; C = \mathbb{1}_N B^T J_p, \; B = J_p C^T \mathbb{1}_N \tag{4.15}$$

where we've used $J_p^{-1} = -J_p$. The relations (1) define the **Hamiltonian matrices**; the basis of the **symplectic Lie algebra** $\mathfrak{sp}(2p; \mathbb{C})$. (2) are clearly those of the Lie algebra of the special orthogonal group, $\mathfrak{so}(N; \mathbb{C})$. Thus $L_0 = \mathfrak{sp}(2p; \mathbb{C}) \oplus \mathfrak{so}(N; \mathbb{C})$. For this reason the correpsonding supergroup $Osp(2p|N)$ is called the **orthosymplectic group**.

It actually turns out that if we set $p = 2$ and take an *Inönu-Wigner contraction* of this group we obtain the superPoincaré group, denoted $\overline{Osp(4|N;\mathbb{C})}$ or simply $\overline{Osp(4|N)}$. The Inönu-Wigner contraction essentially involves re-scaling some of the group generators by a parameter e, finding the new re-scaled commutation relations, and taking the limit $e \to 0$. In general this will reduce some relations to zero thus *contracting* the group algebra. In this contraction we have that $A \to P_\mu, J_{\mu\nu}$, $D \to Z^{IJ}$, $B, C \to Q^I, \bar{Q}^I$. Let's just take a moment to appreciate the consistency of this. First of all $ISO(1,3) \cong \mathcal{R}^{1,3} \rtimes SL(2,\mathbb{C})$ so it corresponds to 4 complex dimensions, 2 for $SL(2,\mathbb{C})$ and 2 more since $\mathcal{R}^{1,3}$ has 4 real dimensions. For each value of I, the fermionic generators $Q^I, \bar{Q}^I$ add 1 complex dimension (since $\bar{Q}^I$ is the conjugate of Q^I). The superalgebra dimension is $4+N$ so this is consistent. If we take $N = 1$ to obtain the unextended superPoincaré algebra, then the D's are 1-dimensional so the only solution to (2) is $D = 0$. Exponentiating this gives 1 which is indeed the only element of $SO(1,\mathbb{C})$. This is consistent with $Z^{IJ} = 0$ for $N = 1$.

4.3 Representations on States

The final section of this chapter is based on material from chapters 2, 3 and 23 of [18], [3] and [7] respectively. Now that we've extended the Poincaré algebra we want to understand how such a symmetry would behave in the real world. Since we concluded chapter 3 by linking our abstract mathematics of symmetry to real world particles, let's begin with the representations on the Hilbert space of Quantum states. The supersymmetry algebra contains the Poincaré algebra as a closed subalgebra, therefore any irrep of the superalgebra must be a (generally reducible) representation of the Poincaré algebra also. Now the particle interpretation of the $ISO(1,3)$ irreps has an interesting consequence: the multiplets of the supersymmetry algebra will contain a collection of different particles - they are **supermultiplets**.

Like in section 3.5, here we will again be using the method of induced representations. We will only consider the cases in which central charges are all zero. To begin, we have a few of considerations. Note that by (4.9) we have $[P_\mu, Q^I_\alpha] = [P_\mu, \bar{Q}^I_{\dot{\alpha}}] = 0$ and thus given a Quantum state $\psi_{p,m}$ we have:

$$0 = [P_\mu, Q^I_\alpha]\psi_{p,m} = P_\mu(Q^I_\alpha \psi_{p,m}) - Q^I_\alpha(P_\mu \psi_{p,m}) = P_\mu(Q^I_\alpha \psi_{p,m}) - pQ^I_\alpha \psi_{p,m}$$

since p is not Grassmann so commutes with Q^I_α. This means $P^2(Q^I_\alpha \psi_{p,m}) = M^2(Q^I_\alpha \psi_{p,m})$ *i.e.* particles in the same supermultiplet share the same mass. Thus P^2 is still a Casimir operator! However, as we can see from (4.9), $[J_{\mu\nu}, Q_\alpha] = i(\sigma_{\mu\nu})_\alpha{}^\beta Q_\beta$ so W^2 is no longer a Casimir operator. This is to be expected; if supersymmetry transforms fermions onto bosons and vice versa then a supermultiplet will contain particles of both integer and half integer spin so W^2 won't be constant over an irrep! More specifically we have the following result:

Proposition 4.5. Any given supermultiplet contains an *equal* number of bosonic and fermionic particles.

Proof. Let $(-1)^{N_F}$ define the **fermion number operator**, where $N_F = 2s$ for s the spin of the particle. By the spin-statistics theorem, this will give 1 for a boson and -1 for a fermion. On a finite supermultiplet, if the trace $\mathrm{Tr}((-1)^{N_F}) = 0$ then the number of bosons and fermions must be equal so as to cancel all the 1's and -1's. Let $|F\rangle$ be an arbitrary fermionic state, and $|B\rangle = Q^I_\alpha|F\rangle$ is it's corresponding bosonic state. We first have that:

$$\{Q^I_\alpha, (-1)^{N_F}\}|F\rangle = Q^I_\alpha(-1)^{N_F}|F\rangle + (-1)^{N_F}Q^I_\alpha|F\rangle = -Q^I_\alpha|F\rangle + (-1)^{N_F}|B\rangle = -|B\rangle + |B\rangle = 0$$

Now using $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$ and $\{Q^I_\alpha, (-1)^{N_F}\} = 0$ we have:

$$\begin{aligned} 0 &= \mathrm{Tr}(Q^I_\alpha(-1)^{N_F}\bar{Q}^J_{\dot\beta}) - \mathrm{Tr}(Q^I_\alpha(-1)^{N_F}\bar{Q}^J_{\dot\beta}) = \mathrm{Tr}((-1)^{N_F}\bar{Q}^J_{\dot\beta}Q^I_\alpha - Q^I_\alpha(-1)^{N_F}\bar{Q}^J_{\dot\beta}) \\ &= \mathrm{Tr}((-1)^{N_F}\bar{Q}^J_{\dot\beta}Q^I_\alpha + (-1)^{N_F}Q^I_\alpha\bar{Q}^J_{\dot\beta}) = \mathrm{Tr}((-1)^{N_F}\{\bar{Q}^J_{\dot\beta}, Q^I_\alpha\}) \equiv \mathrm{Tr}((-1)^{N_F}\{Q^I_\alpha, \bar{Q}^J_{\dot\beta}\}) \end{aligned}$$

But $\mathrm{Tr}((-1)^{N_F}\{Q^I_\alpha, \bar{Q}^J_{\dot\beta}\}) = 2\sigma^\mu_{\alpha\dot\beta}\mathrm{Tr}((-1)^{N_F})P_\mu\delta^{IJ}$. Since this holds for all P_μ, we can arbitrarily assert $P_\mu \neq 0$, thus $\mathrm{Tr}((-1)^{N_F}) \equiv 0$. □

Finally, we will show that the little algebras of Wigner's classification generalize to little *superalgebras* which incorporate the additional supercharge generators. This allows us to use the same familiar standard k-states and then boost to other states in the orbit of k. To remove ambiguity now that spin is no longer constant on a multiplet, we denote the state vectors $|p, j, m\rangle$ where p is the momentum 4-vector, j the max spin of the $ISO(1,3)$ irrep, and m the degeneracy of the irrep. In this notation (3.39) takes the form, $\forall\ (\Lambda, a) \in ISO(1,3)$:

$$U^{\rho_k}[(\Lambda, a)]|p, j, m\rangle = U^{\Lambda p}_{\mathcal{R}}[(\mathbb{1}, a)]\sum_{n=-j}^{j}\Gamma^j[(\tilde\Lambda(\Lambda p, k)^{-1}\Lambda\tilde\Lambda(p,k), 0)]_{mn}|\Lambda p, j, n\rangle \tag{4.16}$$

Now in setting $a = 0$ and $p = k$ we note that $\tilde\Lambda(p,k)^{-1} = \Lambda$ and $\tilde\Lambda(k,k) \in L(k)$, meaning we choose $\tilde\Lambda(k,k) = \mathbb{1}$ w.l.o.g. Thus (4.16) becomes:

$$U^{\rho_k}[(\Lambda, 0)]|k, j, m\rangle = |\Lambda k, j, m\rangle \tag{4.17}$$

Now since the supercharges $Q^I_\alpha, \bar{Q}^I_{\dot\alpha}$ transform as $(\frac{1}{2}, 0)$,$(0, \frac{1}{2})$ they span the representation space of $\mathcal{M} \in SL(2,\mathbb{C})$. Thus, if $U^{\rho_k}[(\Lambda, 0)] \equiv U^{\rho_k}(\Lambda) = \exp(\frac{i}{2}\omega^{\mu\nu}J_{\mu\nu})$ the relation $[J_{\mu\nu}, Q^I_\alpha] = i(\sigma_{\mu\nu})_\alpha{}^\beta Q^I_\beta$ gives:

$$[U^{\rho_k}(\Lambda), Q^I_\alpha] = (\mathcal{M}_L)_\alpha{}^\beta Q^I_\beta$$

i.e. the vector $U^{\rho_k}(\Lambda)(Q^I_\alpha|k, j, m\rangle)$ differs from $Q^I_\alpha|\Lambda k, j, m\rangle$ (given by (4.17)) by the vector given by the action of the *transformed* supercharge $(\mathcal{M}_L)_\alpha{}^\beta Q^I_\beta|k, j, m\rangle$ (the case for $\mathcal{M}_R$ and $\bar{Q}^I_{\dot\beta}$ is analogous). Thus we see that:

$$Q^I_\alpha|\Lambda k, j, m\rangle = U^{\rho_k}(\Lambda)Q^I_\alpha|k, j, m\rangle - (\mathcal{M}_L)_\alpha{}^\beta Q^I_\beta|k, j, m\rangle \tag{4.18}$$

so we do in fact only require $Q^I_\alpha|k, j, m\rangle$ and we can simply boost it using the above expression to obtain the other states $Q^I_\alpha|\Lambda k, j, m\rangle$. The actual supercharges present in the little superalgebra depends on the case at hand.

4.3.1 Massive Case

In the non-super case, the massive case has standard vector $k = (M, 0, 0, 0)$ and the little algebra is $\mathfrak{so}(3)$ as we see from Table 3.1, generated by $L_1 = J_{23}, L_2 = J_{13}, L_3 = J_{12}$. Recall from chapter 3 that these satisfied the $\mathfrak{su}(2)$ Lie algebra relations $[L_i, L_j] = i\epsilon_{ijk}L_k$, and thus using our discussion in section 2.3.3 to construct operators L_+, L_- satisfying (2.11). Let us now look at how our supercharges $(Q^I_\alpha, \bar{Q}^I_{\dot\alpha})$ extend the little algebra. Setting $P_\mu = (M, 0, 0, 0)$ gives $\sigma^\mu P_\mu = M\mathbb{1}$, so we obtain:

$$\{Q^I_\alpha, \bar{Q}^J_{\dot\beta}\} = 2M\delta_{\alpha\dot\beta}\delta^{IJ} \tag{4.19}$$

Now using $[J_{\mu\nu}, Q^I_\alpha] = i(\sigma_{\mu\nu})_\alpha{}^\beta Q^I_\beta$ and $[J_{\mu\nu}, \bar{Q}^{I\dot\alpha}] = i(\bar\sigma_{\mu\nu})^{\dot\alpha}{}_{\dot\beta}\bar{Q}^{I\dot\beta}$ we can see how the supercharges in the little superalgebra act on the states $|k, j, m\rangle$. For example, using the explicit form of the Pauli matrices, and the epsilon tensors $\epsilon^{\dot\alpha\dot\beta}, \epsilon_{\dot\alpha\dot\beta}$, we can calculate for $L_3 \equiv J_{12}$:

$$[L_3, Q^I_\alpha] = \frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)_\alpha{}^\beta Q^I_\beta, \quad [L_3, \bar{Q}^I_{\dot\alpha}] = \frac{1}{2}\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)_{\dot\alpha}{}^{\dot\beta}\bar{Q}^I_{\dot\beta} \tag{4.20}$$

This shows us that $Q^I_1, \bar{Q}^I_{\dot 2}$ raise the spin of a state by $\frac{1}{2}$, whereas the $Q^I_2, \bar{Q}^I_{\dot 1}$ lower it by $\frac{1}{2}$. Our little superalgebra therefore contains the generators $\{L_\pm, L_3, Q^I_1, Q^I_2, \bar{Q}^I_{\dot 1}, \bar{Q}^I_{\dot 2}\}$. To build our irreps we define the set of **creation** and **annihilation** operators:

$$a^I_{1,2} := \frac{1}{\sqrt{2M}}Q^I_{1,2}, \quad (a^\dagger)^I_{1,2} := \frac{1}{\sqrt{2M}}\bar{Q}^I_{\dot 1,\dot 2} \tag{4.21}$$

These satisfy the oscillator algebra $\{a^I, (a^\dagger)^J\} = \delta^{IJ}, \{a^I, a^J\} = \{(a^\dagger)^I, (a^\dagger)^J\} = 0$. We then choose some state $|k, j, m\rangle$ such that:

$$a^I_\alpha|k, j, m\rangle = 0, \;\; \forall\; \alpha = 1, 2, \;\; m = j, j-1, ..., -j \tag{4.22}$$

This condition asserts that the state $|k, j, m\rangle$ carries an irrep of $ISO(1, 3)$. Such a state is called a **Clifford vacuum**. By further use of the relations (**??**) and (4.9) in the same manner as in (4.20), we can deduce that the state space spanned by $|k, j, m\rangle$, $\bar{Q}_{\dot 1}|k, j, m\rangle$, $\bar{Q}^I_{\dot 2}|k, j, m\rangle$, and $\bar{Q}^I_{\dot 1}\bar{Q}^I_{\dot 2}|k, j, m\rangle$, for $m = j, j-1, ..., -j$ is invariant under the action of the little superalgebra. The $4(2j + 1)$ members of this supermultiplet fall into 4 categories:

1. $|k, j + \frac{1}{2}, m\rangle$: highest spin $j' = j + \frac{1}{2}$, and $m = j + \frac{1}{2}, j - \frac{1}{2}, ..., -j - \frac{1}{2}$. There are $2j + 2$ such vectors.

2. $|k, j - \frac{1}{2}, m\rangle$: highest spin $j' = j - \frac{1}{2}$, and $m = j - \frac{1}{2}, j - 1, ..., -j + \frac{1}{2}$. There are $2j$ such vectors.

3. $|k, j, m\rangle_{1,2}$: 2 subsets of highest spin $j' = j$, and $m = j, ..., -j$ each with $2j + 1$ vectors.

This is best seen with an example:

$N = 1$ Massive Supermultiplets

Here we have $a_{1,2} := \frac{1}{\sqrt{2M}} Q_{1,2}$, $a^\dagger_{\dot{1},\dot{2}} := \frac{1}{\sqrt{2M}} \bar{Q}_{\dot{1},\dot{2}}$.

1. **Matter multiplet**: here we choose the clifford vacuum to carry the $(0,0)$ representation, *i.e.* $j = 0$: $|k,0,0\rangle$. The little superalgebra then generates the states: $|k,0,0\rangle$, $|k,0,+\frac{1}{2}\rangle$,$|k,0,-\frac{1}{2}\rangle$, $|k,0,0\rangle$. This has the DoF's of 1 massive complex scalar, and 1 massive Majorana fermion. We write:
$$(-\frac{1}{2}, 0, 0, +\frac{1}{2}) \tag{4.23}$$

2. **Gauge (vector) multiplet**: Here the clifford vacuum is $j = \frac{1}{2}$. We generate the states: $|k,1,-1\rangle$, $|k,1,0\rangle$, $|k,1,1\rangle$; $|k,0,0\rangle$; and $|k,\frac{1}{2},-\frac{1}{2}\rangle_{1,2}$, $|k,\frac{1}{2},+\frac{1}{2}\rangle_{1,2}$. This has the DoF's of 1 vector boson, 1 Dirac fermion, and 1 massive real scalar. We write:
$$(-1, \mathbf{2}\times -\frac{1}{2}, \mathbf{2}\times 0, \mathbf{2}\times +\frac{1}{2}, 1) \tag{4.24}$$

4.3.2 Massless Case

The massless case procedure is similar, with a few key differences to consider:

1. We now have $k_\mu = (\kappa, 0, 0, \kappa)$ so setting $P_\mu = k_\mu$ gives $\sigma^\mu P_\mu = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 2\kappa \end{smallmatrix}\right)$. $\{Q^I_\alpha, \bar{Q}^J_{\dot{\beta}}\} = (\sigma^\mu P_\mu)_{\alpha\dot{\beta}} \delta^{IJ}$ then implies that $\{Q^I_1, \bar{Q}^J_{\dot{1}}\} = 0$. Given a state $|k,j\rangle$ this means:
$$0 = \langle k,j|\{Q_1, \bar{Q}_{\dot{1}}\}|k,j\rangle = \|Q_1|k,j\rangle\| + \|\bar{Q}_{\dot{1}}|k,j\rangle\| \implies \|Q_1|k,j\rangle\| = -\|\bar{Q}_{\dot{1}}|k,j\rangle\|$$
so for positive definite norm we must assert that $Q^I_1, \bar{Q}^J_{\dot{1}}$ are trivial and $Q^I_2, \bar{Q}^I_{\dot{2}}$ are our only non-trivial generators. We define $a_I = \frac{1}{\sqrt{2\kappa}} Q^I_2$, $a^\dagger_I = \frac{1}{\sqrt{2\kappa}} \bar{Q}^I_{\dot{2}}$.

2. The little algebra was $\{A := L_2 + K_1, B := -L_1 + K_2, L_3\}$ with $[L_3, A] = iB$ and $[L_3, B] = -iA$, where the eigenvalue of L_3, j, is now the helicity. By the same procedure as in (4.20) we obtain $[L_3, Q^I_2] = -\frac{1}{2} Q^I_2$ and $[L_3, \bar{Q}^I_{\dot{2}}] = \frac{1}{2} \bar{Q}^I_{\dot{2}}$, so they lower/raise the helicity by $\frac{1}{2}$ respectively.

3. Conjugation, parity and time-reversal all flip the helicity thus the overall CPT transformation flips the helicity. It is widely accepted that CPT is a fundamental symmetry in nature, thus we must add the CPT conjugate to obtain the full multiplet. This will be denoted by $\overset{\text{CPT}}{\oplus}$.

The little superalgebra is given by $\{A, B, L_3, Q^I_2, \bar{Q}^I_{\dot{2}}\}$. Applying our process from 4.3.1, *i.e.* choosing a clifford vacuum $|k,j\rangle$ annihilated by a_I, it's not hard to see we only generate the states $|k,j\rangle$ and $|k,j+\frac{1}{2}\rangle$. The CPT symmetry then adds $|k,-j\rangle$ and $|k,-j-\frac{1}{2}\rangle$. For example:

$N = 1$ **Massless Supermultiplets**

1. **Matter (Chiral) Multiplet**: The clifford vacuum is $|k, 0\rangle$ thus we obtain $|k, 0\rangle$ and $|k, +\frac{1}{2}\rangle$. With CPT this becomes:

$$(0, \frac{1}{2}) \overset{\text{CPT}}{\oplus} (-\frac{1}{2}, 0) \tag{4.25}$$

This corresponds to 1 Weyl fermion and 1 complex scalar.

2. **Gauge (vector) Multiplet**: The clifford vacuum is $|k, \frac{1}{2}\rangle$ thus we obtain $|k, 1\rangle$ and $|k, \frac{1}{2}\rangle$. With CPT this becomes:

$$(\frac{1}{2}, 1) \overset{\text{CPT}}{\oplus} (-1, -\frac{1}{2}) \tag{4.26}$$

i.e. 1 Weyl fermion and 1 vector boson.

3. **The Graviton**: The clifford vacuum is $|k, \frac{3}{2}\rangle$ thus we obtain $|k, \frac{3}{2}\rangle$ and $|k, 2\rangle$. With CPT this becomes:

$$(\frac{3}{2}, 2) \overset{\text{CPT}}{\oplus} (-2, -\frac{3}{2}) \tag{4.27}$$

This corresponds to 1 helicity 2 particle, the Graviton, and a $\frac{3}{2}$ particle called the gravitino.

As an afterthought, we briefly mention some of the advantages of supersymmetry that spurred theoretical physicists on in their early investigations of these strange ideas. After all the disappointment of the exotic theories that led to the Coleman-Mandula theorem, supersymmetry provided a saving grace. The idea that most excited theoreticians was in *grand unification*. The minimally supersymmetric version of the standard model suggests a unified force of nature from which the strong, weak and electromagnetic forces emerge at lower energy scales by *spontaneous symmetry breaking*. Elsewhere, the theoretical *superpartners* of the already known particles could provide an explanation for the mysterious dark matter that is abundantly present yet somehow invisible in our universe. Naturally, the discussion in this chapter suggests that we would have already seen these particles since they should exist in the same mass shells as their non-supersymmetric partners. The fact that we haven't seen them means supersymmetry, if it does exist, is spontaneously broken. A lovely overview of these ideas is given in chapter 1 of [3].

The natural next step is to consider field representations of the superPoincaré algebra such that we may build supersymmetric QFT's. We alluded earlier in the chapter that such fields would depend on additional Grassmann parameters. In the following chapter, it is this detail that will motivate us to diverge from generic discussion of supersymmetric theories, and delve into the world of **homogeneous spaces**.

Chapter 5

Superspace and Superfields

In this chapter we will introduce the fields carrying representations of the general supersymmetry algebra. To accommodate the action of the fermionic gennerators $Q_\alpha, \bar{Q}_{\dot{\alpha}}$, these fields will depend on Grassmann parameters $\theta, \bar{\theta}$ in addition to the usual Minkowski space points x^μ. To assure that this structure is well defined, we turn to the topic of *homogeneous spaces.*

5.1 Homogeneous Spaces

In this section we follow the material in [19]. The study of homogeneous spaces essentially provides us with a way to control the symmetry and transformation properties of a space by constructing it from the desired symmetry group. This not only enables us to build new spaces to serve new purposes, but also gain a deeper understanding of the profound connection between symmetry and space in theoretical physics. The very general definition is as follows:

Definition 5.1. A **homogeneous space** is a smooth manifold M endowed with the smooth, transitive action of a Lie group G. It is sometimes called a ***G*-space**.

Before we proceed, it is worth mentioning that there is some very profound and elegant mathematical machinery at work under the surface in the study of homogeneous spaces. Broadly speaking, this beauty is captured by the notion of *fibre bundles*, however we won't discuss this here.

Definition 5.2. The action of a Lie group G on smooth manifold M is called **free** if the isotropy subgroup of all points $p \in M$ is simply the identity: $G_p \equiv \{e\}$, $\forall p \in M$. It is **proper** if the map $(g, p) \in G \times M \to (g \cdot p, p) \in M \times M$ is a **proper map** *i.e.* the pre-image of a compact set is itself compact.

Let M and N be two smooth manifolds in the following two definitions:

Definition 5.3. Given a point $p \in M$, a differentiable map $f : M \to N$ is a **submersion at p** if it's differential $df_p : T_pM \to T_{f(p)}N$ is a surjective linear map. f is a **submersion** if this holds $\forall\ p \in M$.

Definition 5.4. A map $F : M \to N$ is **equivariant** if $\mathcal{L}_g^N(F(p)) = F(\mathcal{L}_g^M(p))$ *i.e.* if the action on the image is the image of the action.

It is now our goal to prove the following theorem:

Theorem 5.5 (*Homogeneous space construction theorem*)**.** Let G be a Lie group and $H \subset G$ a closed subgroup of G.

1. The left coset space G/H is a topological manifold of dimension $\dim(G) - \dim(H)$ and has unique smooth structure such that the projection map $\pi : G \to G/H$ is a smooth submersion.

2. The left action of G on G/H, $G \curvearrowright G/H$, given by $g_1 \cdot (g_2 H) = (g_1 \cdot g_2)H$, turns G/H into a homogeneous space.

To prove this, we require two further theorems which we will state without proof:

Theorem 5.6 (*Quotient Manifold Theorem*)**.** Let G be a Lie group acting smoothly, freely, and properly on a smooth manifold M. The orbit space M/G is a topological manifold of dimension $\dim(M) - \dim(G)$ and has unique smooth structure such that the projection (quotient) map $\pi : M \to M/G$ is a smooth submersion.

Theorem 5.7. Let M, N, P be smooth manifolds and $\pi : M \to N$ a submersion. If $F : M \to P$ is a smooth map such that $\pi(p) = \pi(q) \implies F(p) = F(q)\ \forall\ p, q \in M$, then there exists a unique smooth map $\theta : N \to P$ such that $\theta \circ \pi = F$.

We now give the proof of theorem (5.5):

Proof. To prove part 1, theorem (5.6) does a lot of the hard work. All we need to do is show that the conditions of theorem (5.5) produce the conditions of theorem (5.6) so that we can apply it. The action we want to consider is the right action of H on G, $G \curvearrowleft H$, given by $G \times H \to G$, $(g, h) \mapsto g \cdot h$. We have the following:

1. Since H is a closed subgroup of G it is a Lie subgroup.

2. The action is smooth since it is simply a restriction of the group action of G on itself which is smooth by definition. *i.e.* If $\mathcal{R} : G \times G \to G$ is the right action of G on itself, then the right action of $H \subset G$ on G is $\mathcal{R}|_H : G \times H \to G$.

3. The action is free since for $g \in G$, $h \in H \subset G$ we have that $gh = g \implies h = e$, where e is the identity. This follows from our group axioms in definition (2.1).

4. The action is proper. For this we use **sequential compactness** which is an equivalent formulation. Roughly speaking, a set is compact if all sequences in it have a convergent subsequence. We won't worry about this too much, but essentially, if the sequences $(g_i)_{i\in\mathbb{N}} \in G$, $(h_i)_{i\in\mathbb{N}} \in H$ are such that $(g_i)_{i\in\mathbb{N}}$ and $(g_i \cdot h_i)_{i\in\mathbb{N}}$ converge in G then this is a compact subset in the image of the right action. Now by smoothness and therefore continuity of the inverse map on G, we

have $h_i = g_i^{-1}(g_i \cdot h_i)$. Since $(g_i)_{i\in\mathbb{N}}$ and $(g_i \cdot h_i)_{i\in\mathbb{N}}$ both converge in G, $(h_i)_{i\in\mathbb{N}}$ must converge in G. But $(h_i)_{i\in\mathbb{N}} \in H$ and H is closed, thus $(h_i)_{i\in\mathbb{N}}$ converges in H. Hence, the preimage $((g_i)_{i\in\mathbb{N}}, (h_i)_{i\in\mathbb{N}}) \in G \times H$ is a compact subset.

Thus H is a Lie group acting smoothly, freely, and properly on a smooth manifold G, so by QMT (5.6), G/H is a manifold with unique submersion $\pi : G \to G/H$. This proves part (1) of theorem (5.5).

We now wish to prove part (2). For this we will consider the following maps:

$$\begin{tikzcd} G \times G \arrow[r, "\mathcal{R}"] \arrow[d, "e_G \times \pi"'] & G \arrow[d, "\pi"] \\ G \times G/H \arrow[r, "\phi"'] & G/H \end{tikzcd}$$

where $\mathcal{R}$ is the right action, π is our smooth submersion courtesy of QMT, $e_G \times \pi$ is a smooth submersion, and $\phi : (g_1, g_2H) \mapsto g_1 \cdot (g_2H) = (g_1g_2)H$, *i.e.* ϕ is the left action of G on G/H. It is through this map ϕ that we hope to show (G,G/H) is a homogeneous space. We first want to show that the left action of G on G/H, ϕ, is smooth and well defined, and for this we require theorem (5.7). The figure above allows us to express ϕ using the maps we are already familiar with: $\phi \equiv \pi \circ \mathcal{R} \circ (e_G \times \pi)^{-1}$. Now to translate the language of theorem (5.7) into our situation: the manifolds are as follows $M \equiv G \times G$, $N \equiv G \times G/H$ and $P \equiv G/H$. The statement of the theorem in our case is then:

If $e_G \times \pi : G \times G \to G \times G/H$ is a submersion, and if $\pi \circ \mathcal{R} : G \times G \to G/H$ is a smooth map such that $e_G \times \pi(g_1, g_2) = e_G \times \pi(g_3, g_4) \implies \pi \circ \mathcal{R}(g_1, g_2) = \pi \circ \mathcal{R}(g_3, g_4)$, then there exists a unique smooth map $\phi : G \times G/H \to G/H$ such that $\phi \circ (e_G \times \pi) = \pi \circ \mathcal{R}$.

Our statement above make the application of this theorem obvious. If we can prove that $e_G \times \pi(g_1, g_2) = e_G \times \pi(g_3, g_4) \implies \pi \circ \mathcal{R}(g_1, g_2) = \pi \circ \mathcal{R}(g_3, g_4)$ then the theorem provides us with the uniqueness and smoothness we require for the action in the definition of a homogeneous space. Let $g_1, g_2 \in G$ such that $(e_G \times \pi)(g_1, g_2) = (g_1, g_2H)$. Then we define:

$$\begin{aligned} (G \times G)_{(g_1,g_2)} := (e_G \times \pi)^{-1}(g_1, g_2H) &\equiv \{(g_1, g) | gH = g_2H\} \\ &\equiv \{(g_1, g) | \exists h \in H : g_2 \cdot h = g\} \end{aligned} \tag{5.1}$$

Clearly we see $e_G \times \pi(g_1, g_2) = e_G \times \pi(g_3, g_4) \implies (G \times G)_{(g_1,g_2)} \equiv (G \times G)_{(g_3,g_4)}$. Now the restriction $\pi \circ \mathcal{R}|_{(G\times G)_{(g_1,g_2)}} : (G \times G)_{(g_1,g_2)} \to G/H$ is the following map $(g_1, g) \mapsto \pi(g_1 \cdot g) = (g_1g)H = g_1(gH) = g_1(g_2H) = (g_1g_2)H \ \forall\ g$. Thus we do indeed have that $e_G \times \pi(g_1, g_2) = e_G \times \pi(g_3, g_4) \implies \pi \circ \mathcal{R}(g_1, g_2) = \pi \circ \mathcal{R}(g_3, g_4)$. With this theorem (5.7) gives us the smooth action of G on G/H. All that's left to prove is that this action is transitive: let $(g_1, g_2) \in G \times G$. Then $g_2g_1^{-1} \in G$ and we have:

$$\pi \circ \mathcal{R}(g_2g_1^{-1}, g_1H) = \pi(g_2g_1^{-1}g_1H) = \pi(g_2H) = g_2H \ \ \forall g_1, g_2 \in$$

so any two points g_1H and g_2H in G/H are connected by the action $g_2g_1^{-1}$. With this we are (finally...) done! $\square$

We will shortly be applying this theorem to the familiar example of Minkowski space and ultimately use it construct superspace. Before we do this, there is one final theorem to prove. This one concerns the case when we have space which is a homogeneous G-space and wish to categorise it as a coset space G/H, for some subgroup $H \subset G$.

Theorem 5.8 (Characterisation of Homogeneous Spaces)**.** Let G be a Lie group and $\mathcal{M}$ a homogeneous G-space with $p \in \mathcal{M}$. The isotropy group G_p is a closed subgroup of G and the map

$$F : G/G_p \to \mathcal{M}, \;\; g \cdot G_p \mapsto g \cdot p$$

is an equivariant diffeomorphism.

To prove this we require the following theorem, which we will state without proof:

Theorem 5.9. Let M, N be smooth manifolds, G a Lie group, and $F : M \to N$ a smooth map that is equivariant with respect to the transative smooth action of G on M, and any smooth action of G on N. Then F has constant rank *i.e.* if it's bijective, it's a diffeomorphism.

We now prove theorem (5.8):

Proof. Using the left action of G on $\mathcal{M}$, we define the **orbit map** of a point $p \in \mathcal{M}$, $\mathcal{L}^p : G \to \mathcal{M}$, by $g \mapsto g \cdot p$. Note that the pre-image of this map is the isotropy subgroup of p: $(\mathcal{L}^p)^{-1}(p) = G_p$. Since it's induced by the left action of G we know it's smooth and continuous and thus p, as a point, being closed means that $(\mathcal{L}^p)^{-1}(p) = G_p$ is closed. We now wish to show that F is well defined, equivariant, smooth, unique and bijective.

- **Well defined**: $g_1G_p = g_2G_p$ iff $\exists\; h \in G_p$ such that $g_2 = g_1 \cdot h$. Then:

  $$F(g_2G_p) = g_2p = (g_1h)p = g_1(h \cdot p) = g_1 \cdot F(h) = g_1 \cdot F(eG_p) = g_1 \cdot p = F(g_1G_p)$$

- **Equivariant**: $F((g \cdot g')G_p) = (gg') \cdot p = g(g' \cdot p) = g \cdot F(g'G_p)$

- **Smooth and Unique**: For this we use theorem (5.7) with $M \equiv G$, $N \equiv G/G_p$, $P \equiv \mathcal{M}$, $\pi : G \to G/G_p$ a smooth submersion, and $\mathcal{L}^p : G \to \mathcal{M}$.

$$\begin{tikzcd} G \arrow[d, "\pi"'] \arrow[dr, "\mathcal{L}^p"] & \\ G/G_p \arrow[r, "F"'] & \mathcal{M} \end{tikzcd}$$

Now $F = \mathcal{L}^p \circ \pi^{-1}$ and is therefore smooth since $\mathcal{L}^p$ and π both are by definition. Furthermore, $\pi(g_1 G_p) = \pi(g_2 G_p) \iff \pi^{-1}(g_1 G_p) = \pi^{-1}(g_2 G_p)$. We have that $\pi^{-1}(g_1 G_p) = \{g | g G_p = g_1 G_p\}$ and thus $\{g | g G_p = g_1 G_p\} = \{g | g G_p = g_2 G_p\}$. This means, for a given $g \in \pi^{-1}(g_1 G_p) = \pi^{-1}(g_2 G_p)$, $\exists\ h_1 \in G_p$ such that $g_1 \cdot h_1 = g$ (since $g \in \pi^{-1}(g_1 G_p)$) and $\exists\ h_2 \in G_p$ such that $g_2 \cdot h_2 = g$ (since $g \in \pi^{-1}(g_2 G_p)$). Then we can write $\mathcal{L}^p(g_1) = g_1 \cdot p = (g h_1^{-1}) \cdot p = (g_2 h_2 h_1^{-1}) \cdot p = g_2 \cdot (h_2 h_1^{-1} \cdot p)$. Since $h_1, h_2 \in G_p$ and G_p is a closed subgroup we have $h_1^{-1}, h_2 \in G_p$ and thus $(h_2 h_1^{-1} \cdot p) = p$. Then $\mathcal{L}^p(g_1) = g_2 \cdot (h_2 h_1^{-1} \cdot p) = g_2 \cdot p = \mathcal{L}^p(g_2)$. Thus we have proved that $\pi(g_1 G_p) = \pi(g_2 G_p) \implies \mathcal{L}^p(g_1) = \mathcal{L}^p(g_2)$ and theorem (5.7) says that F is a unique map with our desired properties.

- **Bijective**: For arbitrary $q \in \mathcal{M}$ we have, by transitivity of the action of G, that $\exists\ g \in G$ such that $g \cdot p = q$ thus $F(g G_p) = g \cdot p = q$. This proves **surjectivity**. Now for $g_1, g_2 \in G$, $F(g_1 G_p) = F(g_2 G_p) \implies g_1 \cdot p = g_2 \cdot p \implies (g_2^{-1} g_1) \cdot p = p \implies (g_2^{-1} g_1) \in G_p$ thus $g_1 G_p = g_2 G_p$. This proves **injectivity**. We thus have **bijectivity**.

Thus F is well defined, equivariant, smooth, unique and bijective and so by theorem (5.9) F is a diffeomorphism. $\square$

It is now time to apply all this to some examples.

5.1.1 Minkowski Space as a Coset

In this subsection we will give the construction of Minkowski space as the homogeneous coset space:

$$\mathcal{M} \equiv \frac{ISO(1,3)}{SO(1,3)} \tag{5.2}$$

First of all, not that the right action of $SO^+(1,3)$ on $ISO(1,3)$ fixes the translation component of the overall $ISO(1,3)$ group element:

$$(\Lambda, x)(L, 0) = (\Lambda L, x) = (\mathbb{1}, x)(\Lambda L, 0),\ \Lambda, L \in SO^+(1,3) \tag{5.3}$$

i.e $SO^+(1,3)$ is the isotropy subgroup for a given pure translation $(\mathbb{1}, x) \in \mathcal{R}^{1,3}$ under the right action of $ISO(1,3)$ on itself. This motivates us to consider the set of left cosets, $(\Lambda, x)SO^+(1,3) = \{(\Lambda, x)(L, 0) | L \in SO^+(1,3)\}$, of $SO^+(1,3)$ in $ISO(1,3)$:

$$ISO(1,3)/SO^+(1,3) = \{(\Lambda_1, x_1)SO^+(1,3), (\Lambda_2, x_2)SO^+(1,3), ...\} \tag{5.4}$$

Now since $SO^+(1,3)$ is a closed Lie subgroup of $ISO(1,3)$ theorem (5.5) tells us that this is indeed a homogeneous space. However, we will continue so as to make the machinery of this theorem explicit and show that the homogeneous space is indeed Minkowski space. Using the group product of $ISO(1,3)$, and since ΛL spans $SO^+(1,3)$ provided L does, we can choose $L = \Lambda^{-1}$ without loss of generality such that:

$$\begin{aligned}(\Lambda, x)SO^+(1,3) &= \{(\Lambda L, x) | L \in SO^+(1,3)\} = \{(\mathbb{1}, x) | L \in SO^+(1,3)\} \\ &= (\mathbb{1}, x)SO^+(1,3) \\ \implies ISO(1,3)/SO^+(1,3) &= \{(\mathbb{1}, x_1)SO^+(1,3), (\mathbb{1}, x_2)SO^+(1,3), ...\}\end{aligned} \tag{5.5}$$

Now by theorem (5.8) we can define an equivariant map:

$$F : ISO(1,3)/SO^+(1,3) \to \mathcal{M}, \quad (\Lambda, x)SO^+(1,3) \mapsto x \tag{5.6}$$

Then we can show that the equivariance property means F induces the transitive action of $\mathcal{R}^{1,3}$ on $\mathcal{M}$,

$$\begin{aligned}(\mathbb{1}, y) \cdot x &= (\mathbb{1}, y)F[(\mathbb{1}, x)SO^+(1,3)] \\ &= F[(\mathbb{1}, x)(\Lambda, y)SO^+(1,3)] \\ &= F[(\mathbb{1}, x+y)SO^+(1,3)] = x + y\end{aligned} \tag{5.7}$$

and the general Poincaré transformation we expect by $ISO(1,3) \curvearrowright \mathcal{M}$:

$$\begin{aligned}(\Lambda, y)x &= (\Lambda, y)F[(\mathbb{1}, x)SO^+(1,3)] \\ &= F[(\Lambda, y)(\mathbb{1}, x)SO^+(1,3)] \\ &= F[(\Lambda, \Lambda x + y)SO^+(1,3)] = \Lambda x + y\end{aligned} \tag{5.8}$$

Thus we have constructed Minkowski space as a homogeneous space. We will now use this construction to motivate the slightly more complicated case of superspace.

5.1.2 $N = 1, d = 4$ Superspace

To apply what we've learned so far to the case of superspace, we need to consider the behaviour of the supergroup we intend to use. We mentioned this supergroup back in chapter 4; it was the very strange looking *orthosymplectic* group $\overline{Osp(1|4)}$. Our examination of this group used complex valued matrices, and while this made the example simpler to deal with, it's not particularly enlightening when it comes to defining the notion of a *supertranslation*. Ideally we want to exponentiate the $N = 1$ superPoincaré algebra in the *fundamental* representation so as to see how the grade-1 elements of the new group interact with each other and the original group elements (Λ, a). The issue here is that the algebra contains anti-commutators and thus we cannot simply take the exponential as we have done before. However, with the introduction of four Grassmann parameters $\theta_\alpha, \bar{\theta}_{\dot{\alpha}}$ we can convert the anti-commutator relations to commutators. Define $\theta Q = \theta^\alpha Q_\alpha = -\theta_\alpha Q^\alpha$ and $\bar{\theta}\bar{Q} = \bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}} = -\bar{\theta}^{\dot{\alpha}}\bar{Q}_{\dot{\alpha}}$. Now using the following identity:

$$[AB, CD] = A\{B, C\}D - AC\{B, D\} + \{A, C\}DB - C\{A, D\}B \tag{5.9}$$

and using the fact that the parameters $\theta_\alpha, \bar{\theta}_{\dot{\alpha}}$ anti-commute with the supercharges $Q_\alpha, \bar{Q}_{\dot{\beta}}$ and each other, we have:

$$\begin{aligned}[\theta Q, \bar{\theta}\bar{Q}] \equiv [\theta^\alpha Q_\alpha, -\bar{\theta}^{\dot{\beta}}\bar{Q}_{\dot{\beta}}] &= +\theta^\alpha\bar{\theta}^{\dot{\beta}}\{Q_\alpha, \bar{Q}_{\dot{\beta}}\} = \theta^\alpha\{Q_\alpha, \bar{Q}_{\dot{\beta}}\}\bar{\theta}^{\dot{\beta}} \\ &= \theta^\alpha(2\sigma^\mu_{\alpha\dot{\beta}}P_\mu)\bar{\theta}^{\dot{\beta}} = 2\theta\sigma^\mu\bar{\theta}P_\mu\end{aligned} \tag{5.10}$$

Ultimately, we obtain the new commutator relations:

$$[\theta Q, \bar{\theta}\bar{Q}] = 2\theta\sigma^\mu\bar{\theta}P_\mu, \quad [\theta Q, \theta Q] = [\bar{\theta}\bar{Q}, \bar{\theta}\bar{Q}] = 0 \tag{5.11}$$

We omit the laborious process of exponentiating this in the fundamental representation explicitly, it's covered in [7]. We would obtain a matrix similar in form to (4.13) with (4.15), but larger in dimension. The group multiplication rule we obtain is then:

$$(\Lambda_1, x_1, \Theta)(\Lambda_2, x_2, \Theta') = (\Lambda_1\Lambda_2, \Lambda_1 x_2 + x_1 + i\bar{\Theta}\gamma^\mu \mathcal{M}(\Lambda_1)\Theta', \Theta + \mathcal{M}(\Lambda_1)\Theta') \tag{5.12}$$

where $\Theta = (\theta_\alpha, \bar{\theta}_{\dot{\alpha}})$, and similarly for Θ', and $\mathcal{M}(\Lambda_i) = \mathcal{M}_L(\Lambda_i) \oplus \mathcal{M}_R(\Lambda_i)$, $i = 1, 2$. Recall that $\bar{\Theta}$ refers to the Dirac adjoint we saw in chapter 3: $\bar{\Theta} = \Theta^\dagger\gamma^0$. As expected given (5.11), the supercharges generate a change in the translation component of the group element as well as in the Grassmann component. Thus we obtain a **supertranslation**:

$$(\mathbb{1}, x_1, \Theta)(\mathbb{1}, x_2, \Theta') = (\mathbb{1}, x_2 + x_1 + i\bar{\Theta}\gamma^\mu\Theta', \Theta + \Theta') \tag{5.13}$$

We will take a moment to make a little more sense of this. Let's begin by considering the following group multiplication in exponential form:

$$(\mathbb{1}, 0, \theta, \bar{\theta})(\mathbb{1}, x^\mu, \theta', \bar{\theta}') = e^{i(\theta Q + \bar{\theta}\bar{Q})} e^{i(x^\mu P_\mu + \theta' Q + \bar{\theta}'\bar{Q})}$$

By the Baker-Campbell-Hausdorff formula (2.9) this gives:

$$\begin{aligned}
& e^{i(x^\mu P_\mu + (\theta+\theta')Q + (\bar{\theta}+\bar{\theta}')\bar{Q}) + \frac{i^2}{2}[\theta Q + \bar{\theta}\bar{Q}, x^\mu P_\mu + \theta' Q + \bar{\theta}'\bar{Q}] + \ldots} \\
\overset{(5.11)}{=} \; & e^{i(x^\mu P_\mu + (\theta+\theta')Q + (\bar{\theta}+\bar{\theta}')\bar{Q}) - \frac{1}{2}([\theta Q, \bar{\theta}'\bar{Q}] - [\theta' Q, \bar{\theta}\bar{Q}])} \\
= \; & e^{i[(x^\mu + i\theta\sigma^\mu\bar{\theta}' - i\theta'\sigma^\mu\bar{\theta})P_\mu + (\theta+\theta')Q + (\bar{\theta}+\bar{\theta}')\bar{Q}]} = (x^\mu + i\theta\sigma^\mu\bar{\theta}' - i\theta'\sigma^\mu\bar{\theta}, \theta + \theta', \bar{\theta} + \bar{\theta}')
\end{aligned}$$

To turn this into the form (5.13) required a little trickery. Using that $\bar{\sigma}^{\mu\dot{\alpha}\beta} = \epsilon^{\beta\delta}\epsilon^{\dot{\alpha}\dot{\gamma}}\sigma^\mu_{\delta\dot{\gamma}}$ we can write:

$$\begin{aligned}
& \theta\sigma^\mu\bar{\theta}' - \theta'\sigma^\mu\bar{\theta} = \theta\sigma^\mu\bar{\theta}' - \theta'_\beta\epsilon^{\beta\delta}\sigma^\mu_{\delta\dot{\gamma}}\epsilon^{\dot{\gamma}\dot{\alpha}}\bar{\theta}_{\dot{\alpha}} = \theta\sigma^\mu\bar{\theta}' - \theta'_\beta\bar{\sigma}^{\mu\beta\dot{\alpha}}\bar{\theta}_{\dot{\alpha}} = \theta\sigma^\mu\bar{\theta}' + \bar{\theta}_{\dot{\alpha}}\bar{\sigma}^{\mu\beta\dot{\alpha}}\theta'_\beta \\
= & \theta\sigma^\mu\bar{\theta}' + \bar{\theta}\bar{\sigma}^\mu\theta' = (\theta \; \bar{\theta})\begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}\begin{pmatrix} \theta' \\ \bar{\theta}' \end{pmatrix} = \bar{\Theta}\gamma^\mu\Theta'
\end{aligned}$$

Thus we recover the more compact form in (5.13). We now note that, like in the Minkowski space case, the supertranslation component of a group element is invariant under the right action of the Lorentz group *i.e.* a generic group element (Λ, x, Θ) has decomposition:

$$(\Lambda, x, \Theta) = (\mathbb{1}, x, \Theta)(\Lambda, 0, 0) \tag{5.14}$$

Motivated by this, we again consider the left coset space:

$$\frac{\overline{Osp(1|4)}}{SO(1,3)} = \{(\mathbb{1}, x_1, \Theta_1)SO(1,3), (\mathbb{1}, x_2, \Theta_2)SO(1,3), ...\} \tag{5.15}$$

Theorem (5.5) confirms this is indeed a homogeneous space, and theorem (5.8) allows us to define an equivariant map $F : \frac{\overline{Osp(1|4)}}{SO(1,3)} \to \mathcal{S}$, with which we define the superspace coordinates:

$$(x^\mu, \Theta) \equiv (x^\mu, \theta, \bar{\theta}) \tag{5.16}$$

This provides us with a rigorous realization of the $N = 1$, $d = 4$ superMinkowski space. It contains four additional Grassmann coordinates $\Theta = (\theta_\alpha, \bar{\theta}_{\dot{\alpha}})$, $\alpha, \dot{\alpha} = 1, 2$.

Having given superspace a meaningful construction, we will now give a brief overview of the *superfield* representations of the superPoincaré algebra.

5.2 SuperFields

This section follows material covered in chapters 3, 4 and 5 in [18], and chapter 4 of [3]. In this section we will introduce the 'off shell' field supermultiplets. We will quickly discover how the formalism of *superfields* provides an elegant framework within which the task is simplified.

5.2.1 Motivating Superfields

Suppose there exists a field analogous to the Clifford vacuum: let's take a scalar field $\phi(x)$ satisfying the condition

$$[\bar{Q}_{\dot{\alpha}}, \phi(x)] = 0 \tag{5.17}$$

For this to lead to a non-trivial construction we take ϕ to be complex, since if $\phi^\dagger = \phi$ then $0 = [\bar{Q}_{\dot{\alpha}}, \phi(x)]^\dagger = [Q_\alpha, \phi(x)]$ and the Jacobi identity

$$[\phi, \{Q_\alpha, \bar{Q}_{\dot{\beta}}\}] + \{Q_\alpha, [\bar{Q}_{\dot{\beta}}, \phi]\} - \{\bar{Q}_{\dot{\beta}}, [\phi, Q_\alpha]\} = 0 \tag{5.18}$$

which implies $2\sigma^\mu[\phi, P_\mu] = 0$. In turn, since $[\phi, P_\mu] = i\partial_\mu \phi$, this implies $\phi(x)$ is constant. So taking $\phi(x)$ complex, we must have that:

$$[Q_\alpha, \phi(x)] = \psi_\alpha(x)$$

We now continue to act with our supersymmetry generators until we obtain no new fields since we will have then found a set of fields which transform amongst themselves. Continuing, we suggest two further general forms: $\{Q_\alpha, \psi_\beta(x)\} = F_{\alpha\beta}$ and $\{\bar{Q}_{\dot{\alpha}}, \psi_\beta(x)\} = X_{\dot{\alpha}\beta}(x)$. Once again, we employ (5.18) to obtain:

$$[\phi(x), 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu] + \{\bar{Q}_{\dot{\beta}}, \psi_\alpha\} = 0 \implies X_{\dot{\alpha}\beta} = \{\psi_\beta, \bar{Q}_{\dot{\alpha}}\} = 2\sigma^\mu_{\beta\dot{\alpha}}[P_\mu, \phi] = -2i\sigma^\mu_{\beta\dot{\alpha}}\partial_\mu \phi(x)$$

thus $X_{\dot{\alpha}\beta}$ is not a new field, as it's a derivative of $\phi(x)$. Using a further Jacobi identity,

$$[\phi(x), \{Q_\beta, Q_\alpha\}] + \{Q_\beta, [Q_\alpha, \phi(x)]\} - \{Q_\alpha, [\phi(x), Q_\beta]\} = 0 \tag{5.19}$$

we have:

$$\{Q_\beta, \psi_\alpha\} - \{Q_\alpha, -\psi_\beta\} = 0 \implies F_{\alpha\beta} + F_{\beta\alpha} = 0 \implies F_{\alpha\beta} = -F_{\beta\alpha} \implies F_{\alpha\beta} = \epsilon_{\alpha\beta} F$$

Thus we have found a new scalar field F. Once again, we suggest the general forms: $[Q_\alpha, F] = \lambda_\alpha$ and $[\bar{Q}_{\dot{\alpha}}, F] = \bar{\chi}_{\dot{\alpha}}$. Now (5.19) (with F instead of ϕ) gives:

$$\{Q_\beta, \lambda_\alpha\} - \{Q_\alpha, -\lambda_\beta\} = 0 \implies \lambda_\alpha = 0 \tag{5.20}$$

This along with (5.18) gives:

$$\begin{aligned} &[F, 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu] - \{Q_\alpha, \bar{\chi}_{\dot{\beta}}\} + \{\bar{Q}_\beta, \lambda_\alpha\} = 0 \overset{(5.20)}{\implies} [F, 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu] = \{Q_\alpha, \bar{\chi}_{\dot{\beta}}\} \\ &\implies \{Q_\alpha, \bar{\chi}_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}} \epsilon^{\gamma\delta}[\{Q_\gamma, \psi_\delta\}, P_\mu] \end{aligned} \tag{5.21}$$

We make use of one final Jacobi identity:

$$[P_\mu, \{Q_\gamma, \psi_\delta\}] + \{Q_\gamma, [\psi_\delta, P_\mu]\} - \{\psi_\delta, [P_\mu, Q_\gamma]\} = 0 \tag{5.22}$$

$$\implies [P_\mu, \{Q_\gamma, \psi_\delta\}] + \{Q_\gamma, [\psi_\delta, P_\mu]\} = 0 \implies [\{Q_\gamma, \psi_\delta\}, P_\mu] = \{Q_\gamma, \partial_\mu \psi_\delta\}$$

This with (5.21) gives $\{Q_\alpha, \bar{\chi}_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}}\epsilon^{\gamma\delta}\{Q_\gamma, \partial_\mu\psi_\delta\}$. Thus $\bar{\chi}_{\dot{\alpha}}$ is a derivative of ψ_α and is thus not a new field. We now have no new fields to consider so we're done. The multiplet (ϕ, ψ, F) carries a representation of the supersymmetry algebra. It contains a scalar field ϕ, a spinor field ψ and an auxiliary field F. This is called a **chiral multiplet** and it's the field equivalent of the chiral multiplet on states. However it appears to contain too many degrees of freedom: it's bosonic DoF's are $\mathrm{Re}(\phi)$, $\mathrm{Im}(\phi)$, $\mathrm{Re}(F)$, $\mathrm{Im}(F)$ and it's fermionic DoF's are $\mathrm{Re}(\psi_1)$, $\mathrm{Im}(\psi_1)$, $\mathrm{Re}(\psi_2)$, $\mathrm{Im}(\psi_2)$. The massive chiral multiplet had $2_F + 2_B$ DoF's not $4_F + 4_B$. This is because fields are *off-shell* whereas states are *on-shell*. Fields becomes on-shell when we assert that they satisfy the equations of motion, which are derived from the energy relations that restrict them to the mass shell. F is an auxiliary field; it's non-dynamical and depends on ϕ and ψ.

Introducing the Superfield

This procedure was a little tedious. It turns out that by considering fields that depend on points in superspace, $z = (x^\mu, \theta_\alpha, \bar{\theta}_{\dot{\alpha}})$, multiplets such as the one above arise in a significantly more elegant and natural manner. Using the Taylor expansion in Grassmann coordinates, the general form of a superfield on the superspace of coordinates $(x^\mu, \theta_\alpha, \bar{\theta}_{\dot{\alpha}})$ with α, $\dot{\alpha} = 1, 2$, is:

$$\begin{aligned} F(x, \theta, \bar{\theta}) &= f(x) + \theta\phi(x) + \bar{\theta}\bar{\chi}(x) + \theta\theta m(x) + \bar{\theta}\bar{\theta}n(x) \\ &\quad + \theta\sigma^\mu\bar{\theta}v_\mu(x) + \theta\theta\bar{\theta}\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\theta\psi(x) + \theta\theta\bar{\theta}\bar{\theta}d(x) \end{aligned} \tag{5.23}$$

where all higher order terms vanish since they must contain the product of θ's or $\bar{\theta}$'s carrying the same $\alpha/\dot{\alpha}$ index, and $\{\theta_\alpha, \theta_\beta\} = 0 \implies \theta_\alpha^2 = 0$ if we set $\alpha = \beta$. So how does an arbitrary supertranslation $(\mathbb{1}, x, \xi, \bar{\xi})$ act on a superfield? We want to convert the action of $(\mathbb{1}, 0, \xi, \bar{\xi}) \equiv e^{i(\xi Q + \bar{\xi}\bar{Q})}$ on superspace coordinates into differential operators on superfields. Again, here our training in Grassmann numbers comes in handy:

$$\begin{aligned} F(x, \theta, \bar{\theta}) &\to F(x + i\xi\sigma^\mu\bar{\theta} - i\theta\sigma^\mu\bar{\xi}, \xi + \theta, \bar{\xi} + \bar{\theta}) \\ &\approx F(x, \theta, \bar{\theta}) + (i\theta\sigma^\mu\bar{\xi} - i\xi\sigma^\mu\bar{\theta})\partial_\mu F + \bar{\xi}\frac{\partial}{\partial\bar{\theta}}F + \xi\frac{\partial}{\partial\theta}F \\ &= F(x, \theta, \bar{\theta}) + \xi_\alpha\Big[\frac{\partial}{\partial\theta_\alpha} - i(\sigma^\mu)^{\alpha\dot{\alpha}}\bar{\theta}_{\dot{\alpha}}\partial_\mu\Big]F + \bar{\xi}_{\dot{\alpha}}\Big[\frac{\partial}{\partial\bar{\theta}_{\dot{\alpha}}} - i\theta_\alpha(\sigma^\mu)^{\alpha\dot{\alpha}}\partial_\mu\Big]F \end{aligned}$$

Now we equate this with:

$$\begin{aligned} &e^{i(\xi Q + \bar{\xi}\bar{Q})}F(x, \theta, \bar{\theta}) \approx (1 + i\xi^\alpha Q_\alpha + i\bar{\xi}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}})F(x, \theta, \bar{\theta}) \\ \implies &-iQ^\alpha = \partial^\alpha - i(\sigma^\mu)^{\alpha\dot{\alpha}}\bar{\theta}_{\dot{\alpha}}\partial_\mu, \quad i\bar{Q}^{\dot{\alpha}} = \bar{\partial}^{\dot{\alpha}} - i\theta_\alpha(\sigma^\mu)^{\alpha\dot{\alpha}}\partial_\mu \\ \implies &Q^\alpha = i\partial^\alpha + (\sigma^\mu)^{\alpha\dot{\alpha}}\bar{\theta}_{\dot{\alpha}}\partial_\mu, \quad \bar{Q}^{\dot{\alpha}} = -i\bar{\partial}^{\dot{\alpha}} - \theta_\alpha(\sigma^\mu)^{\alpha\dot{\alpha}}\partial_\mu \end{aligned}$$

Or by lowering the indices:

$$Q_\alpha = -i\partial_\alpha - (\sigma^\mu)_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_\mu, \quad \bar{Q}_{\dot{\alpha}} = i\bar{\partial}_{\dot{\alpha}} + \theta^\alpha(\sigma^\mu)_{\alpha\dot{\alpha}}\partial_\mu \tag{5.24}$$

Superfields form linear representations of the supersymmetry algebra, however they are in general highly reducible. We can eliminate some of the component field terms by imposing certain restrictions. In this way, the superfield formalism has reduced the task of finding irreducible representations to finding constraints.

Supersymmetric Actions

Since the supersymmetry transformations are translations in superspace, and as we saw in chapter 4 the Grassmann valued integration measure is invariant under a Grassmann translation, our supersymmetric variation of the action passes inside the integral to give:

$$\begin{aligned}\delta_{\xi\bar{\xi}}\int d^4xd^2\theta d^2\bar{\theta}F(x,\theta,\bar{\theta}) &= \int d^4xd^2\theta d^2\bar{\theta}\delta_{\xi\bar{\xi}}F(x,\theta,\bar{\theta}) \\ &= \int d^4xd^2\theta d^2\bar{\theta}[\partial_\mu((i\theta\sigma^\mu\bar{\xi} - i\xi\sigma^\mu\bar{\theta})F) + \bar{\xi}\frac{\partial}{\partial\bar{\theta}}F + \xi\frac{\partial}{\partial\theta}F] = 0\end{aligned}$$

since the $d^2\theta d^2\bar{\theta}$ integration kills off the latter terms, and the first term is a total derivative *i.e.* vanishes on the boundary. In this way, superfields provide us with a simpler way to present actions with manifest supersymmetric invariance. If we were to simply build Lagrangians, this provides us with a fundamental rule for constructing them: *they are to be built from the component fields of superfields which transform into total space derivatives.*

5.2.2 Chiral Superfields

We will now look at our first example of a superfield, the **chiral superfield**. For this we construct the 'super' covariant derivatives:

$$D_\alpha = \partial_\alpha + i\sigma^\mu_{\alpha\dot{\beta}}\bar{\theta}^{\dot{\theta}}\partial_\mu, \quad \bar{D}_{\dot{\alpha}} = \bar{\partial}_{\dot{\alpha}} + i\theta^\beta\sigma^\mu_{\beta\dot{\alpha}}\partial_\mu \tag{5.25}$$

These have been constructed such that $\{D_\alpha, \bar{D}_{\dot{\alpha}}\} = 2i\sigma^\mu_{\alpha\dot{\beta}}\partial_\mu$, and so as to anticommute with the differential operators $Q_\alpha, \bar{Q}_{\dot{\alpha}}$ so that:

$$D_\alpha(\delta_{\xi,\bar{\xi}}F) = D_\alpha((i\xi Q + i\bar{\xi}\bar{Q})F) = (i\xi Q + i\bar{\xi}\bar{Q})D_\alpha F = \delta_{\xi,\bar{\xi}}(D_\alpha F) \tag{5.26}$$

This is necessary to assure $D_\alpha F$ transforms as a superfield provided F does. Ordinary Grassmann derivatives such as ∂_α remove terms from F which means $\partial_\alpha F$ no longer transforms amongst it's component fields. The covariant derivative solves this problem. We now define the **chiral** and **antichiral** superfields with the following constraints:

$$\bar{D}_{\dot{\alpha}}\Phi = 0, \quad D_\alpha\Psi = 0 \tag{5.27}$$

Clearly, if Φ is chiral, then $\bar{\Phi}$ is antichiral. Note that Φ must be complex in the same way we required ϕ to be complex in (5.17). If it were not, then $\Phi = \bar{\Phi}$ implies $0 = \{D_\alpha, \bar{D}_{\dot{\beta}}\}\Phi = 2i\sigma^\mu_{\alpha\dot{\beta}}\partial_\mu\Phi$ which means $\partial_\mu\Phi = 0$ so Φ is trivially a constant.

So how can we solve this constraint to obtain the most general form of a chiral superfield? It is useful to work in a new set of coordinates:

$$y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}, \quad \bar{y}^\mu = x^\mu - i\theta\sigma^\mu\bar{\theta} \tag{5.28}$$

It then follows that:

$$\bar{D}_{\dot{\alpha}}\theta_\beta = 0, \;\; \bar{D}_{\dot{\alpha}}y^\mu = (\bar{\partial}_{\dot{\alpha}} + i\theta^\beta\sigma^\mu_{\beta\dot{\alpha}}\partial_\mu)(x^\mu + i\theta\sigma^\mu\bar{\theta}) = (-i\theta\sigma^\mu + i\theta\sigma^\mu) = 0$$

and similarly $D_\alpha\bar{\theta}_{\dot{\beta}} = D_\alpha\bar{y}^\mu = 0$. Thus if $\bar{D}_{\dot{\alpha}}\Phi = 0$ defines a chiral superfield, then by the above Φ can only depend (y^μ, θ_α). Using the general form of our superfield (5.23) we obtain:

$$\Phi(y,\theta) = \phi(y) + \sqrt{2}\theta\psi(y) - \theta\theta F(y) \tag{5.29}$$

Substituting in (5.28) and Taylor expanding we find:

$$\Phi(x,\theta,\bar{\theta}) = \phi(x) + \sqrt{2}\theta\psi(x) + i\theta\sigma^\mu\bar{\theta}\partial_\mu\phi(x) - \theta\theta F(x) - \frac{i}{\sqrt{2}}\theta\theta\partial_\mu\psi(x)\sigma^\mu\bar{\theta} - \frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\partial^2\phi(x) \tag{5.30}$$

To see how this transforms under $\delta_{\xi,\bar{\xi}}$ it's useful to transform our operators $Q_\alpha, \bar{Q}_{\dot{\alpha}}$ into the new coordinate system:

$$\begin{aligned}
\frac{\partial}{\partial x^\mu} &\to \frac{\partial y^\nu}{\partial x^\mu}\frac{\partial}{\partial y^\nu} + \frac{\partial\theta}{\partial x^\mu}\frac{\partial}{\partial\theta} + \frac{\partial\bar{\theta}}{\partial x^\mu}\frac{\partial}{\partial\bar{\theta}} = \delta^\nu_\mu\frac{\partial}{\partial y^\nu} = \frac{\partial}{\partial y^\mu} \\
\frac{\partial}{\partial\theta} &\to \frac{\partial y^\mu}{\partial\theta}\frac{\partial}{\partial y^\mu} + \frac{\partial}{\partial\theta} + \frac{\partial\bar{\theta}}{\partial\theta}\frac{\partial}{\partial\bar{\theta}} = i\sigma^\mu\bar{\theta}\frac{\partial}{\partial y^\mu} + \frac{\partial}{\partial\theta} \\
\frac{\partial}{\partial\bar{\theta}} &\to \frac{\partial y^\mu}{\partial\bar{\theta}}\frac{\partial}{\partial y^\mu} + \frac{\partial\theta}{\partial\bar{\theta}}\frac{\partial}{\partial\theta} + \frac{\partial}{\partial\bar{\theta}} = -i\theta\sigma^\mu\frac{\partial}{\partial y^\mu} + \frac{\partial}{\partial\bar{\theta}}
\end{aligned}$$

Thus our operators are now: $Q^{new}_\alpha = -i\partial_\alpha$, $\bar{Q}^{new}_{\dot{\alpha}} = i\bar{\partial}_{\dot{\alpha}} + 2\theta^\alpha(\sigma^\mu)_{\alpha\dot{\alpha}}\partial_{y,\mu}$. With this we obtain:

$$\begin{aligned}
\delta_{\xi,\bar{\xi}}\Phi(y,\theta) &= (\xi^\alpha\partial_\alpha + 2i\theta^\alpha\sigma^\mu_{\alpha\dot{\beta}}\bar{\xi}^{\dot{\beta}}\partial_{y,\mu})\Phi(y,\theta) = \sqrt{2}\xi\psi - 2\xi\theta F + 2i\theta\sigma^\mu\bar{\xi}(\partial_{y,\mu}\phi + \sqrt{2}\theta\partial_{y,\mu}\psi) \\
&= \sqrt{2}\xi\psi + \sqrt{2}\theta(-\sqrt{2}\xi F + \sqrt{2}i\sigma^\mu\bar{\xi}\partial_{y,\mu}\phi) - \theta\theta(-i\sqrt{2}\bar{\xi}\sigma^\mu\partial_{y,\mu}\psi)
\end{aligned}$$

This corresponds to $\delta\phi = \sqrt{2}\xi\psi$, $\delta\psi_\alpha = \sqrt{2}i(\sigma^\mu\bar{\xi})_\alpha\partial_\mu\psi - \sqrt{2}\xi_\alpha F$, $\delta F = i\sqrt{2}\partial_\mu\psi\sigma^\mu\bar{\xi}$. This is exactly the multiplet we discussed in 5.2.1! Finally, we mention in passing that using the rules we established for supersymmetric actions, the Lagrangian for this is:

$$\mathcal{L} = \bar{\Phi}_i\Phi_j|_{\theta\theta\bar{\theta}\bar{\theta}} + \Big[\frac{1}{2}m_{ij}\Phi_i\Phi_j + \frac{1}{3}g_{ijk}\Phi_i\Phi_j\Phi_k + \lambda_i\Phi_i)_{\theta\theta} + h.c.\Big] \tag{5.31}$$

where the restrictions are to the component fields with the Grassmann coefficients specified by the subscripts.

5.2.3 Vector Superfields

If we want to build supersymmetric gauge theories we require a constraint on the general superfield such that the vector component is preserved, and asserted to be real. We therefore define a real **vector superfield** with the following *reality* constraint:

$$\overline{V} = V \tag{5.32}$$

Imposing this on our general superfield is simpler than in the chiral case, but produces a larger expression. We obtain the following general form for a real vector superfield:

$$\begin{aligned} V(x,\theta,\bar{\theta}) = {} & C(x) + i\theta\chi(x) - i\bar{\theta}\bar{\chi}(x) + \theta\sigma^\mu\bar{\theta}v_\mu + \frac{i}{2}\theta\theta(M(x)+iN(x)) \\ & - \frac{i}{2}\bar{\theta}\bar{\theta}(M(x)-iN(x)) + i\theta\theta\bar{\theta}(\bar{\lambda}(x) + \frac{i}{2}\bar{\sigma}^\mu\partial_\mu\chi(x)) \\ & - i\bar{\theta}\bar{\theta}\theta(\lambda(x) + \frac{i}{2}\sigma^\mu\partial_\mu\bar{\chi}(x)) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}(D(x) - \frac{1}{2}\partial^2 C(x)) \end{aligned} \tag{5.33}$$

Note that $\Phi+\bar{\Phi}$ is a vector superfield provided Φ is a chiral superfield. This actually provides us with a natural way to define a supersymmetric gauge transformation. Observe that if we do the following transformation:

$$V \to V + \Phi + \bar{\Phi} \tag{5.34}$$

then the vector v_μ transforms as $v_\mu \to v_\mu - \partial_\mu(2\mathrm{Im}(\phi))$. The other component fields of V transform as:

$$\begin{aligned} & C \to C + 2\mathrm{Re}(\phi),\;\; \chi \to \chi - i\sqrt{2}\psi,\;\; M \to M - 2\mathrm{Im}(F) \\ & N \to N + 2\mathrm{Re}(F),\;\; D \to D,\;\; \lambda \to \lambda \end{aligned} \tag{5.35}$$

Therefore, choosing Re(ϕ)= $-\frac{C}{2}$, $\psi = -\frac{i}{\sqrt{2}}\chi$, Re($F$)= $-\frac{N}{2}$ and Im(F)= $\frac{M}{2}$ we can gauge away C, M, N, χ. This is called the **Wess-Zumino Gauge** and we obtain the much simpler field:

$$V_{WZ} = \theta\sigma^\mu\bar{\theta}v_\mu(x) + i\theta\theta\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}\bar{\theta}\theta\lambda(x) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x) \tag{5.36}$$

Ultimately, the form (5.33) has $8_B + 8_F$ off-shell degrees of freedom. The WZ gauge reduces this to $4_B + 4_F$ off-shell. In imposing the EoM's it would turn out that $D(x)$ is an auxiliary field and thus the multiplet $(D, \lambda_\alpha, v_\mu)$ gives us $2_B + 2_F$ on-shell degrees of freedom. In this gauge, each term of V_{WZ} has at least one θ, thus interaction terms V_{WZ}^n for $n \geq 3$ are all zero. This implies this gives us a framework with which to develop a supersymmetric version of **Yang-Mills** theory.

Supersymmetric gauge theory is a very interesting topic, and one I'd love to explore further. As we remarked in chapter 4, when we make supersymmetric transformations depend on local coordinates gravity arises naturally. In our final chapter, we will explore a topic that draws much of the material of the past three chapters into one, elegant study and helps provide us with the machinery to build supergravity theories.

Chapter 6

SuperConclusion

In this final chapter, we will endeavour to neatly tie together the material we've discussed so far, and conclude with some idea of the direction this can be taken in.

6.1 Unexpected Links: The Magic of Fibre Bundles

It turns out their is a very powerful geometrical concept which ties together our discussion of field representations, induced representations and homogeneous spaces: the notion of *fibre bundles.* This section will not serve as any kind of meaningful introduction to fibre bundles, but will hopefully present the interesting link they provide between the topics we've discussed, and provide some direction for where one might go next. We draw on material from section 2.4 of [9] and chapter 9 of [13].

Definition 6.1. A **fibre bundle** is a geometric structure consisting of a smooth manifold P called the **total space**, with $\dim(P) = n+m$, which is *locally* the direct product of two smooth submanifolds: the m-dimensional **base space** M, and the n-dimensional **standard fibre** F. A submersion $\pi : P \to M$ defines the fibre at each point $p \in U_\alpha \subset M$, $F_p := \pi^{-1}(p)$, such that locally $P = U_\alpha \times F_p$.

To be precise, we can define **local trivialisation's** $(\phi_\alpha, U_\alpha)_\alpha$, $U_\alpha \subset M$ open, with $\phi_\alpha : \pi^{-1}(U_\alpha) \subset P \to U_\alpha \times F$. Locally these maps are trivial since locally $P = U_\alpha \times F$ hence the name. Now writing $\phi^{-1}_{\alpha,p}(f) := \phi^{-1}_\alpha(p, f)$, $p \in U_\alpha$, $f \in F$ allows us to define **transition functions**,

$$t_{\alpha\beta} := \phi_{\alpha,p} \circ \phi^{-1}_{\beta,p} : F \to F \tag{6.1}$$

acting on non-trivial intersections $U_\alpha \cap U_\beta$, which are smooth elements of the **structure group** G of the standard fibre. (A Lie group of transformations acting transitively on the F).

So what's this rather complicated definition got to do with our discussion in this report? Let's consider a fibre bundle P with submersion $\pi : P \to M$ such that $\forall$ $p \in M$, the fibre $\pi^{-1}(p)$ is a real vector space. *i.e.* elements of P can be added if the 'lie over' the same element of M: P is "vertically linear w.r.t M" but "horizontally

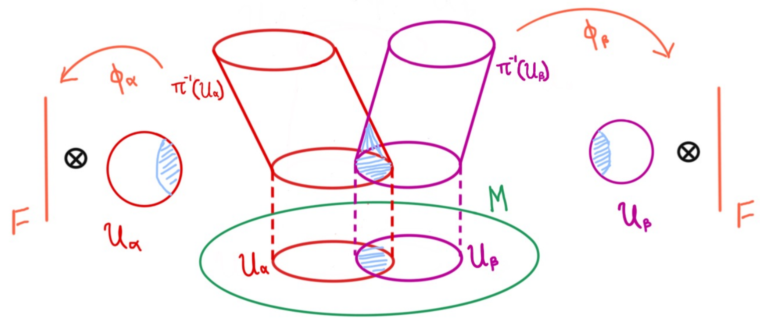

Figure 6.1: A fibre bundle is locally the direct product of a base space M and the standard fibre F. This structure is realised by local trivialisations $(U_\alpha, \phi_\alpha), (U_\beta, \phi_\beta)$.

non-linear". We can define a map $\psi : M \to P$ such that $\psi(p) \in \pi^{-1}(p)$. This is called a **cross section** and since for any given $p \in M$ $\psi(p)$ is an element of a vector space, the *function space* of all cross sections for any given point $p \in M$ defines an infinite dimensional vector space $\mathcal{H}(P)$. Now consider an element of the structure group of P, $g \in G$, mapping a given fibre $\pi^{-1}(p)$ to $\pi^{-1}(gp)$ by linear transform. Then we can define a linear transform $\rho : G \times \mathcal{H}(P) \to \mathcal{H}(P)$ by $\rho(g)(\psi)(p) = g^{-1}\psi(gp)$. So if the set of cross sections $\{\psi_1, ..., \psi_n\}$ forms a basis of $\pi^{-1}(p)$, and $g \in G$ we can write

$$\rho(g)(\psi_i)(p) = A_{ij}(g,p)\psi_i(p) \tag{6.2}$$

This is our first remarkable link: fields whose indices transform under some Lie group G, such as the tensor/spinor field representations of the Lorentz group, correspond to cross sections on the fibres of bundles with structure group G provided F is a vector space!

Let's push this further. Suppose G acts transitively on M. If H is the isotropy subgroup of some point $p_0 \in M$ then by theorem (5.8) $M \cong G/H$. A fibre bundle P with base manifold G/H s.t. the action of G on G/H is the usual group action on a coset space (see (5.5)) then P is called a **homogeneous vector bundle** (HVB). There are typically many HVBs to a given coset manifold G/H, so let's consider how we might determine them.

Let $V = \pi^{-1}(p_0)$. Since H is the isotropy group of p_0, each $h \in H$ acting on P maps V linearly to itself. *i.e.* $h \in H \implies h \in \mathrm{Aut}(V)$. But this is exactly the definition of a representation of H on V: (σ, V), $\sigma : V \to \mathrm{Aut}(V)$. But this goes further, it turns out there is a one-to-one correspondence between cross sections of the vector bundle and maps $\tilde{\psi} : G \to V$, satisfying $\tilde{\psi}(gh) = \sigma(h)\tilde{\psi}(g)$. To prove this, suppose $p = g^{-1}p_0$ is a point in G/H, $g \in G$. For each $\psi \in \mathcal{H}(P)$ we can define a map $\tilde{\psi} : G \to V$:

$$\tilde{\psi}(g) = g\psi(g^{-1}p_0) \tag{6.3}$$

Now for $h \in H$, $\tilde{\psi}(gh) = hg\psi(g^{-1}h^{-1}p_0) = hg\psi(g^{-1}p_0) = \sigma(h)\tilde{\psi}(g)$ by (6.3). Conversely, if $\tilde{\psi}$ is the map $G \to V$ transforming as $\tilde{\psi}(gh) = \sigma(h)\tilde{\psi}(g)$ under the right action of H

then we can assign to it a cross section $\psi \in \mathcal{H}$ by $\psi(g^{-1}p_0) = g^{-1}\tilde{\psi}(g)$. By (6.3) this does indeed transform as we specified for a cross section before (6.2). Thus there is indeed a one-to-one correspondence between the ψ and the $\tilde{\psi}$. Now using the functions $\tilde{\psi}$ we defined given the representation σ of H, the representation ρ of G on $\mathcal{H}(P)$ in (6.3) takes an interesting form. For $g_0, g \in G$ with $p = g^{-1}p_0$ still, we have:

$$\rho(g_0)(\psi)(g^{-1}p_0) = g_0^{-1}\psi(g_0g^{-1}p) = g^{-1}gg_0^{-1}\psi(g_0g^{-1}p) = g^{-1}\tilde{\psi}(g_0g^{-1}) \tag{6.4}$$

We can then define $\tilde{\rho}$ by $\tilde{\rho}(g_0)(\tilde{\psi})(g) := g\rho(g_0)(\psi)(g^{-1}p_0)$, thus

$$\tilde{\rho}(g_0)(\tilde{\psi})(g) = g\rho(g_0)(\psi)(g^{-1}p_0) = g[g^{-1}\tilde{\psi}(g_0g^{-1})] = \tilde{\psi}(g_0g^{-1}) \tag{6.5}$$

Now this should look somewhat reminiscent of the induced representations we saw in section 3.5.2! In fact, the necessity of the fibre bundle has dropped out at the end here and we see that given a representation of $H \subset G$, σ, on the vector space $\mathcal{H}(P)$ we can define the space of $\tilde{\psi}$ functions and subsequently induce a representation of G! This formally shows how the method of induced representations works, and its thanks to fibre bundles over homogeneous spaces! Or more explicitly, HVBs over base manifold $M \cong G/H$!

So before we even formally introduced homogeneous spaces in chapter 5, they were beneath the surface, generating the induced representations we encountered in chapters 3 and 4 through the magic of fibre bundles or, more specifically, homogeneous vector bundles! Wow!

6.2 Conclusion

So there it is: from the field representations of chapter 3, to the induced representations of chapters 3 and 4, and the homogeneous spaces of chapter 5; not only do fibre bundles reveal some remarkable machinery at play beneath the surface of the mathematics of symmetry and introduce some beautiful links between these concepts, but point in a direction which I would loved to have dedicated more time to exploring. They are at the core of developing Einstein gravity as a theory of gauge symmetry. This involves an object called the *vielbein*; a multiplet of *differential one-forms* on the base manifold. It takes values on the sections of fibres and, via homogeneous spaces, it provides a route through which colourful geometries can be built into theories of gravity. In the case of supersymmetry, these notions can be generalised to differential forms on superspace and the *supervielbein* (see chapters 12, 13 and 14 of [18]). In this formalism, a supersymmetric theory of gravity arises completely naturally from the *local* equivalent of the procedure we introduced in 5.1.2. A nice construction of supergravity arising from local superspace is given in chapter 8 of [19]. These theories of supergravity, built on the most advanced applications of some of the concepts we've introduced throughout this report, stand at the forefront of research into a whole array of problems in theoretical physics; from tackling the notorious information problem, to theories of Quantum Gravity, String Theory, and the ADS/CFT correspondence. Maybe one day supersymmetry will find itself at the core of the solution to one of the most profound problems in physics. Now that would be super.

Appendix A

The Structure of Lie Algebras

This material originally preceded the discussion of $SU(2)$ at the end of chapter 2, however it's too long for something that wasn't necessary beyond chapter 2 for material later in the report. However, I was reluctant to remove it entirely; most courses I've looked at discussing the structure of semi-simple Lie algebras seem to spring techniques out of nowhere with little motivation. I think it's enlightening to understand how the notions of root vectors etc arose directly from the adjoint representation. That's not to say this material necessarily does this well, especially since it was removed before the editing stage was complete and thus hasn't received any treatment to improve it from it's first draft, but it was the product of my self-study and I enjoyed reading about it. It is based on material from [8] and [11].

In this appendix I will attempt to give a brief intuitive and historical introduction to the importance of the **Cartan-Weyl basis** of a semi-simple Lie algebra. In the late 19th century, inspired by notions of geometry introduced first by Felix Klein, the mathematician and physicist Wilhelm Killing endeavoured to determine all possible *space forms* - very loosely speaking, all possible "types" of space as categorised by their specific geometric properties. Inspired by Weierstrass, Killing's treatment was highly analytic and used the idea of "motions" generated by infinitesimal transformations. He began by suggesting that finite motions need not commute; that is, the conjugation action of the group of finite transformations on itself is non-trivial.

Definition A.1. The **conjugate,** or **adjoint, action** of a group G is defined to be

$$G \times G \to G, \quad (g, h) \mapsto g \cdot h \cdot g^{-1}$$

Remark. We can see how commutation relations arise from this. Suppose g and h have infinitesimal forms $g \approx (e + \omega)$, $h \approx (e + \epsilon)$. The for $h_1 \neq h_2$ we have

$$gh_1g^{-1} = h_2 \implies (e + \omega)(e + \epsilon_1)(e - \omega) = (e + \epsilon_2) \implies \epsilon_1 + [\omega, \epsilon_1] = \epsilon_2$$

Clearly $h_1 = h_2 \implies \epsilon_1 = \epsilon_2 \implies [\omega, \epsilon_1] = 0$ as we would expect.

In a process similar to that which we followed in section 2.1.4 he produced a set of commutator relations these generators must satisfy, including asserting antisymmetry and reproducing the Jacobi identity, and thus his task to determine all such forms was reduced to the classification of Lie algebras.

Since Killing's goal was to determine space forms, once he'd classified the algebras he needed to be able to integrate the system of equations $[X_i, X_j] = \sum_{k=1}^{n} C_{ij}^{k} X_k$. He very understandably observed that this would be a lot simpler if most of the structure constants were zero, *i.e.* $[X_i, X_j] = cX_k$, for c some real constant. This led him to investigating, for a given $X \in \mathfrak{g}$ ($\mathfrak{g}$ a Lie algebra), solutions $Y \in \mathfrak{g}$ to the problem:

$$[X, Y] = cY, \quad \Longleftrightarrow \quad \mathrm{ad}_X(Y) = cY \tag{A.1}$$

As we clearly see, this is an eigenvalue problem on the adjoint representation of the Lie algebra. It's easy to see why the adjoint representation is useful here: it maps an element $X \in \mathfrak{g}$ to a linear operator (automorphism) $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$. The representation space is the Lie algebra itself - we can explore it's structure via it's action on itself! The characteristic polynomial naturally became the central object of Killing's investigation since solutions of (A.1) are only possible if c is one of it's **roots**:

$$\det(\mathrm{ad}_X - c\mathbb{1}) = (-1)^n [c^n - \psi_1(X)c^{n-1} + \cdots \pm \psi_{n-m}(X)c^m] \tag{A.2}$$

Note that in the above polynomial the root $c = 0$ has multiplicity m. Since $\mathrm{ad}_X(X) = [X, X] = 0$ we know this is at least 1. The multiplicity m of $c = 0$ gives the dimension of an abelian subalgebra $\{Y \in \mathfrak{g} | [X, Y] = 0\}$. This subalgebra has a special name:

Definition A.2. For a Lie algebra $\mathfrak{g}$, the **Cartan subalgebra**[1] (CSA) $\mathfrak{h} \subset \mathfrak{g}$ is a maximal, abelian subalgebra containing **ad-diagonalizable** elements *i.e.* elements whose adjoint representation is diagonalizable.

Remark. What do we mean by maximal? This means it contains all such elements *i.e.* $H \in \mathfrak{h}$, $X \in \mathfrak{g}$ s.t. $[H, X] = 0 \implies X \in \mathfrak{h}$. To guarantee this, Killing chose X in (A.2) such that m was minimal *i.e.* element with least multiplicity of $c = 0$. This makes sure we have that $H_1, H_2 \in \mathfrak{h} \implies [H_1, H_2] = 0$. The minimal zero-multiplicity element, in our case X, is called **regular**. If we chose an element H_p with zero-multiplicity $p > m$ there would exist $p - m$ elements $X_i \in \mathfrak{h}$ s.t $[H_p, X_i] = 0$ but $[X_i, X] \neq 0$ despite $X \in \mathfrak{h}$ thus the CSA would not be closed. Clearly if $\mathfrak{g}$ is simple then the CSA is $\mathfrak{h} \equiv \{0\}$.

Taking X in (A.2) to be the regular element we now turn to the elements of $\mathfrak{g}$ with non-zero roots (eigenvalues) *i.e.* $[X, E_\alpha] = \alpha E_\alpha$, $\alpha \neq 0$. First of all, suppose we choose an ad-diagonal basis for the CSA, $\{H_i\}_i$, such that $[H_i, E_\alpha] = \alpha_i E_\alpha$. Then for any $H = \sum_i a_i H_i \in \mathfrak{h}$ we have that:

$$[H, E_\alpha] = \Big[\sum_i a_i H_i, E_\alpha\Big] = \sum_i a_i [H_i, E_\alpha] = \sum_i a_i \alpha_i E_\alpha \implies [H, E_\alpha] = \alpha(H) E_\alpha$$

[1] Despite being named after Elié Cartan, it was really Killing who first introduce the CSA.

i.e. the roots are linear functionals on the CSA; they live in the **dual space** $\alpha \in \mathfrak{h}^*$. The set of all roots of $\mathfrak{g}$ is called the **root set** denoted Φ. Note that as roots of a the polynomial (A.2) they are necessarily complex numbers, $\alpha \in \mathbb{C}$. This is a very important detail; our basis of root vectors $\{E_\alpha \in \mathfrak{g} | \alpha \in \Phi\}$ exists in the *complexification* of $\mathfrak{g}$: $\mathfrak{g}_\mathbb{C}$.

Now these **root vectors**, $\{E_\alpha\}$, have some very interesting behaviour: suppose E_α, E_β are root vectors, $\alpha, \beta \in \Phi$. Employing the Jacobi identity we obtain:

$$\begin{aligned}[H,[E_\alpha,E_\beta]] &= -[E_\alpha,[E_\beta,H]] - [E_\beta,[H,E_\alpha]] \\ &= -[E_\alpha, -\beta(H)E_\beta] + [\alpha(H)E_\alpha, E_\beta] \\ &= (\alpha(H)+\beta(H))[E_\alpha,E_\beta]\end{aligned} \tag{A.3}$$

So either $[E_\alpha, E_\beta] \equiv 0$, or $\alpha(H) + \beta(H) = 0$, or $[E_\alpha, E_\beta] \propto E_{\alpha+\beta}$. To complete this development we move away from the reasoning of Killing. His work, though brilliant, had it's limitations; despite proving all coefficients ψ_i of (A.2) were invariant under the adjoint group, he had focused his analysis on the $n-1$ coefficient, the linear polynomial ψ_1. It was Elié Cartan who advanced the study with the use of the $n-2$ coefficient, the quadratic form ψ_2. This brings us to our next definition.

Definition A.3. The **Killing Form**[2] of a Lie algebra $\mathfrak{g}$ over a field $\mathbb{F}$ is the bilinear, quadratic form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$ given by:

$$\kappa(X,Y) = \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y)$$

It is **non-degenerate** if $\forall\ X \in \mathfrak{g}\ \exists\ Y \in \mathfrak{g}$ such that $\kappa(X,Y) \neq 0$.

Remark. The Killing form defines a pseudo-inner product on a Lie algebra *i.e.* it is symmetric and bilinear but positivity is not guaranteed. This is key to obtaining a more geometrical picture of the algebras structure. For κ, invariance corresponds to associativity: $\kappa(X, \mathrm{ad}_Y(Z)) = \kappa(X,[Y,Z]) = \kappa([X,Y],Z) = \kappa(\mathrm{ad}_X(Y), Z)$.

The Killing form gives a natural connection between the CSA and it's dual. Let $\{H_i\}$ be a basis for the CSA s.t. $\mathrm{ad}_{H_i} = \mathrm{diag}(0,..0,\alpha_1(H_i),...)$ and $[H_i, E_\alpha] = \alpha(H_i)E_\alpha$. Then

$$\begin{aligned}\kappa(H_i,H_j) = \mathrm{Tr}(\mathrm{ad}_{H_i} \circ \mathrm{ad}_{H_j}) &= \mathrm{Tr}(\mathrm{diag}(0,...,\alpha_1(H_i)\alpha_1(H_j),...)) \\ &= \sum_{\alpha\in\Phi} \alpha(H_i)\alpha(H_j)\end{aligned} \tag{A.4}$$

Now it was with the Killing form that Cartan produced one of his most powerful theorems:

Theorem A.4 (Cartan's Theorem)**.** A Lie algebra is semi-simple iff the Killing form is non-degenerate.

[2] Here Killing has his revenge for the CSA misnomer; despite being known as the Killing form it was of little interest to Killing. It was Cartan who introduced it formally and made proper use of it.

To conclude the discussion of $[E_\alpha, E_\beta]$ and give rigorous structure to semi-simple Lie algebras we prove the following:

Lemma A.5. Let κ be non-degenerate. Then

1. $\kappa(H, E_\alpha) = 0 \;\forall\; H \in \mathfrak{h},\ \alpha \in \Phi$.

2. $\kappa(E_\alpha, E_\beta) = 0 \;\forall\; \alpha, \beta \in \Phi$ s.t. $\alpha \neq -\beta$.

3. $\forall\; H \in \mathfrak{h} \;\exists\; X \in \mathfrak{h}$ s.t. $\kappa(H, X) \neq 0$.

4. $\alpha \in \Phi \implies -\alpha \in \Phi$ and $\kappa(E_\alpha, E_{-\alpha})$

Proof. For all $X \in \mathfrak{h}$, $\alpha(X) \neq 0$, we have that $\alpha(X)\kappa(H, E_\alpha) = \kappa(H, [X, E_\alpha]) = -\kappa([H, X], E_\alpha) = 0$ (since $[H, X] = 0$) $\implies \kappa(H, E_\alpha) = 0$ since $\alpha(X) \neq 0$. This proves 1. Let $\alpha, \beta \in \Phi$. $(\alpha(X) + \beta(X))\kappa(E_\alpha, E_\beta) = \kappa([X, E_\alpha], E_\beta) + \kappa(E_\alpha, [X, E_\beta]) = 0$ by invariance of κ. So $\alpha(X) + \beta(X) \neq 0$ means $\kappa(E_\alpha, E_\beta) = 0$. This proves 2. Now suppose for $H \in \mathfrak{h}$ we have $\kappa(H, X) = 0 \;\forall\; X \in \mathfrak{h}$ then 1 implies $\kappa(H, E_\alpha) = 0$ which in turn means $\kappa(H, Z) = 0 \;\forall\; Z \in \mathfrak{g}$ but this means κ is degenerate. This contradiction proves 3. Finally, 1 and 2 imply κ is degenerate unless $-\alpha$ is a root and $\kappa(E_\alpha, E_{-\alpha}) \neq 0$ thus κ non-degenerate implies the result. $\square$

We now split our discussion into cases:

1. If $\alpha(H) + \beta(H) \notin \Phi$ and $\alpha(H) \neq -\beta(H)$ then $[E_\alpha, E_\beta] \equiv 0$.

2. If $\alpha(H) + \beta(H) = 0$ ($\beta \equiv -\alpha$) then by invariance of the Killing form we have $\kappa([E_\alpha, E_{-\alpha}], H) = \kappa(E_\alpha, [E_{-\alpha}, H]) = \kappa(E_\alpha, \alpha(H)E_{-\alpha}) = \alpha(H)\kappa(E_\alpha, E_{-\alpha}) \neq 0$ by the lemma. Now by (A.4) this proves $[E_\alpha, E_{-\alpha}] \in \mathfrak{h}$. We define $H_\alpha := [E_\alpha, E_{-\alpha}]/\kappa(E_\alpha, E_{-\alpha})$.

3. If $0 \neq \alpha(H) + \beta(H) \in \Phi$ then $[E_\alpha, E_\beta] = N_{\alpha\beta}E_{\alpha+\beta}$ where $N_{\alpha\beta}$ is a normalization constant.

With this we are done, we conclude with the following definition:

Definition A.6. The **Cartan-Weyl Basis** of a semi-simple Lie algebra is $\{\text{CSA}\} \cup \{E_\alpha | \alpha \in \Phi\}$. For a given $\alpha \in \Phi$ it satisfies the following:

$$\begin{aligned}
[H_\alpha, E_\alpha] &= \alpha(H_\alpha)E_\alpha \\
[H_\alpha, E_{-\alpha}] &= -\alpha(H_\alpha)E_\alpha \\
[E_\alpha, E_{-\alpha}] &= H_\alpha
\end{aligned} \tag{A.5}$$

Appendix B

The Poincare Algebra

In this appendix, we give a full detailed derivation of the Poincaré algebra. We begin by considering the conjugation action of a general element (Λ, a) on the infinitesimal group element $(\mathbb{1} + \omega, \epsilon)$:

$$(\Lambda, a)(\mathbb{1} + \omega, \epsilon)(\Lambda, a)^{-1} = (\Lambda, a)(\mathbb{1} + \omega, \epsilon)(\Lambda^{-1}, -\Lambda^{-1}a) = (\mathbb{1} + \Lambda\omega\Lambda^{-1}, -\Lambda\omega\Lambda^{-1}a + \Lambda\epsilon)$$

Expanding both sides in terms of the generators gives:

$$\begin{aligned}(\Lambda, a)(\mathbb{1} + \frac{i}{2}\omega_{\rho\sigma}J^{\rho\sigma} - i\epsilon_\rho P^\rho)(\Lambda, a)^{-1} &= \mathbb{1} + \frac{i}{2}(\Lambda\omega\Lambda^{-1})_{\mu\nu}J^{\mu\nu} - i(\Lambda\epsilon - \Lambda\omega\Lambda^{-1}a)_\mu P^\mu \\ &= \mathbb{1} + \frac{i}{2}\Lambda_\mu{}^\rho\omega_{\rho\sigma}(\Lambda^{-1})^\sigma{}_\nu J^{\mu\nu} - i(\Lambda_\mu{}^\rho\epsilon_\rho - \Lambda_\mu{}^\rho\omega_{\rho\sigma}(\Lambda^{-1})^\sigma{}_\nu a^\nu)P^\mu\end{aligned}$$

Which, using $(\Lambda^{-1})^\sigma{}_\nu = \Lambda_\nu{}^\sigma$ and equating coefficients in ω and ϵ, produces the relations:

$$\frac{1}{2}(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} = \Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma(\frac{1}{2}J^{\mu\nu} + a^\nu P^\mu) \tag{B.1}$$

$$(\Lambda, a)P^\rho(\Lambda, a)^{-1} = \Lambda_\mu{}^\rho P^\mu \tag{B.2}$$

Now (B.2) is fine already, however (B.1) isn't quite right; $J^{\rho\sigma}$ must be anti-symmetric in the indices ρ, σ so we require:

$$(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} = -(\Lambda, a)J^{\sigma\rho}(\Lambda, a)^{-1}$$

By the same procedure as before we obtain:

$$\begin{aligned}\frac{1}{2}(\Lambda, a)J^{\sigma\rho}(\Lambda, a)^{-1} &= \Lambda_\mu{}^\sigma\Lambda_\nu{}^\rho(\frac{1}{2}J^{\mu\nu} + a^\nu P^\mu) \\ &= \Lambda_\nu{}^\sigma\Lambda_\mu{}^\rho(\frac{1}{2}J^{\nu\mu} + a^\mu P^\nu) \\ &= \Lambda_\nu{}^\sigma\Lambda_\mu{}^\rho(-\frac{1}{2}J^{\mu\nu} + a^\mu P^\nu)\end{aligned}$$

Now we can correct the relation:

$$\begin{aligned}
(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} &= \frac{1}{2}(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} + \frac{1}{2}(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} \\
&= \frac{1}{2}(\Lambda, a)J^{\rho\sigma}(\Lambda, a)^{-1} - \frac{1}{2}(\Lambda, a)J^{\sigma\rho}(\Lambda, a)^{-1} \\
&= \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma(\frac{1}{2}J^{\mu\nu} + a^\nu P^\mu) - \Lambda_\nu{}^\sigma \Lambda_\mu{}^\rho(-\frac{1}{2}J^{\mu\nu} + a^\mu P^\nu) \\
&= \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma(J^{\mu\nu} + a^\nu P^\mu - a^\mu P^\nu)
\end{aligned}$$

Now these relations tell us how the Poincaré group acts on it's own generators. We can now expand the group elements in the relations to discover how the generators interact amongst themselves! Neglecting orders $\mathcal{O}(\omega^n \epsilon^m)$ s.t. $m + n \geq 2$ (hence our ability to be slightly cheeky with our labelling of indices) we have:

$$\begin{aligned}
(\Lambda, a)J^{\rho\sigma}(\Lambda^{-1}, -\Lambda^{-1}a) &\approx (\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} - i\epsilon_\mu P^\mu)J^{\rho\sigma}(\mathbb{1} + \frac{i}{2}\omega_{\nu\mu}J^{\nu\mu} + i\epsilon_\mu P^\mu) \\
&\approx J^{\rho\sigma} + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}J^{\rho\sigma} + \frac{i}{2}\omega_{\nu\mu}J^{\rho\sigma}J^{\nu\mu} - i\epsilon_\mu P^\mu J^{\rho\sigma} + iJ^{\rho\sigma}\epsilon_\mu P^\mu
\end{aligned}$$

But on the LHS we have:

$$\begin{aligned}
\Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma(J^{\mu\nu} + \epsilon^\nu P^\mu - \epsilon^\mu P^\nu) &\approx (\delta_\mu{}^\rho + \omega_\mu{}^\rho)(\delta_\nu{}^\sigma + \omega_\nu{}^\sigma)(J^{\mu\nu} + \epsilon^\nu P^\mu - \epsilon^\mu P^\nu) \\
&\approx (\delta_\mu{}^\rho \delta_\nu{}^\sigma + \delta_\mu{}^\rho \omega_\nu{}^\sigma + \delta_\nu{}^\sigma \omega_\mu{}^\rho)(J^{\mu\nu} + \epsilon^\nu P^\mu - \epsilon^\mu P^\nu) \\
&= J^{\rho\sigma} + J^{\rho\nu}\omega_\nu{}^\sigma + J^{\mu\sigma}\omega_\mu{}^\rho - \epsilon^\rho P^\sigma + \epsilon^\sigma P^\rho
\end{aligned}$$

Now using $(\Lambda, a)J^{\rho\sigma}(\Lambda^{-1}, -\Lambda^{-1}a) = \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma(J^{\mu\nu} + \epsilon^\nu P^\mu - \epsilon^\mu P^\nu)$ gives:

$$\begin{aligned}
&\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}J^{\rho\sigma} + \frac{i}{2}\omega_{\nu\mu}J^{\rho\sigma}J^{\nu\mu} - i\epsilon_\mu P^\mu J^{\rho\sigma} + iJ^{\rho\sigma}\epsilon_\mu P^\mu = J^{\rho\nu}\omega_\nu{}^\sigma + J^{\mu\sigma}\omega_\mu{}^\rho - \epsilon^\rho P^\sigma + \epsilon^\sigma P^\rho \\
&\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}J^{\rho\sigma} - \frac{i}{2}\omega_{\mu\nu}J^{\rho\sigma}J^{\mu\nu} - i\epsilon_\mu[P^\mu, J^{\rho\sigma}] = \eta^{\sigma\mu}J^{\rho\nu}\omega_{\nu\mu} + \eta^{\rho\nu}J^{\mu\sigma}\omega_{\mu\nu} - \eta^{\rho\mu}\epsilon_\mu P^\sigma + \eta^{\sigma\mu}\epsilon_\mu P^\rho \\
&\frac{i}{2}\omega_{\mu\nu}[J^{\mu\nu}, J^{\rho\sigma}] - i\epsilon_\mu[P^\mu, J^{\rho\sigma}] = \omega_{\mu\nu}(\eta^{\rho\nu}J^{\mu\sigma} - \eta^{\sigma\mu}J^{\rho\nu}) - \epsilon_\mu(\eta^{\rho\mu}P^\sigma - \eta^{\sigma\mu}P^\rho)
\end{aligned}$$

$$\frac{i}{2}[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\rho\nu}J^{\mu\sigma} - \eta^{\sigma\mu}J^{\rho\nu}, \quad i[P^\mu, J^{\rho\sigma}] = \eta^{\rho\mu}P^\sigma - \eta^{\sigma\mu}P^\rho \tag{B.3}$$

Note that the expression for $[J^{\rho\sigma}, P^\mu]$ is anti-symmetric under $\rho \leftrightarrow \sigma$ however the expression for $[J^{\mu\nu}, J^{\rho\sigma}]$ is neither anti-symmetric under $\rho \leftrightarrow \sigma$ or $\mu \leftrightarrow \nu$. We solve this in a similar fashion to (B.1).

$$\begin{aligned}
&[J^{\mu\nu}, J^{\rho\sigma}] = [-J^{\nu\mu}, J^{\rho\sigma}] = -[J^{\nu\mu}, J^{\rho\sigma}] \\
\implies &\frac{i}{2}[J^{\nu\mu}, J^{\rho\sigma}] = \eta^{\rho\mu}J^{\nu\sigma} - \eta^{\sigma\nu}J^{\rho\mu} \quad \text{by (??).} \\
\implies &i[J^{\mu\nu}, J^{\rho\sigma}] = \frac{i}{2}[J^{\mu\nu}, J^{\rho\sigma}] + \frac{i}{2}[J^{\mu\nu}, J^{\rho\sigma}] \\
&= \frac{i}{2}[J^{\mu\nu}, J^{\rho\sigma}] - \frac{i}{2}[J^{\nu\mu}, J^{\rho\sigma}] \\
&= \eta^{\rho\nu}J^{\mu\sigma} - \eta^{\sigma\mu}J^{\rho\nu} - \eta^{\rho\mu}J^{\nu\sigma} + \eta^{\sigma\nu}J^{\rho\mu}
\end{aligned}$$

This now has the correct symmetry. Finally, expanding the elements in (**??**) in the same manner and equating coefficients in ϵ we obtain:

$$[P^\mu, P^\nu] = 0 \tag{B.4}$$

With this we are done: together, these relations form the **Poincaré Algebra**:

$$\begin{aligned} i[J^{\mu\nu}, J^{\rho\sigma}] &= \eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} - \eta^{\sigma\mu} J^{\rho\nu} + \eta^{\sigma\nu} J^{\rho\mu} \\ i[P^\mu, J^{\rho\sigma}] &= \eta^{\mu\rho} P^\sigma - \eta^{\mu\sigma} P^\rho \\ [P^\mu, P^\nu] &= 0 \end{aligned} \tag{B.5}$$

B.0.1 Discrete Symmetry

Now as we mentioned, the connected components of the Lorentz group are related the discrete symmetry transformations $\mathcal{P}$ and $\mathcal{T}$, **parity** and **time-reversal** respectively. They take the matrix form:

$$\mathcal{P}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \mathcal{T}^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.6}$$

The action of these on the group elements of $ISO(1,3)$ is given in the usual way by $\boldsymbol{P}(\Lambda, a)\boldsymbol{P}^{-1} = (\mathcal{P}\Lambda\mathcal{P}^{-1}, \mathcal{P}a)$ and $\boldsymbol{T}(\Lambda, a)\boldsymbol{T}^{-1} = (\mathcal{T}\Lambda\mathcal{T}^{-1}, \mathcal{T}a)$. We can expand these in identical manner to before to give:

$$\begin{aligned} \boldsymbol{P} iJ^{\rho\sigma}\boldsymbol{P}^{-1} &= i\mathcal{P}_\mu{}^\rho \mathcal{P}_\nu{}^\sigma J^{\mu\nu}, \quad \boldsymbol{P} iP^\rho \boldsymbol{P}^{-1} = i\mathcal{P}_\mu{}^\rho P^\mu \\ \boldsymbol{T} iJ^{\rho\sigma}\boldsymbol{T}^{-1} &= i\mathcal{T}_\mu{}^\rho \mathcal{T}_\nu{}^\sigma J^{\mu\nu}, \quad \boldsymbol{T} iP^\rho \boldsymbol{T}^{-1} = i\mathcal{T}_\mu{}^\rho P^\mu \end{aligned} \tag{B.7}$$

Now setting $\rho = 0$ and using (B.6) tells us $\boldsymbol{P} iH\boldsymbol{P}^{-1} = iH$ and $\boldsymbol{T} iH\boldsymbol{T}^{-1} = -iH$. In order that our theories yield only positive energy solutions we require $H > 0$, thus we must have that $\boldsymbol{P}$ is a linear operator such that $\boldsymbol{P}(cV) = c\boldsymbol{P}(V)$ for $c \in \mathbb{C}$, V arbitrary. Similarly, $\boldsymbol{T}$ must be anti-linear such that $\boldsymbol{T}(cV) = c^*\boldsymbol{T}(V)$ for $c \in \mathbb{C}$, where the star denotes the complex conjugate. This assures $\boldsymbol{P}H\boldsymbol{P}^{-1} = \boldsymbol{T}H\boldsymbol{T}^{-1} = H$. With these properties in mind, using we obtain the following relations:

$$\begin{aligned} \boldsymbol{P} L_i \boldsymbol{P}^{-1} &= +L_i, \quad \boldsymbol{P} K_i \boldsymbol{P}^{-1} = -K_i, \quad \boldsymbol{P} P_i \boldsymbol{P}^{-1} = -P_i \\ \boldsymbol{T} L_i \boldsymbol{T}^{-1} &= -L_i, \quad \boldsymbol{T} K_i \boldsymbol{T}^{-1} = +K_i, \quad \boldsymbol{T} P_i \boldsymbol{T}^{-1} = -P_i \end{aligned} \tag{B.8}$$

Bibliography

[1] M. Ammon and J. Erdmenger. *Gauge/Gravity Duality.* Cambridge University Press, Cambridge CB2 1RP, 2015.

[2] B. Bakker and D. Boer. Group theory in physics: Part ii, lie groups lie algebras, 2010. Accessed: 2020-12-11.

[3] M. Bertolini. Lectures on supersymmetry. `https://people.sissa.it/~bertmat/susycourse.pdf`, 2021. Accessed: 2021-03-11.

[4] R. N. Cahn. Semi-simple lie algebras and their representations. `http://phyweb.lbl.gov/~rncahn/www/liealgebras/texall.pdf`, 1984. Accessed: 2020-10-22.

[5] S. Coleman and J. Mandula. All possible symmetries of the s-matrix. *Physical Review*, 159(5):1251–1256, 1967.

[6] J. F. Cornwell. *Group Theory in Physics, Volume 2.* Academic Press Inc., London, 1984.

[7] J. F. Cornwell. *Group Theory in Physics, Volume 3.* Academic Press Inc., London, 1989.

[8] N. Dorey. Symmetries, fields and particles. `https://kirklin.science/notes/sfp.pdf`, 2015. Accessed: 2020-10-22.

[9] P. G. Fré. *Gravity,a Geometrical Course: Volume 1.* Springer Science+Business Media,LLC, London, 2013.

[10] P. G. Fré. *Gravity,a Geometrical Course: Volume 2.* Springer Science+Business Media,LLC, London, 2013.

[11] T. Hawkins. *Emergence of the Theory of Lie Groups An Essay in the History of Mathematics 1869–1926.* Springer Science+Business Media,LLC, New York, 2000.

[12] H. E. H. Herbi K. Dreiner and S. P. Martin. Two-component spinor techniques and feynman rules for quantum field theory and supersymmetry. `https://arxiv.org/abs/0812.1594v5`, 2010. Accessed: 2021-04-26.

[13] R. Hermann. *Lie Groups for Physicists.* W. A. Benjamin Inc., New York, 1966.

[14] G. W. Mackey. On induced representations of groups. *American Journal of Mathematics*, 73(3):576–592, 1951.

[15] P. Ramond. *Field Theory: A Modern Primer.* Westview Press, Boulder,Colorado,US, 1981.

[16] D. Tong. Quantum field theory. `http://www.damtp.cam.ac.uk/user/tong/qft.html`, 1984. Accessed: 2020-10-26.

[17] S. Weinberg. *The Quantum Theory of Fields: Volume 1 Foundations.* Cambridge University Press, Cambridge CB2 1RP, 1995.

[18] J. Wess and J. Bagger. *Supersymmetry and Supergravity.* Princeton University Press, Princeton, New Jersey, US, 1992.

[19] M. Zikidis. Homogeneous spaces. `https://www.mathi.uni-heidelberg.de/~lee/MenelaosSS16.pdf`, 2021. Accessed: 2021-02-17.