Symmetries and Groups
Michaelmas Term 2008

Hugh Osborn

latest update: 30th November 2010

Books

Books developing group theory by physicists from the perspective of particle physics are

H. F. Jones, *Groups, Representations and Physics,* 2nd ed., IOP Publishing (1998).
A fairly easy going introduction.
H. Georgi, *Lie Algebras in Particle Physics,* Perseus Books (1999).
Describes the basics of Lie algebras for classical groups.
J. Fuchs and C. Schweigert, *Symmetries, Lie Algebras and Representations,* 2nd ed., CUP (2003).
This is more comprehensive and more mathematically sophisticated, and does not describe physical applications in any detail.
Z-Q. Ma, *Group Theory for Physicists,* World Scientific (2007).
Quite comprehensive.

The following books contain useful discussions, in chapter 2 of Weinberg there is a proof of Wigner's theorem and a discussion of the Poincaré group and its role in field theory, and chapter 1 of Buchbinder and Kuzenko has an extensive treatment of spinors in four dimensions.
S. Weinberg, *The Quantum Theory of Fields,* (vol. 1), CUP (2005).
J. Buchbinder and S. Kuzenko, *Ideas and Methods of Supersymmetry and Supergravity, or a Walk Through Superspace,* 2nd ed., Institute of Physics Publishing (1998).

They are many mathematical books with titles containing reference to Groups, Representations, Lie Groups and Lie Algebras. The motivations and language is often very different, and hard to follow, for those with a traditional theoretical physics background. Particular books which may be useful are
B.C. Hall, *Lie Groups, Lie Algebras, and Representations,* Springer (2004), for an earlier version see arXiv:math-ph/0005032.
This focuses on matrix groups.
More accessible than most
W. Fulton and J. Harris, *Representation Theory,* Springer (1991).

Historically the following book, first published in German in 1931, was influential in showing the relevance of group theory to atomic physics in the early days of quantum mechanics. For an English translation
E.P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra,* Academic Press (1959).

Prologue

The following excerpts are from Strange Beauty, by G. Johnson, a biography of Murray Gell-Mann[1], the foremost particle physicist of the 1950's and 1960's who proposed $SU(3)$ as a symmetry group for hadrons and later quarks as the fundamental building blocks. It reflects a time when most theoretical particle physicists were unfamiliar with groups beyond the rotation group, and perhaps also a propensity for some to invent mathematics as they went along.

As it happened, $SU(2)$ could also be used to describe the isospin symmetry- the group of abstract ways in which a nucleon can be "rotated" in isospin space to get a neutron or a proton, or a pion to get negative, positive or neutral versions. These rotations were what Gell-Mann had been calling currents. The groups were what he had been calling algebras.

He couldn't believe how much time he had wasted. He had been struggling in the dark while all these algebras, these groups- these possible classification schemes- had been studied and tabulated decades ago. All he would have to do was to go to the library and look them up.

In Paris, as Murray struggled to expand the algebra of the isospin doublet, $SU(2)$, to embrace all hadrons, he had been playing with a hierarchy of more complex groups, with four, five, six, seven rotations. He now realized that they had been simply combinations of the simpler groups $U(1)$ and $SU(2)$. No wonder they hadn't led to any interesting new revelations. What he needed was a new, higher symmetry with novel properties. The next one in Cartan's catalogue was $SU(3)$, a group that can have eight operators.

Because of the cumbersome way he had been doing the calculations in Paris, Murray had lost the will to try an algebra so complex and inclusive. He had gone all the way up to seven and stopped.

[1]Murray Gell-Mann, 1929-, American, Nobel prize 1969.

Contents

0 Notational Conventions

Hopefully we use standard conventions. For any M_{ij}, i belonging to an ordered set with m elements, not necessarily $1, \dots m$, and similarly j belonging to an ordered set with n elements, $M = [M_{ij}]$ is the corresponding $m \times n$ matrix, with of course i labelling the rows, j the columns. I is the unit matrix, on occasion I_n denotes the $n \times n$ unit matrix.

For any multi-index $T_{i_1 \dots i_n}$ then $T_{(i_1 \dots i_n)}$, $T_{[i_1 \dots i_n]}$ denote the symmetric, antisymmetric parts, obtained by summing over all permutations of the order of the indices in $T_{i_1 \dots i_n}$, with an additional -1 for odd permutations in the antisymmetric case, and then dividing by $n!$. Thus for $n = 2$,

$$T_{(ij)} = \tfrac{1}{2}(T_{ij} + T_{ji}) \, , \qquad T_{[ij]} = \tfrac{1}{2}(T_{ij} - T_{ji}) \, . \tag{0.1}$$

We use μ, ν, σ, ρ as space-time indices, i, j, k are spatial indices while α, β, γ are spinorial indices.

For a set of elements x then $\{x : P\}$ denotes the subset satisfying a property P.

A vector space $\mathcal{V}$ may be defined in terms of linear combinations of basis vectors $\{v_r\}$, $r = 1, \dots, \dim \mathcal{V}$ so that an arbitrary vector can be expressed as $\sum_r a_r v_r$. For two vector spaces $\mathcal{V}_1$, $\mathcal{V}_2$ with bases $\{v_{1r}\}$, $\{v_{2s}\}$ we may define the tensor product space $\mathcal{V}_1 \otimes \mathcal{V}_2$ in terms of the basis of pairs of vectors $\{v_{1r} v_{2s}\}$ for all r, s. An arbitrary vector in $\mathcal{V}_1 \otimes \mathcal{V}_2$ is a linear combination $v = \sum_{r,s} a_{rs} \, v_{1r} v_{2s}$ and $\dim(\mathcal{V}_1 \otimes \mathcal{V}_2) = \dim \mathcal{V}_1 \, \dim \mathcal{V}_2$. The direct sum $\mathcal{V}_1 \oplus \mathcal{V}_2$ is defined so that if $v \in \mathcal{V}_1 \oplus \mathcal{V}_2$ then $v = v_1 + v_2$ with $v_i \in \mathcal{V}_i$. It has a basis $\{v_{1r}, v_{2s}\}$ and $\dim(\mathcal{V}_1 \oplus \mathcal{V}_2) = \dim \mathcal{V}_1 + \dim \mathcal{V}_2$.

1 Introduction

There are nowadays very few papers in theoretical particle physics which do no not mention groups or Lie algebras and correspondingly make use of the mathematical language and notation of group theory, and in particular of that for Lie groups. Groups are relevant whenever there is a symmetry of a physical system, symmetry transformations correspond to elements of a group and the combination of one symmetry transformation followed by another corresponds to group multiplication. Associated with any group there are sets of matrices which are in one to one correspondence with each element of the group and which obey the same the same multiplication rules. Such a set a of matrices is called a representation of the group. An important mathematical problem is to find or classify all groups within certain classes and then to find all possible representations. How this is achieved for Lie groups will be outlined in these lectures although the emphasis will be on simple cases. Although group theory can be considered in the abstract, in theoretical physics finding and using particular matrix representations are very often the critical issue. In fact large numbers of groups are defined in terms of particular classes of matrices.

Group theoretical notions are relevant in all areas of theoretical physics but they are particularly important when quantum mechanics is involved. In quantum theory physical systems are associated with vectors belonging to a vector space and symmetry transformations of the system are associated with linear transformations of the vector space. With a choice of basis these correspond to matrices so that directly we may see why group representations are so crucial. Historically group theory as an area of mathematics particularly relevant in theoretical physics first came to the fore in the 1930's directly because of its applications in quantum mechanics (or matrix mechanics as the Heisenberg formulation was then sometimes referred to). At that time the symmetry group of most relevance was that for rotations in three dimensional space, the associated representations, which are associated with the quantum mechanical treatment of angular momentum, were used to classify atomic energy levels. The history of nuclear and particle physics is very much a quest to find symmetry groups. Initially the aim was to find a way of classifying particles with nearly the same mass and initially involved isospin symmetry. This was later generalised to the symmetry group $SU(3)$, the eightfold way, and famously led to the prediction of a new particle the Ω^-. The representations of $SU(3)$ are naturally interpreted in terms of more fundamental particles the quarks which are now the basis of our understanding of particle physics.

Apart from symmetries describing observed particles, group theory is of fundamental importance in gauge theories. All field theories which play a role in high energy physics are gauge field theories which are each associated with a particular gauge group. Gauge groups are Lie groups where the group elements depend on the space-time position and the gauge fields correspond to a particular representation, the adjoint representation. To understand such gauge field theories it is essential to know at least the basic ideas of Lie group theory, although active research often requires going considerably further.

1.1 Basic Definitions and Terminology

A group G is a set of elements $\{g_i\}$ (here we suppose the elements are labelled by a discrete index i but the definitions are easily extended to the case where the elements depend on continuously varying parameters) with a product operation such that

$$g_i, g_j \in G \quad \Rightarrow \quad g_i g_j \in G\,. \tag{1.1}$$

Further we require that there is an *identity* $e \in G$ such that for any $g \in G$

$$eg = ge = g\,, \tag{1.2}$$

and also g has an *inverse* g^{-1} so that

$$gg^{-1} = g^{-1}g = e\,. \tag{1.3}$$

Furthermore the product must satisfy *associativity*

$$g_i(g_j g_k) = (g_i g_j)g_k \quad \text{for all} \quad g_i, g_j, g_k \in G\,, \tag{1.4}$$

so that the order in which a product is evaluated is immaterial. A group is *abelian* if

$$g_i g_j = g_j g_i \quad \text{for all} \quad g_i, g_j \in G\,. \tag{1.5}$$

For a discrete group with n elements then $n = |G|$ is the *order* of the group.

Two groups $G = \{g_i\}$ and $G' = \{g'_i\}$ are *isomorphic*, $G \simeq G'$, if there is a one to one correspondence between the elements consistent with the group multiplication rules.

For any group G a *subgroup* $H \subset G$ is naturally defined as a set of elements belonging to G which is also a group. For any subgroup H we may an equivalence relation between g_i, g'_i,

$$g_i \sim g'_i \quad \Leftrightarrow \quad g_i = g'_i h \quad \text{for} \quad h \in H\,. \tag{1.6}$$

Each equivalence class defines a *coset* and has $|H|$ elements. The cosets form the *coset space* G/H,

$$G/H \simeq G/\sim\,, \qquad \dim G/H = |G|/|H|\,. \tag{1.7}$$

In general G/H is not a group since $g_i \sim g'_i, g_j \sim g'_j$ does not imply $g_i g_j \sim g'_i g'_j$.

A *normal* or *invariant subgroup* is a subgroup $H \subset G$ such that

$$gHg^{-1} = H \quad \text{for all} \quad g \in G\,. \tag{1.8}$$

In this case G/H becomes a group since for $g'_i = g_i h_i, g'_j = g_j h_j$, with $h_i, h_j \in H$, then $g'_i g'_j = g_i g_j h$ for some $h \in G$. For an abelian group all subgroups are normal subgroups.

The *centre* of a group G, $\mathcal{Z}(G)$, is the set of elements which commute with all elements of G. This is clearly an abelian normal subgroup. For an abelian group $\mathcal{Z}(G) \simeq G$.

For two groups G_1, G_2 we may define a *direct product* group $G_1 \otimes G_2$ formed by pairs of elements $\{(g_{1i}, g_{2k})\}$, belonging to (G_1, G_2), with the product rule $(g_{1i}, g_{2k})(g_{1j}, g_{2l}) = (g_{1i}g_{1j}, g_{2k}g_{2l})$. Clearly the identity element is (e_1, e_2) and $(g_{1i}, g_{2k})^{-1} = (g_{1i}^{-1}, g_{2k}^{-1})$. So long as it is clear which elements belong to G_1 and which to G_2 we may write the elements of $G_1 \otimes G_2$ as just $g_1 g_2 = g_2 g_1$. For finite groups $|G_1 \otimes G_2| = |G_1|\,|G_2|$.

1.2 Particular Examples

It is worth describing some particular finite discrete groups which appear frequently.

The group $\mathbb{Z}_n$ is defined by integers $0, 1, \dots n-1$ with the group operation addition modulo n and the identity 0. Alternatively the group may be defined by the complex numbers $e^{2\pi i r/n}$, of modulus one, under multiplication. Clearly it is abelian. Abstractly it consists of elements a^r with $a^0 = a^n = e$ and may be generated just from a single element a satisfying $a^n = e$.

The dihedral group D_n, of order $2n$, is the symmetry group for a regular n-sided polygon and is formed by rotations through angles $2\pi r/n$ together with reflections. The elements are then $\{a^r, ba^r\}$ where $a^n = e, b^2 = e$ and we require $ba = a^{n-1}b$. For $n > 2$ the group is non abelian, note that $D_2 \simeq \mathbb{Z}_2 \otimes \mathbb{Z}_2$.

The remaining frequently occurring group is the permutation group $\mathcal{S}_n$ on n objects. It is easy to see that the order of $\mathcal{S}_n$ is $n!$.

1.3 Further Definitions

Here we give some supplementary definitions connected with groups which are often notationally convenient.

If $g_j = g g_i g^{-1}$ for some $g \in G$ then g_j is *conjugate* to g_i, $g_j \sim g_i$. The equivalence relation $\sim$ divides G into *conjugacy classes* $\mathcal{C}_r$. Clearly the identity is in a conjugacy class by itself, for an abelian group all elements have their own conjugacy class.

For a subgroup $H \subset G$ then the elements $g \in G$ such that $ghg^{-1} \in H$ for all $h \in H$, or $gHg^{-1} = H$, form a subgroup of G, which contains H itself, called the *normaliser* of H in G, written $N_G(H)$. If H is a normal subgroup, $N_G(H) = G$.

An *automorphism* of a group $G = \{g_i\}$ is defined as a mapping between elements, $g_i \to g'_i$, such that the product rule is preserved, i.e.

$$g'_i g'_j = (g_i g_j)' \quad \text{for all} \quad g_i, g_j \in G\,, \tag{1.9}$$

so that $G' = \{g'_i\} \simeq G$. Clearly we must have $e' = e$. In general for any fixed $g \in G$ we may define an *inner automorphism* by $g'_i = g g_i g^{-1}$. It is straightforward to see that the set of all automorphisms of G itself forms a group $\operatorname{Aut} G$ which must include $G/\mathcal{Z}(G)$ as a normal subgroup.

If $H \subset \operatorname{Aut} G$, so that for any $h \in H$ and any $g \in G$ we have $g \underset{h}{\to} g^h$ with ${g_1}^h {g_2}^h = (g_1 g_2)^h$ and $(g^{h_1})^{h_2} = g^{h_1 h_2}$, we may define a new group called the semi-direct product of H with G, denoted $H \ltimes G$. As with the direct product this is defined in terms of pairs of elements (h, g) belonging to (H, G) but with the rather less trivial product rule $(h, g)(h', g') = (hh', gg'^h)$. Note that $(h, g)^{-1} = (h^{-1}, (g^{-1})^{h^{-1}})$. It is often convenient to write the elements of $H \ltimes G$ as simple products so that $(h, g) \to hg = g^h h$. For the semi-direct product $H \ltimes G$, G is a normal subgroup since $hgh^{-1} = g^h \in G$ and hence $H \simeq H \ltimes G/G$.

As a simple illustration we have $D_n \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_n$ where $\mathbb{Z}_2 = \{e, b\}$ with $b^2 = e$ and $\mathbb{Z}_n = \{a^r : r = 0, \ldots n-1\}$ with $a^n = e$ and we define, for any $g = a^r \in \mathbb{Z}_n$, $g^b = g^{-1}$.

1.4 Representations

For any group G a *representation* is a set of non singular (i.e. non zero determinant) square matrices $\{D(g)\}$, for all $g \in G$, such that

$$D(g_1)D(g_2) = D(g_1 g_2)\,, \tag{1.10}$$

$$D(e) = I\,, \tag{1.11}$$

$$D(g^{-1}) = D(g)^{-1}\,, \tag{1.12}$$

where I denotes the unit matrix. If the matrices $D(g)$ are $n \times n$ the representation has *dimension* n.

The representation is *faithful* if $D(g_1) \neq D(g_2)$ for $g_1 \neq g_2$. There is always a *trivial representation* or *singlet representation* in which $D(g) = 1$ for all g. If the representation is not faithful then if $D(h) = I$ for $h \in H$ it is easy to see that H must be a subgroup of G, moreover it is a normal subgroup.

For complex matrices the *conjugate representation* is defined by the matrices $D(g)^*$. The matrices $(D(g)^{-1})^T$ also define a representation.

Since

$$\det\big(D(g_1)D(g_2)\big) = \det D(g_1)\,\det D(g_2)\,, \quad \det I = 1\,, \quad \det D(g)^{-1} = \big(\det D(g)\big)^{-1}\,, \tag{1.13}$$

$\{\det D(g)\}$ form a one-dimensional representation of G which may be trivial and in general is not faithful.

Two representations of the same dimension $D(g)$ and $D'(g)$ are *equivalent* if

$$D'(g) = S D(g) S^{-1} \ \text{ for all } \ g \in G\,. \tag{1.14}$$

For any finite group $G = \{g_i\}$ of order n we may define the dimension n *regular representation* by considering the action of the group on itself

$$g g_i = \sum_j g_j D_{ji}(g)\,, \tag{1.15}$$

where $[D_{ji}(g)]$ are representation matrices with a 1 in each column and row and with all other elements zero. As an example for $D_3 = \{e, a, a^2, b, ba, ba^2\}$, where $a^3 = b^2 = e, ab = ba^2$, then

$$D_{\rm reg}(a) = \begin{pmatrix} 0&0&1&0&0&0\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 0&0&0&1&0&0 \end{pmatrix}\,, \qquad D_{\rm reg}(b) = \begin{pmatrix} 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0 \end{pmatrix}\,. \tag{1.16}$$

A representation of dimension n acts on an associated n-dimensional vector space $\mathcal{V}$, the *representation space*. For any vector $v \in \mathcal{V}$ we may define a group transformation acting on v by

$$v \underset{g}{\to} v^g = D(g)v \,. \tag{1.17}$$

Transformations as in (1.14) correspond to a change of basis for $\mathcal{V}$. A representation is *reducible* if there is a subspace $\mathcal{U} \subset \mathcal{V}$, $\mathcal{U} \neq \mathcal{V}$, such that

$$D(g)u \in \mathcal{U} \quad \text{for all} \quad u \in \mathcal{U} \,, \tag{1.18}$$

otherwise it is an *irreducible representation* or *irrep.* For a reducible representation we may define a representation of lower dimension by restricting to the invariant subspace. More explicitly with a suitable choice of basis we may write, corresponding to (1.18),

$$D(g) = \begin{pmatrix} \hat{D}(g) & B(g) \\ 0 & C(g) \end{pmatrix} \quad \text{for} \quad u = \begin{pmatrix} \hat{u} \\ 0 \end{pmatrix} , \tag{1.19}$$

where the matrices $\hat{D}(g)$ form a representation of G. If, for any invariant subspace, we may restrict the representation matrices to the form shown in (1.19) with $B(g) = 0$ for all g the representation is *completely reducible.*

For an abelian group G all irreducible representations are one-dimensional since all matrices $D(g)$ commute for all $g \in G$ and they may be simultaneously diagonalised. For the n-dimensional translation group T_n, defined by n-dimensional vectors under addition (with 0 as the unit), then for a representation it necessary, for $a \in \mathbb{R}^n$, $a \to D(a)$ satisfying $D(a_1)D(a_2) = D(a_1 + a_2)$. Irreducible representations are all of the form $D(a) = e^{b \cdot a}$, for any n-vector b dual to a.

Representations need not be completely reducible, if $\{R\}$ are $n \times n$ matrices forming a group G_R and a is a n-component column vector then we may define a group in terms of the matrices

$$D(R, a) = \begin{pmatrix} R & a \\ 0 & 1 \end{pmatrix} , \tag{1.20}$$

with the group multiplication rule

$$D(R_1, a_1)D(R_2, a_2) = D(R_1 R_2, R_1 a_2 + a_1) \,, \tag{1.21}$$

which has the abelian subgroup T_n for $R = I$. The group defined by (1.21) is then $G_R \ltimes T_n$ with $a^R = Ra$.

In general for a completely reducible representation the representation space $\mathcal{V}$ decomposes into a direct sum of invariant spaces $\mathcal{U}_r$ which are not further reducible, $\mathcal{V} \simeq \oplus_{r=1}^{k} \mathcal{U}_r$, and hence there is a matrix S such that

$$SD(g)S^{-1} = \begin{pmatrix} D_1(g) & 0 & & \\ 0 & D_2(g) & & \\ & & \ddots & \\ & & & D_k(g) \end{pmatrix} , \tag{1.22}$$

and where $D_r(g)$ form irreducible representations for each r. Writing $\mathcal{R}$ for the representation given by the matrices $D(g)$ and $\mathcal{R}_r$ for the irreducible representation matrices $D_r(g)$ then (1.22) is written as

$$\mathcal{R} = \mathcal{R}_1 \oplus \cdots \oplus \mathcal{R}_k \,. \tag{1.23}$$

Useful results, which follow almost directly from the definition of irreducibility, characterising irreducible representations are:

Schur's Lemmas. If $D_1(g), D_2(g)$ form two irreducible representations then *(i)*

$$SD_1(g) = D_2(g)S \,, \tag{1.24}$$

for all g requires that the two representation are equivalent or $S = 0$. Also *(ii)*

$$SD(g) = D(g)S \,, \tag{1.25}$$

for all g for an irreducible representation $D(g)$ then $S \propto I$.

To prove *(i)* suppose $\mathcal{V}_1, \mathcal{V}_2$ are the representation spaces corresponding to the representations given by the matrices $D_1(g), D_2(g)$, so that $\mathcal{V}_1 \underset{S}{\to} \mathcal{V}_2$. Then the image of S, $\operatorname{Im} S = \{v : v = Su, u \in \mathcal{V}_1\}$, is an invariant subspace of $\mathcal{V}_2$, $D_2(g) \operatorname{Im} S = \operatorname{Im} S\, D_2(g)$, by virtue of (1.24). Similarly the kernel of S, $\operatorname{Ker} S = \{u : Su = 0, u \in \mathcal{V}_1\}$ forms an invariant subspace of $\mathcal{V}_1$, both sides of (1.24) giving zero. For both representations to be irreducible we must have $\operatorname{Im} S = \mathcal{V}_2$, $\operatorname{Ker} S = 0$, so that S is invertible, $\det S \neq 0$, (this is only possible if $\dim \mathcal{V}_2 = \dim \mathcal{V}_1$). Since then $D_2(g) = SD_1(g)S^{-1}$ for all g the two representations are equivalent.

To prove *(ii)* suppose the eigenvectors of S with eigenvalue λ span a space $\mathcal{V}_\lambda$. Applying (1.25) to $\mathcal{V}_\lambda$ shows that $D(g)\mathcal{V}_\lambda$ are also eigenvectors of S with eigenvalue λ so that $D(g)\mathcal{V}_\lambda \subset \mathcal{V}_\lambda$ and consequently $\mathcal{V}_\lambda$ is an invariant subspace unless $\mathcal{V}_\lambda = \mathcal{V}$ and then $S = \lambda I$.

1.4.1 Induced Representations

A representation of a group G also gives a representation when restricted to a subgroup H. Conversely for a subgroup $H \subset G$ then it is possible to obtain representations of G in terms of those for H by constructing the *induced representation.* Assume

$$v \underset{h}{\to} D(h)v \,, \qquad h \in H \,, \quad v \in \mathcal{V} \,, \tag{1.26}$$

with $\mathcal{V}$ the representation space for this representation of H. For finite groups the cosets forming G/H may be labelled by an index i so that for each coset we may choose an element $g_i \in G$ such that all elements belonging to the i'th coset can be expressed as $g_i h$ for some $h \in H$. The choice of g_i is arbitrary to the extent that we may let $g_i \to g_i h_i$ for some fixed $h_i \in H$. For any $g \in G$ then

$$gg_i = g_j h \qquad \text{for some} \quad h \in H \,, \quad i,j = 1,\ldots,N \,, \quad N = |G|/|H| \,. \tag{1.27}$$

Assuming (1.27) determines h the induced representation is defined so that that under the action of a group transformation $g \in G$,

$$v_i \underset{g}{\rightarrow} D(h)v_j \,, \qquad v_i = (g_i, v) \,, \quad D(h)v_j = (g_j, D(h)v) \,. \tag{1.28}$$

In (1.28) h depends on i as well as g and $v_i \in \mathcal{V}_i$ which is isomorphic to $\mathcal{V}$ for each i so that the representation space for the induced representation is the N-fold tensor product $\mathcal{V}^{\otimes N}$. The representation matrices for the induced representation are then given by $N \times N$ matrices whose elements are $D(h)$ for some $h \in H$,

$$D_{ji}(g) = \begin{cases} D(h)\,, & g_j^{-1} g\, g_i = h \in H\,, \\ 0\,, & g_j^{-1} g\, g_i \notin H\,. \end{cases} \tag{1.29}$$

To show that (1.28) is in accord with the group multiplication rule we consider a subsequent transformation g' so that

$$v_i \underset{g}{\rightarrow} D(h)v_j \underset{g'}{\rightarrow} D(h')D(h)v_k = D(h'h)v_k \quad \text{for} \quad g'g_j = g_k h' \quad \Rightarrow \quad (g'g)g_i = g_k h'h \,. \tag{1.30}$$

If $H = \{e\}$, forming a trivial subgroup of G, and $D(h) \to 1$, the induced representation is identical with the regular representation for finite groups.

As a simple example we consider $G = D_n$ generated by elements a, b with $a^n = b^2 = e, ab = ba^{n-1}$. H is chosen to be the abelian subgroup $\mathbb{Z}_n = \{a^r : r = 0, \dots, n-1\}$. This has one-dimensional representations labelled by $k = 0, 1, \dots, n-1$ defined by

$$v \underset{a}{\rightarrow} e^{\frac{2\pi ik}{n}} v \,. \tag{1.31}$$

With this choice for H there are two cosets belonging to $D_n/\mathbb{Z}_n$ labelled by $i = 1, 2$ and we may take $g_1 = e$, $g_2 = b$. Then for $v_1 = (e, v)$ transforming as in (1.31) then with $v_2 = (b, v)$ (1.28) requires, using $ab = ba^{-1}$,

$$(v_1, v_2) \underset{a}{\rightarrow} (e^{\frac{2\pi ki}{n}} v_1, e^{-\frac{2\pi ki}{n}} v_2) = (v_1, v_2)A\,, \qquad (v_1, v_2) \underset{b}{\rightarrow} (v_2, v_1) = (v_1, v_2)B\,, \tag{1.32}$$

for 2×2 matrices A, B,

$$A = \begin{pmatrix} e^{\frac{2\pi ik}{n}} & 0 \\ 0 & e^{-\frac{2\pi ik}{n}} \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{1.33}$$

which satisfy $A^n = I$, $B^2 = I$, $AB = BA^{n-1}$ and so give a two dimensional representation of D_n for each k. By considering $A \to BAB$ it is clear that the representation for $k \to n-k$ is equivalent to that in (1.33).

1.4.2 Unitary Representations

For application in quantum mechanics we are almost always interested in *unitary representations* where the matrices are require to satisfy

$$D(g)^\dagger = D(g^{-1}) = D(g)^{-1} \,. \tag{1.34}$$

For such representation then the usual scalar product on $\mathcal{V}$ is invariant, for transformations as in (1.17) ${v_1{}^g}^\dagger v_2{}^g = {v_2}^\dagger v_1$. If $\mathcal{U}$ is an invariant subspace then the orthogonal subspace $\mathcal{U}_\perp$, as defined by the scalar product, is also an invariant subspace. Hence unitary representations are always completely reducible.

Theorem: For a finite group all representations are equivalent to unitary representations.

To show this define

$$S = \sum_i D(g_i)^\dagger D(g_i)\,, \tag{1.35}$$

where the sum is over all elements of the group $G = \{g_i\}$. Noting that for any g, $\{g_i g\} = \{g_i\}$, since if $g_j g = g_i g$ then $g_j = g_i$, we have

$$\begin{aligned} SD(g)^{-1} = SD(g^{-1}) &= \sum_i D(g_i)^\dagger D(g_i g^{-1}) \\ &= \sum_i D(g_i g)^\dagger D(g_i) \\ &= D(g)^\dagger \sum_i D(g_i)^\dagger D(g_i) = D(g)^\dagger S\,, \end{aligned} \tag{1.36}$$

using that $D(g)$ form a representation and also $(AB)^\dagger = B^\dagger A^\dagger$. Hence if we define $\langle v_1, v_2\rangle = {v_1}^\dagger S v_2$ then we have $\langle v_1, D(g^{-1})v_2\rangle = (D(g)v_1, v_2)$ or $\langle v_1{}^g, v_2{}^g\rangle = \langle v_1, v_2\rangle$. With respect to this scalar product $D(g)$ is unitary (or we may define $D'(g) = S^{\frac{1}{2}}D(g)S^{-\frac{1}{2}}$ and then show $D'(g)^\dagger D'(g) = I$).

1.4.3 Orthogonality Relations

Schur's lemmas have an important consequence in that the matrices for irreducible representations obey an orthogonality relation. To derive this let

$$\mathcal{A}^{(\mathcal{R}',\mathcal{R})} = \sum_i D^{(\mathcal{R}')}(g_i{}^{-1})AD^{(\mathcal{R})}(g_i)\,, \qquad \mathcal{B}^{(\mathcal{R},\mathcal{R}')} = \sum_i D^{(\mathcal{R})}(g_i)BD^{(\mathcal{R}')}(g_i{}^{-1})\,, \tag{1.37}$$

where $D^{(\mathcal{R})}(g), D^{(\mathcal{R}')}(g)$ are the matrices corresponding to the irreducible representation $\mathcal{R}, \mathcal{R}'$, and A, B are arbitrary matrices of the appropriate dimension. Then

$$\mathcal{A}^{(\mathcal{R}',\mathcal{R})}D^{(\mathcal{R})}(g) = D^{(\mathcal{R}')}(g)\mathcal{A}^{(\mathcal{R}',\mathcal{R})}\,, \qquad D^{(\mathcal{R})}(g)\mathcal{B}^{(\mathcal{R},\mathcal{R}')} = \mathcal{B}^{(\mathcal{R},\mathcal{R}')}D^{(\mathcal{R})}(g')\,, \tag{1.38}$$

for any $g \in G$. The proof of (1.38) follows exactly in the fashion as in (1.36), essentially since $\{g_i\} = \{g_i g\}$. Schur's lemmas then require that $\mathcal{A}^{(\mathcal{R}',\mathcal{R})}, \mathcal{B}^{(\mathcal{R},\mathcal{R}')} = 0$ unless $\mathcal{R}' = \mathcal{R}$ when both $\mathcal{A}^{(\mathcal{R}',\mathcal{R})}, \mathcal{B}^{(\mathcal{R},\mathcal{R}')}$ are proportional to the identity. Hence we must have

$$S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = \sum_i D^{(\mathcal{R}')}_{rv}(g_i{}^{-1})D^{(\mathcal{R})}_{us}(g_i) = \frac{|G|}{n_\mathcal{R}}\,\delta_{\mathcal{R}'\mathcal{R}}\,\delta_{rs}\,\delta_{uv}\,, \tag{1.39}$$

where $n_\mathcal{R}$ is the dimension of the representation $\mathcal{R}$. The constant in (1.39) is determined by considering $S^{(\mathcal{R},\mathcal{R})}_{ru,us} = \sum_i D^{(\mathcal{R})}_{rs}(e) = |G|\,\delta_{rs}$.

1.4.4 Characters

For any representation $\mathcal{R}$ the *character* is defined by

$$\chi_{\mathcal{R}}(g) = \operatorname{tr}\big(D^{(\mathcal{R})}(g)\big) \,. \tag{1.40}$$

Since traces are unchanged under cyclic permutations $\chi_{\mathcal{R}}(g'gg'^{-1}) = \chi_{\mathcal{R}}(g)$ so that the character depends only on the conjugacy classes of each element. Similarly the character is unchanged when calculated for any representations related by an equivalence transformation as in (1.14). Since for a finite group any representation is equivalent to a unitary one we must also have

$$\chi_{\mathcal{R}}(g^{-1}) = \chi_{\mathcal{R}}(g)^* \,. \tag{1.41}$$

As a consequence of the orthogonality relations, (1.37) and (1.39), then using (1.41) for two irreducible representations $\mathcal{R}, \mathcal{R}'$

$$\sum_i \chi_{\mathcal{R}'}(g_i)^* \, \chi_{\mathcal{R}}(g_i) = |G| \; \delta_{\mathcal{R}'\mathcal{R}} \,. \tag{1.42}$$

For an induced representation as in (1.29) if for the subgroup representation

$$\chi(h) = \operatorname{tr}\big(D(h)\big) \,, \tag{1.43}$$

then

$$\chi_{\text{induced rep.}}(g) = \sum_i \chi(g_i{}^{-1} g \, g_i) \, \big|_{g_i{}^{-1} g \, g_i \in H} \,. \tag{1.44}$$

If this is applied to the case when $H = \{e\}$ giving the regular representation we get

$$\chi_{\text{regular rep.}}(g) = \begin{cases} |G| \,, & g = e \\ 0 \,, & g \neq e \,. \end{cases} \tag{1.45}$$

1.4.5 Tensor Products

If $\mathcal{V}_1, \mathcal{V}_2$ are representation spaces for representations $\mathcal{R}_1, \mathcal{R}_2$, given by matrices $D_1(g)$, $D_2(g)$, for a group G then we may define a *tensor product representation* $\mathcal{R}_1 \otimes \mathcal{R}_2$ in terms of the matrices $D(g) = D_1(g) \otimes D_2(g)$ acting on the tensor product space $\mathcal{V}_1 \otimes \mathcal{V}_2$ where $D(g)v = \sum_{r,s} a_{rs} \, D_1(g) v_{1r} D_2(g) v_{2s}$. Since $\dim \mathcal{V} = \dim \mathcal{V}_1 \; \dim \mathcal{V}_2$ the tensor product matrices have dimensions which are the products of the dimensions of the matrices forming the tensor product. If $D_1(g), D_2(g)$ are unitary then so is $D(g)$.

In general the tensor product representation $\mathcal{R}_1 \otimes \mathcal{R}_2$ for two representations $\mathcal{R}_1, \mathcal{R}_2$ is reducible and may be decomposed into irreducible ones. If the irreducible representations are listed as $\mathcal{R}_r$ then in general for the product of any two irreducible representations

$$\mathcal{R}_r \otimes \mathcal{R}_s \simeq \mathcal{R}_s \otimes \mathcal{R}_r \simeq \bigoplus_t n_{rs,t} \, \mathcal{R}_t \,, \tag{1.46}$$

where $n_{rs,t}$ are integers, which may be zero, and $n_{rs,t} > 1$ if the representation $\mathcal{R}_t$ occurs more than once. For non finite groups there are infinitely many irreducible representations but the sum in (1.46) is finite for finite dimensional representations. The trace of a tensor product of matrices is the product of the traces of each individual matrix, in consequence $\mathrm{tr}_{\mathcal{V}_r \otimes \mathcal{V}_s}\big(D^{(\mathcal{R}_r)}(g) \otimes D^{(\mathcal{R}_s)}(g)\big) = \mathrm{tr}_{\mathcal{V}_r}\big(D^{(\mathcal{R}_r)}(g)\big)\,\mathrm{tr}_{\mathcal{V}_s}\big(D^{(\mathcal{R}_s)}(g)\big)$, so that, in terms of the characters $\chi_{\mathcal{R}_r}(g) = \mathrm{tr}_{\mathcal{V}_r}\big(D^{(\mathcal{R}_r)}(g)\big)$, (1.46) is equivalent to

$$\chi_{\mathcal{R}_r}(g)\chi_{\mathcal{R}_s}(g) = \sum_t n_{rs,t}\,\chi_{\mathcal{R}_t}(g)\,. \tag{1.47}$$

Using (1.42) the coefficients $n_{rs,t}$ can be determined by

$$n_{rs,t} = \frac{1}{|G|}\sum_i \chi_{\mathcal{R}_t}(g_i)^*\chi_{\mathcal{R}_r}(g_i)\chi_{\mathcal{R}_s}(g_i)\,. \tag{1.48}$$

The result (1.46) is exactly equivalent to the decomposition of the associated representation spaces, with the same expansion for $\mathcal{V}_r \otimes \mathcal{V}_s$ into a direct sum of irreducible spaces $\mathcal{V}_t$. If $\mathcal{R}_r \otimes \mathcal{R}_s$ contains the trivial or singlet representation then it is possible to construct a scalar product $\langle v, v'\rangle$ between vectors $v \in \mathcal{V}_r, v' \in \mathcal{V}_j$ which is invariant under group transformations, $\langle D^{(\mathcal{R}_i)}(g)v, D^{(\mathcal{R}_j)}(g)v'\rangle = \langle v, v'\rangle$.

1.5 Matrix Groups

It is easy to see that any set of non singular matrices which are closed under matrix multiplication form a group since they satisfy (1.2),(1.3),(1.4) with the identity e corresponding to the unit matrix and the inverse of any element given by the matrix inverse, requiring that the matrix is non singular so that the determinant is non zero. Many groups are defined in terms of matrices. Thus $Gl(n,\mathbb{R})$ is the set of all real $n \times n$ non singular matrices, $Sl(n,\mathbb{R})$ are those with unit determinant and $Gl(n,\mathbb{C})$, $Sl(n,\mathbb{C})$ are the obvious extensions to complex numbers. Since $\det(M_1M_2) = \det M_1 \det M_2$ and $\det M^{-1} = (\det M)^{-1}$ the matrix determinants form an invariant abelian subgroup unless the the conditions defining the matrix group require unit determinant for all matrices.

Matrix groups of frequent interest are

$O(n)$, *real orthogonal* $n \times n$ matrices $\{M\}$, so that

$$M^T M = I\,. \tag{1.49}$$

This set of matrices is closed under multiplication since $(M_1M_2)^T = M_2{}^T M_1{}^T$. For $SO(n)$ $\det M = 1$. A general $n\times n$ real matrix has n^2 real parameters while a symmetric matrix has $\frac{1}{2}n(n+1)$. M^TM is symmetric so that (1.49) provides $\frac{1}{2}n(n+1)$ conditions. Hence $O(n)$, and also $SO(n)$, have $\frac{1}{2}n(n-1)$ parameters. If v, v' belong to the n-dimensional representation space for $O(n)$ or $SO(n)$ then scalar product $v'^T v$ is invariant under $v \to Mv$, $v' \to Mv'$.

For n even $\pm I \in SO(n)$ and these form the centre of the group so long as $n > 2$. Thus $\mathcal{Z}(SO(2n)) \simeq \mathbb{Z}_2$, $n = 2, 3, \dots$, while $\mathcal{Z}(SO(2n+1)) = \{I\}$, $n = 1, 2, \dots$ is trivial although $\mathcal{Z}(O(2n+1)) = \{\pm I\} \simeq \mathbb{Z}_2$, $n = 1, 2, \dots$.

$U(n)$, *complex unitary* $n \times n$ matrices, so that

$$M^\dagger M = I \,. \tag{1.50}$$

Closure follows from $(M_1 M_2)^\dagger = {M_2}^\dagger {M_1}^\dagger$. For $SU(n)$ $\det M = 1$. A general $n \times n$ complex matrix has $2n^2$ real parameters while a hermitian matrix has n^2. $M^\dagger M$ is hermitian so that $U(n)$ has n^2 parameters. (1.50) requires $|\det M| = 1$ so imposing $\det M = 1$ now provides one additional condition so that $SU(n)$ has $n^2 - 1$ parameters. The $U(n)$ invariant scalar product for complex n-dimensional vectors v, v' is $v'^\dagger v$.

The centre of $U(n)$ or $SU(n)$ consists of all elements proportional to the identity, by virtue of Schur's lemma, so that $\mathcal{Z}(SU(n)) = \{e^{2\pi r i/n} I : r = 0, \dots n-1\} \simeq \mathbb{Z}_n$, while $\mathcal{Z}(U(n)) = \{e^{i\alpha} I : 0 \le \alpha < 2\pi\} \simeq U(1)$.

Note that $SO(2) \simeq U(1)$ since a general $SO(2)$ matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} , \qquad 0 \le \theta < 2\pi \,, \tag{1.51}$$

is in one to one correspondence with a general element of $U(1)$,

$$e^{i\theta} \,, \qquad 0 \le \theta < 2\pi \,. \tag{1.52}$$

$Sp(2n, \mathbb{R})$ and $Sp(2n, \mathbb{C})$, *symplectic* $2n \times 2n$ real or complex matrices satisfying

$$M^T J M = J \,, \tag{1.53}$$

where J is a $2n \times 2n$ antisymmetric matrix with the standard form

$$J = \begin{pmatrix} {\scriptstyle \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}} & 0 & & \\ 0 & {\scriptstyle \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}} & & \\ & & \ddots & \\ & & & {\scriptstyle \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}} \end{pmatrix} . \tag{1.54}$$

In this case $M^T J M$ is antisymmetric so that (1.53) provides $n(2n-1)$ conditions and hence $Sp(2n, \mathbb{R})$ has $n(2n+1)$ parameters. For symplectic transformations there is an antisymmetric invariant form $\langle v', v \rangle = -\langle v, v' \rangle = v'^T J v$.

The condition (1.53) requires $\det M = 1$ so there are no further restrictions as for $O(n)$ and $U(n)$. To show this we define the *Pfaffian*[1] for $2n \times 2n$ antisymmetric matrices A by

$$\mathrm{Pf}(A) = \frac{1}{2^n n!} \, \varepsilon_{i_1 \dots i_{2n}} A_{i_1 i_2} \dots A_{i_{2n-1} i_{2n}} \,, \tag{1.55}$$

with $\varepsilon_{i_1 \dots i_{2n}}$ the $2n$-dimensional antisymmetric symbol, $\varepsilon_{1 \dots 2n} = 1$. The Pfaffian is essentially the square root of the usual determinant since

$$\det A = \mathrm{Pf}(A)^2 \,, \tag{1.56}$$

[1] Johann Friedrich Pfaff, 1765-1825, German.

and it is easy to see that

$$\mathrm{Pf}(J) = 1 \,. \tag{1.57}$$

The critical property here is that

$$\mathrm{Pf}(M^T A M) = \det M \, \mathrm{Pf}(A) \quad \text{since} \quad \varepsilon_{i_1 \dots i_{2n}} M_{i_1 j_1} \dots M_{i_{2n} j_{2n}} = \det M \, \varepsilon_{j_1 \dots j_{2n}} \,. \tag{1.58}$$

Applying (1.58) with $A = J$ to the definition of symplectic matrices in (1.53) shows that we must have $\det M = 1$.

Since both $\pm I$ belong to $Sp(2n, \mathbb{R})$ then the centre $\mathcal{Z}(Sp(2n, \mathbb{R})) \simeq \mathbb{Z}_2$.

The matrix groups $SO(n)$ and $SU(n)$ are *compact*, which will be defined precisely later but which for the moment can be taken to mean that the natural parameters vary over a finite range. $Sp(2n, \mathbb{R})$ is not compact, which is evident since matrices of the form

$$\begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} , \qquad -\infty < \theta < \infty \,. \tag{1.59}$$

belong to $Sp(2, \mathbb{R})$.

A compact Sp group, denoted $Sp(n)$ or sometimes $USp(2n)$, can be obtained by considering matrices belonging to both $U(2n)$ and $Sp(2n, \mathbb{C})$. An alternative characterisation of $Sp(n)$ is in terms of $n \times n$ quaternionic unitary matrices. A basis for quaternionic numbers, denoted $\mathbb{H}$ after Hamilton[2] and extending $\mathbb{C}$, is provided by the unit imaginary *quaternions* i, j, k, satisfying $i^2 = j^2 = k^2 = -1$ and $ij = k$, $jk = i$, $ki = j$, together with the real 1. A general quaternion is a linear combination $q = x1 + yi + uj + vk$, for $x, y, u, v \in \mathbb{R}$, and the conjugate $\bar{q} = x1 - yi - uj - vk$, $\bar{q}q = |q|^2 1$. A $n \times n$ quaternionic matrix M has the form

$$M = a1 + bi + cj + dk \,, \quad a, b, c, d \;\; \text{real } n \times n \;\; \text{matrices} \,, \tag{1.60}$$

and the adjoint is

$$\bar{M} = a^T 1 - b^T i - c^T j - d^T k \,. \tag{1.61}$$

$Sp(n) \simeq U(n, \mathbb{H})$ is defined in terms of $n \times n$ quaternion matrices with the property

$$\bar{M} M = I_n \, 1 \,, \tag{1.62}$$

for I_n the unit $n \times n$ matrix. A general quaternionic $n \times n$ M then has $4n^2$ parameters whereas $U = \bar{M} M = \bar{U}$ is a hermitian quaternion matrix which has n real diagonal elements and $\frac{1}{2} n(n-1)$ independent off diagonal quaternionic numbers giving $n(2n-1)$ parameters altogether. Hence (1.62) provides $n(2n-1)$ conditions so that $Sp(n)$ has $n(2n+1)$ parameters.

To show the correspondence of $U(n, \mathbb{H})$ with $USp(2n)$ we replace the quaternions by 2×2 matrices according to

$$1 \to \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) , \quad i \to \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right) , \quad j \to \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) , \quad k \to \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right) , \tag{1.63}$$

[2] William Rowan Hamilton, 1805-65, Irish.

so that any $n \times n$ quaternion matrix M becomes a $2n \times 2n$ complex matrix $\mathcal{M}$,

$$M \to \mathcal{M}\,, \quad \bar{M} \to \mathcal{M}^\dagger\,, \quad I_n 1 \to I_{2n}\,, \quad I_n j \to J \quad \Rightarrow \quad \mathcal{M}^\dagger = -J\mathcal{M}^T J\,. \tag{1.64}$$

(1.62) then ensures $\mathcal{M}^\dagger \mathcal{M} = I_{2n}$ so that $\mathcal{M} \in U(2n)$ and furthermore (1.64) requires also that $\mathcal{M}$ obeys (1.53).

There are also various extensions which also arise frequently in physics. Suppose g is the diagonal $(n+m) \times (n+m)$ matrix defined by

$$g = \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix}\,, \tag{1.65}$$

then the *pseudo-orthogonal* groups $O(n,m)$, and hence $SO(n,m)$, are defined by real matrices M such that

$$M^T g M = g\,. \tag{1.66}$$

The invariant form in this case is $v'^T g v$. Similarly we may define $U(n,m)$ and $SU(n,m)$. It is easy to see that $O(n,m) \simeq O(m,n)$ and similarly for other analogous cases. The parameter count for these groups is the same as for the corresponding $O(n{+}m)$ or $U(n{+}m)$, $SU(n+m)$. Note that matrices belonging to $SO(1,1)$ are just those given in (1.59).

For each matrix group the definition of course provides a representation which is termed the *fundamental representation*.

1.6 Symmetries and Quantum Mechanics

A symmetry of a physical system is defined as a set of transformations acting on the system such that the physical observables are invariant. In quantum mechanics the state of a particular physical system is represented by a vector $|\psi\rangle$ belonging to a vector (or Hilbert) space $\mathcal{H}$. The essential observables are then the probabilities, given that the system is in a state $|\psi\rangle$, of finding, under some appropriate measurement, the system in a state $|\phi\rangle$. Assuming $|\psi\rangle, |\phi\rangle$ are both normalised this probability is $|\langle\phi|\psi\rangle|^2$. For a symmetry transformation $|\psi\rangle \to |\psi'\rangle$ we must require

$$|\langle\phi|\psi\rangle|^2 = |\langle\phi'|\psi'\rangle|^2 \quad \text{for all} \quad |\psi\rangle, |\phi\rangle \in \mathcal{H}\,. \tag{1.67}$$

Any quantum state vector is arbitrary up to a complex phase $|\psi\rangle \sim e^{i\alpha}|\psi\rangle$. Making use of this potential freedom Wigner[3] proved that there is an operator U such that

$$U|\psi\rangle = |\psi'\rangle\,, \tag{1.68}$$

and either $\langle\phi'|\psi'\rangle = \langle\phi|\psi\rangle$ with U linear

$$U\big(a_1|\psi_1\rangle + a_2|\psi_2\rangle\big) = a_1 U|\psi_1\rangle + a_2 U|\psi_2\rangle\,, \tag{1.69}$$

or $\langle\phi'|\psi'\rangle = \langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$ with U anti-linear

$$U\big(a_1|\psi_1\rangle + a_2|\psi_2\rangle\big) = a_1{}^* U|\psi_1\rangle + a_2{}^* U|\psi_2\rangle\,. \tag{1.70}$$

[3]Eugene Paul Wigner, 1902-1995, Hungarian until 1937, then American. Nobel Prize 1962.

Thus U is *unitary linear* or *unitary anti-linear*. Mostly the anti-linear case is not relevant, if U is continuously connected to the identity it must be linear. For the discrete symmetry linked to *time reversal* $t \to -t$ the associated operator T must be anti-linear, in order for the Schrödinger equation $i\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle$ to be invariant when $THT^{-1} = H$ (we must exclude the alternative possibility $THT^{-1} = -H$ since energies should be positive or bounded below).

For a symmetry group $G = \{g\}$ then we must have unitary operators $U[g]$ where we require $U[e] = 1$, $U[g^{-1}] = U[g]^{-1}$. Because of the freedom of complex phases we may relax the product rule and require only

$$U[g_i]U[g_j] = e^{i\gamma(g_i,g_j)}U[g_ig_j]\,. \tag{1.71}$$

If the phase factor $e^{i\gamma}$ is present this gives rise to a *projective representation*. However the associativity condition (1.4) ensures $\gamma(g_i, g_j)$ must satisfy consistency conditions,

$$\gamma(g_i, g_jg_k) + \gamma(g_j, g_k) = \gamma(g_ig_j, g_k) + \gamma(g_i, g_j)\,. \tag{1.72}$$

There are always solutions to (1.72) of the form

$$\gamma(g_i, g_j) = \alpha(g_ig_j) - \alpha(g_i) - \alpha(g_j)\,, \tag{1.73}$$

for any arbitrary $\alpha(g)$ depending on $g \in G$. However such solutions are trivial since in this case we may let $e^{i\alpha(g)}U[g] \to U[g]$ to remove the phase factor in (1.71). For most groups there are no non trivial solutions for $\gamma(g_i, g_j)$ so the extra freedom allowed by (1.71) may be neglected so there is no need to consider projective representations, although there are some cases when it is essential.

If G is a symmetry for a physical system with a Hamiltonian H we must require

$$U[g]HU[g]^{-1} = H \quad \text{for all} \quad g \in G\,. \tag{1.74}$$

If H has energy levels with degeneracy so that

$$H|\psi_r\rangle = E|\psi_r\rangle\,, \quad r = 1, \dots, n\,, \tag{1.75}$$

then it is easy to see that

$$H\,U[g]|\psi_r\rangle = E\,U[g]|\psi_r\rangle\,. \tag{1.76}$$

Hence we must have

$$U[g]|\psi_r\rangle = \sum_{s=1}^{n} |\psi_s\rangle D_{sr}(g)\,, \tag{1.77}$$

and furthermore the matrices $[D_{sr}(g)]$ form a n-dimensional representation of G. If $\{|\psi_r\rangle\}$ are orthonormal, $\langle\psi_r|\psi_s\rangle = \delta_{rs}$, then the matrices are unitary. The representation need not be irreducible but, unless there are additional symmetries not taken into account or there is some accidental special choice for the parameters in H, in realistic physical examples only irreducible representations are relevant.

2 Rotations and Angular Momentum, $SO(3)$ and $SU(2)$

Symmetry under rotations in three dimensional space is an essential part of general physical theories which is why they are most naturally expressed in vector notation. The fundamental property of rotations is that the lengths, and scalar products, of vectors are invariant.

Rotations correspond to orthogonal matrices, since acting on column vectors v, they are the most general transformations leaving $v^T v$ invariant, for real v the length $|v|$ is given by $|v|^2 = v^T v$. For any real orthogonal matrix M then if v is an eigenvector, in general complex, $Mv = \lambda v$ we also have $Mv^* = \lambda^* v^*$, so that if λ is complex both λ, λ^* are eigenvalues, and $(Mv^*)^T Mv = |\lambda|^2 v^\dagger v = v^\dagger v$ so that we must have $|\lambda|^2 = 1$.

2.1 Three Dimensional Rotations

Rotations in three dimensions are then determined by matrices $R \in O(3)$ and hence satisfying

$$R^T R = I \,. \tag{2.1}$$

The eigenvalues of R can only be $e^{i\theta}, e^{-i\theta}$ and 1 or -1 so that a general R can therefore be reduced, by a real transformation S, to the form

$$SRS^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \pm 1 \end{pmatrix} . \tag{2.2}$$

For $\det R = 1$, so that $R \in SO(3)$, we must have the $+1$ case when

$$\mathrm{tr} R = 2\cos\theta + 1 \,. \tag{2.3}$$

Acting on a spatial vector $\mathbf{x}$ the matrix R induces a linear transformation

$$\mathbf{x} \underset{R}{\rightarrow} \mathbf{x}' = \mathbf{x}^R \,, \tag{2.4}$$

where, for i, j, three dimensional indices, we have

$$x'_i = R_{ij} x_j \,, \tag{2.5}$$

For $\det R = -1$ the transformation involves a reflection.

A general $R \in SO(3)$ has 3 parameters which may be taken as the rotation angle θ and the unit vector n, which is also be specified by two angles, and is determined by $Rn = n$. n defines the axis of the rotation. The matrix may then be expressed in general as

$$R_{ij}(\theta, n) = \cos\theta\, \delta_{ij} + (1 - \cos\theta) n_i n_j - \sin\theta\, \varepsilon_{ijk} n_k \,, \tag{2.6}$$

where ε_{ijk} is the three dimensional antisymmetric symbol, $\varepsilon_{123} = 1$. The parameters (θ, n) cover all rotations if

$$n \in S^2 \,, \quad 0 \le \theta \le \pi \,, \qquad (\pi, n) \simeq (\pi, -n) \,, \tag{2.7}$$

with S^2 the two-dimensional unit sphere. For an infinitesimal rotation $R(\delta\theta, n)$ acting on a vector $\mathbf{x}$ and using standard vector notation we then have

$$\mathbf{x} \xrightarrow[R(\delta\theta,n)]{} \mathbf{x}' = \mathbf{x} + \delta\theta\, \mathbf{n} \times \mathbf{x} \,. \tag{2.8}$$

It is easy to see that $\mathbf{x}'^2 = \mathbf{x}^2 + \mathrm{O}(\delta\theta^2)$.

For a vector product $(\mathbf{n} \times \mathbf{x})^R = \mathbf{n}^R \times \mathbf{x}^R$ so that making use of (2.8)

$$\mathbf{x} \xrightarrow[RR(\delta\theta,n)R^{-1}]{} \mathbf{x}' = \mathbf{x} + \delta\theta\, \mathbf{n}^R \times \mathbf{x} \,, \tag{2.9}$$

so that we must have

$$RR(\delta\theta, n)R^{-1} = R(\delta\theta, Rn) \,. \tag{2.10}$$

Furthermore we must then have $RR(\theta, n)R^{-1} = R(\theta, Rn)$, so that all rotations with the same θ belong to a single conjugacy class.

2.2 Isomorphism of $SO(3)$ and $SU(2)/\mathbb{Z}_2$

$SO(3) \simeq SU(2)/\mathbb{Z}_2$, where $\mathbb{Z}_2$ is the centre of $SU(2)$ which is formed by the 2×2 matrices $I, -I$, is of crucial importance in understanding the role of spinors under rotations. To demonstrate this we introduce the standard *Pauli*[4] *matrices*, a set of three 2×2 matrices which have the explicit form

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.11}$$

These matrices satisfy the algebraic relations

$$\sigma_i\sigma_j = \delta_{ij}\, I + i\, \epsilon_{ijk}\sigma_k \,, \tag{2.12}$$

and also are traceless and hermitian. Adopting a vector notation $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, so that (2.12) is equivalent to $\mathbf{a}\cdot\boldsymbol{\sigma}\; \mathbf{b}\cdot\boldsymbol{\sigma} = \mathbf{a}\cdot\mathbf{b}\, I + i\, \mathbf{a}\times\mathbf{b}\cdot\boldsymbol{\sigma}$, we have

$$\boldsymbol{\sigma}^\dagger = \boldsymbol{\sigma} \,, \qquad \mathrm{tr}(\boldsymbol{\sigma}) = 0 \,. \tag{2.13}$$

Using (2.12) then gives

$$\mathrm{tr}(\sigma_i\sigma_j) = 2\delta_{ij} \,, \tag{2.14}$$

which ensures that any 2×2 matrix A can be expressed in the form

$$A = \tfrac{1}{2}\mathrm{tr}(A)\, I + \tfrac{1}{2}\mathrm{tr}(\boldsymbol{\sigma} A)\cdot\boldsymbol{\sigma} \,, \tag{2.15}$$

since the Pauli matrices form a complete set of traceless and hermitian 2×2 matrices.

The Pauli matrices ensure that there is a one to one correspondence between real three vectors and hermitian traceless 2×2 matrices, given explicitly by

$$\mathbf{x} \to \mathbf{x}\cdot\boldsymbol{\sigma} = (\mathbf{x}\cdot\boldsymbol{\sigma})^\dagger \,, \qquad \mathbf{x} = \tfrac{1}{2}\mathrm{tr}(\boldsymbol{\sigma}\, \mathbf{x}\cdot\boldsymbol{\sigma}) \,, \tag{2.16}$$

[4]Wolfgang Ernst Pauli, 1900-58, Austrian. Nobel prize 1945.

Furthermore $\mathbf{x}\cdot\boldsymbol{\sigma}$ satisfies the matrix equation

$$(\mathbf{x}\cdot\boldsymbol{\sigma})^2 = \mathbf{x}^2 I\,. \tag{2.17}$$

From (2.17) and (2.13) the eigenvalues of $\mathbf{x}\cdot\boldsymbol{\sigma}$ must be $\pm\sqrt{\mathbf{x}^2}$ and in consequence we have

$$\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = -\mathbf{x}^2\,. \tag{2.18}$$

For any $A \in SU(2)$ we can then define a linear transformation $\mathbf{x} \to \mathbf{x}'$ by

$$\mathbf{x}'\cdot\boldsymbol{\sigma} = A\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger\,, \tag{2.19}$$

since we may straightforwardly verify that $A\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger$ is hermitian and is also traceless, using the invariance of any trace of products of matrices under cyclic permutations and

$$AA^\dagger = I\,. \tag{2.20}$$

With, $\mathbf{x}'$ defined by (2.19) and using (2.18),

$$\mathbf{x}'^2 = -\det(\mathbf{x}'\cdot\boldsymbol{\sigma}) = -\det(A\,\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger) = -\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = \mathbf{x}^2\,, \tag{2.21}$$

using $\det(XY) = \det X \ \det Y$ and from (2.20) $\det A \ \det A^\dagger = 1$. Hence, since this shows that $|\mathbf{x}'| = |\mathbf{x}|$,

$$x'_i = R_{ij}x_j\,, \tag{2.22}$$

with $[R_{ij}]$ an orthogonal matrix. Furthermore since as $A \to I$, $R_{ij} \to \delta_{ij}$ we must have $\det[R_{ij}] = 1$. Explicitly from (2.19) and (2.14)

$$\sigma_i R_{ij} = A\sigma_j A^\dagger \quad \Rightarrow \quad R_{ij} = \tfrac{1}{2}\,\mathrm{tr}(\sigma_i A\sigma_j A^\dagger)\,. \tag{2.23}$$

To show the converse then from (2.23), using (note $\sigma_j\sigma_i\sigma_j = -\sigma_i$) $\sigma_j A^\dagger\sigma_j = 2\,\mathrm{tr}(A^\dagger)I - A^\dagger$, we obtain

$$R_{jj} = |\mathrm{tr}(A)|^2 - 1\,, \qquad \sigma_i R_{ij}\sigma_j = 2\,\mathrm{tr}(A^\dagger)A - I\,. \tag{2.24}$$

For $A \in SU(2)$, $\mathrm{tr}(A) = \mathrm{tr}(A^\dagger)$ is real (the eigenvalues of A are $e^{\pm i\alpha}$ giving $\mathrm{tr}(A) = 2\cos\alpha$) so that (2.24) may be solved for $\mathrm{tr}(A)$ and then A,

$$A = \pm\,\frac{I + \sigma_i R_{ij}\sigma_j}{2(1+R_{jj})^{\frac{1}{2}}}\,. \tag{2.25}$$

The arbitrary sign, which cancels in (2.23), ensures that in general $\pm A \leftrightarrow R_{ij}$.

For a rotation through an infinitesimal angle as in (2.8) then from (2.6)

$$R_{ij} = \delta_{ij} - \delta\theta\,\varepsilon_{ijk}n_k\,, \tag{2.26}$$

and it is easy to obtain, assuming $A \to I$ as $\delta\theta \to 0$,

$$A = I - \tfrac{1}{2}\delta\theta\, i\mathbf{n}\cdot\boldsymbol{\sigma}\,. \tag{2.27}$$

Note that since $\det(I+X) = 1+\mathrm{tr}X$, to first order in X, for any matrix then the tracelessness of the Pauli matrices is necessary for (2.27) to be compatible with $\det A = 1$. For a finite rotation angle θ then, with (2.3), (2.24) gives $|\mathrm{tr}(A)| = 2|\cos\frac{1}{2}\theta|$ and the matrix A can be found by exponentiation, where corresponding to (2.6),

$$A(\theta, n) = e^{-\frac{1}{2}i\theta\,\mathbf{n}\cdot\boldsymbol{\sigma}} = \cos\tfrac{1}{2}\theta\; I - \sin\tfrac{1}{2}\theta\; i\mathbf{n}\cdot\boldsymbol{\sigma}\,. \tag{2.28}$$

The parameters (θ, n) cover all $SU(2)$ matrices for

$$n \in S^2\,, \qquad 0 \le \theta < 2\pi\,, \tag{2.29}$$

in contrast to (2.7).

2.3 Infinitesimal Rotations and Generators

To analyse the possible representation spaces for the rotation group it is sufficient to consider rotations which are close to the identity as in (2.8). If consider two infinitesimal rotations $R_1 = R(\delta\theta_1, n_1)$ and $R_2 = R(\delta\theta_2, n_2)$ then it is easy to see that

$$R = R_2{}^{-1}R_1{}^{-1}R_2R_1 = I + \mathrm{O}(\delta\theta_1\delta\theta_2)\,. \tag{2.30}$$

Acting on a vector $\mathbf{x}$ and using (2.8) and keeping only terms which are $\mathrm{O}(\delta\theta_1\delta\theta_2)$ we find

$$\begin{aligned}\mathbf{x} \underset{R}{\to} \mathbf{x}' &= \mathbf{x} + \delta\theta_1\delta\theta_2\big(\mathbf{n}_2 \times (\mathbf{n}_1 \times \mathbf{x}) - \mathbf{n}_1 \times (\mathbf{n}_2 \times \mathbf{x})\big)\\ &= \mathbf{x} + \delta\theta_1\delta\theta_2\,(\mathbf{n}_2 \times \mathbf{n}_1) \times \mathbf{x}\,,\end{aligned} \tag{2.31}$$

using standard vector product identities.

Acting on a quantum mechanical vector space the corresponding unitary operators are assumed to be of the form

$$U[R(\delta\theta, n)] = 1 - i\delta\theta\,\mathbf{n}\cdot\mathbf{J}\,, \tag{2.32}$$

$\mathbf{J}$ are the *generators* of the rotation group. Since $U[R(\delta\theta, n)]^{-1} = 1 + i\delta\theta\,\mathbf{n}\cdot\mathbf{J} + \mathrm{O}(\delta\theta^2)$ the condition for U to be a unitary operator becomes

$$\mathbf{J}^\dagger = \mathbf{J}\,, \tag{2.33}$$

or each J_i is hermitian. If we consider the combined rotations as in (2.30) in conjunction with (2.31) and (2.32) we find

$$\begin{aligned}U[R] &= 1 - i\delta\theta_1\delta\theta_2\,(\mathbf{n}_2 \times \mathbf{n}_1)\cdot\mathbf{J}\\ &= U[R_2]^{-1}U[R_1]^{-1}U[R_2]U[R_1]\\ &= 1 - \delta\theta_1\delta\theta_2\,[\mathbf{n}_2\cdot\mathbf{J}\,,\, \mathbf{n}_1\cdot\mathbf{J}]\,,\end{aligned} \tag{2.34}$$

where it is only necessary to keep $\mathrm{O}(\delta\theta_1\delta\theta_2)$ contributions as before. Hence we must have

$$[\mathbf{n}_2\cdot\mathbf{J}\,,\, \mathbf{n}_1\cdot\mathbf{J}] = i\,(\mathbf{n}_2 \times \mathbf{n}_1)\cdot\mathbf{J}\,, \tag{2.35}$$

or equivalently

$$[J_i,\, J_j] = i\,\varepsilon_{ijk}J_k\,. \tag{2.36}$$

Although (2.32) expresses U in terms of $\mathbf{J}$ for infinitesimal rotations it can be extended to finite rotations since

$$U[R(\theta, n)] = \exp(-i\theta\, \mathbf{n}\cdot\mathbf{J}) = \lim_{N\to\infty} \left(1 - i\,\frac{\theta}{N}\,\mathbf{n}\cdot\mathbf{J}\right)^{N} . \tag{2.37}$$

Under rotations $\mathbf{J}$ is a vector since, from (2.10), $U[R]U[R(\delta\theta, n)]U[R]^{-1} = U[R(\delta\theta, Rn)]$ which in turn from (2.32) implies

$$U[R]J_i U[R]^{-1} = (R^{-1})_{ij} J_j \,. \tag{2.38}$$

For a physical system the vector operator, rotation group generator, $\mathbf{J}$ is identified as that corresponding to the total angular momentum of the system and then (2.36) are the fundamental angular momentum commutation relations. It is important to recognise that rotational invariance of the Hamiltonian is equivalent to conservation of angular momentum since

$$U[R]HU[R]^{-1} = H \qquad \Leftrightarrow \qquad [\mathbf{J}, H] = 0\,. \tag{2.39}$$

This ensures that the degenerate states for each energy must belong to a representation space for a representation of the rotation group.

2.4 Representations of Angular Momentum Commutation Relations

We here describe how the commutation relations (2.36) can be directly analysed to determine possible representation spaces $\mathcal{V}$ on which the action of the operators $\mathbf{J}$ is determined. First we define

$$J_\pm = J_1 \pm iJ_2 \,, \tag{2.40}$$

and then (2.36) is equivalent to

$$[J_3, J_\pm] = \pm J_\pm \,, \tag{2.41a}$$

$$[J_+, J_-] = 2J_3 \,. \tag{2.41b}$$

The hermeticity conditions (2.33) then become

$$J_+{}^\dagger = J_- \,, \qquad J_3{}^\dagger = J_3 \,. \tag{2.42}$$

A basis for a space on which a representation for the angular momentum commutation relations is defined in terms of eigenvectors of J_3. Let

$$J_3|m\rangle = m|m\rangle \,. \tag{2.43}$$

Then from (2.41a) it is easy to see that

$$J_\pm|m\rangle \propto |m \pm 1\rangle \;\text{ or }\; 0\,, \tag{2.44}$$

so that the possible J_3 eigenvalues form a sequence $\ldots, m-1, m, m+1 \ldots$.

If the states $|m \pm 1\rangle$ are non zero we define

$$J_-|m\rangle = |m-1\rangle\,, \qquad J_+|m\rangle = \lambda_m|m+1\rangle\,, \tag{2.45}$$

and hence

$$J_+J_-|m\rangle = \lambda_{m-1}|m\rangle\,, \qquad J_-J_+|m\rangle = \lambda_m|m\rangle\,. \tag{2.46}$$

By considering $[J_+, J_-]|m\rangle$ we have from (2.41b), if $|m \pm 1\rangle$ are non zero,

$$\lambda_{m-1} - \lambda_m = 2m\,. \tag{2.47}$$

This can be solved for any m by

$$\lambda_m = j(j+1) - m(m+1)\,, \tag{2.48}$$

for some constant written as $j(j+1)$. For sufficiently large positive or negative m we clearly have $\lambda_m < 0$. The hermeticity conditions (2.42) require that J_+J_- and J_-J_+ are of the form $O^\dagger O$ and so must have positive eigenvalues with zero possible only if J_- or respectively J_+ annihilates the state ($\langle\psi|O^\dagger O|\psi\rangle \geq 0$, if 0 then $O|\psi\rangle = 0$). Hence there must be both a maximum $m_{\max}$ and a minimum $m_{\min}$ for m requiring

$$J_+|m_{\max}\rangle = 0 \quad \Rightarrow \quad \lambda_{m_{\max}} = (j - m_{\max})(j + m_{\max} + 1) = 0\,, \tag{2.49a}$$

$$J_-|m_{\min}\rangle = 0 \quad \Rightarrow \quad \lambda_{m_{\min}-1} = (j + m_{\min})(j - m_{\min} + 1) = 0\,, \tag{2.49b}$$

where also

$$m_{\max} - m_{\min} = 0, 1, 2, \dots\,. \tag{2.50}$$

Taking $j \geq 0$ the result (2.48) then requires

$$m_{\max} = j\,, \qquad m_{\min} = -j\,. \tag{2.51}$$

For this to be possible we must have

$$j \in \{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots\}\,, \tag{2.52}$$

and then for each value of j

$$m \in \{-j, -j+1, \dots j-1, j\}\,. \tag{2.53}$$

The corresponding states $|m\rangle$ form a basis for a $(2j+1)$-dimensional representation space $\mathcal{V}_j$.

2.5 The $|j\,m\rangle$ basis

It is more convenient to define an orthonormal basis for $\mathcal{V}_j$ in terms of states $\{|j\,m\rangle\}$, with j, m as in (2.52) and (2.53), satisfying

$$\langle j\,m|j\,m'\rangle = \delta_{mm'}\,. \tag{2.54}$$

These are eigenvectors of J_3 as before

$$J_3|j\,m\rangle = m|j\,m\rangle\,. \tag{2.55}$$

and j may be defined as the maximum value of m so that

$$J_+|j\,j\rangle = 0\,. \tag{2.56}$$

A state satisfying both (2.55) and (2.56) is called a highest weight state. In this case the action of $J_\pm$ gives

$$J_\pm|j\,m\rangle = N^\pm_{jm}|j\,m{\pm}1\rangle\,, \tag{2.57}$$

where $N^\pm_{jm}$ are determined by requiring (2.54) to be satisfied. From (2.46) and (2.48) we must then have

$$|N^+_{jm}|^2 = \lambda_m = (j-m)(j+m+1)\,, \qquad |N^-_{jm}|^2 = \lambda_{m-1} = (j+m)(j-m+1)\,. \tag{2.58}$$

By convention $N^\pm_{jm}$ are chosen to real and positive so that

$$N^\pm_{jm} = \sqrt{(j\mp m)(j\pm m+1)}\,. \tag{2.59}$$

In general we may then define the the states $\{|j\,m\rangle\}$ in terms of the highest weight state by

$$(J_-)^n|j\,j\rangle = \Big(\tfrac{n!\,(2j)!}{(2j-n)!}\Big)^{\frac{1}{2}}|j\,j-n\rangle\,, \quad n=0,1,\dots,2j\,. \tag{2.60}$$

An alternative prescription for specifying the states $|j\,m\rangle$ is to consider the operator $\mathbf{J}^2 = {J_1}^2 + {J_2}^2 + {J_3}^2$. In terms of $J_\pm, J_3$ this can be expressed in two alternative forms

$$\mathbf{J}^2 = \begin{cases} J_-J_+ + J_3^2 + J_3\,, \\ J_+J_- + J_3^2 - J_3\,. \end{cases} \tag{2.61}$$

With the first form in (2.61) and using (2.56) we then get acting on the highest weight state

$$\mathbf{J}^2|j\,j\rangle = j(j+1)|j\,j\rangle\,. \tag{2.62}$$

Moreover $\mathbf{J}^2$ is a rotational scalar and satisfies

$$[\mathbf{J}^2\,,\,J_i] = 0\,, \quad i=1,2,3\,. \tag{2.63}$$

In particular J_- commutes with $\mathbf{J}^2$ so that the eigenvalue is the same for all m. Hence the states $|j\,m\rangle$ satisfy

$$\mathbf{J}^2|j\,m\rangle = j(j+1)|j\,m\rangle\,, \tag{2.64}$$

as well as (2.55). Nevertheless we require (2.57), with (2.59), to determine the relative phases of all states.

2.5.1 Representation Matrices

Using the $|j\,m\rangle$ basis it is straightforward to define corresponding representation matrices for each j belonging to (2.52). For the angular momentum operator

$$\mathbf{J}^{(j)}{}_{m'm} = \langle j\,m'|\mathbf{J}|j\,m\rangle \tag{2.65}$$

or alternatively

$$\mathbf{J}|j\,m\rangle = \sum_{m'} |j\,m'\rangle\, \mathbf{J}^{(j)}{}_{m'm}\,. \tag{2.66}$$

The $(2j+1)\times(2j+1)$ matrices $\mathbf{J}^{(j)} = [\mathbf{J}^{(j)}{}_{m'm}]$ then satisfy the angular momentum commutation relations (2.36). From (2.55) and (2.57)

$$J_3^{(j)}{}_{m'm} = m\,\delta_{m',m}\,, \qquad J_\pm^{(j)}{}_{m'm} = \sqrt{(j\mp m)(j\pm m+1)}\,\delta_{m',m\pm1}\,. \tag{2.67}$$

For R a rotation then

$$D^{(j)}_{m'm}(R) = \langle j\,m'|U[R]|j\,m\rangle\,, \tag{2.68}$$

defines $(2j+1)\times(2j+1)$ matrices $D^{(j)}(R) = [D^{(j)}_{m'm}(R)]$ forming a representation of the the rotation group corresponding to the representation space $\mathcal{V}_j$,

$$U[R]|jm\rangle = \sum_{m'} |jm'\rangle D^{(j)}_{m'm}(R)\,. \tag{2.69}$$

Note that $D^{(0)}(R) = 1$ is the trivial representation and for an infinitesimal rotation as in (2.8)

$$D^{(j)}\big(R(\delta\theta, n)\big) = I_{2j+1} - i\delta\theta\,\mathbf{n}{\cdot}\mathbf{J}^{(j)}\,. \tag{2.70}$$

To obtain explicit forms for the rotation matrices it is convenient to parameterise a rotation in terms of Euler angles ψ,θ,ϕ when

$$R = R(\phi, e_3)R(\theta, e_2)R(\psi, e_3)\,, \tag{2.71}$$

for e_2, e_3 corresponding to unit vectors along the $2,3$ directions Then

$$U[R] = e^{-i\phi J_3}\, e^{-i\theta J_2}\, e^{-i\psi J_3}\,, \tag{2.72}$$

so that in (2.68)

$$D^{(j)}_{m'm}(R) = e^{-im'\phi - im\psi}\, d^{(j)}_{m'm}(\theta)\,, \qquad d^{(j)}_{m'm}(\theta) = \langle j\,m'|e^{-i\theta J_2}|j\,m\rangle\,. \tag{2.73}$$

For the special cases of $\theta = \pi, 2\pi$,

$$d^{(j)}_{m'm}(\pi) = (-1)^{j-m}\delta_{m',-m}\,, \qquad d^{(j)}_{m'm}(2\pi) = (-1)^{2j}\delta_{m',m}\,. \tag{2.74}$$

In general $D^{(j)}(R(2\pi, n)) = (-1)^{2j} I_{2j+1}$, which for j a $\frac{1}{2}$-integer is not the identity. For representations of $SO(3)$ it would be necessary to take j to be an integer but in quantum mechanics any j given by (2.52) is allowed since we require representations only up to a phase factor. From the result for $\theta = \pi$ we have

$$e^{-i\pi J_2}|j\, m\rangle = (-1)^{j-m}|j\, {-m}\rangle \,. \tag{2.75}$$

Using this and $e^{-i\pi J_3}|j\, m\rangle = e^{-i\pi m}|j\, m\rangle$ with $e^{-i\pi J_3}J_2 e^{i\pi J_3} = -J_2$ we must have from the definition in (2.73)

$$d^{(j)}_{m'm}(\theta) = (-1)^{m'-m} d^{(j)}_{-m'\, -m}(\theta) = (-1)^{m'-m} d^{(j)}_{m'm}(-\theta) = (-1)^{m'-m} d^{(j)}_{mm'}(\theta) \,. \tag{2.76}$$

For the simplest case $j = \frac{1}{2}$, it is easy to see from (2.67) that

$$J_+^{(\frac{1}{2})} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \qquad J_-^{(\frac{1}{2})} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} , \qquad J_3^{(\frac{1}{2})} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{2.77}$$

and hence we have

$$\mathbf{J}^{(\frac{1}{2})} = \tfrac{1}{2}\boldsymbol{\sigma} \,, \tag{2.78}$$

where σ_i, $i = 1, 2, 3$ are the Pauli matrices as given in (2.11). It is clear that $\frac{1}{2}\sigma_i$ must satisfy the commutation relations (2.36). The required commutation relations are a consequence of (2.12). For $j = \frac{1}{2}$ we also have

$$d^{(\frac{1}{2})}(\theta) = \begin{pmatrix} \cos\frac{1}{2}\theta & -\sin\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix} . \tag{2.79}$$

With the definition of characters in (1.40) the rotation group characters

$$\chi_j(\theta) = \mathrm{tr}\big(D^{(j)}(R(\theta, n))\big) \,, \tag{2.80}$$

depend only on the rotation angle θ. Hence they may be easily calculated by considering

$$\chi_j(\theta) = \sum_{m=-j}^{j} \langle j\, m| e^{-i\theta J_3} |j\, m\rangle = \sum_{m=-j}^{j} e^{-im\theta} = \frac{\sin(j+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta} \,. \tag{2.81}$$

2.6 Tensor Products and Angular Momentum Addition

The representation space $\mathcal{V}_j$, which has the orthonormal basis $\{|j\, m\rangle\}$, determines an irreducible representation of $SU(2)$ and also the commutation relations (2.36) of the generators or physically the angular momentum operators. The tensor product $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ of two representation spaces $\mathcal{V}_{j_1}, \mathcal{V}_{j_2}$ has a basis

$$|j_1\, m_1\rangle_1 |j_2\, m_2\rangle_2 \,. \tag{2.82}$$

Associated with $\mathcal{V}_{j_1}, \mathcal{V}_{j_2}$ there are two independent angular operators $\mathbf{J}_1, \mathbf{J}_2$ both satisfying the commutation relations (2.36)

$$\begin{aligned} [J_{1,i}\, , \, J_{1,j}] &= i\varepsilon_{ijk} J_{1,k} \,, \\ [J_{2,i}\, , \, J_{2,j}] &= i\varepsilon_{ijk} J_{2,k} \,. \end{aligned} \tag{2.83}$$

They may be extended to act on $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ since with the basis (2.82)

$$\begin{aligned} &\mathbf{J}_1 \equiv \mathbf{J}_1 \otimes 1_2\,, \qquad \mathbf{J}_1\big(|j_1 m_1\rangle_1\,|j_2\, m_2\rangle_2\big) = \mathbf{J}_1|j_1 m_1\rangle_1\,|j_2\, m_2\rangle_2\,, \\ &\mathbf{J}_2 \equiv 1_1 \otimes \mathbf{J}_2\,, \qquad \mathbf{J}_2\big(|j_1 m_1\rangle_1\,|j_2\, m_2\rangle_2\big) = |j_1 m_1\rangle_1\,\mathbf{J}_2|j_2\, m_2\rangle_2\,. \end{aligned} \tag{2.84}$$

With this definition it is clear that they commute

$$[J_{1,i}\,,\, J_{2,j}] = 0\,. \tag{2.85}$$

The generator for the tensor product representation, or the total angular momentum operator, is then defined by

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2\,. \tag{2.86}$$

It is easy to see that this has the standard commutation relations (2.36).

In the space $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ we may construct states which are standard basis states for the total angular momentum $|JM\rangle$ labelled by the eigenvalues of $\mathbf{J}^2$, J_3,

$$\begin{aligned} J_3|JM\rangle &= M|JM\rangle\,, \\ \mathbf{J}^2|JM\rangle &= J(J+1)|JM\rangle\,. \end{aligned} \tag{2.87}$$

These states are chosen to be orthonormal so that

$$\langle J'M'|JM\rangle = \delta_{J'J}\delta_{M'M}\,, \tag{2.88}$$

and satisfy (2.57). All states in $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ must be linear combinations of the basis states (2.82) so that we may write

$$|JM\rangle = \sum_{m_1,m_2} |j_1 m_1\rangle_1 |j_2 m_2\rangle_2\ \langle j_1 m_1\, j_2 m_2|JM\rangle\,. \tag{2.89}$$

Here

$$\langle j_1 m_1\, j_2 m_2|JM\rangle\,, \tag{2.90}$$

are *Clebsch-Gordan coefficients*[5].

As $J_3 = J_{1,3} + J_{2,3}$ Clebsch-Gordan coefficients must vanish unless $M = m_1 + m_2$. To determine the possible values of J it is sufficient to find all highest weight states $|JJ\rangle$ in $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ such that

$$J_3|JJ\rangle = J|JJ\rangle\,, \qquad J_+|JJ\rangle = 0\,. \tag{2.91}$$

We may then determine the states $|JM\rangle$ by applying J_- as in (2.60). There is clearly a unique highest weight state with $J = j_1 + j_2$ given by

$$|j_1{+}j_2\, j_1{+}j_2\rangle = |j_1 j_1\rangle_1\,|j_2\, j_2\rangle_2\,, \tag{2.92}$$

from which $|j_1{+}j_2\, M\rangle$ is obtained as in (2.60). We may then construct the states $|JM\rangle$ for $J = j_1 + j_2, j_1 + j_2 - 1, \dots$ iteratively. Defining $\mathcal{V}^{(M)} \subset \mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ to be the subspace for which J_3 has eigenvalue M then, since it has a basis as in (2.82) for all $m_1 + m_2 = M$, we have, assuming $j_1 \ge j_2$, $\dim \mathcal{V}^{(M)} = j_1 + j_2 - M + 1$ for $M \ge j_1 - j_2$ and $\dim \mathcal{V}^{(M)} = 2j_2 + 1$

[5]Rudolf Friedrich Alfred Clebsch, 1833-1872, German. Paul Albert Gordan, 1837-1912, German.

for $M \le j_1 - j_2$. Assume all states $|J'M\rangle$ have been found as in (2.89) for $j_1+j_2 \ge J' > J$. For $j_1+j_2 > J \ge j_1 - j_2$ there is a one dimensional subspace in $\mathcal{V}^{(J)}$ which is orthogonal to all states $|J'J\rangle$ for $J < J' \le j_1+j_2$. This subspace must be annihilated by J_+, as otherwise there would be too many states with $M = J+1$, and hence there is a highest weight state $|JJ\rangle$. If $M < j_1 - j_2$ it is no longer possible to construct further highest weight states. Hence we have shown, since the results must be symmetric in j_1, j_2, that in $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ there exists exactly one vector subspace $\mathcal{V}_J$, of dimension $(2J+1)$, for each J-value in the range

$$J \in \{j_1+j_2, j_1+j_2-1, \ldots, |j_1-j_2|+1, |j_1-j_2|\}\,, \tag{2.93}$$

or

$$\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2} = \bigoplus_{J=|j_1-j_2|}^{j_1+j_2} \mathcal{V}_J\,. \tag{2.94}$$

If $j_1 \ge j_2$ we can easily check that

$$\begin{aligned}\sum_{J=j_1-j_2}^{j_1+j_2} (2J+1) &= \sum_{J=j_1-j_2}^{j_1+j_2} \big((J+1)^2 - J^2\big)\\ &= (j_1+j_2+1)^2 - (j_1-j_2)^2 = (2j_1+1)(2j_2+1)\,,\end{aligned} \tag{2.95}$$

so that the basis $\{|JM\rangle\}$ has the correct dimension to span the vector space $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$.

Alternatively in terms of the characters given in (2.81)

$$\begin{aligned}\chi_{j_1}(\theta)\,\chi_{j_2}(\theta) = \chi_{j_1}(\theta) \sum_{m=-j_2}^{j_2} e^{-im\theta} &= \frac{1}{2i\,\sin\frac{1}{2}\theta} \sum_{m=-j_2}^{j_2} \big(e^{(j_1+m+\frac{1}{2})\theta} - e^{(-j_1+m+\frac{1}{2})\theta}\big)\\ = \sum_{j=j_1-j_2}^{j_1+j_2} \chi_j(\theta) &= \sum_{j=|j_1-j_2|}^{j_1+j_2} \chi_j(\theta)\,,\end{aligned} \tag{2.96}$$

where if $j_2 > j_1$ we use $\chi_{-j}(\theta) = -\chi_{j-1}(\theta)$ to show all contributions to the sum for $j < j_2 - j_1$ cancel. Comparing with (1.47) the result of this character calculation of course matches the tensor product decomposition given in (2.95).

The construction of $|JM\rangle$ states described above allows the Clebsch-Gordan coefficients to be iteratively determined starting from $J = j_1 + j_2$ and then progressively for lower J as in (2.93). By convention they are chosen to be real and for each J there is a standard choice of the overall sign. With standard conventions

$$\langle j_1m_1\, j_2m_2|JM\rangle = (-1)^{j_1+j_2-J} \langle j_2m_2\, j_1m_1|JM\rangle\,. \tag{2.97}$$

Since the original basis (2.82) and $\{|JM\rangle\}$ are both orthonormal we have the orthogonality/completeness conditions

$$\begin{aligned}\sum_{m_1,m_2} \langle j_1m_1\, j_2m_2|JM\rangle \langle j_1m_1\, j_2m_2|J'M'\rangle &= \delta_{JJ'}\,\delta_{MM'}\,,\\ \sum_{JM} \langle j_1m_1\, j_2m_2|JM\rangle \langle j_1m_1'\, j_2m_2'|JM\rangle &= \delta_{m_1m_1'}\,\delta_{m_2m_2'}\,.\end{aligned} \tag{2.98}$$

For the tensor product representation defined on the tensor product space $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ we may use the Clebsch-Gordan coefficients as in (2.89) to give the decomposition into irreducible representations for each J allowed by (2.93)

$$\sum_{m_1{}',m_1} \sum_{m_2{}',m_2} D^{(j_1)}_{m_1{}'m_1}(R)\, D^{(j_2)}_{m_2{}'m_2}(R)\, \langle j_1 m_1{}'\, j_2 m_2{}' | J'M'\rangle \langle j_1 m_1\, j_2 m_2 | JM\rangle = \delta_{J'J}\, D^{(J)}_{M'M}(R)\,. \tag{2.99}$$

2.7 Examples of the calculation of Clebsch-Gordan coefficients

In the case in which $j_1 = 1$ and $j_2 = \frac{1}{2}$, there are in $\mathcal{V}_1 \otimes \mathcal{V}_{\frac{1}{2}}$ six basis states $|1\, m_1\rangle\, |\frac{1}{2}\, m_2\rangle$.

$$\begin{array}{rcccccc} m_1 & 1 & 1 & 0 & 0 & -1 & -1 \\ m_2 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ M = m_1 + m_2 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \end{array} \,.$$

In the basis $|JM\rangle$ for the total angular momentum there are also six states $|\frac{3}{2}\, M\rangle$ and $|\frac{1}{2}\, M\rangle$. Since there is only one state with $M = \frac{3}{2}$, it follows that we may identify

$$|\tfrac{3}{2}\, \tfrac{3}{2}\rangle = |1\, 1\rangle\, |\tfrac{1}{2} \tfrac{1}{2}\rangle\,. \tag{2.100}$$

Now action of the lowering operator $J_- = J_{1-} + J_{2-}$, allows all the states $|\frac{3}{2}\, M\rangle$ to be given as linear combinations of product states $|1\, m_1\rangle\, |\frac{1}{2}\, m_2\rangle$. Applying $J_- = J_{1-} + J_{2-}$ to (2.100) gives

$$J_-|\tfrac{3}{2}\, \tfrac{3}{2}\rangle = J_{1-}|1\, 1\rangle\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle + |1\, 1\rangle\, J_{2-}|\tfrac{1}{2}\, \tfrac{1}{2}\rangle\,, \tag{2.101}$$

or

$$\sqrt{3}\, |\tfrac{3}{2}\, \tfrac{1}{2}\rangle = \sqrt{2}\, |1\, 0\rangle\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle + |1\, 1\rangle\, |\tfrac{1}{2} -\tfrac{1}{2}\rangle\,, \tag{2.102}$$

giving

$$|\tfrac{3}{2}\, \tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}\, |1\, 0\rangle\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{3}}\, |1\, 1\rangle\, |\tfrac{1}{2} -\tfrac{1}{2}\rangle\,. \tag{2.103}$$

This result then gives explicit numerical values for two Clebsch-Gordan coefficients. Repeating this process twice gives similar expressions for the states $|\frac{3}{2} -\frac{1}{2}\rangle$ and $|\frac{3}{2} -\frac{3}{2}\rangle$. The last step provides a check: to within a sign at least one should find that

$$|\tfrac{3}{2} -\tfrac{3}{2}\rangle = |1 -1\rangle\, |\tfrac{1}{2} -\tfrac{1}{2}\rangle\,, \tag{2.104}$$

because there is only one possible state with $M = -\frac{3}{2}$ in $\mathcal{V}_1 \otimes \mathcal{V}_{\frac{1}{2}}$.

Turning next to the $J = \frac{1}{2}$ multiplet, we use the fact that the state $|\frac{1}{2}\, \frac{1}{2}\rangle$ is orthogonal to the state $|\frac{3}{2}\, \frac{1}{2}\rangle$ constructed above. It follows that from (2.103) that we may write

$$|\tfrac{1}{2}\, \tfrac{1}{2}\rangle = -\sqrt{\tfrac{1}{3}}|1\, 0\rangle\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}|1\, 1\rangle\, |\tfrac{1}{2} -\tfrac{1}{2}\rangle\,. \tag{2.105}$$

This result is uniquely determined to within an overall phase which we have taken in accordance with the so-called Condon and Shortley phase convention, *i.e.* we have chosen

$$\langle 1 1\, \tfrac{1}{2} -\tfrac{1}{2} | \tfrac{1}{2} \tfrac{1}{2}\rangle\,, \tag{2.106}$$

to be real and positive, in fact here equal to $+\sqrt{\frac{2}{3}}$. To summarise we have shown that the total angular momentum basis states $|jm\rangle$ are given in terms of the product states $|1m_1\rangle\,|\frac{1}{2}m_2\rangle$ by

$$\begin{aligned}
|\tfrac{3}{2}\,\tfrac{3}{2}\rangle &= |1\,1\rangle\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle \\
|\tfrac{3}{2}\,\tfrac{1}{2}\rangle &= \sqrt{\tfrac{2}{3}}\,|1\,0\rangle\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{3}}\,|1\,1\rangle\,|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle \\
|\tfrac{3}{2}\,{-\tfrac{1}{2}}\rangle &= \sqrt{\tfrac{1}{3}}\,|1\,{-1}\rangle\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}\,|1\,0\rangle\,|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle \\
|\tfrac{3}{2}\,{-\tfrac{3}{2}}\rangle &= |1\,{-1}\rangle\,|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle\,,
\end{aligned} \tag{2.107}$$

and

$$\begin{aligned}
|\tfrac{1}{2}\,\tfrac{1}{2}\rangle &= -\sqrt{\tfrac{1}{3}}\,|1\,0\rangle\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}\,|1\,1\rangle\,|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle \\
|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle &= -\sqrt{\tfrac{2}{3}}\,|1\,{-1}\rangle\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{3}}\,|1\,0\rangle\,|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle\,.
\end{aligned} \tag{2.108}$$

All the Clebsch-Gordan coefficients

$$\langle 1m_1\,\tfrac{1}{2}m_2|JM\rangle\,, \tag{2.109}$$

can then be read off from (2.107) and (2.108).

2.7.1 Construction of Singlet States

A special example of decomposition of tensor products is the construction of the singlet states $|0\,0\rangle$, which corresponds to the one-dimensional trivial representation and so is invariant under rotations. For $\mathcal{V}_{j_1}\otimes\mathcal{V}_{j_2}$, as is clear from (2.93) this is only possible for $j_1 = j_2 = j$ and the singlet state must have the general form

$$|0\,0\rangle = \sum_m a_m\,|j\,m\rangle_1|j\,{-m}\rangle_2\,. \tag{2.110}$$

Requiring $J_+|0\,0\rangle = 0$ gives $a_m = -a_{m-1}$ so that, imposing the normalisation condition,

$$|0\,0\rangle = \frac{1}{\sqrt{2j+1}}\sum_{n=0}^{2j}(-1)^n|j\,j-n\rangle_1|j\,{-j+n}\rangle_2\,. \tag{2.111}$$

This determines the Clebsch-Gordan coefficients $\langle jm\,j{-m}|00\rangle$. Note that $|0\,0\rangle$ is symmetric, antisymmetric under $1\leftrightarrow 2$ according to whether $2j$ is even, odd.

If we consider the extension to three spins for the tensor product space $\mathcal{V}_{j_1}\otimes\mathcal{V}_{j_2}\otimes\mathcal{V}_{j_3}$, then we may couple $|j_1\,m_1\rangle_1|j_2\,m_2\rangle_2$ to form a vector with $J = j_3$ and then use (2.111). The result may be expressed as

$$|0\,0\rangle = \sum_{m_1,m_2,m_3}\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}|j_1\,m_1\rangle_1|j_2\,m_2\rangle_2|j_3\,m_3\rangle_3\,, \tag{2.112}$$

where,
$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{j_1-j_2-m_3}}{\sqrt{2j_3+1}} \langle j_1 m_1\, j_2 m_2 | j_3 \, {-m_3} \rangle \, . \tag{2.113}$$

(2.113) defines the $3j$-symbols which are more symmetric than Clebsch-Gordan coefficients, under the interchange any pair of the j, m's the $3j$-symbol is invariant save for an overall factor $(-1)^{j_1+j_2+j_3}$. They are non zero only if $m_1 + m_2 + m_3 = 0$ and j_1, j_2, j_3 satisfy triangular inequalities $|j_1 - j_2| \le j_3 \le j_1 + j_2$.

2.8 $SO(3)$ Tensors

In the standard treatment of rotations vectors and tensors play an essential role. For $R = [R_{ij}]$ and $SO(3)$ rotation then a vector is required to transform as

$$V_i \underset{R}{\to} V'_i = R_{ij} V_j \, . \tag{2.114}$$

Vectors then give a three dimensional representation space $\mathcal{V}$. A *rank n tensor* $T_{i_1 \dots i_n}$ is then defined as belonging to the n-fold tensor product $\mathcal{V} \otimes \cdots \otimes \mathcal{V}$ and hence satisfy the transformation rule

$$T_{i_1 \dots i_n} \underset{R}{\to} T'_{i_1 \dots i_n} = R_{i_1 j_1} \dots R_{i_n j_n} T_{j_1 \dots j_n} \, . \tag{2.115}$$

It is easy to see the dimension of the representation space, $\mathcal{V}(\otimes \mathcal{V})^{n-1}$, formed by rank n tensors, is 3^n. For $n = 0$ we have a scalar which is invariant and $n = 1$ corresponds to a vector. The crucial property of rotational tensors is that they be multiplied to form tensors of higher rank, for two vectors U_i, V_i then $U_i V_j$ is a rank two tensor, and also that contraction of indices preserves tensorial properties essential because for any two vectors $U_i V_i$ is a scalar and invariant under rotations, $U'_i V'_i = U_i V_i$. The rank n tensor vector space then has an invariant scalar product $T \cdot S$ formed by contracting all indices on any pair of rank n tensors $T_{i_1 \dots i_n}, S_{i_1 \dots i_n}$.

In tensorial analysis *invariant tensors*, satisfying $I'_{i_1 \dots i_n} = I_{i_1 \dots i_n}$, are of critical importance. For rotations we have the Kronecker delta δ_{ij}

$$\delta'_{ij} = R_{ik} R_{jk} = \delta_{ij} \, , \tag{2.116}$$

as a consequence of the orthogonality property (2.1), and also the ε-symbol

$$\varepsilon'_{ijk} = R_{ij} R_{jm} R_{kn} \, \varepsilon_{lmn} = \det R \; \varepsilon_{ijk} = \varepsilon_{ijk} \, , \tag{2.117}$$

if $R \in SO(3)$. Any higher rank invariant tensor is formed in terms of Kronecker deltas and ε-symbols, for rank $2n$ we may use n Kronecker deltas and for rank $2n + 3$, n Kronecker deltas and one ε-symbol, since two ε-symbols can always be reduced to combinations of Kronecker deltas.

Using δ_{ij} and ε_{ijk} we may reduce tensors to ones of lower rank. Thus for a rank two tensor T_{ij}, $T_{ii} = \delta_{ij} T_{ij}$, which corresponds to the trace of the associated matrix, is rank zero and thus a scalar, and $V_i = \frac{1}{2} \varepsilon_{ijk} T_{jk}$ is a vector. Hence the 9 dimensional space formed by rank two tensors contains invariant, under rotations, subspaces of dimension one and dimension three formed by these scalars and vectors. In consequence rank 2 tensors do not form an irreducible representation space for rotations.

To demonstrate the decomposition of rank 2 tensors into irreducible components we write it as a sum of symmetric and antisymmetric tensors and re-express the latter as a vector. Separating out the trace of the symmetric tensor then gives

$$T_{ij} = S_{ij} + \varepsilon_{ijk} V_k + \tfrac{1}{3}\delta_{ij}\, T_{kk}\,, \tag{2.118}$$

for

$$S_{ij} = T_{(ij)} - \tfrac{1}{3}\delta_{ij}\, T_{kk}\,, \qquad V_i = \tfrac{1}{2}\varepsilon_{ijk} T_{jk}\,. \tag{2.119}$$

Each term in (2.118) transforms independently under rotations, so that for $T_{ij} \to T'_{ij}$, $S_{ij} \to S'_{ij}$, $V_k \to V'_k$, $T_{kk} \to T'_{kk} = T_{kk}$. The tensors S_{ij} are symmetric and traceless, $S_{kk} = 0$, and it is easy to see that they span a space of dimension 5.

These considerations may be generalised to higher rank but it is necessary to identify for each n those conditions on rank n tensors that ensure they form an irreducible space. If $S_{i_1\dots i_n}$ is to be irreducible under rotations then all lower rank tensors formed using invariant tensors must vanish. Hence we require

$$\delta_{i_r i_s} S_{i_1\dots i_n} = 0\,, \quad \varepsilon_{j i_r i_s} S_{i_1\dots i_n} = 0\,, \quad \text{for all} \quad r,s,\ 1 \le r < s \le n\,. \tag{2.120}$$

These conditions on the tensor S are easy to solve, it is necessary only that it is symmetric

$$S_{i_1\dots i_n} = S_{(i_1\dots i_n)}\,, \tag{2.121}$$

and also traceless on any pair of indices. With the symmetry condition (2.121) it is sufficient to require just

$$S_{i_1\dots i_{n-2}jj} = 0\,. \tag{2.122}$$

Such tensors then span a space $\mathcal{V}_n$ which is irreducible.

To count the dimension of $\mathcal{V}_n$ we first consider only symmetric tensors satisfying (2.121), belonging to the symmetrised n-fold tensor product, $\mathrm{sym}(\mathcal{V} \otimes \cdots \otimes \mathcal{V})$. Because of the symmetry not all tensors are independent of course, any tensor with r indices 1, s indices 2 and t indices 3 will be equal to

$$S_{\underbrace{1\dots1}_{r}\underbrace{2\dots2}_{s}\underbrace{3\dots3}_{t}} \quad \text{where} \quad r,s,t \ge 0\,,\ r+s+t = n\,. \tag{2.123}$$

Independent rank n symmetric tensors may then be counted by counting all r, s, t satisfying the conditions in (2.123), hence this gives

$$\dim\big(\mathrm{sym}(\underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{n})\big) = \tfrac{1}{2}(n+1)(n+2)\,. \tag{2.124}$$

To take the traceless conditions (2.122) into account it is sufficient, since taking the trace of rank n symmetric tensors gives rank $n-2$ symmetric tensors spanning a space of dimension $\frac{1}{2}(n-1)n$, to subtract the dimension for rank $n-2$ symmetric tensors giving

$$\dim \mathcal{V}_n = \tfrac{1}{2}(n+1)(n+2) - \tfrac{1}{2}(n-1)n = 2n+1\,. \tag{2.125}$$

Thus this irreducible space $\mathcal{V}_n$ may be identified with the representation space $j = n$, with n an integer.

For rank n symmetric traceless tensors an orthonormal basis $S^{(m)}_{i_1\dots i_n}$, labelled by m taking $2n+1$ values, satisfies $S^{(m)} \cdot S^{(m')} = \delta^{mm'}$. Such a basis may be used to define a corresponding set of spherical harmonics, depending on a unit vector $\hat{\mathbf{x}}$, by

$$Y_{nm}(\hat{\mathbf{x}}) = S^{(m)}_{i_1\dots i_n} \hat{x}_{i_1} \dots \hat{x}_{i_n} \,. \tag{2.126}$$

With two symmetric traceless tensors $S_{1,i_1\dots i_n}$ and $S_{2,i_1\dots i_m}$ then their product can be decomposed into symmetric traceless tensors by using the invariant tensors δ_{ij}, ε_{ijk}, generalising (2.118) and (2.119). Assuming $n \geq m$, and using only one ε-symbol since two may be reduced to Kronecker deltas, we may construct the following symmetric tensors

$$\begin{aligned} & S_{1,(i_1\dots i_{n-r}\, j_1\dots j_r}\, S_{2,i_{n-r+1}\dots i_{n+m-2r})\, j_1\dots j_r} \,, && r = 0, \dots m \,, \\ & \varepsilon_{jk(i_1}\, S_{1,i_2\dots i_{n-r}\, j_1\dots j_r\, j}\, S_{2,i_{n-r+1}\dots i_{n+m-1-2r})\, j_1\dots j_r\, k} \,, && r = 0, \dots m-1 \,. \end{aligned} \tag{2.127}$$

For each symmetric tensor there is a corresponding one which is traceless obtained by subtracting appropriate combinations of lower order tensors in conjunction with Kronecker deltas, as in (2.119) for the simplest case of rank two. Hence the product of the two symmetric tensors of rank n, m decomposes into irreducible tensors of rank $n + m - r$, $r = 0, 1, \dots, m$, in accord with general angular momentum product rules.

In quantum mechanics we may extend the notion of a tensor to operators acting on the quantum mechanical vector space. For a vector operator we require

$$U[R] V_i U[R]^{-1} = (R^{-1})_{ij} V_j \,, \tag{2.128}$$

as in (2.38), while for a rank n tensor operator

$$U[R] T_{i_1\dots i_n} U[R]^{-1} = (R^{-1})_{i_1 j_1} \dots (R^{-1})_{i_n j_n} T_{j_1\dots j_n} \,. \tag{2.129}$$

These may be decomposed into irreducible tensor operators as above. For infinitesimal rotations as in (2.8), with $U[R]$ correspondingly given by (2.32), then (2.128) gives

$$[J_i, V_j] = i\, \varepsilon_{ijk} V_k \,, \tag{2.130}$$

which is an alternative definition of a vector operator. From (2.129) we similarly get

$$[J_i, T_{j_1 j_2 \dots j_n}] = i\, \varepsilon_{ij_1k} T_{kj_2\dots j_n} + i\, \varepsilon_{ij_2k} T_{j_1 k\dots j_n} + \dots + i\, \varepsilon_{ij_nk} T_{j_1 j_2\dots k} \,. \tag{2.131}$$

The operators $\mathbf{x}, \mathbf{p}$ are examples of vector operators for the angular momentum operator given by $\mathbf{L} = \mathbf{x} \times \mathbf{p}$.

2.8.1 Spherical Harmonics

Rank n symmetric traceless tensors are directly related to *spherical harmonics*. If we choose an orthonormal basis for such tensors $S^{(m)}_{i_1\dots i_n}$, labelled by m taking $2n+1$ values and satisfying $S^{(m)} \cdot S^{(m')} = \delta^{mm'}$, then the basis may be used to define a corresponding complete set of orthogonal spherical harmonics on the unit sphere, depending on a unit vector $\hat{\mathbf{x}} \in S^2$, by

$$Y_{nm}(\hat{\mathbf{x}}) = S^{(m)}_{i_1 \dots i_n} \hat{x}_{i_1} \dots \hat{x}_{i_n} \, . \tag{2.132}$$

To discuss the scalar product for such harmonics we consider the three dimensional integrals

$$\int \mathrm{d}^3 x \; e^{-\mathbf{x}^2 + \mathbf{k}\cdot\mathbf{x}} = \pi^{\frac{3}{2}} \, e^{\frac{1}{4}\mathbf{k}^2} \quad \Rightarrow \quad \int \mathrm{d}^3 x \; e^{-\mathbf{x}^2} (\mathbf{k}\cdot\mathbf{x})^{2n} = \pi^{\frac{3}{2}} \, \frac{(2n)!}{2^{2n} n!} \, (\mathbf{k}^2)^n \, . \tag{2.133}$$

Since $\mathrm{d}^3 x = r^2 \mathrm{d}r \, \mathrm{d}\Omega$ and using $\int_0^\infty \mathrm{d}r \, r^{2n+2} e^{-r^2} = \frac{1}{2}\Gamma(n+\frac{3}{2})$ we obtain

$$\int_{S^2} \mathrm{d}\Omega \; (\mathbf{k}\cdot\hat{\mathbf{x}})^{2n} = 4\pi \, \frac{(2n)!}{2^{2n} n!} \, \frac{1}{(\frac{3}{2})_n} \, (\mathbf{k}^2)^n \, , \qquad (\tfrac{3}{2})_n = \tfrac{3}{2} \cdot \tfrac{5}{2} \dots (\tfrac{3}{2} + n - 1) \, . \tag{2.134}$$

If now $\mathbf{k} = \mathbf{t} + \bar{\mathbf{t}}$ with $\mathbf{t}^2 = \bar{\mathbf{t}}^2 = 0$ then

$$\int_{S^2} \mathrm{d}\Omega \; (\mathbf{t}\cdot\hat{\mathbf{x}})^n (\bar{\mathbf{t}}\cdot\hat{\mathbf{x}})^n = 4\pi \, \frac{n!}{2^n (\frac{3}{2})_n} \, (\mathbf{t}\cdot\bar{\mathbf{t}})^n \, . \tag{2.135}$$

Since $\mathbf{t}^2 = 0$ then $t_{i_1} \dots t_{i_n}$ defines a symmetric traceless tensor so that $(\mathbf{t}\cdot\hat{\mathbf{x}})^n$ represents a spherical harmonic. Applying the integral (2.135) then gives

$$\int_{S^2} \mathrm{d}\Omega \; Y_{nm}(\hat{\mathbf{x}}) \, Y_{nm'}(\hat{\mathbf{x}}) = 4\pi \, \frac{n!}{2^n (\frac{3}{2})_n} \, \delta^{mm'} \, . \tag{2.136}$$

2.9 Irreducible Tensor Operators

An alternative basis for irreducible tensor operators is achieved by requiring them to transform similarly to the angular momentum states $|j\, m\rangle$. An irreducible tensor operator in the standard angular momentum basis satisfies

Definition: The set of $(2k+1)$ operators $\{T_{kq}\}$ for

$$k \in \{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots \} \, , \tag{2.137}$$

and

$$q \in \{-k, -k+1, \dots, k-1, k\} \, , \tag{2.138}$$

for each k in (2.137), constitute a *tensor operator* of rank k if they satisfy the commutation relations

$$\begin{aligned} [J_3 \, , \, T_{kq}] &= q \, T_{kq} \, , \\ [J_\pm \, , \, T_{kq}] &= N^\pm_{kq} \, T_{kq\pm 1} \, , \end{aligned} \tag{2.139}$$

with $N^\pm_{kq}$ given by (2.59). This definition is of course modelled exactly on that for the $|j\, m\rangle$ states in (2.55) and (2.57) and ensures that we may treat it, from the point of view of its angular moment properties, just like a state $|k\, q\rangle$.

Examples:

If $k=0$ then $q=0$ and hence $[\mathbf{J}\,,\,T_{00}]=0$, *i.e.* T_{00} is just a scalar operator.

If $k=1$ then setting

$$V_{1\pm 1} = \mp\sqrt{\tfrac{1}{2}}(V_1 \pm iV_2)\,, \qquad V_{10} = V_3\,, \tag{2.140}$$

ensures that V_{1q} satisfy (2.139) for $k=1$ as a consequence of (2.130).

If $k=2$ we may form an irreducible tensor operator T_{2q} from two vectors V_i, U_i using Clebsch-Gordan coefficients

$$T_{2q} = \sum_{m,m'} V_{1m} U_{1m'}\, \langle 1m\, 1m' | 2q \rangle\,, \tag{2.141}$$

with $V_{1m}, U_{1m'}$ defined as in (2.140). This gives

$$\begin{aligned} T_{22} = V_{11}U_{11}\,, \quad T_{21} &= \sqrt{\tfrac{1}{2}}\big(V_{11}U_{10} + V_{10}U_{11}\big)\,, \\ T_{20} &= \sqrt{\tfrac{1}{6}}\big(V_{11}U_{1-1} + 2V_{10}U_{10} + V_{1-1}U_{11}\big)\,, \\ T_{2-1} &= \sqrt{\tfrac{1}{2}}\big(V_{10}U_{1-1} + V_{1-1}U_{10}\big)\,, \quad T_{2-2} = V_{1-1}U_{1-1}\,. \end{aligned} \tag{2.142}$$

The individual T_{2q} may all be expressed in terms of components of the symmetric traceless tensor $S_{ij} = V_{(i}U_{j)} - \frac{1}{3}\delta_{ij}\, V_k U_k$.

For irreducible tensor operators T_{kq} their matrix elements with respect to states $|\alpha, j\, m\rangle$, where α are any extra labels necessary to specify the states in addition to jm, are constrained by the theorem:

Wigner-Eckart Theorem.

$$\langle \alpha', j'\, m' | T_{kq} | \alpha, j\, m \rangle = \langle jm\, kq | j'\, m' \rangle\; C\,, \tag{2.143}$$

with $\langle jm\, kq | j'\, m' \rangle$ a Clebsch-Gordan coefficient. The crucial features of this result are:

(i) The dependence of the matrix element on m, q and m' is contained in the Clebsch-Gordan coefficient, and so is known completely. This ensures that the matrix element is non zero only if $j' \in \{j+k, j+k-1, \dots, |j-k|+1, |j-k|\}$.

(ii) The coefficient C depends only on j, j', k and on the particular operator and states involved. It may be written as

$$C = \langle \alpha' j' || T_k || \alpha j \rangle\,, \tag{2.144}$$

and is referred to as a reduced matrix element.

The case $k=q=0$ is an important special case. If $[\mathbf{J}\,,\,T_{00}]=0$, then T_{00} is scalar operator and we we have

$$\begin{aligned} \langle \alpha', j'\, m' | T_{00} | \alpha, j\, m \rangle &= \langle jm\, 00 | j' m' \rangle\; \langle \alpha' j' || T_0 || \alpha j \rangle \\ &= \delta_{jj'}\, \delta_{mm'}\; \langle \alpha' j' || T_0 || \alpha j \rangle\,, \end{aligned} \tag{2.145}$$

with reduced matrix-element independent of m.

To prove the Wigner-Eckart theorem we first note that $T_{kq}|\alpha, j\, m\rangle$ transforms under the action of the angular momentum operator $\mathbf{J}$ just like the product state $|k\,q\rangle_1|j\,m\rangle_2$ under the combined $\mathbf{J}_1+\mathbf{J}_2$. Hence

$$\sum_{q,m} T_{kq}|\alpha, j\, m\rangle\langle kq\, jm|JM\rangle = |J\,M\rangle \tag{2.146}$$

defines a set of states $\{|J\,M\rangle\}$ satisfying, by virtue of the definition of Clebsch-Gordan coefficients in (2.89),

$$J_3|J\,M\rangle = M|J\,M\rangle\,, \qquad J_\pm|J\,M\rangle = N^\pm_{J,M}|J\,M{\pm}1\rangle\,. \tag{2.147}$$

Although the states $|J\,M\rangle$ are not normalised, it follows then that

$$\langle\alpha', j'm'|J\,M\rangle = C_J\,\delta_{j'J}\delta_{m'M}\,, \tag{2.148}$$

defines a constant C_J which is independent of m', M. To verify this we note

$$\begin{aligned}\langle\alpha', JM|J\,M\rangle N^+_{JM-1} &= \langle\alpha', JM|J_+|J\,M{-}1\rangle \\ &= \langle\alpha', JM|J_-{}^\dagger|J\,M{-}1\rangle = \langle\alpha', JM{-}1|J\,M{-}1\rangle N^-_{JM}\,.\end{aligned} \tag{2.149}$$

Since $N^+_{JM-1} = N^-_{JM}$ we then have $\langle\alpha', JM|J\,M\rangle = \langle\alpha', JM{-}1|J\,M{-}1\rangle$ so that, for $m' = M$, (2.148) is independent of M. Inverting (2.146)

$$T_{kq}|\alpha, j\,m\rangle = \sum_{JM}|J\,M\rangle\langle kq\,jm|JM\rangle\,, \tag{2.150}$$

and then taking the matrix element with $\langle\alpha', j'm'|$ gives the Wigner-Eckart theorem, using (2.148), with $C_{j'} = \langle\alpha' j'||T_k||\alpha j\rangle$.

2.10 Spinors

For the rotation groups there are spinorial representations as well as those which can be described in terms of tensors, which are essentially all those which can be formed from multiple tensor products of vectors. For $SO(3)$, spinorial representations involve j being half integral and are obtained from the fundamental representation for $SU(2)$.

For the moment we generalise to $A = [A_\alpha{}^\beta] \in SU(r)$, satisfying (2.20), and consider a vector η belonging to the r-dimensional representation space for the fundamental representation and transforming as

$$\eta_\alpha \underset{A}{\to} \eta'_\alpha = A_\alpha{}^\beta\eta_\beta\,. \tag{2.151}$$

The extension to a tensor with n indices is straightforward

$$T_{\alpha_1\ldots\alpha_n} \underset{A}{\to} T'_{\alpha_1\ldots\alpha_n} = A_{\alpha_1}{}^{\beta_1}\ldots A_{\alpha_n}{}^{\beta_n}T_{\beta_1\ldots\beta_n}\,, \tag{2.152}$$

Since A is unitary

$$(A_\alpha{}^\beta)^* = (A^{-1})_\beta{}^\alpha\,. \tag{2.153}$$

The complex conjugation of (2.151) defines a transformation corresponding to the conjugate representation. If we define

$$\bar{\eta}^{\alpha} = (\eta_{\alpha})^{*} \,, \tag{2.154}$$

then using (2.153) allows the conjugate transformation rule to be written as

$$\bar{\eta}^{\alpha} \underset{A}{\to} \bar{\eta}'^{\alpha} = \bar{\eta}^{\beta}(A^{-1})_{\beta}{}^{\alpha} \,. \tag{2.155}$$

It is clear then that $\bar{\eta}^{\alpha}\eta_{\alpha}$ is a scalar. A general tensor may have both upper and lower indices, of course each upper index transforms as (2.151), each lower one as (2.155).

As with the previous discussion of tensors it is critical to identify the invariant tensors. For the case when $A \in SU(2)$ and $\alpha, \beta = 1, 2$ we have the two-dimensional ε-symbols, $\varepsilon^{\alpha\beta} = -\varepsilon^{\beta\alpha}$, $\varepsilon^{12} = 1$, and $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$, where it is convenient to take $\varepsilon_{12} = -1$. To verify $\varepsilon_{\alpha\beta}$ is invariant under the transformation corresponding to A we use

$$\varepsilon'_{\alpha\beta} = A_{\alpha}{}^{\gamma} A_{\beta}{}^{\delta} \varepsilon_{\gamma\delta} = \det A \, \varepsilon_{\alpha\beta} = \varepsilon_{\alpha\beta} \quad \text{for} \quad A \in SU(2) \,, \tag{2.156}$$

and similarly for $\varepsilon^{\alpha\beta}$. The Kronecker delta also forms an invariant tensor if there is one lower and one upper index since,

$$\delta'_{\alpha}{}^{\beta} = A_{\alpha}{}^{\gamma} \delta_{\gamma}{}^{\delta} (A^{-1})_{\delta}{}^{\beta} = \delta_{\alpha}{}^{\beta} \,. \tag{2.157}$$

For this two-dimensional case, with the preceding conventions, we have the relations

$$\varepsilon_{\alpha\beta}\,\varepsilon^{\gamma\delta} = -\delta_{\alpha}{}^{\gamma}\delta_{\beta}{}^{\delta} + \delta_{\alpha}{}^{\delta}\delta_{\beta}{}^{\gamma} \,, \qquad \varepsilon_{\alpha\gamma}\varepsilon^{\gamma\beta} = \delta_{\alpha}{}^{\beta} \,. \tag{2.158}$$

Rank n tensors as in (2.152) here span a vector space of dimension 2^n. To obtain an irreducible vector space under $SU(2)$ transformations we require that contractions with invariant tensors of lower rank give zero. For $S_{\alpha_1\dots\alpha_n}$ it is sufficient to impose $\varepsilon^{\alpha_r\alpha_s} S_{\alpha_1\dots\alpha_n} = 0$ for all $r < s$. The invariant tensors must then be totally symmetric $S_{\alpha_1\dots\alpha_n} = S_{(\alpha_1\dots\alpha_n)}$. To count these we may restrict to those of the form

$$S_{\underbrace{1\dots1}_{r}\,\underbrace{2\dots2}_{s}} \quad \text{where} \quad r = 0, \dots, n \,, \ r + s = n \,. \tag{2.159}$$

Hence there are $n + 1$ independent symmetric tensors $S_{\alpha_1\dots\alpha_n}$ so that the representation corresponds to $j = \frac{1}{2}n$.

The $SU(2)$ vectors η_{α} and also $\bar{\eta}^{\alpha}$ form $SO(3)$ spinors. For this case the two index invariant tensors $\varepsilon^{\alpha\beta}$ and $\varepsilon_{\alpha\beta}$ may be used to raise and lower indices. Hence we may define

$$\eta^{\alpha} = \varepsilon^{\alpha\beta}\,\eta_{\beta} \,, \tag{2.160}$$

which transforms as in (2.155) and correspondingly

$$\bar{\eta}_{\alpha} = \varepsilon_{\alpha\beta}\,\bar{\eta}^{\beta} \,, \tag{2.161}$$

As a consequence of (2.158) raising and then lowering an index leaves the spinors η_{α} unchanged, and similarly for $\bar{\eta}^{\alpha}$. In general the freedom to lower indices ensures that only $SU(2)$ tensors with lower indices, as in (2.152), need be considered.

For an infinitesimal $SU(2)$ transformation, with A as in (2.27), the corresponding change in a spinor arising from the transformation (2.151) is

$$\delta\eta_\alpha = -i\delta\theta\,\tfrac{1}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})_\alpha{}^\beta\eta_\beta\,. \tag{2.162}$$

For a tensor then correspondingly from (2.152)

$$\delta T_{\alpha_1\ldots\alpha_n} = -i\delta\theta\sum_{r=1}^{n}\tfrac{1}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})_{\alpha_r}{}^\beta T_{\alpha_1\ldots\alpha_{r-1}\beta\alpha_{r+1}\ldots\alpha_n}\,, \tag{2.163}$$

where there is a sum over contributions for each separate index.

Making use of (2.158) we have

$$\varepsilon^{\alpha\gamma}\varepsilon_{\beta\delta}\,\boldsymbol{\sigma}_\gamma{}^\delta = \boldsymbol{\sigma}_\beta{}^\alpha\,, \tag{2.164}$$

since $\mathrm{tr}(\boldsymbol{\sigma})=0$. From (2.164) we get

$$\varepsilon^{\alpha\gamma}\boldsymbol{\sigma}_\gamma{}^\beta = \varepsilon^{\beta\gamma}\boldsymbol{\sigma}_\gamma{}^\alpha\,, \tag{2.165}$$

showing that $(\varepsilon\boldsymbol{\sigma})^{\alpha\beta}$ form a set of three symmetric 2×2 matrices. Similar considerations also apply to $(\boldsymbol{\sigma}\varepsilon)_{\alpha\beta}$. The completeness relations for Pauli matrices can be expressed as

$$(\boldsymbol{\sigma}\varepsilon)_{\alpha\beta}\cdot(\varepsilon\boldsymbol{\sigma})^{\gamma\delta} = \delta_\alpha{}^\gamma\,\delta_\beta{}^\delta + \delta_\alpha{}^\delta\,\delta_\beta{}^\gamma\,, \qquad (\varepsilon\boldsymbol{\sigma})^{\alpha\beta}\cdot(\varepsilon\boldsymbol{\sigma})^{\gamma\delta} = -\varepsilon^{\alpha\gamma}\,\varepsilon^{\beta\delta} - \varepsilon^{\alpha\delta}\,\varepsilon^{\beta\gamma}\,. \tag{2.166}$$

The Pauli matrices allow symmetric spinorial tensors to be related to equivalent irreducible vectorial tensors. Thus we may define, for an even number of spinor indices, the tensor

$$T_{i_1\ldots i_n} = (\varepsilon\sigma_{i_1})^{\alpha_1\beta_1}\ldots(\varepsilon\sigma_{i_n})^{\alpha_n\beta_n}\,S_{\alpha_1\ldots\alpha_n\beta_1\ldots\beta_n}\,, \tag{2.167}$$

where it is easy to see that $T_{i_1\ldots i_n}$ is symmetric and also zero on contraction of any pair of indices, as a consequence of (2.166). For an odd number of indices we may further define

$$T_{\alpha\, i_1\ldots i_n} = (\varepsilon\sigma_{i_1})^{\alpha_1\beta_1}\ldots(\varepsilon\sigma_{i_n})^{\alpha_n\beta_n}\,S_{\alpha\alpha_1\ldots\alpha_n\beta_1\ldots\beta_n}\,, \tag{2.168}$$

where $T_{\alpha\, i_1\ldots i_n}$ is symmetric and traceless on the vectorial indices and satisfies the constraint

$$(\sigma_j)_\alpha{}^\beta T_{\beta\, i_1\ldots i_{n-1}j} = 0\,. \tag{2.169}$$

For two symmetric spinorial tensors $S_{1,\alpha_1\ldots\alpha_n}, S_{2,\beta_1\ldots\beta_m}$ their product are decomposed into symmetric rank $(n+m-2r)$-tensors, for $r=0,\ldots m$ if $n\geq m$, where for each r,

$$\varepsilon^{\beta_1\gamma_1}\ldots\varepsilon^{\beta_r\gamma_r}S_{1,(\alpha_1\ldots\alpha_{n-r}\,\beta_1\ldots\beta_r}S_{2,\alpha_{n-r+1}\ldots\alpha_{n+m-2r})\gamma_1\ldots\gamma_r}\,, \qquad r=0,\ldots,m\,. \tag{2.170}$$

For two spinors $\eta_{1\alpha},\eta_{2\alpha}$ the resulting decomposition into irreducible representation spaces is given by

$$\eta_{1\alpha}\,\eta_{2\beta} = \eta_{1(\alpha}\eta_{2\beta)} + \varepsilon_{\alpha\beta}\,\tfrac{1}{2}\eta_1{}^\gamma\eta_{2\gamma}\,, \tag{2.171}$$

where $\eta_{1(\alpha}\eta_{2\beta)}$ may be re-expressed as a vector using (2.167). This result demonstrates the decomposition of the product of two spin-$\frac{1}{2}$ representations into $j=0,1$, scalar, vector, irreducible components which are respectively antisymmetric, symmetric under interchange.

3 Isospin

The symmetry which played a significant role in the early days of nuclear and particle physics is isospin, which initially was based on the symmetry between neutrons and protons as far as nuclear forces were concerned. The symmetry group is again $SU(2)$ with of course the same mathematical properties as discussed in its applications to rotations, but with a very different physical interpretation. In order to distinguish this $SU(2)$ group from various others which arise in physics it is convenient to denote it as $SU(2)_I$.

From a modern perspective this symmetry arises since the basic QCD lagrangian depends on the Dirac u and d quark fields only in terms of

$$q = \begin{pmatrix} u \\ d \end{pmatrix} , \qquad \bar{q} = \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix} , \tag{3.1}$$

in such a way that it is invariant under $q \to Aq$, $\bar{q} \to \bar{q}A^{-1}$ for $A \in SU(2)$. This symmetry is violated by quark mass terms since $m_u \neq m_d$, although they are both tiny in relation to other mass scales, and also by electromagnetic interactions since u, d have different electric charges.

Neglecting such small effects there exist conserved charges $I_\pm, I_3$ which obey the $SU(2)$ commutation relations

$$[I_3, I_\pm] = \pm I_\pm \, , \quad [I_+, I_-] = 2I_3 \quad \text{or} \quad [I_a, I_b] = i\,\varepsilon_{abc} I_c \, , \tag{3.2}$$

as in (2.41a),(2.41b) or (2.36), and also commute with the Hamiltonian

$$[I_a, H] = 0 \, . \tag{3.3}$$

The particle states must then form multiplets, with essentially the same mass, which transform according to some $SU(2)_I$ representations. Each particle is represented by an isospin state $|I\, I_3\rangle$ which form the basis states for a representation of dimension $2I + 1$.

The simplest example is the proton and neutron which have $I = \frac{1}{2}$ and $I_3 = \frac{1}{2}, -\frac{1}{2}$ respectively. Neglecting other momentum and spin variables, the proton, neutron states are a doublet $(|p\rangle, |n\rangle)$ and we must have

$$I_3|p\rangle = \tfrac{1}{2}|p\rangle \, , \quad I_3|n\rangle = -\tfrac{1}{2}|n\rangle \, , \quad I_-|p\rangle = |n\rangle \, , \quad I_+|n\rangle = |p\rangle \, . \tag{3.4}$$

Other examples of $I = \frac{1}{2}$ doublets are the kaons $(|K^+\rangle, |K^0\rangle)$ and $(|\bar{K}^0\rangle, |K^-\rangle)$. The pions form a $I = 1$ triplet $(|\pi^+\rangle, |\pi^0\rangle, |\pi^-\rangle)$ so that

$$I_3(|\pi^+\rangle, |\pi^0\rangle, |\pi^-\rangle) = (|\pi^+\rangle, 0, -|\pi^-\rangle) \, , \qquad I_-|\pi^+\rangle = \sqrt{2}|\pi^0\rangle \, , \quad I_-|\pi^0\rangle = \sqrt{2}|\pi^-\rangle \, . \tag{3.5}$$

Another such triplet are the Σ baryons $(|\Sigma^+\rangle, |\Sigma^0\rangle, |\Sigma^-\rangle)$. Finally we note that the spin-$\frac{3}{2}$ baryons form a $I = \frac{3}{2}$ multiplet $(|\Delta^{++}\rangle, |\Delta^+\rangle, |\Delta^0\rangle, |\Delta^-\rangle)$. Low lying nuclei also belong to isospin multiplets, sometimes with quite high values of I. For each multiplet the electric charge for any particle is given by $Q = Q_0 + I_3$, where Q_0 has the same value for all particles in the multiplet.

Isospin symmetry has implications beyond that of just classification of particle states since the interactions between particles is also invariant. The fact that the isospin generators I_a are conserved, (3.3), constrains dynamical processes such as scattering. Consider a scattering process in which two particles, represented by isospin states $|I_1\, m_1\rangle, |I_2\, m_2\rangle$, scatter to produce two potentially different particles, with isospin states $|I_3\, m_3\rangle, |I_4\, m_4\rangle$. The scattering amplitude is $\langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle$ and to the extent that the dynamics are invariant under $SU(2)_I$ isospin transformations this amplitude must transform covariantly, i.e.

$$\sum_{m'_3,m'_4,m'_1,m'_2} D^{(I_3)}_{m'_3 m_3}(R)^* D^{(I_4)}_{m'_4 m_4}(R)^* D^{(I_1)}_{m'_1 m_1}(R) D^{(I_2)}_{m'_2 m_2}(R)\, \langle I_3 m'_3, I_4 m'_4|T|I_1 m'_1, I_2 m'_2\rangle$$
$$= \langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle\,. \qquad (3.6)$$

This condition is solved by decomposing the initial and final states into states $|IM\rangle$ with definite total isospin using Clebsch-Gordan coefficients,

$$\begin{aligned} |I_1 m_1, I_2 m_2\rangle &= \sum_{I,M} |IM\rangle\, \langle I_1 m_1, I_2 m_2\, |IM\rangle\,, \\ \langle I_3 m_3, I_4 m_4| &= \sum_{I,M} \langle I_3 m_3, I_4 m_4|IM\rangle \langle IM|\,, \end{aligned} \qquad (3.7)$$

since then, as in (2.145),

$$\langle I'M'|T|IM\rangle = A_I\, \delta_{I'I}\delta_{M'M}\,, \qquad (3.8)$$

as a consequence of T being an isospin singlet operator. Hence we have

$$\langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle = \sum_I A_I\, \langle I_3 m_3, I_4 m_4|IM\rangle \langle I_1 m_1, I_2 m_2\, |IM\rangle\,. \qquad (3.9)$$

The values of I which appear in this sum are restricted to those which can be formed by states with isospin I_1, I_2 and also I_3, I_4. The observed scattering cross sections depend only on $|\langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle|^2$.

As an illustration we consider πN scattering for $N = p, n$. In this case we can write

$$\begin{aligned} |\pi^+ p\rangle &= |\tfrac{3}{2}\, \tfrac{3}{2}\rangle\,, & |\pi^0 p\rangle &= \sqrt{\tfrac{2}{3}}|\tfrac{3}{2}\, \tfrac{1}{2}\rangle - \sqrt{\tfrac{1}{3}}|\tfrac{1}{2}\, \tfrac{1}{2}\rangle\,, \\ |\pi^0 n\rangle &= \sqrt{\tfrac{2}{3}}|\tfrac{3}{2}\, {-\tfrac{1}{2}}\rangle + \sqrt{\tfrac{1}{3}}|\tfrac{1}{2}\, {-\tfrac{1}{2}}\rangle\,, & |\pi^- p\rangle &= \sqrt{\tfrac{1}{3}}|\tfrac{3}{2}\, {-\tfrac{1}{2}}\rangle - \sqrt{\tfrac{2}{3}}|\tfrac{1}{2}\, {-\tfrac{1}{2}}\rangle\,, \end{aligned} \qquad (3.10)$$

using the Clebsch-Gordan coefficients which have been calculated in (2.107) and (2.108). Hence we have the results for the scattering amplitudes

$$\begin{aligned} \langle \pi^+ p|T|\pi^+ p\rangle &= A_{\frac{3}{2}}\,, \\ \langle \pi^- p|T|\pi^- p\rangle &= \tfrac{1}{3}\, A_{\frac{3}{2}} + \tfrac{2}{3}\, A_{\frac{1}{2}}\,, \\ \langle \pi^0 n|T|\pi^- p\rangle &= \tfrac{\sqrt{2}}{3}\big(A_{\frac{3}{2}} - A_{\frac{1}{2}}\big)\,, \end{aligned} \qquad (3.11)$$

so that three observable processes are reduced to two complex amplitudes $A_{\frac{3}{2}}, A_{\frac{1}{2}}$. For the observable cross sections

$$\sigma_{\pi^+ p\to\pi^+ p} = k\big|A_{\frac{3}{2}}\big|^2\,, \quad \sigma_{\pi^- p\to\pi^- p} = \tfrac{1}{9}k\big|A_{\frac{3}{2}} + 2A_{\frac{1}{2}}\big|^2\,, \quad \sigma_{\pi^- p\to\pi^0 n} = \tfrac{2}{9}k\big|A_{\frac{3}{2}} - A_{\frac{1}{2}}\big|^2\,, \qquad (3.12)$$

for k some isospin independent constant. There is no immediate algebraic relation between the cross sections since A_I are complex. However at the correct energy $A_{\frac{3}{2}}$ is large due to the $I = \frac{3}{2}$ Δ resonance, then the cross sections are in the ratios $1 : \frac{1}{9} : \frac{2}{9}$.

An example with more precise predictions arises with $NN \to \pi d$ scattering, where d is the deuteron, a pn bound state with $I = 0$. Hence the πd state has only $I = 1$. Decomposing NN states into states $|IM\rangle$ with $I = 1, 0$ we have $|pp\rangle = |11\rangle$, $|pn\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |00\rangle)$. Using this we obtain $\sigma_{pn\to\pi^0 d}/\sigma_{pp\to\pi^+ d} = \frac{1}{2}$.

The examples of isospin symmetry described here involve essentially low energy processes. Although it now appears rather fortuitous, depending on the lightness of the u, d quarks in comparison with the others, it was clearly the first step in the quest for higher symmetry groups in particle physics.

3.0.1 G-parity

G-parity is a discrete quantum number obtained by combining isospin with *charge conjugation*. Charge conjugation is a discrete $\mathbb{Z}_2$ symmetry where the unitary charge conjugation operator $\mathcal{C}$ acts on a particle state to give the associated anti-particle state with opposite charge. If these are different any associated phase factor is unphysical, since it may be absorbed into a redefinition of the states. In consequence the charge conjugation parity is well defined only for particle states with all conserved charges zero. For pions we have without any arbitrariness just

$$\mathcal{C}|\pi^0\rangle = |\pi^0\rangle \,. \tag{3.13}$$

The associated charged pion states are obtained, with standard isospin conventions, by $I_\pm|\pi^0\rangle = \sqrt{2}|\pi^\pm\rangle$. Since charge conjugation reverses the sign of all charges we must take $\mathcal{C}I_3\mathcal{C}^{-1} = -I_3$ and we require also $\mathcal{C}I_\pm\mathcal{C}^{-1} = -I_\mp$ (more generally if $\mathcal{C}I_+\mathcal{C}^{-1} = -e^{i\alpha}I_-$, $\mathcal{C}I_-\mathcal{C}^{-1} = -e^{-i\pi\alpha}I_+$ the dependence on α can be absorbed in a redefinition of $I_\pm$). By calculating $\mathcal{C}I_\pm|\pi^0\rangle$ we then determine unambiguously

$$\mathcal{C}|\pi^\pm\rangle = -|\pi^\mp\rangle \,. \tag{3.14}$$

G-parity is defined by combining C with an isospin rotation,

$$G = \mathcal{C}e^{-i\pi I_2} \,. \tag{3.15}$$

The action of $e^{-i\pi I_2}$ on an isospin multiplet is determined for any representation by (2.75). In this case we have

$$e^{-i\pi I_2}|\pi^+\rangle = |\pi^-\rangle \,, \quad e^{-i\pi I_2}|\pi^0\rangle = -|\pi^0\rangle \,, \quad e^{-i\pi I_2}|\pi^-\rangle = |\pi^+\rangle \,, \tag{3.16}$$

and hence on any pion state

$$G|\pi\rangle = -|\pi\rangle \,. \tag{3.17}$$

Conservation of G-parity ensures that in any $\pi\pi$ scattering process only even numbers of pions are produced. The notion of G-parity can be extended to other particles such as the spin one meson ω, with $I = 0$, and $\rho^\pm, \rho^0$, with $I = 1$. The neutral states have negative parity under charge conjugation so the G-parity of ω and the ρ's is respectively 1 and -1. This constrains various possible decay processes.

4 Relativistic Symmetries, Lorentz and Poincaré Groups

Symmetry under rotations plays a crucial role in atomic physics, isospin is part of nuclear physics but it is in high energy particle physics that relativistic Lorentz[6] transformations, forming the Lorentz group, have a vital importance. Extending Lorentz transformations by translations, in space and time, generates the Poincaré[7] group. Particle states can be considered to be defined as belonging to irreducible representations of the Poincaré Group.

4.1 Lorentz Group

For space-time coordinates $x^\mu = (x^0, x^i) \in \mathbb{R}^4$ then the *Lorentz group* is defined to be the group of transformations $x^\mu \to x'^\mu$ leaving the relativistic interval

$$x^2 \equiv g_{\mu\nu} x^\mu x^\nu \, , \qquad g_{00} = 1 \, , \, g_{0i} = g_{i0} = 0 \, , \, g_{ij} = -\delta_{ij} \, , \tag{4.1}$$

invariant. Assuming linearity a Lorentz transformation $x^\mu \to x'^\mu$

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \, , \tag{4.2}$$

ensures

$$x'^2 = x^2 \, , \tag{4.3}$$

which requires, for arbitrary x

$$g_{\sigma\rho} \Lambda^\sigma{}_\mu \Lambda^\rho{}_\nu = g_{\mu\nu} \, . \tag{4.4}$$

Alternatively in matrix language

$$\Lambda^T g \Lambda = g \, , \qquad \Lambda = [\Lambda^\mu{}_\nu] \, , \quad g = [g_{\mu\nu}] = \begin{pmatrix} 1 & 0 \\ 0 & -I_3 \end{pmatrix} \, . \tag{4.5}$$

Matrices satisfying (4.5) belong to the group $O(1,3) \simeq O(3,1)$.

In general we define contravariant and covariant vectors, V^μ and U_μ, under Lorentz transformations by

$$V^\mu \underset{\Lambda}{\to} V'^\mu = \Lambda^\mu{}_\nu V^\nu \, , \qquad U_\mu \underset{\Lambda}{\to} U'_\mu = U_\nu (\Lambda^{-1})^\nu{}_\mu \, . \tag{4.6}$$

It is easy to see, using (4.4) or (4.5), $V'^T g = V^T \Lambda^T g = V^T g \Lambda^{-1}$, that we may use $g_{\mu\nu}$ to lower indices, so that $g_{\mu\nu} V^\nu$ is a covariant vector. Defining the inverse $g^{\mu\nu}$, so that $g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu{}_\nu$, we may also raise indices, $g^{\mu\nu} U_\nu$ is a contravariant vector.

4.1.1 Proof of Linearity

We here demonstrate that the only transformations which satisfy (4.3) are linear. We rewrite (4.3) in the form

$$g_{\mu\nu} \mathrm{d}x'^\mu \mathrm{d}x'^\nu = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \, , \tag{4.7}$$

[6] Hendrik Antoon Lorentz, 1853-1928, Dutch. Nobel prize 1902.
[7] Jules Henri Poincaré, 1853-1912, French.

and consider infinitesimal transformations

$$x'^\mu = x^\mu + f^\mu(x)\,, \qquad \mathrm{d}x'^\mu = \mathrm{d}x^\mu + \partial_\sigma f^\mu(x)\,\mathrm{d}x^\sigma\,. \tag{4.8}$$

Substituting (4.8) into (4.7) and requiring this to hold for any infinitesimal $\mathrm{d}x^\mu$ gives

$$g_{\mu\sigma}\partial_\nu f^\sigma + g_{\sigma\nu}\partial_\mu f^\sigma = 0\,, \tag{4.9}$$

or, with $f_\mu = g_{\mu\sigma} f^\sigma$, we have the Killing equation,

$$\partial_\mu f_\nu + \partial_\nu f_\mu = 0\,. \tag{4.10}$$

Then we write

$$\partial_\omega(\partial_\mu f_\nu + \partial_\nu f_\mu) + \partial_\mu(\partial_\nu f_\omega + \partial_\omega f_\nu) - \partial_\nu(\partial_\omega f_\mu + \partial_\mu f_\omega) = 2\partial_\omega\partial_\mu f_\nu = 0\,. \tag{4.11}$$

The solution is obviously linear in x,

$$f_\mu(x) = a_\mu + \omega_{\mu\nu}x^\nu\,, \tag{4.12}$$

and then substituting back in (4.10) gives

$$\omega_{\mu\nu} + \omega_{\nu\mu} = 0\,. \tag{4.13}$$

For $a_\mu = 0$, (4.12) corresponds to an infinitesimal version of (4.2) with

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu\,, \qquad \omega^\mu{}_\nu = g^{\mu\sigma}\omega_{\sigma\nu}\,. \tag{4.14}$$

4.1.2 Structure of Lorentz Group

Taking the determinant of (4.5) gives

$$(\det\Lambda)^2 = 1 \quad \Rightarrow \quad \det\Lambda = \pm 1\,. \tag{4.15}$$

By considering the 00'th component we also get

$$(\Lambda^0{}_0)^2 = 1 + \textstyle\sum_i (\Lambda^0{}_i)^2 \geq 1 \quad \Rightarrow \quad \Lambda^0{}_0 \geq 1 \ \text{ or } \ \Lambda^0{}_0 \leq -1\,. \tag{4.16}$$

The Lorentz group has four components according to the signs of $\det\Lambda$ and $\Lambda^0{}_0$ since no continuous change in Λ can induce a change in these signs. For the component connected to the identity we have $\det\Lambda = 1$ and also $\Lambda^0{}_0 \geq 1$. This connected subgroup is denoted $SO(3,1)^\uparrow$.

Rotations form a subgroup of the Lorentz group, which is obtained by imposing $\Lambda^T\Lambda = I$ as well as (4.5). In this case the Lorentz transform matrix has the form,

$$\Lambda_R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}\,, \qquad R^T R = I_3\,, \tag{4.17}$$

where $R \in O(3)$, $\det R = \pm 1$, represents a three dimensional rotation or reflection, obviously $\Lambda_R\Lambda_{R'} = \Lambda_{RR'}$ forming a reducible representation of this subgroup.

Another special case is when

$$\Lambda = \Lambda^T \, . \tag{4.18}$$

To solve the constraint (4.5) we first write

$$\Lambda = \begin{pmatrix} \cosh\theta & \sinh\theta\, n^T \\ \sinh\theta\, n & \mathcal{B} \end{pmatrix} , \qquad \mathcal{B}^T = \mathcal{B} \, , n^T n = 1 \, , \tag{4.19}$$

where n is a 3-dimensional column vector, and then

$$\Lambda^T g \Lambda = \begin{pmatrix} 1 & \sinh\theta(\cosh\theta\, n^T - n^T\mathcal{B}) \\ \sinh\theta(\cosh\theta\, n - \mathcal{B}n) & \sinh^2\theta\, nn^T - \mathcal{B}^2 \end{pmatrix} . \tag{4.20}$$

Hence (4.5) requires

$$\mathcal{B}n = \cosh\theta\, n \, , \qquad \mathcal{B}^2 - \sinh^2\theta\, nn^T = I_3 \, . \tag{4.21}$$

The solution is just

$$\mathcal{B} = I_3 + (\cosh\theta - 1) nn^T \, . \tag{4.22}$$

The final expression for a general symmetric Lorentz transformation defining a boost is then

$$B(\theta, n) = \begin{pmatrix} \cosh\theta & \sinh\theta\, n^T \\ \sinh\theta\, n & I_3 + (\cosh\theta - 1) nn^T \end{pmatrix} , \tag{4.23}$$

where the parameter θ has an infinite range. Acting on x^μ, using vector notation,

$$\begin{aligned} x'^0 = {} & \cosh\theta\, x^0 + \sinh\theta\, \mathbf{n}\cdot\mathbf{x} \, , \\ \mathbf{x}' = {} & \mathbf{x} + (\cosh\theta - 1)\mathbf{n}\, \mathbf{n}\cdot\mathbf{x} + \sinh\theta\, \mathbf{n}\, x^0 \, . \end{aligned} \tag{4.24}$$

This represents a Lorentz boost with velocity $\mathbf{v} = \tanh\theta\; \mathbf{n}$.

Boosts do not form a subgroup since they are not closed under group composition, in general the product of two symmetric matrices is not symmetric, although there is a one parameter subgroup for n fixed and θ varying which is isomorphic to $SO(1,1)$ with matrices as in (1.59). With Λ_R as in (4.17) then for B as in (4.23)

$$\Lambda_R B(\theta, n) \Lambda_R{}^{-1} = B(\theta, Rn) \, , \tag{4.25}$$

gives the rotated Lorentz boost. Any Lorentz transformation can be written as at of a boost followed by a rotation. To show this we note that $\Lambda^T\Lambda$ is symmetric and positive so we may define $B = \sqrt{\Lambda^T\Lambda} = B^T$, corresponding to a boost. Then ΛB^{-1} defines a rotation since $(\Lambda B^{-1})^T \Lambda B^{-1} = B^{-1}\Lambda^T\Lambda B^{-1} = I$ and so $\Lambda B^{-1} = \Lambda_R$, or $\Lambda = \Lambda_R B$, with Λ_R of the form in (4.17).

4.2 Infinitesimal Lorentz Transformations and Commutation Relations

General infinitesimal Lorentz transformations have already been found in (4.14) with $\omega^\mu{}_\nu$ satisfying the conditions in (4.13). For two infinitesimal Lorentz transformations

$$\Lambda_1{}^\mu{}_\nu = \delta^\mu{}_\nu + \omega_1{}^\mu{}_\nu \, , \qquad \Lambda_2{}^\mu{}_\nu = \delta^\mu{}_\nu + \omega_2{}^\mu{}_\nu \, , \tag{4.26}$$

then

$$\Lambda^\mu{}_\nu = (\Lambda_2{}^{-1}\Lambda_1{}^{-1}\Lambda_2\,\Lambda_1)^\mu{}_\nu = \delta^\mu{}_\nu + [\omega_2,\omega_1]^\mu{}_\nu\,, \tag{4.27}$$

where it is clear that $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu$ if either $\omega_1{}^\mu{}_\nu$ or $\omega_2{}^\mu{}_\nu$ are zero.

For a relativistic quantum theory there must be unitary operators $U[\Lambda]$ acting on the associated vector space for each Lorentz transformation Λ which define a representation,

$$U[\Lambda_2]U[\Lambda_1] = U[\Lambda_2\Lambda_1]\,. \tag{4.28}$$

For an infinitesimal Lorentz transformation as in (4.13) we require

$$U[\Lambda] = 1 - i\,\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu}\,, \qquad M_{\mu\nu} = -M_{\nu\mu}\,. \tag{4.29}$$

$M_{\mu\nu}$ are the Lorentz group generators. Since we also have $U[\Lambda^{-1}] = 1 + i\,\frac{1}{2}\omega^{\mu\nu}M_{\mu\nu}$ (4.27) requires

$$\begin{aligned} U[\Lambda] &= 1 - i\,[\omega_2,\omega_1]^{\mu\nu}M_{\mu\nu} \\ &= U[\Lambda_2{}^{-1}]U[\Lambda_1{}^{-1}]U[\Lambda_2]U[\Lambda_1] \\ &= 1 - \big[\tfrac{1}{2}\,\omega_2{}^{\mu\nu}M_{\mu\nu},\,\tfrac{1}{2}\,\omega_1{}^{\sigma\rho}M_{\sigma\rho}\big]\,, \end{aligned} \tag{4.30}$$

or

$$\big[\tfrac{1}{2}\,\omega_2{}^{\mu\nu}M_{\mu\nu},\,\tfrac{1}{2}\,\omega_1{}^{\sigma\rho}M_{\sigma\rho}\big] = i\,[\omega_2,\omega_1]^{\mu\nu}M_{\mu\nu}\,, \quad [\omega_2,\omega_1]^{\mu\nu} = g_{\sigma\rho}(\omega_2{}^{\mu\sigma}\omega_1{}^{\rho\nu} - \omega_1{}^{\mu\sigma}\omega_2{}^{\rho\nu})\,. \tag{4.31}$$

Since this is valid for any ω_1, ω_1 we must have the commutation relations

$$[M_{\mu\nu}, M_{\sigma\rho}] = i\big(g_{\nu\sigma}\,M_{\mu\rho} - g_{\mu\sigma}\,M_{\nu\rho} - g_{\nu\rho}\,M_{\mu\sigma} + g_{\mu\rho}\,M_{\nu\sigma}\big)\,, \tag{4.32}$$

where the four terms on the right side are essentially dictated by antisymmetry under $\mu \leftrightarrow \nu$, $\sigma \leftrightarrow \rho$. For a unitary representation we must have

$$M_{\mu\nu}{}^\dagger = M_{\mu\nu}\,. \tag{4.33}$$

Just as in (2.128) we may define contravariant and covariant vector operators by requiring

$$U[\Lambda]V^\mu U[\Lambda]^{-1} = (\Lambda^{-1})^\mu{}_\nu V^\nu\,, \qquad U[\Lambda]U_\mu U[\Lambda]^{-1} = U_\nu\Lambda^\nu{}_\mu\,. \tag{4.34}$$

For an infinitesimal transformation, with Λ as in (4.14) and $U[\Lambda]$ as in (4.29), this gives

$$[M_{\mu\nu}, V^\sigma] = -i(\delta^\sigma{}_\mu V_\nu - \delta^\sigma{}_\nu V_\mu)\,, \qquad [M_{\mu\nu}, U_\sigma] = -i(g_{\mu\sigma}U_\nu - g_{\nu\sigma}U_\mu)\,. \tag{4.35}$$

To understand further the commutation relations (4.32) we decompose it into a purely spatial part and a part which mixes time and space (like magnetic and electric fields for the field strength $F_{\mu\nu}$. For spatial indices (4.32) becomes

$$[M_{ij}, M_{kl}] = -i(\delta_{jk}\,M_{il} - \delta_{ik}\,M_{jl} - \delta_{jl}\,M_{ik} + \delta_{il}\,M_{jk})\,. \tag{4.36}$$

Defining

$$J_m = \tfrac{1}{2}\varepsilon_{mij}M_{ij} \quad \Rightarrow \quad M_{ij} = \varepsilon_{ijm}J_m\,, \tag{4.37}$$

and similarly $J_n = \frac{1}{2}\varepsilon_{nkl}M_{kl}$ we get

$$[J_m, J_n] = -i\,\varepsilon_{mij}\varepsilon_{nkl}\,M_{il} = \tfrac{1}{2}i\,\varepsilon_{mnj}\varepsilon_{ilj}\,M_{il} = i\,\varepsilon_{mnj}J_j\,. \tag{4.38}$$

The commutation relations are identical with those obtained in (2.36) which is unsurprising since purely spatial Lorentz transformations reduce to the subgroup of rotations. As previously, $\mathbf{J} = (J_1, J_2, J_3)$ are identified with the angular momentum operators.

Besides the spatial commutators we consider also

$$[M_{ij}, M_{0k}] = -i(\delta_{jk}\,M_{0i} - \delta_{ik}\,M_{0j})\,, \tag{4.39}$$

and

$$[M_{0i}, M_{0j}] = -i\,M_{ij}\,. \tag{4.40}$$

Defining now

$$K_i = M_{0i}\,, \qquad K_i{}^\dagger = K_i\,, \tag{4.41}$$

and, using (4.37), (4.39) and (4.40) become

$$[J_i, K_j] = i\,\varepsilon_{ijk}K_k\,, \tag{4.42}$$

and

$$[K_i, K_j] = -i\,\varepsilon_{ijk}J_k\,. \tag{4.43}$$

The commutator (4.43) shows that $\mathbf{K} = (K_1, K_2, K_3)$ is a vector operator, as in (2.130). The $-$ sign in the commutator is (4.43) reflects the non compact structure of the Lorentz group $SO(3,1)$, if the group were $SO(4)$ then $g_{\mu\nu} \to \delta_{\mu\nu}$ and there would be a $+$.

For $\delta x^\mu = \omega^\mu{}_\nu x^\nu$ letting $\omega_{ij} = \varepsilon_{ijk}\theta_k$ and $\omega^0{}_i = \omega^i{}_0 = v_i$ then we have, for $t = x^0$ and $\mathbf{x} = (x^1, x^2, x^3)$,

$$\delta t = \mathbf{v}\cdot\mathbf{x}\,, \qquad \delta\mathbf{x} = \boldsymbol{\theta}\times\mathbf{x} + \mathbf{v}t\,, \tag{4.44}$$

representing an infinitesimal rotation and Lorentz boost. Using (4.29) with (4.37) and (4.41) gives correspondingly

$$U[\Lambda] = 1 - i\,\boldsymbol{\theta}\cdot\mathbf{J} + i\,\mathbf{v}\cdot\mathbf{K}\,, \tag{4.45}$$

which shows that $\mathbf{K}$ is associated with boosts in the same way as $\mathbf{J}$ is with rotations, as demonstrated by (2.32).

The commutation relations (4.38), (4.42) and (4.43) can be rewritten more simply by defining

$$J_i^\pm = \tfrac{1}{2}(J_i \pm iK_i)\,, \qquad \mathbf{J}^{+\dagger} = \mathbf{J}^-\,, \tag{4.46}$$

when they become

$$[J_i^+, J_j^+] = i\,\varepsilon_{ijk}J_k^+\,, \qquad [J_i^-, J_j^-] = i\,\varepsilon_{ijk}J_k^-\,, \qquad [J_i^+, J_j^-] = 0\,. \tag{4.47}$$

The commutation relations are then two commuting copies of the standard angular momentum commutation relations although the operators $\mathbf{J}^\pm$ are not hermitian.

4.3 Lorentz Group and Spinors

For $SO(3,1)$ there are corresponding spinorial representations just as for $SO(3)$. For $SO(3)$ a crucial role was played by the three Pauli matrices $\boldsymbol{\sigma}$. Here we define a four dimensional extension by

$$\sigma_\mu = (I, \boldsymbol{\sigma}) = {\sigma_\mu}^\dagger \,, \qquad \bar{\sigma}_\mu = (I, -\boldsymbol{\sigma}) = {\bar{\sigma}_\mu}^\dagger \,. \tag{4.48}$$

Both σ_μ and $\bar{\sigma}_\mu$ form a complete set of hermitian 2×2 matrices. As a consequence of (2.12) we have

$$\sigma_\mu \, \bar{\sigma}_\nu + \sigma_\nu \, \bar{\sigma}_\mu = 2g_{\mu\nu} \, I \,, \qquad \bar{\sigma}_\mu \, \sigma_\nu + \bar{\sigma}_\nu \, \sigma_\mu = 2g_{\mu\nu} \, I \,, \tag{4.49}$$

and also

$$\mathrm{tr}(\sigma_\mu \bar{\sigma}_\nu) = 2g_{\mu\nu} \,. \tag{4.50}$$

Hence for a 2×2 matrix A we may write $A = \frac{1}{2}\, \mathrm{tr}(\bar{\sigma}^\mu A)\, \sigma_\mu$.

4.3.1 Isomorphism $SO(3,1) \simeq Sl(2,\mathbb{C})/\mathbb{Z}_2$

The relation of $SO(3,1)$ to the group of 2×2 complex matrices with determinant one is an extension of the isomorphism $SO(3) \simeq SU(2)/\mathbb{Z}_2$. To demonstrate this we first describe the one to one correspondence between real 4-vectors x_μ and hermitian 2×2 matrices x where

$$x^\mu \to \mathrm{x} = \sigma_\mu x^\mu = \mathrm{x}^\dagger \,, \qquad x^\mu = \tfrac{1}{2}\, \mathrm{tr}(\bar{\sigma}^\mu \mathrm{x}) \,. \tag{4.51}$$

With the standard conventions in (2.11)

$$\mathrm{x} = \begin{pmatrix} x^0 + x^3 & x^1 - i x^2 \\ x^1 + i x^2 & x^0 - x^3 \end{pmatrix} \,. \tag{4.52}$$

Hence

$$\det \mathrm{x} = x^2 \equiv g_{\mu\nu} x^\mu x^\nu \,. \tag{4.53}$$

Defining

$$\bar{\mathrm{x}} = \bar{\sigma}_\mu x^\mu \,, \tag{4.54}$$

then (4.49) are equivalent to

$$\mathrm{x}\, \bar{\mathrm{x}} = x^2 I \,, \qquad \bar{\mathrm{x}} \mathrm{x} = x^2 I \,. \tag{4.55}$$

For any $A \in Sl(2,\mathbb{C})$ we may then define a linear transformation $x^\mu \to x'^\mu$ by

$$\mathrm{x} \underset{A}{\to} \mathrm{x}' = A \mathrm{x} A^\dagger = \mathrm{x}'^\dagger \,. \tag{4.56}$$

where, using $\det A = \det A^\dagger = 1$,

$$\det x' = \det x \quad \Rightarrow \quad x'^2 = x^2 \,. \tag{4.57}$$

Hence this must be a real Lorentz transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \,. \tag{4.58}$$

From (4.56) this requires

$$\sigma_\mu \Lambda^\mu{}_\nu = A\sigma_\nu A^\dagger \,, \qquad \Lambda^\mu{}_\nu = \tfrac{1}{2}\,\mathrm{tr}(\bar\sigma^\mu A\sigma_\nu A^\dagger)\,. \tag{4.59}$$

To establish the converse we may use $\sigma_\nu A^\dagger \bar\sigma^\nu = 2\,\mathrm{tr}(A^\dagger)\, I$ to give

$$\Lambda^\mu{}_\mu = |\mathrm{tr}(A)|^2 \,, \qquad \sigma_\mu \Lambda^\mu{}_\nu \bar\sigma^\nu = 2\,\mathrm{tr}(A^\dagger)\, A \,, \tag{4.60}$$

and hence, for $\mathrm{tr}A = e^{i\alpha}|\mathrm{tr}A|$,

$$A = e^{i\alpha}\, \frac{\sigma_\mu \Lambda^\mu{}_\nu \bar\sigma^\nu}{2\sqrt{\Lambda^\mu{}_\mu}} \,, \tag{4.61}$$

where the phase $e^{i\alpha}$ may be determined up to ± 1 by imposing $\det A = 1$. Hence for any $A \in Sl(2,\mathbb{C})$, $\pm A \leftrightarrow \Lambda$ for any $\Lambda \in SO(3,1)$.

As special cases if $A^\dagger = A^{-1}$, so that $A \in SU(2)$, it is easy to see that $x'^0 = x^0$ in (4.56) and this is just a rotation of $\mathbf{x}$ as given by (2.19) and (2.22). If $A^\dagger = A$ then Λ, given by (4.59), is symmetric so this is a boost. Taking

$$A_B(\theta, n) = \cosh \tfrac{1}{2}\theta \; I + \sinh \tfrac{1}{2}\theta \; \mathbf{n}\cdot\boldsymbol{\sigma} \,, \tag{4.62}$$

corresponds to the Lorentz boost in (4.23).

For a general infinitesimal Lorentz transformation as in (4.14) then, using $\Lambda^\mu{}_\mu = 4$ to this order and $\sigma_\mu \bar\sigma^\mu = 4\, I$, (4.61) gives

$$A = I + \tfrac{1}{4}\,\omega^{\mu\nu}\sigma_\mu \bar\sigma_\nu \,, \tag{4.63}$$

setting $\alpha = 0$, since $\mathrm{tr}(\omega^{\mu\nu}\sigma_\mu\bar\sigma_\nu) = 0$ as a consequence of $\omega^{\mu\nu} = -\omega^{\nu\mu}$. From (4.63)

$$A^\dagger = I - \tfrac{1}{4}\,\omega^{\mu\nu}\bar\sigma_\mu \sigma_\nu \,. \tag{4.64}$$

Alternatively, with these expressions for $A, A^\dagger$,

$$A\sigma_\rho A^\dagger = \sigma_\rho + \tfrac{1}{4}\,\omega^{\mu\nu}\big(\sigma_\mu\bar\sigma_\nu\sigma_\rho - \sigma_\rho\bar\sigma_\mu\sigma_\nu\big) = \sigma_\rho + \tfrac{1}{2}\,\omega^{\mu\nu}\big(g_{\nu\rho}\,\sigma_\mu - g_{\rho\mu}\,\sigma_\nu\big)\,, \tag{4.65}$$

using, from (4.49),

$$\sigma_\mu\bar\sigma_\nu\sigma_\rho = g_{\nu\rho}\,\sigma_\mu - \sigma_\mu\bar\sigma_\rho\sigma_\nu \,, \qquad \sigma_\rho\bar\sigma_\mu\sigma_\nu = 2g_{\rho\mu}\,\sigma_\nu - \sigma_\mu\bar\sigma_\rho\sigma_\nu \,, \tag{4.66}$$

and therefore (4.65) verifies $A\sigma_\nu A^\dagger = \sigma_\mu\Lambda^\mu{}_\nu$ with $\Lambda^\mu{}_\nu$ given by (4.14).

In general (4.63),(4.64) may be written as

$$A = I - i\,\tfrac{1}{2}\omega^{\mu\nu}s_{\mu\nu} \,, \quad A^\dagger = I + i\,\tfrac{1}{2}\omega^{\mu\nu}\bar s_{\mu\nu} \,, \qquad s_{\mu\nu} = \tfrac{1}{2}i\,\sigma_{[\mu}\bar\sigma_{\nu]} \,, \quad \bar s_{\mu\nu} = \tfrac{1}{2}i\,\bar\sigma_{[\mu}\sigma_{\nu]} \,, \tag{4.67}$$

where $s_{\mu\nu}, \bar s_{\mu\nu} = s_{\mu\nu}{}^\dagger$ are matrices each obeying the same commutation rules as $M_{\mu\nu}$ in (4.32). To verify this it is sufficient to check

$$s_{\mu\nu}\,\sigma_\rho - \sigma_\rho\,\bar s_{\mu\nu} = i(g_{\nu\rho}\,\sigma_\mu - g_{\mu\rho}\,\sigma_\nu)\,, \quad \bar s_{\mu\nu}\,\bar\sigma_\rho - \bar\sigma_\rho\, s_{\mu\nu} = i(g_{\nu\rho}\,\bar\sigma_\mu - g_{\mu\rho}\,\bar\sigma_\nu)\,. \tag{4.68}$$

4.3.2 Spinors, Dotted and Undotted Indices

In a similar fashion to the discussion in section 2.10 spinors are defined to transform under the action of the $Sl(2,\mathbb{C})$ matrix A. Fundamental spinors ψ, χ are required to transform as

$$\psi_\alpha \underset{A}{\to} A_\alpha{}^\beta \psi_\beta \,, \qquad \chi^\alpha \underset{A}{\to} \chi^\beta (A^{-1})_\beta{}^\alpha \,, \qquad \alpha, \beta = 1, 2 \,. \tag{4.69}$$

We may also, as hitherto, raise and lower spinor indices with the ε-symbols $\varepsilon^{\alpha\beta}, \varepsilon_{\alpha\beta}$, where $\varepsilon^{12} = \varepsilon_{21} = 1$, so that the representations defined by ψ_α, χ^α in (4.69) are equivalent

$$\psi^\alpha = \varepsilon^{\alpha\beta}\psi_\beta \,, \qquad \chi_\alpha = \varepsilon_{\alpha\beta}\,\chi^\beta \,, \tag{4.70}$$

as, since $\det A = 1$[8],

$$(A^{-1})_\beta{}^\alpha = \varepsilon^{\alpha\gamma} A_\gamma{}^\delta \,\varepsilon_{\delta\beta} \,. \tag{4.71}$$

The crucial difference between spinors for the Lorentz group $SO(3,1)$ and those for $SO(3)$ is that conjugation now defines an inequivalent representation. Hence there are two inequivalent two-component fundamental spinors. It is convenient to adopt the notational convention that the conjugate spinors obtained from ψ_α, χ^α have dotted indices, $\dot\alpha = 1, 2$. In general complex conjugation interchanges dotted and undotted spinor indices. For ψ, χ conjugation then defines the conjugate representation spinors

$$\bar\psi_{\dot\alpha} = (\psi_\alpha)^* \,, \qquad \bar\chi^{\dot\alpha} = (\chi^\alpha)^* \,, \tag{4.72}$$

which have the transformation rules, following from (4.69),

$$\bar\psi_{\dot\alpha} \underset{A}{\to} \bar\psi_{\dot\beta} (\bar A^{-1})^{\dot\beta}{}_{\dot\alpha} \,, \qquad \bar\chi^{\dot\alpha} \underset{A}{\to} \bar A^{\dot\alpha}{}_{\dot\beta}\, \bar\chi^{\dot\beta} \,, \tag{4.73}$$

for

$$(\bar A^{-1})^{\dot\alpha}{}_{\dot\beta} = (A_\beta{}^\alpha)^* \quad \text{or} \quad \bar A^{-1} = A^\dagger \,. \tag{4.74}$$

Both $A, \bar A \in Sl(2,\mathbb{C})$ and obey the same group multiplication rules, since $\overline{A_1 A_2} = \bar A_1 \bar A_2$. The corresponding ε-symbols, $\varepsilon^{\dot\alpha\dot\beta}, \varepsilon_{\dot\alpha\dot\beta}$, allow dotted indices to raised and lowered,

$$\bar\psi^{\dot\alpha} = \varepsilon^{\dot\alpha\dot\beta}\bar\psi_{\dot\beta} \,, \qquad \bar\chi_{\dot\alpha} = \varepsilon_{\dot\alpha\dot\beta}\,\bar\chi^{\dot\beta} \,, \tag{4.75}$$

in accord with the conjugation of (4.70).

In terms of these conventions the hermitian 2×2 matrices defined in (4.48) are written in terms of spinor index components as

$$(\sigma_\mu)_{\alpha\dot\alpha} \,, \qquad (\bar\sigma_\mu)^{\dot\alpha\alpha} \,, \tag{4.76}$$

where

$$(\bar\sigma_\mu)^{\dot\alpha\alpha} = \varepsilon^{\dot\alpha\dot\beta}\varepsilon^{\alpha\beta}(\sigma_\mu)_{\beta\dot\beta} \,, \qquad (\sigma_\mu)_{\alpha\dot\alpha} = \varepsilon_{\alpha\beta}\varepsilon_{\dot\alpha\dot\beta}(\bar\sigma_\mu)^{\dot\beta\beta} \,. \tag{4.77}$$

[8]Using (2.158), $\varepsilon^{\alpha\gamma} A_\gamma{}^\delta \,\varepsilon_{\delta\beta} = \delta_\beta{}^\alpha \mathrm{tr}(A) - A_\beta{}^\alpha = (A^{-1})_\beta{}^\alpha$, since for any 2×2 matrix the characteristic equation requires $A^2 - \mathrm{tr}(A)\; A + \det A\, I = 0$, so that if $\det A = 1$ then $A^{-1} = \mathrm{tr}(A)\, I - A$.

With the definitions in (4.51) and (4.54) then (4.77) requires $\mathrm{tr}(\mathrm{x}\bar{\mathrm{x}}) = 2\det x = 2x^2$. Using the definition of $\bar{A}$ we may rewrite (4.59) in the form

$$A\,\sigma_\nu \bar{A}^{-1} = \sigma_\mu \Lambda^\mu{}_\nu\,, \qquad \bar{A}\,\bar{\sigma}_\nu A^{-1} = \bar{\sigma}_\mu \Lambda^\mu{}_\nu\,, \tag{4.78}$$

showing the essential symmetry under $A \leftrightarrow \bar{A}$.

The independent fundamental spinors ψ, χ and their conjugates $\bar{\psi}, \bar{\chi}$ can be combined as a single 4-component *Dirac*[9] *spinor* together with its conjugate in the form

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}, \qquad \bar{\Psi} = \begin{pmatrix} \chi^\alpha & \bar{\psi}_{\dot{\alpha}} \end{pmatrix}, \tag{4.79}$$

where $\bar{\Psi} = \Psi^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Correspondingly there are 4×4 *Dirac matrices*

$$\gamma_\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix}. \tag{4.80}$$

These satisfy, by virtue of (4.49), the Dirac algebra

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2 g_{\mu\nu}\, I_4\,. \tag{4.81}$$

For these Dirac matrices

$$\gamma_0 \gamma_\mu \gamma_0 = \gamma_\mu{}^\dagger \qquad \text{since} \qquad \gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.82}$$

and from (4.77)

$$C\gamma_\mu C^{-1} = -\gamma_\mu{}^T \qquad \text{for} \quad C = \begin{pmatrix} \varepsilon^{\alpha\beta} & 0 \\ 0 & \varepsilon_{\dot{\alpha}\dot{\beta}} \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} \varepsilon_{\alpha\beta} & 0 \\ 0 & \varepsilon^{\dot{\alpha}\dot{\beta}} \end{pmatrix}. \tag{4.83}$$

4.3.3 Tensorial Representations

Both vector and spinor tensors are naturally defined in terms of the tensor products of vectors satisfying (4.6) and correspondingly spinors satisfying (4.69) or (4.73). Thus for a purely contragredient rank n tensor

$$T^{\mu_1\dots\mu_n} \underset{\Lambda}{\to} \Lambda^{\mu_1}{}_{\nu_1} \dots \Lambda^{\mu_n}{}_{\nu_n}\, T^{\nu_1\dots\nu_n}\,. \tag{4.84}$$

For a general spinor with $2j$ lower undotted indices and $2\bar{\jmath}$ lower dotted indices

$$\Upsilon_{\alpha_1\dots\alpha_{2j},\dot{\alpha}_1\dots\dot{\alpha}_{2\bar{\jmath}}} \underset{A}{\to} A_{\alpha_1}{}^{\beta_1} \dots A_{\alpha_{2j}}{}^{\beta_{2j}}\, \Upsilon_{\beta_1\dots\beta_{2j},\dot{\beta}_1\dots\dot{\beta}_{2\bar{\jmath}}}\, (\bar{A}^{-1})^{\dot{\beta}_1}{}_{\dot{\alpha}_1} \dots (\bar{A}^{-1})^{\dot{\beta}_{2\bar{\jmath}}}{}_{\dot{\alpha}_{2\bar{\jmath}}}\,. \tag{4.85}$$

The invariant tensors are just those already met together with the 4-index ε-symbol,

$$g^{\mu\nu}\,, \quad \varepsilon^{\mu\nu\sigma\rho}\,, \qquad \varepsilon_{\alpha\beta}\,, \quad \varepsilon_{\dot{\alpha}\dot{\beta}}\,, \tag{4.86}$$

[9] Paul Adrian Maurice Dirac, 1902-84, English. Nobel prize, 1933.

as well as all those derived from these by raising or lowering indices. Here $\varepsilon^{0123} = 1$ while $\varepsilon_{0123} = -1$.

To obtain irreducible tensors it is sufficient to consider spinorial tensors as in (4.85) which are totally symmetric in each set of indices

$$\Upsilon_{\alpha_1\ldots\alpha_{2j},\dot\alpha_1\ldots\dot\alpha_{2\bar\jmath}} = \Upsilon_{(\alpha_1\ldots\alpha_{2j}),(\dot\alpha_1\ldots\dot\alpha_{2\bar\jmath})}\,. \tag{4.87}$$

The resulting irreducible spinorial representation of $SO(3,1)$ is labelled $(j,\bar\jmath)$. Under complex conjugation $(j,\bar\jmath) \to (\bar\jmath, j)$. Extending the counting in the $SO(3)$ case, it is easy to see that the dimension of the space of such tensors is $(2j+1)(2\bar\jmath+1)$. The fundamental spinors transform according to the $(\frac12,0)$ and $(0,\frac12)$ representations while the Dirac spinor corresponds to $(\frac12,0)\oplus(0,\frac12)$. These representations are not unitary since there is no positive group invariant scalar product, for the simplest cases of a vector or a $(\frac12,0)$ spinor the scalar products $g_{\mu\nu}V^\mu V^\nu$ or $\varepsilon^{\beta\alpha}\psi_\alpha\psi_\beta$ clearly have no definite sign.

The tensors products of irreducible tensors as in (4.87) may be decomposed just as for $SO(3)$ spinors giving

$$(j_1,\bar\jmath_1)\otimes(j_2,\bar\jmath_2) \simeq \bigoplus_{\substack{|j_1-j_2|\le j\le j_1+j_2\\ |\bar\jmath_1-\bar\jmath_2|\le\bar\jmath\le\bar\jmath_1+\bar\jmath_2}} (j,\bar\jmath)\,. \tag{4.88}$$

Rank n vectorial tensors are related to spinorial tensors as in (4.85) for $2j = 2\bar\jmath = n$ by

$$T_{\mu_1\ldots\mu_n} = \Upsilon_{\alpha_1\ldots\alpha_n,\dot\alpha_1\ldots\dot\alpha_n}(\bar\sigma_{\mu_1})^{\dot\alpha_1\alpha_1}\ldots(\bar\sigma_{\mu_n})^{\dot\alpha_n\alpha_n}\,. \tag{4.89}$$

If Υ is irreducible, as in (4.87), corresponding to the $(\frac12 n,\frac12 n)$ real representation, then $T_{\mu_1\ldots\mu_n}$ is symmetric and traceless.

A corollary of $\varepsilon^{\mu\nu\sigma\rho}$ being an invariant tensor is, from (4.78),

$$A\,\varepsilon^{\mu\nu\sigma\rho}\sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho A^{-1} = \varepsilon^{\mu\nu\sigma\rho}\sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho\,, \quad \bar A\,\varepsilon^{\mu\nu\sigma\rho}\bar\sigma_\mu\sigma_\nu\bar\sigma_\sigma\sigma_\rho\bar A^{-1} = \varepsilon^{\mu\nu\sigma\rho}\bar\sigma_\mu\sigma_\nu\bar\sigma_\sigma\sigma_\rho\,. \tag{4.90}$$

By virtue of Schur's lemma these products of σ-matrices must be proportional to the identity. With (2.12) we get

$$\tfrac{1}{24}\,\varepsilon^{\mu\nu\sigma\rho}\sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho = \sigma_0\bar\sigma_1\sigma_2\bar\sigma_3 = i\,I\,, \quad \tfrac{1}{24}\,\varepsilon^{\mu\nu\sigma\rho}\bar\sigma_\mu\sigma_\nu\bar\sigma_\sigma\sigma_\rho = \bar\sigma_0\sigma_1\bar\sigma_2\sigma_3 = -i\,I\,, \tag{4.91}$$

using $(\sigma_0\bar\sigma_1\sigma_2\bar\sigma_3)^2 = \sigma_0\bar\sigma_1\sigma_2\bar\sigma_3\sigma_3\bar\sigma_2\sigma_1\bar\sigma_0 = -I$, and similarly $(\bar\sigma_0\sigma_1\bar\sigma_2\sigma_3)^2 = -I$, by virtue of (4.49). The two identities in (4.91) are related by conjugation. In terms of the Dirac matrices defined in (4.80)

$$\tfrac{1}{24}\,\varepsilon^{\mu\nu\sigma\rho}\gamma_\mu\gamma_\nu\gamma_\sigma\gamma_\rho = \gamma_0\gamma_1\gamma_2\gamma_3 = i\gamma_5\,, \qquad \gamma_5 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2\end{pmatrix}. \tag{4.92}$$

As a consequence of (4.91) we may further obtain[10]

$$\tfrac12\,\varepsilon^{\mu\nu\sigma\rho}\sigma_\sigma\bar\sigma_\rho = -i\,\sigma^{[\mu}\bar\sigma^{\nu]}\,, \qquad \tfrac12\,\varepsilon^{\mu\nu\sigma\rho}\bar\sigma_\sigma\sigma_\rho = i\,\bar\sigma^{[\mu}\sigma^{\nu]}\,. \tag{4.93}$$

[10] For a somewhat convoluted demonstration note, that since the indices only take four values, $\varepsilon^{\mu\nu\sigma\rho}\sigma_{[\mu}\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho\sigma_{\lambda]} = \frac15\varepsilon^{\mu\nu\sigma\rho}(\sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho\sigma_\lambda - \sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\lambda\sigma_\rho + \sigma_\mu\bar\sigma_\nu\sigma_\lambda\bar\sigma_\sigma\sigma_\rho - \sigma_\mu\bar\sigma_\lambda\sigma_\nu\bar\sigma_\sigma\sigma_\rho + \sigma_\lambda\bar\sigma_\mu\sigma_\nu\bar\sigma_\sigma\sigma_\rho) = 0$. Then using (4.49) move σ_λ or $\bar\sigma_\lambda$ to the right giving $\varepsilon^{\mu\nu\sigma\rho}\sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho\sigma_\lambda + 4\,\varepsilon_\lambda{}^{\nu\sigma\rho}\sigma_\nu\bar\sigma_\sigma\sigma_\rho = 0$. Hence, with (4.91), $i\sigma_\lambda = -\frac16\varepsilon_\lambda{}^{\nu\sigma\rho}\sigma_\nu\bar\sigma_\sigma\sigma_\rho$. Similarly $i\bar\sigma_\mu = \frac16\varepsilon_\mu{}^{\nu\sigma\rho}\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho$. Using these results, $i(\sigma_\lambda\bar\sigma_\mu - \sigma_\mu\bar\sigma_\sigma) = -\frac16\varepsilon_\lambda{}^{\nu\sigma\rho}(\sigma_\nu\bar\sigma_\sigma\sigma_\rho\bar\sigma_\mu + \sigma_\mu\bar\sigma_\nu\sigma_\sigma\bar\sigma_\rho)$. The right hand side may be simplified using (4.49) again and leads to just (4.93).

Since $\mathrm{tr}(\sigma_{[\mu}\bar\sigma_{\nu]}) = \mathrm{tr}(\bar\sigma_{[\mu}\sigma_{\nu]}) = 0$, $(\varepsilon\sigma_{[\mu}\bar\sigma_{\nu]})^{\alpha\beta}$, $(\bar\sigma_{[\mu}\sigma_{\nu]}\varepsilon)^{\dot\alpha\dot\beta}$ are symmetric in $\alpha \leftrightarrow \beta$, $\dot\alpha \leftrightarrow \dot\beta$ respectively so that for $(1,0)$ or $(0,1)$ representations there are associated antisymmetric tensors

$$f_{\mu\nu} = \tfrac{1}{2}(\varepsilon\sigma_{[\mu}\bar\sigma_{\nu]})^{\alpha\beta}\,\Upsilon_{\alpha\beta}\,, \qquad \bar f_{\mu\nu} = \tfrac{1}{2}(\bar\sigma_{[\mu}\sigma_{\nu]}\varepsilon)^{\dot\alpha\dot\beta}\,\bar\Upsilon_{\dot\alpha\dot\beta}\,, \tag{4.94}$$

which satisfy $f_{\mu\nu} = \frac{1}{2}i\,\varepsilon_{\mu\nu}{}^{\sigma\rho}f_{\sigma\rho}$, $\bar f_{\mu\nu} = -\frac{1}{2}i\,\varepsilon_{\mu\nu}{}^{\sigma\rho}\bar f_{\sigma\rho}$. Only $f_{\mu\nu} + \bar f_{\mu\nu}$ is a real tensor.

4.4 Poincaré Group

The complete space-time symmetry group includes translations as well as Lorentz transformations. For a Lorentz transformation Λ and a translation a the combined transformation denoted by (Λ, a) gives

$$x^\mu \underset{(\Lambda,a)}{\to} x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\nu\,. \tag{4.95}$$

These transformations form a group since

$$(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2)\,, \qquad (\Lambda, a)^{-1} = (\Lambda^{-1}, -\Lambda^{-1}a)\,, \tag{4.96}$$

with identity $(I, 0)$. The corresponding group is the *Poincaré group*, sometimes denoted as $ISO(3,1)$, if $\det\Lambda = 1$. It contains the translation group T_4, formed by (I, a), as a normal subgroup and also the Lorentz group, formed by $(\Lambda, 0)$. A general element may be written as $(\Lambda, a) = (I, a)(\Lambda, 0)$ and the Poincaré Group can be identified with the semi-direct product $O(3,1) \ltimes T_4$.

If we define

$$(\Lambda, a) = (\Lambda_2, a_2)^{-1}(\Lambda_1, a_1)^{-1}(\Lambda_2, a_2)(\Lambda_1, a_1)\,, \tag{4.97}$$

then direct calculation gives

$$\Lambda = \Lambda_2{}^{-1}\Lambda_1{}^{-1}\Lambda_2\,\Lambda_1\,, \qquad a = \Lambda_2{}^{-1}\Lambda_1{}^{-1}(\Lambda_2 a_1 - \Lambda_1 a_2 - a_1 + a_2)\,. \tag{4.98}$$

For infinitesimal transformations as in (4.26) we then have

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + [\omega_2, \omega_1]^\mu{}_\nu\,, \qquad a^\mu = \omega_2{}^\mu{}_\nu a_1{}^\nu - \omega_1{}^\mu{}_\nu a_2{}^\nu\,. \tag{4.99}$$

In a quantum theory there are associated unitary operators $U[\Lambda, a]$ such that

$$U[\Lambda_2, a_2]U[\Lambda_1, a_1] = U[\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2]\,. \tag{4.100}$$

For an infinitesimal Lorentz transformation as in (4.14) and also for infinitesimal a we require

$$U[\Lambda, a] = 1 - i\,\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu} + i\,a^\mu P_\mu\,, \qquad P_\mu{}^\dagger = P_\mu\,, \tag{4.101}$$

defining the generators P_μ in addition to $M_{\mu\nu} = -M_{\nu\mu}$ discussed in section 4.2. To derive the commutation relations we extend (4.30) to give

$$\begin{aligned} U[\Lambda, a] &= 1 - i\,[\omega_2, \omega_1]^{\mu\nu}M_{\mu\nu} + i\,(\omega_2 a_1 - \omega_1 a_2)^\mu P_\mu \\ &= U[\Lambda_2, a_2]^{-1}U[\Lambda_1, a_1]^{-1}U[\Lambda_2, a_2]U[\Lambda_1, a_1] \\ &= 1 - \big[\tfrac{1}{2}\,\omega_2{}^{\mu\nu}M_{\mu\nu} - a_2{}^\mu P_\mu, \tfrac{1}{2}\,\omega_1{}^{\sigma\rho}M_{\sigma\rho} - a_1{}^\sigma P_\sigma\big]\,. \end{aligned} \tag{4.102}$$

Hence, in addition to the $[M, M]$ commutators which are given in (4.31) and (4.32), we must have

$$\left[\tfrac{1}{2}\,\omega_1{}^{\sigma\rho}M_{\sigma\rho},\, a_2{}^{\mu}P_\mu\right] = i\,(\omega_1 a_2)^\mu P_\mu\,, \qquad \left[a_2{}^{\mu}P_\mu,\, a_1{}^{\sigma}P_\sigma\right] = 0\,, \tag{4.103}$$

or

$$[M_{\mu\nu}, P_\sigma] = i(g_{\nu\sigma}\,P_\mu - g_{\mu\sigma}\,P_\nu)\,, \qquad [P_\mu, P_\sigma] = 0\,. \tag{4.104}$$

This agrees with general form in (4.35) and shows that P_μ is a covariant 4-vector operator. Since $(\Lambda, 0)(I, a)(\Lambda, 0)^{-1} = (I, \Lambda a)$ and using $(\Lambda a)^\mu P_\mu = a^\mu (P\Lambda)_\mu$ we have for finite Lorentz transformations

$$U[\Lambda, 0]\, P_\mu\, U[\Lambda, 0]^{-1} = P_\nu \Lambda^\nu{}_\mu\,. \tag{4.105}$$

If we decompose

$$P^\mu = (H, \mathbf{P})\,, \qquad P_\mu = (H, -\mathbf{P})\,, \tag{4.106}$$

then using (4.37) and (4.41) the commutation relations become

$$[J_i, H] = 0\,, \qquad [J_i, P_j] = i\,\varepsilon_{ijk}P_k\,, \tag{4.107}$$

and

$$[K_i, H] = i\,P_i\,, \qquad [K_i, P_j] = i\,\delta_{ij}\,H\,. \tag{4.108}$$

4.5 Irreducible Representations of the Poincaré Group

It is convenient to write

$$U[\Lambda, a] = T[a]U[\Lambda]\,, \qquad U[\Lambda, 0] = U[\Lambda]\,, \quad T[a] = U[I, a]\,, \tag{4.109}$$

where $T[a]$ are unitary operators corresponding to the abelian translation group T_4. In general

$$T[a] = e^{ia^\mu P_\mu}\,. \tag{4.110}$$

As a consequence of (4.100)

$$U[\Lambda]T[a] = T[\Lambda a]U[\Lambda]\,. \tag{4.111}$$

The irreducible representations of the the translation subgroup T_4 of the Poincaré Group are one-dimensional and are defined in terms of vector $|p\rangle$ such that

$$P_\mu|p\rangle = p_\mu|p\rangle\,, \qquad T[a]|p\rangle = e^{ia^\mu p_\mu}|p\rangle\,, \tag{4.112}$$

for any real 4-vector p_μ which labels the representation. As a consequence of (4.105)

$$P_\mu U[\Lambda]|p\rangle = (p\Lambda^{-1})_\mu U[\Lambda]|p\rangle\,, \tag{4.113}$$

so that $U[\Lambda]$ acting on the states $\{|p\rangle\}$ generates a vector space $\mathcal{V}$ such that $|p'\rangle, |p\rangle$ belong to $\mathcal{V}$ if $p'{}_\mu = (p\Lambda^{-1})_\mu$ for some Lorentz transformation Λ. All such p', p satisfy $p'^2 = p^2$ and conversely for any p', p satisfying this there is a Lorentz transformation linking p', p. The physically relevant cases arise for $p^2 \geq 0$ and also we require, restricting $\Lambda \in SO(3,1)^\uparrow$, $p_0, p'_0 \geq 0$.

The construction of representations of the Poincaré group is essentially identical with the method of induced representations described in 1.4.1 for $G = SO(3,1)^{\uparrow} \ltimes T_4$. A subgroup H is identified by choosing a particular momentum $\mathring{p}$ and then defining

$$G_{\mathring{p}} = \{\Lambda : \Lambda\mathring{p} = \mathring{p}\}\,, \tag{4.114}$$

the *stability group* or *little group* for $\mathring{p}$, the subgroup of $SO(3,1)^{\uparrow}$ leaving $\mathring{p}$ invariant. For a space $\mathcal{V}_{\mathring{p}}$ formed by states $\{|\mathring{p}\rangle\}$ (additional labels are here suppressed) where

$$P_\mu|\mathring{p}\rangle = \mathring{p}_\mu|\mathring{p}\rangle\,, \qquad T[a]|\mathring{p}\rangle = e^{ia^\mu \mathring{p}_\mu}|p\rangle\,, \tag{4.115}$$

then $\mathcal{V}_{\mathring{p}}$ must form a representation space for $G_{\mathring{p}}$ since $U[\Lambda]\,|\mathring{p}\rangle \in \mathcal{V}_{\mathring{p}}$ for any $\Lambda \in G_{\mathring{p}}$ by virtue of (4.114). Hence $\mathcal{V}_{\mathring{p}}$ defines a representation for $H = G_{\mathring{p}} \otimes T_4$. The cosets G/H are then labelled, for all p such that $p^2 = \mathring{p}^2$, by any $L(p) \in SO(3,1)^{\uparrow}$ where

$$p_\mu = (\mathring{p}L(p)^{-1})_\mu\,, \quad \text{or equivalently} \quad p^\mu = L(p)^\mu{}_\nu \mathring{p}^\nu\,, \tag{4.116}$$

and, following the method of induced representations, a representation space for a representation of G is then defined in terms of a basis

$$|p\rangle = U[L(p)]\,|\mathring{p}\rangle \in \mathcal{V}_p\,, \qquad \text{for all} \quad |\mathring{p}\rangle \in \mathcal{V}_{\mathring{p}}\,. \tag{4.117}$$

Finding a representation of the Poincaré group then requires just the determination of $U[\Lambda]\,|p\rangle$ for arbitrary Λ. Clearly, by virtue of (4.113), $U[\Lambda]\,|p\rangle$ must be a linear combination of all states $\{|p'\rangle\}$ where $p'^\mu = \Lambda^\mu{}_\nu p^\nu$. Since $p'^\mu = L(p')^\mu{}_\nu \mathring{p}^\nu$ we have

$$\big(L(p')^{-1}\Lambda L(p)\big)^\mu{}_\nu \mathring{p}^\nu = \mathring{p}^\mu\,. \tag{4.118}$$

It follows that

$$L(\Lambda p)^{-1}\Lambda L(p) = \mathring{\Lambda}_p \in G_{\mathring{p}}\,, \tag{4.119}$$

and hence

$$U[\Lambda]\,|p\rangle = U[L(\Lambda p)]\,U[\mathring{\Lambda}_p]\,|\mathring{p}\rangle \in \mathcal{V}_{\Lambda p}\,, \tag{4.120}$$

where $U[\mathring{\Lambda}_p]|\mathring{p}\rangle$ is determined by the representation of $G_{\mathring{p}}$ on $\mathcal{V}_{\mathring{p}}$.

For physical interest there are two distinct cases to consider.

4.5.1 Massive Representations

Here we assume $p^2 = m^2 > 0$. It is simplest to choose for $\mathring{p}$ the particular momentum

$$\mathring{p}^\mu = (m, \mathbf{0})\,, \tag{4.121}$$

and, since $\mathring{p}$ has no spatial part, then

$$G_{\mathring{p}} \simeq SO(3)\,, \tag{4.122}$$

since the condition $\Lambda\mathring{p} = \mathring{p}$ restricts Λ to the form given in (4.17). As in (4.116) $L(p)$, for any p such that $p^2 = m^2$, $p_0 > 0$, is then a Lorentz transformation such that $p^\mu = L(p)^\mu{}_\nu \mathring{p}^\nu$. With (4.17) defining Λ_R for any $R \in SO(3)$, then (4.119) requires

$$L(\Lambda p)^{-1}\Lambda L(p) = \Lambda_{\mathcal{R}(p,\Lambda)}\,, \qquad \mathcal{R}(p,\Lambda) \in SO(3)\,. \tag{4.123}$$

$\mathcal{R}(p,\Lambda)$ is a *Wigner rotation.* (4.123) ensures that

$$U[L(\Lambda p)]^{-1}U[\Lambda]\,|p\rangle = U[\Lambda_{\mathcal{R}(p,\Lambda)}]\,|\mathring{p}\rangle\,. \tag{4.124}$$

For any R, $U[\Lambda_R]|\mathring{p}\rangle$ is an eigenvector of P^μ with eigenvalue $\mathring{p}^\mu$ and so is a linear combination of all states $\{|\mathring{p}\rangle\}$. In this case $\mathcal{V}_{\mathring{p}}$ must form a representation space for $SO(3)$. For irreducible representations $\mathcal{V}_{\mathring{p}}$ then has a basis, as described in section 2.5, which here we label by $s = 0, \frac{1}{2}, 1\ldots$ and $s_3 = -s, -s+1, \ldots, s$. Hence, assuming $\{|\mathring{p}, s\, s_3\rangle\}$ forms such an irreducible space,

$$U[\Lambda_R]\,|\mathring{p}, s\, s_3\rangle = \sum_{s'_3} |\mathring{p}, s\, s'_3\rangle\, D^{(s)}_{s'_3\, s_3}(R)\,, \tag{4.125}$$

with $D^{(s)}(R)$ standard $SO(3)$ rotation matrices. Extending the definition (4.117) to define a corresponding basis for any p

$$|p, s\, s_3\rangle = U[L(p)]\,|\mathring{p}, s\, s_3\rangle\,, \tag{4.126}$$

then applying (4.125) in (4.124) gives

$$U[\Lambda]\,|p, s\, s_3\rangle = \sum_{s'_3} |\Lambda p, s\, s'_3\rangle\, D^{(s)}_{s'_3\, s_3}\big(\mathcal{R}(p,\Lambda)\big)\,. \tag{4.127}$$

The states $\{|p, s\, s_3\rangle : p^2 = m^2, p_0 > 0\}$ then provide a basis for an irreducible representation space $\mathcal{V}_{m,s}$ for $SO(3,1)^\uparrow$. The representation extends to the full Poincaré group since for translations, from (4.112),

$$T[a]\,|p, s\, s_3\rangle = e^{ip_\mu a^\mu}|p, s\, s_3\rangle\,. \tag{4.128}$$

The states $|p, ss_3\rangle$ are obviously interpreted as single particle states for a particle with mass m and spin s.

In terms of these states there is a group invariant scalar product

$$\langle p', s\, s'_3|p, s\, s_3\rangle = (2\pi)^3 2p^0\, \delta^3(\mathbf{p}' - \mathbf{p})\, \delta_{s_3 s'_3}\,, \tag{4.129}$$

which is positive so the representation is unitary.

The precise definition of the representation depends on the choice of $L(p)$ satisfying (4.116). This doest not specify $L(p)$ uniquely since if $L(p)$ is one solution so is $L(p)\Lambda$ for any $\Lambda \in G_{\mathring{p}}$. One definite choice is to take

$$L(p) = B(\alpha, n)\,, \quad \text{for} \quad p^\mu = m(\cosh\alpha, \sinh\alpha\, \mathbf{n})\,, \tag{4.130}$$

where $B(\alpha, n)$ is the boost Lorentz transformation defined in (4.23). An alternative prescription is

$$L(p) = \Lambda_{R(n)}B(\alpha, e_3)\,, \quad \text{for} \quad R(n) = R(\phi, e_3)R(\theta, e_2)R(-\phi, e_3)\,, \tag{4.131}$$

where α is determined by p^0 as in (4.130), Λ_R, as in (4.17), corresponds to a rotation R, and θ, ϕ are the polar angles specifying n, so that $R(n)$ rotates e_3 into n, $R_{i3}(n) = n_i$. The two definitions, (4.130) and (4.131), give different but equivalent bases for $\mathcal{V}_{m,s}$.

If we consider a rotation Λ_R and use the definition (4.130) then $L(\Lambda_R p) = B(\alpha, Rn)$ and, by virtue of (4.25),

$$B(\alpha, Rn)^{-1}\Lambda_R B(\alpha, n) = \Lambda_R \,. \tag{4.132}$$

The Wigner rotation given by (4.123) hence becomes, with this definition of $L(p)$, just the original rotation

$$\mathcal{R}(p, \Lambda_R) = R \,, \tag{4.133}$$

so that (4.125) extends to any momentum p.

4.5.2 Massless Representations

The construction of representations for the massless case can be carried out in a similar fashion to that just considered. When $p^2 = 0$ then the method requires choosing a particular momentum $\mathring{p}$ satisfying this from which all other momenta with $p^2 = 0$ can be obtained by a Lorentz transformation. There is no rest frame as in (4.121) and we now take

$$\mathring{p}^\mu = \mathring{\omega}(1, 0, 0, 1) \,, \qquad \mathring{\omega} > 0 \,, \tag{4.134}$$

with $\mathring{\omega}$ some arbitrary fixed choice. It is then necessary to identify the little group in this case as defined by (4.114). To achieve this we consider infinitesimal Lorentz transformations as in (4.14) when the necessary requirement reduces to

$$\omega^\mu{}_\nu \mathring{p}^\nu = 0 \,, \qquad \omega^{\mu\nu} = -\omega^{\nu\mu} \,. \tag{4.135}$$

This linear equation is easy to solve giving

$$\omega^0{}_3 = 0 \,, \qquad \omega^1{}_0 = -\omega^1{}_3 \,, \qquad \omega^2{}_0 = -\omega^2{}_3 \,, \qquad \omega^3{}_0 = 0 \,. \tag{4.136}$$

These reduce the six independent $\omega^{\mu\nu} = -\omega^{\nu\mu}$ to three so that

$$\tfrac{1}{2}\omega^{\mu\nu} M_{\mu\nu} = \omega^{12} M_{12} + \omega^{01}(M_{01} + M_{31}) + \omega^{02}(M_{02} + M_{32}) \,. \tag{4.137}$$

Identifying the operators

$$J_3 = M_{12} \,, \qquad E_1 = M_{01} + M_{31} = K_1 + J_2 \,, \qquad E_2 = M_{02} + M_{32} = K_2 - J_1 \,, \tag{4.138}$$

we find the commutators from (4.32), or from (2.36), (4.42) and (4.43),

$$[J_3, E_1] = iE_2 \,, \qquad [J_3, E_2] = -iE_1 \,, \qquad [E_1, E_2] = 0 \,. \tag{4.139}$$

A unitary operator corresponding to finite group elements of $G_{\mathring{p}}$ is then

$$e^{-i(a_1 E_1 + a_2 E_2)} e^{-i\Theta J_3} \,, \tag{4.140}$$

Noting that

$$e^{-i\Theta J_3}(a_1 E_1 + a_2 E_2)e^{i\Theta J_3} = {a_1}^\Theta E_1 + {a_2}^\Theta E_2 \,, \qquad \begin{pmatrix} {a_1}^\Theta \\ {a_2}^\Theta \end{pmatrix} = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \,, \quad (4.141)$$

then if (4.140) corresponds to a group element (Θ, a_1, a_2), with Θ an angle with period 2π, we have the group multiplication rule

$$(\Theta', a'_1, a'_2)(\Theta, a_1, a_2) = (\Theta' + \Theta, {a_1}^{\Theta'} + a'_1, {a_2}^{\Theta'} + a'_2) \,. \qquad (4.142)$$

The group multiplication rule (4.142) is essentially identical to (4.96). The group is then isomorphic with the group formed by rotations and translations on two dimensional space, so that for the massless case we have the little group

$$G_{\mathring{p}} \simeq ISO(2) \simeq SO(2) \ltimes T_2 \,. \qquad (4.143)$$

The representations of this group can be obtained in a very similar fashion to that of the Poincaré group. Define vectors $|a_1, a_2\rangle$ such that

$$(E_1, E_2)|b_1, b_2\rangle = (b_1, b_2)|b_1, b_2\rangle \,, \qquad (4.144)$$

and then we assume, consistency with the group multiplication (4.142),

$$e^{-i\Theta J_3}|b_1, b_2\rangle = e^{-ih\Theta}|{b_1}^\Theta, {b_2}^\Theta\rangle \,, \qquad (4.145)$$

linking all (b_1, b_2) with constant $c = {b_1}^2 + {b_2}^2$. This irreducible representation of $ISO(2)$, labelled by c, h, is infinite dimensional. However there are one-dimensional representations, corresponding to taking $c = 0$, generated from a vector $|h\rangle$ such that

$$E_1|h\rangle = E_2|h\rangle = 0 \,, \qquad J_3|h\rangle = h|h\rangle \,, \qquad (4.146)$$

so that the essential group action is

$$e^{-i\Theta J_3}|h\rangle = e^{-ih\Theta}|h\rangle \,. \qquad (4.147)$$

For applications to representations of the Poincaré group $e^{-i\Theta J_3}$ corresponds to a subgroup of the $SO(3)$ rotation group so it is necessary to require in (4.146) and (4.147)

$$h = 0, \pm\tfrac{1}{2}, \pm 1, \dots \,. \qquad (4.148)$$

For the associated Lorentz transformations then a general element corresponding to the little group is $\Lambda_{(a_1,a_2)}\Lambda_\Theta$ where

$$\Lambda_{(a_1,a_2)} = \begin{pmatrix} 1 + \frac{1}{2}({a_1}^2 + {a_2}^2) & a_1 & a_2 & -\frac{1}{2}({a_1}^2 + {a_2}^2) \\ a_1 & 1 & 0 & -a_1 \\ a_2 & 0 & 1 & -a_2 \\ \frac{1}{2}({a_1}^2 + {a_2}^2) & a_1 & a_2 & 1 - \frac{1}{2}({a_1}^2 + {a_2}^2) \end{pmatrix} \,, \qquad (4.149)$$

and

$$\Lambda_\Theta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\Theta & -\sin\Theta & 0 \\ 0 & \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \qquad (4.150)$$

It is easy to see that $\Lambda_{(a_1,a_2)}\mathring{p} = \Lambda_\Theta \mathring{p} = \mathring{p}$ with $\mathring{p}$ as in (4.134).

The construction of the representation space $\mathcal{V}_h$ when $p^2 = 0$ proceeds in a very similar fashion as in the massive case. Neglecting infinite dimensional representations of the little group, then starting from a vector $|\mathring{p}, h\rangle$ satisfying

$$P^\mu |\mathring{p}, h\rangle = \mathring{p}^\mu |\mathring{p}, h\rangle \, , \qquad J_3 |\mathring{p}, h\rangle = h |\mathring{p}, h\rangle \, , \qquad (4.151)$$

a basis $\{|p, h\rangle : p^2 = 0, p_0 > 0\}$, for $\mathcal{V}_h$ is formed by

$$|p, h\rangle = U[L(p)] \, |\mathring{p}, h\rangle \, , \quad \text{for} \quad p^\mu = L(p)^\mu{}_\nu \mathring{p}^\nu \, , \qquad (4.152)$$

where $L(p)$ is assumed to be determined uniquely by p. Using

$$L(\Lambda p)^{-1} \Lambda L(p) = \Lambda_{(a_1,a_2)} \Lambda_\Theta \in G_{\mathring{p}} \, , \quad \text{for} \quad a_{1,2}(p,\Lambda) \, , \ \Theta(p,\Lambda) \, , \qquad (4.153)$$

and

$$U[\Lambda_{(a_1,a_2)}] U[\Lambda_\Theta] \, |\mathring{p}, h\rangle = |\mathring{p}, h\rangle \, e^{-ih\Theta} \, , \qquad (4.154)$$

then, for any $\Lambda \in SO(3,1)^\uparrow$, the action of the corresponding unitary operator on $\mathcal{V}_h$ is given by

$$U[\Lambda] \, |p, h\rangle = |\Lambda p, h\rangle \, e^{-ih\Theta(p,\Lambda)} \, . \qquad (4.155)$$

For $\mathring{p}$ as in (4.134), and

$$p^\mu = \omega(1, \mathbf{n}) \, , \qquad \omega > 0 \, , \qquad (4.156)$$

then $L(p)$, satisfying (4.116), is determined by assuming it is given by the expression (4.131) with now $e^\alpha = \omega/\mathring{\omega}$ and $R(n)$ the same rotation depending on θ, ϕ, the spherical polar angles specifying $\mathbf{n}$. Since $J_3 U[B(\alpha, e_3)] \, |\mathring{p}, h\rangle = h \, U[B(\alpha, e_3)] \, |\mathring{p}, h\rangle$, from (4.151), and $U[\Lambda_{R(n)}] J_3 U[\Lambda_{R(n)}]^{-1} = J_i R_{i3}(n) = \mathbf{n} \cdot \mathbf{J}$ then

$$\hat{\mathbf{p}} \cdot \mathbf{J} \, |p, h\rangle = h \, |p, h\rangle \, . \qquad (4.157)$$

The component of the angular momentum along the direction of motion, or *helicity*, is then h, taking values as in (4.148).

The irreducible representations of the Poincaré group for massless particles require only a single helicity h. If the symmetry group is extended to include *parity*, corresponding to spatial reflections, then it is necessary for there to be particle states with both helicities $\pm h$. When parity is a symmetry there is an additional unitary operator $\mathcal{P}$ with the action on the Poincaré group generators

$$\mathcal{P}\mathbf{J}\mathcal{P}^{-1} = \mathbf{J} \, , \quad \mathcal{P}\mathbf{K}\mathcal{P}^{-1} = -\mathbf{K} \, , \quad \mathcal{P}H\mathcal{P}^{-1} = H \, , \quad \mathcal{P}\mathbf{P}\mathcal{P}^{-1} = -\mathbf{P} \, . \qquad (4.158)$$

In consequence $\mathcal{P}\mathbf{P} \cdot \mathbf{J}\mathcal{P}^{-1} = -\mathbf{P} \cdot \mathbf{J}$ so that, from (4.157), $\mathcal{P}|p, h\rangle$ must have helicity $-h$, so we must have $\mathcal{P}|p, h\rangle = \eta |p, -h\rangle$, for some phase η, usually $\eta = \pm 1$. Thus photons have helicity ± 1 and gravitons ± 2. However neutrinos, if they were exactly massless, which is no longer compatible with experiment, need only have helicity $-\frac{1}{2}$ since their weak interactions do not conserve parity and experimentally only involve $-\frac{1}{2}$ helicity.

4.5.3 Spinorial Treatment

Calculations involving Lorentz transformations are almost always much simpler in terms of $Sl(2,\mathbb{C})$ matrices, making use of the isomorphism described in section 4.3, rather than working out products of 4×4 matrices Λ. As an illustration we re-express some of the above discussion for massless representations in terms of spinors.

Defining $p_{\alpha\dot\alpha} = p_\mu(\sigma^\mu)_{\alpha\dot\alpha}$, as in (4.51), then since $p^2 = 0$, by virtue of (4.53),

$$\det[p_{\alpha\dot\alpha}] = 0 \quad \Rightarrow \quad p_{\alpha\dot\alpha} = \lambda_\alpha \bar\lambda_{\dot\alpha} \,. \tag{4.159}$$

The spinor λ_α and its conjugate $\bar\lambda_{\dot\alpha}$ are arbitrary up to the $U(1)$ transformation given by $\lambda_\alpha \to \lambda_\alpha e^{-i\theta}$, $\bar\lambda_{\dot\alpha} \to \bar\lambda_{\dot\alpha} e^{i\theta}$. To determine λ_α precisely we choose the phase so that for $\mathring{p}$ given by (4.134), since $[\mathring{p}_{\alpha\dot\alpha}] = 2\mathring{\omega} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, we take

$$\mathring{\lambda} = \sqrt{2\mathring{\omega}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{4.160}$$

Then for any $p = L(p)\mathring{p}$ we define a unique spinor satisfying (4.159) by

$$\lambda_p = A(p)\mathring{\lambda} \quad \text{where} \quad L(p) \underset{SO(3,1)\to Sl(2,\mathbb{C})}{\longrightarrow} A(p) \,, \quad A(\mathring{p}) = I \,. \tag{4.161}$$

From (4.149) and (4.150) we have correspondingly

$$\begin{aligned} \Lambda_{(a_1,a_2)} &\underset{SO(3,1)\to Sl(2,\mathbb{C})}{\longrightarrow} A_{(a_1,a_2)} = \begin{pmatrix} 1 & a_1 - ia_2 \\ 0 & 1 \end{pmatrix} , \\ \Lambda_\Theta &\underset{SO(3,1)\to Sl(2,\mathbb{C})}{\longrightarrow} A_\Theta = \begin{pmatrix} e^{-\frac{1}{2}i\Theta} & 0 \\ 0 & e^{\frac{1}{2}i\Theta} \end{pmatrix} . \end{aligned} \tag{4.162}$$

For any Lorentz transformation $\Lambda \to A_\Lambda$ then (4.153) becomes equivalently

$$A(\Lambda p)^{-1} A_\Lambda A(p) = A_{(a_1,a_2)} A_\Theta \,, \tag{4.163}$$

and with the definition (4.161) we get

$$A_\Lambda \lambda_p = \lambda_{\Lambda P}\, e^{-\frac{1}{2}i\Theta(p,\Lambda)} \,. \tag{4.164}$$

This provides a more convenient method of calculating $\Theta(p,\Lambda)$ if required.

4.6 Casimir Operators

For the rotation group then from the generators $\mathbf{J}$ it is possible to construct an invariant operator $\mathbf{J}^2$ which commutes with all generators, as in (2.63), so that all vectors belonging to any irreducible representation space have the same eigenvalue, for $\mathcal{V}_j$, $j(j+1)$. Such operators, which are quadratic or possibly higher order in the generators, are generically called *Casimir*[11] *operators.* Of course only algebraically independent Casimir operators are of interest.

For the Lorentz group, $SO(3,1)$, there are two basic Casimir operators which can be formed from $M_{\mu\nu}$ using the invariant tensors

$$\tfrac{1}{4}M^{\mu\nu}M_{\mu\nu}\,, \qquad \tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho}M_{\mu\nu}M_{\sigma\rho}\,. \tag{4.165}$$

In terms of the generators $\mathbf{J},\mathbf{K}$, defined in (4.37),(4.41), and then $\mathbf{J}^{\pm}$, defined in (4.46),

$$\begin{aligned}\tfrac{1}{4}M^{\mu\nu}M_{\mu\nu} &= \tfrac{1}{2}(\mathbf{J}^2-\mathbf{K}^2) = \mathbf{J}^{+2}+\mathbf{J}^{-2}\,,\\ \tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho}M_{\mu\nu}M_{\sigma\rho} &= \mathbf{J}\cdot\mathbf{K} = -i\big(\mathbf{J}^{+2}-\mathbf{J}^{-2}\big)\,.\end{aligned} \tag{4.166}$$

Since $\mathbf{J}^{\pm}$ both obey standard angular momentum commutation relations, as in (4.47), then for finite dimensional irreducible representations

$$\mathbf{J}^{+2}\to j(j+1)\,I\,, \quad \mathbf{J}^{-2}\to \bar{\jmath}(\bar{\jmath}+1)\,I\,, \qquad j,\bar{\jmath}=0,\tfrac{1}{2},1,\tfrac{3}{2},\dots\,. \tag{4.167}$$

For the fundamental spinor representation the generators $s_{\mu\nu}=\frac{1}{2}i\,\sigma_{[\mu}\bar{\sigma}_{\nu]}$, as in (4.67), the associated Casimir operators become

$$\begin{aligned}\tfrac{1}{4}s^{\mu\nu}s_{\mu\nu} &= -\tfrac{1}{32}\,\sigma^{\mu}\bar{\sigma}^{\nu}(\sigma_{\mu}\bar{\sigma}_{\nu}-\sigma_{\nu}\bar{\sigma}_{\nu}) = \tfrac{3}{4}\,I\,,\\ \tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho}s_{\mu\nu}s_{\sigma\rho} &= \tfrac{1}{32}\,\varepsilon^{\mu\nu\sigma\rho}\sigma_{\mu}\bar{\sigma}_{\nu}\sigma_{\sigma}\bar{\sigma}_{\rho} = -\tfrac{3}{4}i\,I\,,\end{aligned} \tag{4.168}$$

using (4.49) and (4.91). As expected this is in accord with (4.166) and (4.167) for $j=\frac{1}{2}$, $\bar{\jmath}=0$. Conversely for $\bar{s}_{\mu\nu}$ the role of j and $\bar{\jmath}$ are interchanged since this is the conjugate representation.

For the Poincaré group then (4.165) no longer provides Casimir operators because they fail to commute with P_{μ}. There is now only a single quadratic Casimir

$$P^2 = P^{\mu}P_{\mu}\,, \tag{4.169}$$

whose eigenvalues acting on the irreducible spaces $\mathcal{V}_{m,s},\mathcal{V}_s$, corresponding to the spaces of relativistic single particle states, give the invariant m^2 in the massive case or zero in the massless case. However the irreducible representations are also characterised by a spin label s, helicity in the massless case. To find an invariant characterisation of this we introduce the *Pauli-Lubanski vector*,

$$W^{\mu} = \tfrac{1}{2}\,\varepsilon^{\mu\nu\sigma\rho}P_{\nu}M_{\sigma\rho} = \tfrac{1}{2}\,\varepsilon^{\mu\nu\sigma\rho}M_{\sigma\rho}P_{\nu}\,. \tag{4.170}$$

Using $\varepsilon^{\mu\nu\sigma\rho}P_{\nu}P_{\sigma}=0$ we have

$$[W^{\mu},P_{\nu}] = 0\,. \tag{4.171}$$

Since $\varepsilon^{\mu\nu\sigma\rho}$ is an invariant tensor then W^{μ} should be a contravariant 4-vector, to verify this we may use

$$\begin{aligned}[][W^{\mu},\tfrac{1}{2}\omega^{\sigma\rho}M_{\sigma\rho}] &= -\tfrac{1}{2}i\,\varepsilon^{\mu\nu\sigma\rho}\big(P_{\lambda}\omega^{\lambda}{}_{\nu}\,M_{\sigma\rho}+P_{\nu}\,M_{\lambda\rho}\omega^{\lambda}{}_{\sigma}+P_{\nu}\,M_{\sigma\lambda}\omega^{\lambda}{}_{\sigma}\big)\\ &= \tfrac{1}{2}i\,\omega^{\mu}{}_{\lambda}\varepsilon^{\lambda\nu\sigma\rho}P_{\nu}M_{\sigma\rho} = i\,\omega^{\mu}{}_{\lambda}W^{\lambda}\,,\end{aligned} \tag{4.172}$$

[11] Hendrik Brugt Gerhard Casimir, 1909-2000, Dutch.

to obtain

$$[W^\mu, M_{\sigma\rho}] = i(\delta^\mu{}_\sigma W_\rho - \delta^\mu{}_\rho W_\sigma) \,. \tag{4.173}$$

With (4.171) and (4.173) we may then easily derive

$$[W^\mu, W^\nu] = i\,\varepsilon^{\mu\nu\sigma\rho} P_\sigma W_\rho \,. \tag{4.174}$$

It follows from (4.171) and (4.173) that

$$W_\mu W^\mu \,, \tag{4.175}$$

is a scalar commuting with $P_\nu, , M_{\sigma\rho}$ and so providing an additional Casimir operator.

For the massive representations then, for $\mathring{p}$ as in (4.121),

$$W^0|\mathring{p}, s\, s_3\rangle = 0 \,, \qquad W^i|\mathring{p}, s\, s_3\rangle = -m\,\varepsilon_{ijk} M_{jk}|\mathring{p}, s\, s_3\rangle = -m\, J_i|\mathring{p}, s\, s_3\rangle \,, \tag{4.176}$$

so that

$$W_\mu W^\mu|\mathring{p}, s\, s_3\rangle = -m^2\, \mathbf{J}^2|\mathring{p}, s\, s_3\rangle = -m^2 s(s+1)|\mathring{p}, s\, s_3\rangle \,. \tag{4.177}$$

Hence $W_\mu W^\mu$ has the eigenvalue $-m^2 s(s+1)$ for all vectors in the representation space $\mathcal{V}_{m,s}$.

For the massless representations then, for $\mathring{p}$ as in (4.134),

$$\begin{aligned} W^1|\mathring{p}, h\rangle &= \mathring{\omega}\, E_2|\mathring{p}, h\rangle = 0 \,, & W^2|\mathring{p}, h\rangle &= -\mathring{\omega}\, E_1|\mathring{p}, h\rangle = 0 \,, \\ W^0|\mathring{p}, h\rangle &= -\,\mathring{\omega}\, J_3|\mathring{p}, h\rangle = -\mathring{\omega}\, h|\mathring{p}, h\rangle \,, & W^3|\mathring{p}, h\rangle &= -\mathring{\omega}\, J_3|\mathring{p}, h\rangle = -\mathring{\omega}\, h|\mathring{p}, h\rangle \,, \end{aligned} \tag{4.178}$$

using (4.146). Since W^μ, P^μ are both contravariant 4-vectors the result (4.178) requires

$$(W^\mu + h\, P^\mu)|p, h\rangle = 0 \,, \tag{4.179}$$

for all vectors providing a basis for $\mathcal{V}_h$. This provides an invariant characterisation of the helicity h on this representation space.

4.7 Quantum Fields

To construct a relativistic quantum mechanics compatible with the general principles of quantum mechanics it is essentially inevitable to use quantum field theory. The quantum fields are required to have simple transformation properties under the symmetry transformations belonging to the Poincaré group. For a simple scalar field, depending on the space-time coordinates x^μ, this is achieved by

$$U[\Lambda, a]\phi(x)U[\Lambda, a]^{-1} = \phi(\Lambda x + a) \,, \tag{4.180}$$

where $U[\Lambda, a]$ are the unitary operators satisfying (4.100). For an infinitesimal transformation, with Λ as in (4.14) and U as in (4.101), this gives

$$-i\big[\tfrac{1}{2}\omega^{\mu\nu} M_{\mu\nu} - a^\mu P_\mu, \phi(x)\big] = (\omega^\mu{}_\nu \partial_\nu + a^\mu)\partial_\mu \phi(x) \,, \tag{4.181}$$

or

$$[M_{\mu\nu}, \phi(x)] = -L_{\mu\nu}\phi(x)\,, \quad L_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu)\,, \qquad [P_\mu, \phi(x)] = -i\partial_\mu \phi(x)\,. \qquad (4.182)$$

$L_{\mu\nu}$ and $i\partial_\mu$ obey the same commutation relations as $M_{\mu\nu}$ and P_μ in (4.32) and (4.104). Note that, with (4.106), $[\mathbf{P}, \phi] = i\,\boldsymbol{\nabla}\phi$.

To describe particles with spin the quantum fields are required to transform according to a finite dimensional representation of the Lorentz group so that (4.180) is extended to

$$U[\Lambda, a]\phi(x)U[\Lambda, a]^{-1} = D(\Lambda)^{-1}\phi(\Lambda x + a)\,, \qquad (4.183)$$

regarding ϕ now as a column vector and suppressing matrix indices. For an infinitesimal Lorentz transformation then assuming

$$D(\Lambda) = I - i\,\tfrac{1}{2}\omega^{\mu\nu}S_{\mu\nu}\,, \qquad S_{\mu\nu} = -S_{\nu\mu}\,, \qquad (4.184)$$

the commutator with $M_{\mu\nu}$ in (4.182) is extended to

$$[M_{\mu\nu}, \phi(x)] = -\big(L_{\mu\nu} + S_{\mu\nu}\big)\phi(x)\,. \qquad (4.185)$$

The matrix generators $S_{\mu\nu}$ obey the same commutators as $M_{\mu\nu}$ in (4.32).

The relation of the quantum fields to the particle state representations considered in 4.5 is elucidated by considering, considering first $\mathcal{V}_{m,s}$,

$$\langle 0|\phi(x)|p, s\, s_3\rangle = u(p, s_3)\, e^{-ip\cdot x}\,, \qquad p^2 = m^2\,. \qquad (4.186)$$

Here $|0\rangle$ is the vacuum state, which is just a singlet under the Poincaré group, $U[\Lambda, a]|0\rangle = |0\rangle$. It is easy to check that (4.186) is accord with translation invariance using (4.128). Using (4.183), for $a = 0$, $\Lambda \to \Lambda^{-1}$, with (4.127) we get

$$D(\Lambda)\, u(p, s_3) = \sum_{s'_3} u(\Lambda p, s'_3)\, D^{(s)}_{s'_3\, s_3}\big(\mathcal{R}(p, \Lambda)\big)\,, \qquad (4.187)$$

which is directly analogous to (4.127) but involves the finite dimensional representation matrix $D(\Lambda)$. $u(p, s_3)$ thus allows the complicated Wigner rotation of spin indices given by $\mathcal{R}(p, \Lambda)$ to be replaced by a Lorentz transformation, in some representation, depending just on Λ. To determine $u(p, s_3)$ precisely so as to be in accord with (4.187) it is sufficient to follow the identical route to that which determined the states $|p, s\, s_3\rangle$ in 4.5.1. Thus it is sufficient to require, as in (4.125),

$$D(\Lambda_R)\, u(\mathring{p}, s_3) = \sum_{s'_3} u(\mathring{p}, s'_3)\, D^{(s)}_{s'_3\, s_3}(R)\,, \qquad (4.188)$$

and then define, as in (4.126),

$$u(p, s_3) = D\big(L(p)\big)\, u(\mathring{p}, s_3)\,. \qquad (4.189)$$

For Λ reduced to a rotation Λ_R, as in (4.17), the representation given by the matrices $D(\Lambda_R)$ decomposes into a direct sum of irreducible $SO(3)$ representations $D^{(j)}(R)$. For (4.188) to be possible this decomposition must include, by virtue of Schur's lemmas, the irreducible representation $j = s$, with any other $D^{(j)}$, $j \neq s$, annihilating $u(\mathring{p}, s_3)$.

For the zero mass case the discussion is more involved so we focus on a particular case when the helicity $h = 1$ and the associated quantum field is a 4-vector A^μ. Replacing (4.186) we require

$$\langle 0|A^\mu(x)|p,1\rangle = \epsilon^\mu(p)\, e^{-ip\cdot x}\,, \qquad p^2 = 0\,. \tag{4.190}$$

$\epsilon^\mu(p)$ is referred to as a *polarisation vector*. For 4-vectors there is an associated representation of the Lorentz group which is just given, of course, by the Lorentz transformation matrices Λ themselves. When $p = \mathring{p}$ as in (4.134) then from the little group transformations as in (4.154) we require, for $h = 1$,

$$\Lambda_\Theta \epsilon(\mathring{p}) = \epsilon(\mathring{p})\, e^{-i\Theta}\,. \tag{4.191}$$

Using (4.150) this determines $\epsilon(\mathring{p})$ to be

$$\epsilon^\mu(\mathring{p}) = \tfrac{1}{\sqrt{2}}(0,1,i,0)\,, \tag{4.192}$$

with a normalisation $\epsilon^* \cdot \epsilon = -1$. Using the explicit form for $\Lambda_{(a_1,a_2)}$ in (4.149) we then obtain

$$\Lambda_{(a_1,a_2)}\epsilon(\mathring{p}) = \epsilon(\mathring{p}) + c\,\mathring{p}\,, \qquad c = \tfrac{1}{\sqrt{2}}(a_1 + a_2)\,. \tag{4.193}$$

For general momentum $p = \omega(1,\mathbf{n})$, $p^2 = 0$, as in (4.156), we may define, for $L(p)$ given by (4.131),

$$\epsilon(p) = L(p)\,\epsilon(\mathring{p}) = \Lambda_{R(n)}\,\epsilon(\mathring{p})\,, \tag{4.194}$$

since $B(\alpha, e_3)\epsilon(\mathring{p}) = \epsilon(\mathring{p})$, and where the rotation $R(n)$ is determined by $\mathbf{n}$ just as in (4.131). With the definition (4.194)

$$p_\mu \epsilon^\mu(p) = \mathring{p}_\mu \epsilon^\mu(\mathring{p}) = 0\,. \tag{4.195}$$

For a general Lorentz transformation Λ then from (4.153) and (4.191),(4.192)

$$\Lambda\epsilon(p) = \big(\epsilon(\Lambda p) + c\,\Lambda p\big)\, e^{-i\Theta(p,\Lambda)}\,, \tag{4.196}$$

for some c depending on p, Λ. This matches (4.155), for $h = 1$, save for the inhomogeneous term proportional to c (for $h = -1$ it is sufficient to take $\epsilon(p) \to \epsilon(p)^*$). (4.196) shows that $\epsilon(p)$ does not transform in a Lorentz covariant fashion. Homogeneous Lorentz transformations are obtained if, instead of considering just $\epsilon(p)$, we consider the equivalence classes polarisation vectors $\{\epsilon(p) :\sim\}$ with the equivalence relation

$$\epsilon(p) \sim \epsilon(p) + c\,p\,, \quad \text{for arbitrary} \quad c\,. \tag{4.197}$$

This is the same as saying that the polarisation vectors $\epsilon(p)$ are arbitrary up to the addition of any multiple of the momentum vector p. It is important to note that, because of (4.195), that scalar products of polarisation vectors depend only on their equivalence classes so that

$$\epsilon'(p)^* \cdot \epsilon'(p) = \epsilon(p)^* \cdot \epsilon(p) \quad \text{for} \quad \epsilon'(p) \sim \epsilon(p)\,. \tag{4.198}$$

The *gauge freedom* in (4.197) is a reflection of *gauge invariance* which is a necessary feature of field theories when massless particles are described by quantum fields transforming in a Lorentz covariant fashion.

In general Lorentz covariant fields contain more degrees of freedom than those for the associated particle which are labelled by the spin or helicity in the massless case. It is then necessary to impose supplementary conditions to reduce the number of degrees of freedom, e.g. for a massive 4-vector field ϕ^μ, associated with a spin one particle, requiring $\partial_\mu \phi^\mu = 0$. For the massless case then there are gauge transformations belonging to a gauge group which eliminate degrees of freedom so that just two helicities remain. Although this can be achieved for free particles of arbitrary spin there are inconsistencies when interactions are introduced for higher spins, beyond spin one in the massive case with spin two also allowed for massless particles.

5 Lie Groups and Lie Algebras

Although many discussions of groups emphasise finite discrete groups the groups of most widespread relevance in high energy physics are Lie groups which depend continuously on a finite number of parameters. In many ways the theory of Lie[12] groups is more accessible than that for finite discrete groups, the classification of the former was completed by Cartan[13] over 100 years ago while the latter was only finalised in the late 1970's and early 1980's.

A *Lie Group* is of course a group but also has the structure of a differentiable manifold, so that some of the methods of differential geometry are relevant. It is important to recognise that abstract group elements cannot be added, unlike matrices, so the notion of derivative needs some care. For a Lie group G, with an associated n-dimensional differential manifold $\mathcal{M}_G$, then for an arbitrary element

$$g(a) \in G \,, \quad a = (a^1, \dots, a^n) \in \mathbb{R}^n \quad \text{coordinates on} \quad \mathcal{M}_G \,. \tag{5.1}$$

n is the dimension of the Lie group G. For any interesting $\mathcal{M}_G$ no choice of coordinates is valid on the whole of $\mathcal{M}_G$, it is necessary to choose different coordinates for various subsets of $\mathcal{M}_G$, which collectively cover the whole of $\mathcal{M}_G$ and form a corresponding set of coordinate charts, and then require that there are smooth transformations between coordinates on the overlaps between coordinate charts. Such issues are generally mentioned here only in passing.

For group multiplication we then require

$$g(a)g(b) = g(c) \quad \Rightarrow \quad c^r = \varphi^r(a,b) \,, \quad r = 1, \dots, n \,, \tag{5.2}$$

where φ^r is continuously differentiable. It is generally convenient to choose the origin of the coordinates to be the identity so that

$$g(0) = e \quad \Rightarrow \quad \varphi^r(0,a) = \varphi^r(a,0) = a^r \,, \tag{5.3}$$

[12]Marius Sophus Lie, 1842-1899, Norwegian.
[13]Élie Joseph Cartan, 1869-1951, French.

and then for the inverse

$$g(a)^{-1} = g(\bar{a}) \quad \Rightarrow \quad \varphi^r(\bar{a}, a) = \varphi^r(a, \bar{a}) = 0 \,. \tag{5.4}$$

The crucial associativity condition is then

$$g(a)\big(g(b)g(c)\big) = \big(g(a)g(b)\big)g(c) \quad \Rightarrow \quad \varphi^r(a, \varphi(b,c)) = \varphi^r(\varphi(a,b), c) \,. \tag{5.5}$$

A Lie group may be identified with the associated differentiable manifold $\mathcal{M}_G$ together with a map $\varphi : \mathcal{M}_G \times \mathcal{M}_G \to \mathcal{M}_G$, where φ satisfies (5.3), (5.4) and (5.5).

For an abelian group $\varphi(a,b) = \varphi(b,a)$ and it is possible to choose coordinates such that

$$\varphi^r(a,b) = a^r + b^r \,, \tag{5.6}$$

and in general if we Taylor expand φ we must have

$$\varphi^r(a,b) = a^r + b^r + c^r{}_{st}\, a^s b^t + \mathrm{O}(a^2 b, ab^2) \,, \quad \bar{a}^r = -a^r + c^r{}_{st}\, a^s a^t + \mathrm{O}(a^3) \,. \tag{5.7}$$

As will become apparent the coefficients $c^r{}_{st}$, or rather $f^r{}_{st} = c^r{}_{[st]}$, which satisfy conditions arising from the associativity condition (5.5), essentially determine the various possible Lie groups.

As an illustration we return again to $SU(2)$. For 2×2 matrices A we may express them in terms of the Pauli matrices by

$$A = u_0\, I + i\, \mathbf{u} \cdot \boldsymbol{\sigma} \,, \qquad A^\dagger = u_0\, I - i\, \mathbf{u} \cdot \boldsymbol{\sigma} \,. \tag{5.8}$$

Requiring $u_0, \mathbf{u}$ to be real then

$$A^\dagger A = ({u_0}^2 + \mathbf{u}^2)\, I \,, \qquad \det A = {u_0}^2 + \mathbf{u}^2 \,. \tag{5.9}$$

Hence

$$A \in SU(2) \quad \Rightarrow \qquad {u_0}^2 + \mathbf{u}^2 = 1 \,. \tag{5.10}$$

The condition ${u_0}^2 + \mathbf{u}^2 = 1$ defines the three dimensional sphere S^3 embedded in $\mathbb{R}^4$, so that $\mathcal{M}_{SU(2)} \simeq S^3$. In terms of differential geometry all points on S^3 are equivalent but here the pole $u_0 = 1$, $\mathbf{u} = \mathbf{0}$ is special as it corresponds to the identity. For $SO(3)$ then, since $\pm A$ correspond to the same element of $SO(3)$, we must identify $(u_0, \mathbf{u})$ and $-(u_0, \mathbf{u})$, *i.e.* antipodal points at the ends of any diameter on S^3. In the hemisphere $u_0 \geq 0$ we may use $\mathbf{u}$, $|\mathbf{u}| \leq 1$ as coordinates for $SU(2)$, since then $u_0 = \sqrt{1 - \mathbf{u}^2}$. Then group multiplication defines $\boldsymbol{\varphi}(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \mathbf{v} - \mathbf{u} \times \mathbf{v} + \dots$.

For $A \in Sl(2, \mathbb{C})$ then if $A^\dagger A = e^{2V}$, for $V^\dagger = V$, $R = A\, e^{-V}$ satisfies $R^\dagger R = I$. Since then $\det R = e^{i\alpha}$ while $\det e^V = e^{\mathrm{tr}(V)}$ is real, $\det A = 1$ requires both $\det R = 1$ and $\mathrm{tr}(V) = 0$. Hence there is a unique decomposition $A = R\, e^V$ with $V = V_i \sigma_i$ so that the group manifold $\mathcal{M}_{Sl(2,\mathbb{C})} = S^3 \times \mathbb{R}^3$.

5.0.1 Vector Fields, Differential Forms and Lie Brackets

For any differentiable n-dimensional manifold $\mathcal{M}$, with coordinates x^i, then scalar functions $f : \mathcal{M} \to \mathbb{R}$ are defined in terms of these coordinates by $f(x)$ such that under a change of coordinates $x^i \to x'^i$ we have $f(x) = f'(x')$. Vector fields are defined in terms of differential operators acting on scalar functions

$$X(x) = X^i(x)\,\frac{\partial}{\partial x^i}\,, \tag{5.11}$$

where for the $x \to x'$ change in coordinates we require

$$X^j(x)\,\frac{\partial x'^i}{\partial x^j} = X'^i(x')\,. \tag{5.12}$$

For each x the vector fields belong to a linear vector space $T_x(\mathcal{M})$ of dimension n, the tangent space at the point specified by x.

For two vector fields X, Y belonging to $T_x(\mathcal{M})$ the *Lie bracket*, or *commutator*, defines a further vector field

$$[X, Y] = -[Y, X]\,, \tag{5.13}$$

where

$$[X, Y]^i(x) = X(x)Y^i(x) - Y(x)X^i(x)\,, \tag{5.14}$$

since, for a change $x \to x'$ and using (5.12),

$$[X, Y]' = [X', Y']\,, \tag{5.15}$$

as a consequence of $\frac{\partial^2 x'^i}{\partial x^j \partial x^k} = \frac{\partial^2 x'^i}{\partial x^k \partial x^j}$. The Lie bracket is clearly linear, so that for any $X, Y, Z \in T_x(\mathcal{M})$

$$[\alpha X + \beta Y, Z] = \alpha[X, Y] + \beta[Y, Z]\,, \tag{5.16}$$

as in necessary for the Lie bracket to be defined on the vector space $T_x(\mathcal{M})$, and it also satisfies crucially the *Jacobi*[14] *identity*, which requires

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0\,. \tag{5.17}$$

This follows directly from the definition of the Lie Bracket as a commutator of differential operators.

Dual to vector fields are *one-forms*, belonging to $T_x(\mathcal{M})^*$,

$$\omega(x) = \omega_i(x)\,\mathrm{d}x^i\,, \tag{5.18}$$

where $\langle \mathrm{d}x^i, \partial_j\rangle = \delta^i{}_j$. For $x \to x'$ now

$$\omega_j(x)\,\frac{\partial x^j}{\partial x'^i} = \omega'_i(x')\,. \tag{5.19}$$

[14] Carl Gustav Jacob Jacobi, 1804-1851, German.

For p-forms

$$\rho(x) = \tfrac{1}{p!}\,\rho_{i_1\dots i_p}(x)\,\mathrm{d}x^{i_1}\wedge\dots\wedge\mathrm{d}x^{i_p}\,, \qquad \mathrm{d}x^i\wedge\mathrm{d}x^j = -\mathrm{d}x^j\wedge\mathrm{d}x^i\,, \tag{5.20}$$

so that $\rho_{i_1\dots i_p} = \rho_{[i_1\dots i_p]}$. The transformations $\rho\to\rho'$ for a change of coordinates $x\to x'$ are the natural multi-linear extension of (5.19). For an n-dimensional space $\mathrm{d}x'^{i_1}\wedge\dots\wedge\mathrm{d}x'^{i_n} = \det\big[\frac{\partial x'^i}{\partial x^j}\big]\,\mathrm{d}x^{i_1}\wedge\dots\wedge\mathrm{d}x^{i_n}$ and we may require

$$\mathrm{d}x^{i_1}\wedge\dots\wedge\mathrm{d}x^{i_n} = \varepsilon^{i_1\dots i_n}\,\mathrm{d}^n x \tag{5.21}$$

with $\varepsilon^{i_1\dots i_n}$ the n-dimensional antisymmetric symbol and $\mathrm{d}^n x$ the corresponding volume element. If ρ is a n-form and $\mathcal{M}_n$ a n-dimensional manifold this allows the definition of the integral

$$\int_{\mathcal{M}_n}\rho\,. \tag{5.22}$$

The *exterior derivative* d acts on p-forms to give $(p+1)$-forms, $\mathrm{d}\rho = \mathrm{d}x^i\wedge\partial_i\rho$. For the one-form in (5.18) the corresponding two-form is then given by

$$(\mathrm{d}\omega)_{ij}(x) = \partial_i\omega_j(x) - \partial_j\omega_i(x)\,. \tag{5.23}$$

Of course $(\mathrm{d}\omega)' = \mathrm{d}'\omega'$ with $\mathrm{d}' = \mathrm{d}x'^i\partial'_i$. In general $\mathrm{d}^2 = 0$. If ρ is a *closed* p-form then

$$\mathrm{d}\rho = 0\,. \tag{5.24}$$

A trivial solution of (5.24) is provided by

$$\rho = \mathrm{d}\omega\,, \tag{5.25}$$

for some $(p-1)$-form ω. In this case ρ is *exact*. If the n-form ρ in (5.22) is exact and if also if $\mathcal{M}_n$ is closed then the integral is zero.

5.1 Lie Algebras

The additional structure associated with a differential manifold $\mathcal{M}_G$ corresponding to a Lie group G ensures that the tangent spaces $T_g(\mathcal{M}_G)$, for a point on the manifold for which the group element is g, can be related by group transformations. In particular the tangent space at the origin $T_e(\mathcal{M}_G)$ plays a special role and together with the associated Lie bracket $[\,,\,]$ defines the *Lie algebra* $\mathfrak{g}$ for the Lie group. For all points on $\mathcal{M}_G$ there is a space of vector fields which are invariant in a precise fashion under the action of group transformations and which belong to a Lie algebra isomorphic to $\mathfrak{g}$. There are also corresponding invariant one-forms.

To demonstrate these results we consider how a group element close to the identity generates a small change in an arbitrary group element $g(b)$ when multiplied on the right,

$$g(b+\mathrm{d}b) = g(b)g(\theta)\,,\quad \theta\text{ infinitesimal}\quad\Rightarrow\quad b^r+\mathrm{d}b^r = \varphi^r(b,\theta)\,, \tag{5.26}$$

so that

$$\mathrm{d}b^r = \theta^a {\mu_a}^r(b)\,, \qquad {\mu_a}^r(b) = \frac{\partial}{\partial\theta^a}\varphi^r(b,\theta))\Big|_{\theta=0}\,. \tag{5.27}$$

Here we use a, b, c as indices referring to components for vectors or one-forms belonging to $T_e(\mathcal{M}_G)$ or its dual (which must be distinguished from their use as coordinates) and r, s, t for indices at an arbitrary point. To consider the group action on the tangent spaces we analyse the infinitesimal variation of (5.2) for fixed $g(a)$,

$$g(c + \mathrm{d}c) = g(a)\, g(b + \mathrm{d}b) = g(c)\, g(\theta)\,, \tag{5.28}$$

so that, for fixed $g(a)$,

$$\mathrm{d}c^r = \theta^a\, {\mu_a}^r(c) = \mathrm{d}b^s\, {\lambda_s}^a(b)\, {\mu_a}^r(c)\,, \tag{5.29}$$

using (5.27) and defining $\lambda(b)$ as the matrix inverse of $\mu(b)$,

$$[{\lambda_s}^a(b)] = [{\mu_a}^s(b)]^{-1}\,, \qquad {\lambda_s}^a(b)\, {\mu_a}^r(b) = \delta_s{}^r\,. \tag{5.30}$$

Hence from from (5.29)

$$\boxed{\frac{\partial c^r}{\partial b^s} = {\lambda_s}^a(b)\, {\mu_a}^r(c)\,.} \tag{5.31}$$

If near the identity we assume (5.7) then ${\mu_a}^s(0) = \delta_a{}^s$.

By virtue of (5.31)

$$T_a(b) = {\mu_a}^s(b)\, \frac{\partial}{\partial b^s} = {\mu_a}^s(b)\, \frac{\partial c^r}{\partial b^s}\, \frac{\partial}{\partial c^r} = T_a(c)\,, \tag{5.32}$$

define a basis $\{T_a : a = 1, \dots, n\}$ of *left-invariant* vector fields belonging to $T(\mathcal{M}_G)$, since they are unchanged as linear differential operators under transformations corresponding to $g(b) \to g(c) = g(a)g(b)$. Furthermore the corresponding vector space, formed by constant linear combinations $\mathfrak{g} = \{\theta^a T_a\}$, is closed under taking the Lie bracket for any two vectors belonging to $\mathfrak{g}$ and defines the Lie algebra.

To verify closure we consider the second derivative of $c^r(b)$ where from (5.31) and (5.32)

$$\begin{aligned} {\mu_a}^s(b){\mu_b}^t(b)\, \frac{\partial^2 c^r}{\partial b^s \partial b^t} &= {\mu_a}^s(b) T_b(b) \big({\lambda_s}^a(b)\, {\mu_a}^r(c)\big) \\ &= {\mu_a}^s(b) \big(T_b(b){\lambda_s}^c(b)\, {\mu_c}^r(c) + {\lambda_s}^c(b)\, T_b(c){\mu_c}^r(c)\big)\,. \end{aligned} \tag{5.33}$$

For any matrix $\delta X^{-1} = -X^{-1}\delta X\, X^{-1}$ so that from (5.30)

$$T_b(b){\lambda_s}^c(b) = -{\lambda_s}^d(b)\big(T_b(b){\mu_d}^u(b)\big){\lambda_u}^c(b)\,, \tag{5.34}$$

which allows (5.33) to be written as

$${\mu_a}^s(b){\mu_b}^t(b)\, \frac{\partial^2 c^r}{\partial b^s \partial b^t} = -T_b(b){\mu_a}^u(b)\, {\lambda_u}^c(b){\mu_c}^r(c) + T_b(c){\mu_c}^r(c)\,, \tag{5.35}$$

or, transporting all indices so as to refer to the identity tangent space,

$${\mu_a}^s(b){\mu_b}^t(b)\, \frac{\partial^2 c^r}{\partial b^s \partial b^t}\, {\lambda_r}^c(c) = -\big(T_b(b){\mu_a}^r(b)\big)\, {\lambda_r}^c(b) + \big(T_b(c){\mu_a}^r(c)\big){\lambda_r}^c(c)\,. \tag{5.36}$$

Since

$$\frac{\partial^2 c^r}{\partial b^s \partial b^t} = \frac{\partial^2 c^r}{\partial b^t \partial b^s}\,, \tag{5.37}$$

the right hand side of (5.36) must be symmetric in a, b. Imposing that the antisymmetric part vanishes requires

$$\big(T_a(b)\mu_b{}^r(b) - T_b(b)\mu_a{}^r(b)\big)\lambda_r{}^c(b) = f^c{}_{ab}\,, \tag{5.38}$$

where $f^c{}_{ab}$ are the *structure constants* for the Lie algebra. They are constants since (5.36) requires that (5.38) is invariant under $b \to c$. Clearly

$$\boxed{f^c{}_{ab} = -f^c{}_{ba}\,.} \tag{5.39}$$

From (5.30), (5.38) can be equally written just as first order differential equations in terms of μ,

$$\boxed{T_a\,\mu_b{}^r - T_b\,\mu_a{}^r = f^c{}_{ab}\,\mu_c{}^r\,,} \tag{5.40}$$

or more simply it determines the Lie brackets of the vector fields in (5.32)

$$[T_a, T_b] = f^c{}_{ab}\,T_c\,, \tag{5.41}$$

ensuring that the Lie algebra is closed.

The Jacobi identity (5.17) requires

$$[T_a,[T_b,T_c]] + [T_c,[T_a,T_b]] + [T_b,[T_c,T_a]] = 0\,, \tag{5.42}$$

or in terms of the structure constants

$$\boxed{f^e{}_{ad}f^d{}_{bc} + f^e{}_{cd}f^d{}_{ab} + f^e{}_{bd}f^d{}_{ca} = 0\,.} \tag{5.43}$$

(5.43) is a necessary integrability condition for (5.40) which in turn is necessary for the integrability of (5.31).

The results (5.31), (5.40) with (5.42) and (5.39) are the contents of Lie's fundamental theorems for Lie groups.

Alternatively from (5.33) using

$$\frac{\partial}{\partial c^t}\,\mu_a{}^r(c) = -\mu_a{}^u(c)\frac{\partial}{\partial c^t}\lambda_u{}^c(c)\,\mu_c{}^r(c)\,, \tag{5.44}$$

we may obtain

$$\mu_a{}^s(b)\mu_b{}^t(b)\,\frac{\partial^2 c^r}{\partial b^s \partial b^t}\lambda_r{}^c(c) = \mu_a{}^s(b)\mu_b{}^t(b)\,\frac{\partial}{\partial b^t}\lambda_s{}^c(b) - \mu_b{}^t(c)\mu_a{}^u(c)\,\frac{\partial}{\partial c^t}\lambda_u{}^c(c)\,. \tag{5.45}$$

In a similar fashion as before this leads to

$$\mu_a{}^s(b)\mu_b{}^t(b)\,\frac{\partial}{\partial b^t}\lambda_s{}^c(b) - \mu_b{}^s(b)\mu_a{}^t(b)\,\frac{\partial}{\partial b^t}\lambda_s{}^c(b) = f^c{}_{ab}\,, \tag{5.46}$$

which is equivalent to (5.38), or

$$\frac{\partial}{\partial b^r}\lambda_s{}^c(b) - \frac{\partial}{\partial b^s}\lambda_r{}^c(b) = -f^c{}_{ab}\,\lambda_r{}^a(b)\lambda_s{}^b(b)\,. \tag{5.47}$$

Defining the *left invariant one-forms*

$$\omega^a(b) = \mathrm{d}b^r\lambda_r{}^a(b)\,, \tag{5.48}$$

the result is expressible more succinctly, as consequence of (5.23), by

$$\mathrm{d}\omega^a = -\tfrac{1}{2}f^b{}_{bc}\,\omega^b\wedge\omega^c\,. \tag{5.49}$$

Note that, using $\mathrm{d}(\omega^b\wedge\omega^c) = \mathrm{d}\omega^b\wedge\omega^c - \omega^b\wedge\mathrm{d}\omega^c$, $\mathrm{d}^2\omega^a = -\frac{1}{2}f^a{}_{b[c}f^b{}_{de]}\,\omega^c\wedge\omega^d\wedge\omega^e = 0$ by virtue of the Jacobi identity (5.43).

In general a n-dimensional manifold for which there are n vector fields which are linearly independent and non zero at each point is parallelisable. Examples are the circle S^1 and the 3-sphere S^3. A Lie group defines a parallelisable manifold since a basis for non zero vector fields is given by the left invariant fields in (5.32), the group $U(1)$ corresponnds to S^1 and $SU(2)$ to S^3.

5.2 Lie Algebra Definitions

In general a Lie algebra is a vector space $\mathfrak{g}$ with a commutator $[\,,\,] : \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ satisfying (5.13), (5.16) and (5.17), or in terms of a basis $\{T_a\}$, satisfying (5.41), with (5.39), and (5.42) or (5.43). Various crucial definitions, which are often linked to associated definitions for groups, are given below.

Two Lie algebras $\mathfrak{g}, \mathfrak{g}'$ are *isomorphic*, $\mathfrak{g}\simeq\mathfrak{g}'$, if there is a mapping between elements of the Lie algebras $X\leftrightarrow X'$ such that $[X,Y]' = [X',Y']$. If $\mathfrak{g}=\mathfrak{g}'$ the map is an *automorphism* of the Lie algebra. For any $\mathfrak{g}$ automorphisms form a group, the *automorphism group* of $\mathfrak{g}$.

The Lie algebra is *abelian*, corresponding to an abelian Lie group, if all commutators are zero, $[X,Y]=0$ for all $X,Y\in\mathfrak{g}$.

A *subalgebra* $\mathfrak{h}\subset\mathfrak{g}$ forms a Lie algebra itself and so is closed under commutation. If $H\subset G$ is a Lie group then its Lie algebra $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.

An *invariant subalgebra* or *ideal* $\mathfrak{h}\subset\mathfrak{g}$ is such that

$$[X,Y]\in\mathfrak{h} \qquad \text{for all} \quad Y\in\mathfrak{h}\,,\ X\in\mathfrak{g}\,. \tag{5.50}$$

If H is a normal Lie subgroup then its Lie algebra forms an ideal. Note that

$$\mathfrak{i} = [\,\mathfrak{g},\mathfrak{g}\,] = \{[X,Y] : X,Y\in\mathfrak{g}\}\,, \tag{5.51}$$

forms an ideal $\mathfrak{i}\subset\mathfrak{g}$, since $[Z,[X,Y]]\in\mathfrak{i}$ for all $Z\in\mathfrak{g}$. $\mathfrak{i}$ is called the *derived algebra*.

The *centre* of a Lie algebra $\mathfrak{g}$, $\mathcal{Z}(\mathfrak{g}) = \{Y : [X,Y]=0 \text{ for all } X\in\mathfrak{g}\}$.

A Lie algebra is *simple* if it does not contain any invariant subalgebra.

A Lie algebra is *semi-simple* if it does not contain any invariant abelian subalgebra.

Using the notation in (5.51) and we may define in a similar fashion a sequence of successive invariant subalgebras $\mathfrak{g}^{(n)}$, $n = 1, 2, \dots$, by

$$\mathfrak{g}^{(n+1)} = [\,\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}\,]\,, \qquad \mathfrak{g}^{(1)} = [\,\mathfrak{g}, \mathfrak{g}\,]\,. \tag{5.52}$$

A Lie algebra $\mathfrak{g}$ is *solvable* if $\mathfrak{g}^{(n+1)} = 0$ for some n, and so $\mathfrak{g}^{(n)}$ is abelian and the sequence terminates.

Solvable and semi-simple Lie algebras are clearly mutually exclusive. Lie algebras may be neither solvable nor semi-simple but in general they may be decomposed in terms of such Lie algebras.

The *direct sum* of two Lie algebras, $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 = \{X_1 + X_2 : X_1 \in \mathfrak{g}_1, X_2 \in \mathfrak{g}_2\}$, with the commutator

$$[X_1 + X_2, Y_1 + Y_2] = [X_1, X_2] + [Y_1, Y_2]\,. \tag{5.53}$$

It is easy to see that the direct sum $\mathfrak{g}$ contains $\mathfrak{g}_1$ and $\mathfrak{g}_2$ as invariant subalgebras so that $\mathfrak{g}$ is not simple. The Lie algebra for the direct product of two Lie groups $G = G_1 \otimes G_2$ is the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$.

If a Lie algebra $\mathfrak{g}$ can be defined to act linearly on a Lie algebra $\mathfrak{h}$ such that

$$Y \underset{X}{\to} Y^X\,, \quad (Y^X)^{X'} - (Y^{X'})^X = Y^{[X',X]} \qquad \text{for all} \quad Y \in \mathfrak{h}\,,\ X, X' \in \mathfrak{g}\,, \tag{5.54}$$

then we may define the *semi-direct sum* Lie algebra $\mathfrak{g} \oplus_s \mathfrak{h} = \{X + Y : X \in \mathfrak{h}, Y \in \mathfrak{h}\}$ with commutators $[X + Y, X' + Y'] = [X, X'] + Y'^X - Y^{X'} + [Y, Y']$. $\mathfrak{h}$ forms an invariant subalgebra of $\mathfrak{g} \oplus_s \mathfrak{h}$. The semi-direct sum of Lie algebras arises from the semi-direct product of Lie groups.

5.3 Lie Algebras for Matrix Lie Groups

The definition of the Lie algebra is more straightforward for matrix Lie groups. For a matrix group there are matrices $D(a)$, depending on the parameters a^r, realising the basic group multiplication rule (5.2),

$$D(a)D(b) = D(c)\,. \tag{5.55}$$

For group elements close to the identity with infinitesimal parameters θ^a we can now write

$$D(\theta) = I + \theta^a t_a\,, \tag{5.56}$$

which defins a set of matrices $\{t_a\}$ forming the *generators* for this matrix group. Writing

$$D(b + \mathrm{d}b) = D(b) + \mathrm{d}b^r \frac{\partial}{\partial b^r} D(b)\,, \tag{5.57}$$

then (5.26) becomes

$$\mathrm{d}b^r \frac{\partial}{\partial b^r} D(b) = \theta^a T_a D(b) = D(b)\theta^a t_a\,, \tag{5.58}$$

using (5.27) along with (5.32). Clearly

$$T_a D(b) = D(b) t_a \,, \tag{5.59}$$

and it then follows from (5.41) that

$$[t_a, t_b] = f^c{}_{ab} t_c \,. \tag{5.60}$$

The matrix generators $\{t_a\}$ hence obey the same Lie algebra commutation relations as $\{T_a\}$, and may be used to directly define the Lie algebra instead of the more abstract treatment in terms of vector fields.

5.3.1 $SU(2)$ Example

As a particular illustration we revisit $SU(2)$ and following (5.8) and (5.10) write

$$A(\mathbf{u}) = u_0 \, I + i \, \mathbf{u} \cdot \boldsymbol{\sigma} \qquad u_0 = \sqrt{1 - \mathbf{u}^2} \,. \tag{5.61}$$

This parameterisation is valid for $u_0 \geq 0$. With, for infinitesimal $\boldsymbol{\theta}$, $A(\boldsymbol{\theta}) = I + i \, \boldsymbol{\theta} \cdot \boldsymbol{\sigma}$ we get, using the standard results (2.12) to simplify products of Pauli matrices,

$$A(\mathbf{u} + \mathrm{d}\mathbf{u}) = A(\mathbf{u}) A(\boldsymbol{\theta}) = u_0 - \mathbf{u} \cdot \boldsymbol{\theta} + i(\mathbf{u} + u_0 \, \boldsymbol{\theta} - \mathbf{u} \times \boldsymbol{\theta}) \cdot \boldsymbol{\sigma} \,, \tag{5.62}$$

and hence

$$\mathrm{d}\mathbf{u} = a_0 \boldsymbol{\theta} - \mathbf{u} \times \boldsymbol{\theta} \,, \quad \text{or} \quad \mathrm{d}u_i = \theta_j \mu_{ji}(\mathbf{u}) \,, \quad \mu_{ji}(\mathbf{u}) = u_0 \, \delta_{ji} + u_k \, \varepsilon_{jki} \,. \tag{5.63}$$

The vector fields forming a basis for the Lie algebra $\mathfrak{s}u(2)$ are then

$$T_j(\mathbf{u}) = \mu_{ji}(\mathbf{u}) \frac{\partial}{\partial u_i} \qquad \Rightarrow \qquad \mathbf{T} = u_0 \, \boldsymbol{\nabla}_{\mathbf{u}} + \mathbf{u} \times \boldsymbol{\nabla}_{\mathbf{u}} \,. \tag{5.64}$$

Since

$$\mathbf{T} A(\mathbf{u}) = A(\mathbf{u}) \, i \boldsymbol{\sigma} \,, \tag{5.65}$$

and $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$, the Lie bracket must be

$$[T_i, T_j] = -2 \, \varepsilon_{ijk} \, T_k \,. \tag{5.66}$$

5.3.2 Representations and Lie Algebras

There is an intimate relation between representations of Lie algebras and Lie Groups. Just as described for groups in 1.4, a representation of a Lie algebra $\mathfrak{g}$ is of course such that for any $X \in \mathfrak{g}$ there are corresponding matrices $D(X)$ such that $D([X,Y]) = [D(X), D(Y)]$, where $[D(X), D(Y)]$ is the matrix commutator. For convenience we may take $D(T_a) = t_a$ where $\{t_a\}$ form a basis of matrices in the representation satisfying (5.60), following from (5.41). As for groups an irreducible representation of the Lie algebra is when there are no invariant subspaces of the corresponding representation space $\mathcal{V}$ under the action of all the

Lie algebra generators on $\mathcal{V}$. Just as for groups there is always a trivial representation by taking $D(X) = 0$.

The generators may be defined in terms of the representation matrices for group elements which are close to the identity,

$$D(g(\theta)) = I + \theta^a t_a + \mathrm{O}(\theta^2)\,, \qquad D(g(\theta))^{-1} = I - \theta^a t_a + \mathrm{O}(\theta^2)\,. \tag{5.67}$$

For unitary representations, as in (1.34), the matrix generators are then anti-hermitian,

$$t_a{}^\dagger = -t_a\,. \tag{5.68}$$

If the representation matrices have unit determinant, since $\det(I+\epsilon X) = 1+\epsilon\,\mathrm{tr}(X)+\mathrm{O}(\epsilon^2)$, we must also have

$$\mathrm{tr}(t_a) = 0\,. \tag{5.69}$$

In a physics context it is commonplace to redefine the matrix generators so that $t_a = -i\hat{t}_a$ so that, instead of (5.68), the generators $\hat{t}_a$ are hermitian and satisfy the commutation relations $[\hat{t}_a, \hat{t}_b] = i f^c{}_{ab}\hat{t}_c$.

Two representations of a Lie algebra $\{t'_a\}$ and $\{t_a\}$ are equivalent if, for some non singular S,

$$t'_a = S\,t_a S^{-1}\,. \tag{5.70}$$

For both representations to be unitary then S must be unitary. If the representation is irreducible then, by applying Schur's lemma,

$$t_a = S\,t_a S^{-1} \;\text{ or }\; [S, t_a] = 0 \quad\Rightarrow\quad S \propto I\,. \tag{5.71}$$

The complex conjugate of a representation is also a representation, in general it is inequivalent. If it is equivalent then, for some C,

$$t_a{}^* = C\,t_a C^{-1}\,, \tag{5.72}$$

or for a unitary representation, assuming (5.68),

$$C\,t_a C^{-1} = -t_a{}^T\,. \tag{5.73}$$

By considering the transpose we get $C^{-1T}C\,t_a\,C^{-1}C^T = t_a$ so that for an irreducible representation

$$C^{-1\,T}C = c\,I \quad\Rightarrow\quad C = c\,C^T \quad\Rightarrow\quad c = \pm 1\,. \tag{5.74}$$

If $C = C^T$ then, by a transformation $C \to S^T C S$ together with $t_a \to S^{-1}t_a S$, we can take $C = I$ and the representation is *real*. If $C = -C^T$ the representation is *pseudo-real*. For $\det C \neq 0$ the representation must be even dimensional, $2n$. By a transformation we may take $C = J$, $J^2 = -I$, where J is defined in (1.54). The representation matrices then satisfy $D(g(\theta))^\dagger = -JD(g(\theta))J$, which is just as in (1.64). This is sufficient to ensure that the pseudo-real representation formed by $\{D(g(\theta))\}$ can be expressed in terms of $n\times n$ matrices of quaternions, and so such representations are also referred to as quaternionic.

The $SO(3)$ spinor representation described in section 2.10 is pseudo-real since

$$C\,\boldsymbol{\sigma}\,C^{-1} = -\boldsymbol{\sigma}^T \qquad \text{for} \quad C = i\sigma_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \tag{5.75}$$

which is equivalent to (2.165).

A corollary of (5.73) is that,for real or pseudo-real representations,

$$\mathrm{tr}\big(t_{(a_1}\dots t_{a_n)}\big) = 0 \qquad \text{for} \ \ n \ \ \text{odd}\,. \tag{5.76}$$

For $n = 3$ this has important consequences in the discussion of anomalies in quantum field theories.

5.4 Relation of Lie Algebras to Lie Groups

The Lie algebra of a Lie group is determined by those group elements close to the identity. Nevertheless the Lie group can be reconstructed from the Lie algebra subject to various topological caveats. Firstly the group must be connected, for elements $g \in G$ there is a continuous path $g(s)$ with $g(0) = e$ and $g(1) = g$. Thus we must exclude reflections so that $SO(3)$ and $SO(3,1)^{\uparrow}$ are the connected groups corresponding to rotations and Lorentz transformations. Secondly for a Lie group G having a centre $\mathcal{Z}(G)$ which is a discrete abelian group, then for any subgroup $H_{\mathcal{Z}}(G) \subset \mathcal{Z}(G)$, where $H_{\mathcal{Z}}(G) = \{h\}$ with $gh = hg$ for all $g \in G$, the group $G/H_{\mathcal{Z}}(G)$, defined by $g \sim gh$, is also a Lie group with the same Lie algebra as G. As an example $SO(3)$ and $SU(2)$ have the the same Lie algebra although $SO(3) \simeq SU(2)/\mathbb{Z}_2$ where $\mathbb{Z}_2 = \mathcal{Z}(SU(2))$.

5.4.1 One-Parameter Subgroups

For any element $\theta^a T_a \in \mathfrak{g}$ there is a *one-parameter subgroup* of the associated Lie Group G corresponding to a path in $\mathcal{M}_G$ whose tangent at the identity is $\theta^a T_a$. With coordinates a^r the path is defined by a_s^r, with $s \in \mathbb{R}$, where

$$\frac{\mathrm{d}}{\mathrm{d}s}\, a_s^r = \theta^a {\mu_a}^r(a_s)\,, \quad a_0^r = 0\,, \qquad \text{or} \quad \frac{\mathrm{d}}{\mathrm{d}s}\, g(a_s) = \theta^a T_a(a_s)\, g(a_s)\,. \tag{5.77}$$

To verify that this forms a subgroup consider $g(c) = g(a_t)g(a_s)$ where from (5.2)

$$c^r = \varphi^r(a_t, a_s)\,. \tag{5.78}$$

Using (5.77) with (5.31) we get

$$\frac{\partial}{\partial s}\, c^r = \theta^a {\mu_a}^u(a_s)\, {\lambda_u}^b(a_s)\, {\mu_b}^r(c) = \theta^b {\mu_b}^r(c)\,, \qquad c^r\big|_{s=0} = a_t^r\,. \tag{5.79}$$

The equation is then identical with (5.77), save for the initial condition at $s = 0$, and the solution then becomes

$$c^r = a_{s+t}^r \quad \Rightarrow \quad g(a_t)g(a_s) = g(a_{s+t})\,. \tag{5.80}$$

Since

$$g(a_s)^{-1} = g(a_{-s}) \,, \tag{5.81}$$

then $\{g(a_s)\}$ forms an abelian subgroup of G depending on the parameter s. We may then define an *exponential map*

$$\exp : \mathfrak{g} \to G \,, \tag{5.82}$$

by

$$g(a_s) = \exp(s\,\theta^a T_a) \,. \tag{5.83}$$

For any representation we have

$$D(g(a_s)) = e^{s\,\theta^a t_a} \,, \tag{5.84}$$

where t_a are the matrix generators and the matrix exponential may be defined as an infinite power series, satisfying of course $e^{tX} e^{sX} = e^{(s+t)X}$ for any matrix X.

5.4.2 Baker Cambell Hausdorff Formula

In order to complete the construction of the Lie group G from the Lie algebra $\mathfrak{g}$ it is necessary to show how the group multiplication rules for elements belonging to different one-parameter groups may be determined, i.e for any $X, Y \in \mathfrak{g}$ we require

$$\exp(tX)\,\exp(tY) = \exp(Z(t)) \,, \qquad Z(t) \in \mathfrak{g} \,. \tag{5.85}$$

The *Baker Cambell Hausdorff*[15] *formula* gives an infinite series for $Z(t)$ in powers of t whose first terms are of the form

$$Z(t) = t(X+Y) + \tfrac{1}{2}t^2[X,Y] + \tfrac{1}{12}t^3\big([X,[X,Y]] - [Y,[X,Y]]\big) + \mathrm{O}(t^4) \,, \tag{5.86}$$

where the higher order terms involve further nested commutators of X and Y and so are determined by the Lie algebra $\mathfrak{g}$. For an abelian group we just have $Z(t) = t(X+Y)$. The higher order terms do not have a unique form since they can be rearranged using the Jacobi identity. Needless to say the general expression is virtually never a practical method of calculating group products, for once existence is more interesting than the final explicit formula.

We discuss here the corresponding matrix identity rather than consider the result for an abstract Lie algebra. It is necessary in the derivation to show how matrix exponentials can be differentiated so we first consider the matrix expression

$$f(s) = e^{s(Z+\delta Z)}\, e^{-sZ} \,, \tag{5.87}$$

and then

$$\frac{\mathrm{d}}{\mathrm{d}s} f(s) = e^{s(Z+\delta Z)}\delta Z\, e^{-sZ} = e^{sZ}\delta Z\, e^{-sZ} + \mathrm{O}(\delta Z^2) \,. \tag{5.88}$$

Solving this equation

$$f(1) = I + \int_0^1 \mathrm{d}s\; e^{sZ}\delta Z\, e^{-sZ} + \mathrm{O}(\delta Z^2) \,, \tag{5.89}$$

[15] Henry Frederick Baker, 1866-1956, British, senior wrangler 1887. John Edward Cambell, 1862-1924, Irish. Felix Hausdorff, 1868-1942, German.

so that

$$e^{Z+\delta Z} - e^{Z} = \int_0^1 \mathrm{d}s\, e^{sZ}\delta Z\, e^{(1-s)Z} + \mathrm{O}(\delta Z^2)\,. \tag{5.90}$$

Hence for any $Z(t)$ we have the result for the derivative of its exponential

$$\frac{\mathrm{d}}{\mathrm{d}t}\, e^{Z(t)} = \int_0^1 \mathrm{d}s\; e^{sZ(t)}\, \frac{\mathrm{d}}{\mathrm{d}t} Z(t)\, e^{(1-s)Z(t)}\,. \tag{5.91}$$

If, instead of (5.85), we suppose,

$$e^{tX}\, e^{tY} = e^{Z(t)}\,, \tag{5.92}$$

then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tX}\, e^{tY}\right) e^{-tY}\, e^{-tX} &= X + e^{tX}\, Y\, e^{-tX} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\, e^{Z(t)}\, e^{-Z(t)} = \int_0^1 \mathrm{d}s\; e^{sZ(t)}\, \frac{\mathrm{d}}{\mathrm{d}t}\, Z(t)\, e^{-sZ(t)}\,. \end{aligned} \tag{5.93}$$

With the initial condition $Z(0) = 0$ this equation then allows $Z(t)$ to be determined. To proceed further, using the formula for the exponential expansion

$$e^{A}\, B\, e^{-A} = B + [A,B] + \tfrac{1}{2}[A,[A,B]] + \dots\,, \tag{5.94}$$

(5.93) can be rewritten as an expansion in multiple commutators

$$X + e^{tX} Y\, e^{-tX} = \frac{\mathrm{d}}{\mathrm{d}t}\, Z(t) + \sum_{n=1}^{\infty} \frac{1}{(n+1)!}\, \underbrace{\big[Z(t), \dots \big[Z(t),}_{n} \frac{\mathrm{d}}{\mathrm{d}t}\, Z(t)\big] \dots \big]\,, \tag{5.95}$$

which may be solved iteratively by writing $Z(t) = \sum_{n=1}^{\infty} Z_n t^n$.

The results may be made somewhat more explicit if we adopt the notation

$$f(X^{\mathrm{ad}})Y = \sum_{n=0}^{\infty} f_n\, \underbrace{\big[X, \dots \big[X,}_{n} Y\big] \dots \big] \qquad \text{for} \qquad f(x) = \sum_{n=0}^{\infty} f_n x^n\,, \tag{5.96}$$

so that (5.94) becomes $e^{A}\, B\, e^{-A} = e^{A^{\mathrm{ad}}} B$. Then, since $\int_0^1 \mathrm{d}s\, e^{sz} = (e^z - 1)/z$, (5.93) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\, Z(t) = f\big(e^{Z(t)^{\mathrm{ad}}}\big)\big(X + e^{tX^{\mathrm{ad}}} Y\big)\,, \tag{5.97}$$

for, using the standard series expansion of $\ln(1+x)$,

$$f(x) = \frac{\ln x}{x-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}\, (x-1)^n\,. \tag{5.98}$$

Since

$$e^{Z(t)^{\mathrm{ad}}} U = e^{Z(t)}\, U\, e^{-Z(t)} = e^{tX} e^{tY} U\, e^{-tY} e^{-tX} = e^{tX^{\mathrm{ad}}} e^{tY^{\mathrm{ad}}} U\,, \tag{5.99}$$

we may replace $e^{Z(t)^{\rm ad}} \to e^{tX^{\rm ad}} e^{tY^{\rm ad}}$ on the right hand side of (5.97). With some intricate combinatorics (5.97) may then be expanded as a power series in t which on integration gives a series expansion for $Z(t)$ (a formula can be found on Wikipedia).

A simple corollary of these results is

$$e^{-tX}e^{-tY}e^{tX}e^{tY} = e^{t^2[X,Y]+{\rm O}(t^3)}\,, \tag{5.100}$$

so this combination of group elements isolates the commutator $[X, Y]$ as $t \to 0$.

5.5 Simply Connected Lie Groups and Covering Groups

For a connected topological manifold $\mathcal{M}$ then for any two points $x_1, x_2 \in \mathcal{M}$ there are continuous paths $p_{x_1\to x_2}$ linking x_1 and x_2 defined by functions $p_{x_1\to x_2}(s)$, $0 \le s \le 1$, where $p_{x_1\to x_2}(0) = x_1$, $p_{x_1\to x_2}(1) = x_2$. For three points x_1, x_2, x_3 a composition rule for paths linking x_1, x_2 and x_2, x_3 is given by

$$(p_{x_1\to x_2}\circ p_{x_2\to x_3})(s) = \begin{cases} p_{x_1\to x_2}(2s)\,, & 0 \le s \le \frac{1}{2}\,,\\ p_{x_2\to x_3}(2s-1)\,, & \frac{1}{2} \le s \le 1\,. \end{cases} \tag{5.101}$$

For any $p_{x_1\to x_2}$ the corresponding inverse, and also the trivial identity path, are defined by

$$p^{-1}_{x_2\to x_1}(s) = p_{x_1\to x_2}(1-s)\,, \qquad p^{\rm id}_{x\to x}(s) = x\,. \tag{5.102}$$

The set of paths give topological information about $\mathcal{M}$ by restricting to equivalence, or *homotopy*, classes $[p_{x_1\to x_2}] = \{p'_{x_1\to x_2} : p'_{x_1\to x_2} \sim p_{x_1\to x_2}\}$, where the homotopy equivalence relation requires that $p'_{x_1\to x_2}(s)$ can be continuously transformed to $p_{x_1\to x_2}(s)$. These homotopy classes inherit the composition rule $[p_{x_1\to x_2}]\circ[p_{x_2\to x_3}] = [p_{x_1\to x_2}\circ p_{x_2\to x_3}]$. The *fundamental group* for $\mathcal{M}$ is defined in terms of homotopy classes of closed paths starting and ending at an arbitrary point $x \in \mathcal{M}$,

$$\pi_1(\mathcal{M}) = \big\{[p_{x\to x}]\big\}\,. \tag{5.103}$$

This defines a group using the composition rule for group multiplication and for the identity $e = [p^{\rm id}_{x\to x}]$ and for the inverse $[p_{x\to x}]^{-1} = [p^{-1}_{x\to x}]$. For $\mathcal{M}$ connected $\pi_1(\mathcal{M})$ is independent of the point x chosen in (5.103). $\mathcal{M}$ is *simply connected* if $\pi_1(\mathcal{M})$ is trivial, so that $p_{x\to x} \sim p^{\rm id}_{x\to x}$ for all closed paths. If $\pi_1(\mathcal{M})$ is non trivial then $\mathcal{M}$ is multiply connected, if $\dim \pi_1(\mathcal{M}) = n$ there are n homotopy classes $[p_{x_1\to x_2}]$ for any x_1, x_2.

For Lie groups we can then define $\pi_1(G) \equiv \pi_1(\mathcal{M}_G)$. In many examples this is non trivial. For the rotation group $SO(3)$, as described earlier, $\mathcal{M}_{SO(3)} \simeq S^3/\mathbb{Z}_2$ where antipodal points, at the end of diameters, are identified. Alternatively, by virtue of (2.7), $\mathcal{M}_{SO(3)}$ may be identified with a ball of radius π in three dimensions with again antipodal points on the boundary S^2 identified. There are then closed paths, starting and finishing at the same point, which involve a jump between two antipodal points on S^3, or the surface of the ball, and which therefore cannot be contracted to the trivial constant path. For two antipodal jumps then by smoothly moving the corresponding diameters to coincide the closed path can be contracted to the trivial path. Hence

$$\pi_1(SO(3)) \simeq \mathbb{Z}_2\,. \tag{5.104}$$

As another example we may consider the group $U(1)$, as in (1.52), where it is clear that $\mathcal{M}_{U(1)} \simeq S^1$, the unit circle. For S^1 there are paths which wind round the circle n-times which are homotopically distinct for different n so that homotopy classes belonging to $\pi_1(U(1))$ are labelled by integers n. Under composition it is straightforward to see that the winding number is additive so that

$$\pi_1(U(1)) \simeq \mathbb{Z} \,, \tag{5.105}$$

which is an infinite discrete group in this case.

5.5.1 Covering Group

For a non simply connected Lie group G there is an associated simply connected Lie group $\overline{G}$, the *covering group*, with the same Lie algebra since G and $\overline{G}$ are identical near the identity. Assuming $\pi_1(G)$ has n elements then for any $g \in G$ we associate paths $p_{i,e\to g}$ where

$$p_{i,e\to g}(s) = g_i(s) \,, \quad g_i(0) = e \,, \quad g_i(1) = g \,, \qquad i = 0, \dots, n-1 \,, \tag{5.106}$$

corresponding to the n homotopically distinct paths from the identity e to any g. The elements of $\pi_1(G)$ can be identified with $[p_{i,e\to e}]$. We then define $\overline{G}$ such that the group elements are

$$g_i = \big(g, [p_{i,e\to g}]\big) \in \overline{G} \qquad \text{for all} \quad g \in G \,, \ i = 0, \dots, n-1 \,, \tag{5.107}$$

with a corresponding group product

$$g_{1i}\, g_{2j} = g_k \,, \quad \text{for} \quad g = g_1 g_2 \,, \quad [p_{k,e\to g_1 g_2}] = [p_{i,e\to g_1} \circ g_1 p_{j,e\to g_2}] \,, \tag{5.108}$$

using the path composition as in (5.101) and noting that $g_1 p_{j,e\to g_2}$ defines a path from g_1 to $g = g_1 g_2$. For the inverse and identity elements we have, with the definitions in (5.102),

$$g_i{}^{-1} = \big(g^{-1}, [g^{-1} p_{i,g\to e}^{-1}]\big) \,, \qquad e_0 = \big(e, [p_{0,e\to e}]\big) \,, \quad p_{0,e\to e} = p^{\text{id}}_{e\to e} \,. \tag{5.109}$$

These definitions satisfy the group properties although associativity requires some care. $\overline{G}$ contains the normal subgroup given by

$$\big\{e_i : i = 0, \dots, n-1\big\} \simeq \pi_1(G) \,, \qquad e_i = \big(e, [p_{i,e\to e}]\big) \,. \tag{5.110}$$

Any discrete normal subgroup H of a connected Lie group G must be moreover a subgroup of the centre $\mathcal{Z}(G)$, since if $h \in H$ then $ghg^{-1} \in H$ for any $g \in G$, by the definition of a normal subgroup. Since we may g vary continuously over all G, if G is a connected Lie group, and since H is discrete we must then have $ghg^{-1} = h$ for all g, which is sufficient to ensure that $h \in \mathcal{Z}(G)$.

The construction described above then ensures that the covering group $\overline{G}$ is simply connected and we have therefore demonstrated that

$$G \simeq \overline{G}/\pi_1(G) \,, \qquad \pi_1(G) \subset \mathcal{Z}(\overline{G}) \,. \tag{5.111}$$

As an application we consider the examples of $SO(3)$ and $U(1)$. For $SO(3)$ we consider rotation matrices $R(\theta, n)$ as in (2.6) but allow the rotation angle range to be extended to $0 \to 2\pi$. Hence, instead of (2.7), we have

$$n \in S^2 \, , \quad 0 \le \theta \le 2\pi \, , \qquad (\theta, n) \simeq (2\pi - \theta, -n) \, . \tag{5.112}$$

There are two homotopically inequivalent paths linking the identity to $R(\theta, n)$, $0 \le \theta \le \pi$, which may be defined, with the conventions in (5.112), by

$$p_{0,I\to R(\theta,n)}(s) = R(s\theta, n) \, , \quad p_{1,I\to R(\theta,n)}(s) = R(s(2\pi - \theta), -n) \, , \qquad 0 \le s \le 1 \, , \tag{5.113}$$

since $p_{1,I\to R(\theta,n)}$ involves a jump between antipodal points. The construction of the covering group then defines group elements $R(\theta, n)_i$, for $i = 0, 1$. For rotations about the same axis the group product rule then requires

$$R(\theta, n)_i \, R(\theta', n)_j = \begin{cases} R(\theta + \theta', n)_{\, i+j \bmod 2} \, , & 0 \le \theta + \theta' \le \pi \, , \\ R(\theta + \theta', n)_{\, i+j+1 \bmod 2} \, , & \pi \le \theta + \theta' \le 2\pi \, , \end{cases} \qquad 0 \le \theta, \theta' \le \pi \, . \tag{5.114}$$

It is straightforward to see that this is isomorphic to $SU(2)$, by taking $R(\theta, n)_0 \to A(\theta, n)$, $R(\theta, n)_1 \to -A(\theta, n)$, and hence $\overline{SO(3)} \simeq SU(2)$. For $U(1)$ with group elements as in (1.52) we may define

$$p_{n,1\to e^{i\theta}}(s) = e^{is(\theta + 2n\pi)} \, , \qquad 0 \le s \le 1 \, , \quad n \in \mathbb{Z} \, , \tag{5.115}$$

which are paths with winding number n. Writing the elements of the covering group $\overline{U(1)}$ as $g_n\big(e^{i\theta}\big)$ we have the product rule

$$g_n\big(e^{i\theta}\big) \, g_{n'}\big(e^{i\theta'}\big) = \begin{cases} g_{n+n'}\big(e^{i(\theta+\theta')}\big) \, , & 0 \le \theta + \theta' \le 2\pi \, , \\ g_{n+n'+1}\big(e^{i(\theta+\theta')}\big) \, , & 2\pi \le \theta + \theta' \le 4\pi \, , \end{cases} \qquad 0 \le \theta, \theta' \le 2\pi \, . \tag{5.116}$$

It is straightforward to see that effectively the group action is extended to all real θ, θ' so that $\overline{U(1)} \simeq \mathbb{R}$.

5.5.2 Projective Representations

For a non simply connected Lie group G then in general representations of the covering group $\overline{G}$ generate projective representations of G. Suppose $\{D(g_i)\}$ are representation matrices for $\overline{G}$, where $D(g_{1i}) \, D(g_{2j}) = D(g_k)$ for $g_{1i}, g_{2j}, g_k \in \overline{G}$ satisfying the group multiplication rule in (5.108). To restrict the representation to G it is necessary to restrict to a particular path, say i, since there is then a one to one correspondence $g_i \to g \in G$. Then, assuming $g_{1i} \, g_{2i} = g_j$ for some j,

$$D(g_{1i}) \, D(g_{2i}) = D(g_j) = D(g_j \, g_i^{-1}) \, D(g_i) = D(e_k) \, D(g_i) \, , \tag{5.117}$$

where, by virtue of (5.110) and (5.111),

$$g_j \, g_i^{-1} = e_k \in \mathcal{Z}(\overline{G}) \quad \text{for some } k \, . \tag{5.118}$$

Since e_k belongs to the centre, $D(e_k)$ must commute with $D(g_i)$ for any $g_i \in \overline{G}$ and so, for an irreducible representation must, by Schur's lemma, be proportional to the identity. Hence, for a unitary representation,

$$D(e_k) = e^{i\gamma_k} I \,, \tag{5.119}$$

where $\{e^{i\gamma_k} : k = 0, \dots, n-1\}$ form a one dimensional representation of $\pi_1(G)$. Combining (5.117) and (5.119) illustrates that $\{D(g_i)\}$, for i fixed, provide a projective representation of G as in (1.71).

For $SO(3)$ we have just $e^{i\gamma_k} = \pm 1$. For $U(1)$ then there are one-dimensional projective representations given by $e^{i\alpha\theta}$, for any real α, where we restrict $0 \le \theta < 2\pi$ which corresponds to a particular choice of path in the covering group. Then the multiplication rules become

$$e^{i\alpha\theta}\, e^{i\alpha\theta'} = \begin{cases} e^{i\alpha(\theta+\theta')}\,, & 0 \le \theta+\theta' \le 2\pi\,, \\ e^{2\pi i\alpha}\, e^{i\alpha(\theta+\theta'-2\pi)}\,, & 2\pi \le \theta+\theta' \le 4\pi\,. \end{cases} \tag{5.120}$$

5.6 Lie Algebra and Projective Representations

The possibility of different Lie groups for the same Lie algebra, as has been just be shown, can lead to projective representations with discrete phase factors. There are also cases when the phase factors vary continuously which can be discussed directly using the Lie algebra. We wish to analyse then possible solutions of the consistency conditions (1.72) modulo trivial solutions of the form (1.73) and show how this may lead to a modified Lie algebra.

For simplicity we write the phase factors γ which may appear in a projective representation of a Lie group G, as in (1.71), directly as functions on $\mathcal{M}_G \times \mathcal{M}_G$ so that, in terms of the group parameters in (5.1), we take $\gamma(g(a), g(b)) \equiv \gamma(a,b)$. The consistency condition (1.72) is then analysed with $g_i \to g(a)$, $g_j \to g(b)$, $g_k \to g(\theta)$ with θ infinitesimal and, with the same notation as in (5.26) and (5.28), this becomes

$$\gamma(c,\theta) + \gamma(a,b) = \gamma(a, b+\mathrm{d}b) + \gamma(b,\theta)\,. \tag{5.121}$$

Defining

$$\gamma_a(b) = \frac{\partial}{\partial\theta^a}\,\gamma(b,\theta)\bigg|_{\theta=0}\,, \tag{5.122}$$

and with (5.27) and the definition (5.32) then (5.121) becomes

$$T_a(b)\,\gamma(a,b) = \gamma_a(c) - \gamma_a(b)\,. \tag{5.123}$$

This differential equation for $\gamma(a,b)$ has integrability conditions obtained by considering

$$[T_a(b), T_b(b)]\,\gamma(a,b) = f^c{}_{ab} T_c(b)\,\gamma(a,b) \tag{5.124}$$

which applied to (5.123) and using $T_a(b) = T_a(c)$ from (5.32) leads to a separation of the dependence on b and c so each part must be constant. This gives

$$T_a(b)\,\gamma_b(b) - T_b(b)\,\gamma_a(b) - f^c{}_{ab}\,\gamma_c(b) = h_{ab} = -h_{ba}\,, \tag{5.125}$$

with h_{ab} a constant. Applying $T_c(b)$ and antisymmetrising the indices a, b, c gives, with (5.41),

$$0 = T_c\, h_{ab} + T_b\, h_{ca} + T_a\, h_{bc} = f^d{}_{ab}(T_d\gamma_c - T_c\gamma_d) + f^d{}_{bc}(T_d\gamma_a - T_a\gamma_d) + f^d{}_{ca}(T_d\gamma_b - T_b\gamma_d)\,, \quad (5.126)$$

and hence, with (5.125) and (5.43), there is then a constraint on h_{ab},

$$f^d{}_{ab}\, h_{dc} + f^d{}_{bc}\, h_{da} + f^d{}_{ca}\, h_{db} = 0\,. \quad (5.127)$$

As was discussed in 1.6 there are trivial solutions of the consistency conditions which are given by (1.73), and which, in the context of the Lie group considered here, are equivalent to taking $\gamma(a, b) = \alpha(c) - \alpha(a) - \alpha(b)$ for α any function on $\mathcal{M}_G$. From (5.26) we then have $\gamma(b, \theta) = \alpha(b + \mathrm{d}b) - \alpha(b) - \alpha(\theta)$ so that (5.122) gives

$$\gamma_a(b) = T_a(b)\, \alpha(b) - c_a\,, \qquad c_a = \frac{\partial}{\partial\theta^a}\, \alpha(\theta)\bigg|_{\theta=0}\,, \quad (5.128)$$

and then substituting in (5.125)

$$h_{ab} = f^c{}_{ab}\, c_c\,. \quad (5.129)$$

It is easy to verify that (5.128) and (5.129) satisfy (5.125) and (5.127)[16].

If there are unitary operators $U(a)$, corresponding to $g(a) \in G$, realising the Lie group G as a symmetry group in quantum mechanics then (1.71) requires

$$U(b)U(\theta) = e^{i\gamma_a(b)\theta^a}\, U(b + \mathrm{d}b)\,, \quad (5.130)$$

for infinitesimal θ^a. Assuming

$$U(\theta) = 1 - i\theta^a \hat{T}_a\,, \quad (5.131)$$

for hermitian operators $\hat{T}_a$, then, since $U(b + \mathrm{d}b) = U(b) + \theta^a T_a(b)U(b)$, we have

$$T_a(b)U(b) = -iU(b)\big(\hat{T}_a + \gamma_a(b)\big)\,. \quad (5.132)$$

By considering $[T_a, T_b]U(b)$ and using (5.125) then this requires that the hermitian operators $\{\hat{T}_a\}$ satisfy a modified Lie algebra

$$[\hat{T}_a, \hat{T}_b] = i f^c{}_{ab}\, \hat{T}_c - i\, h_{ab}\, 1\,. \quad (5.133)$$

The additional term involving h_{ab} is a *central extension* of the Lie algebra, it is the coefficient of the identity operator which commutes with all elements in the Lie algebra. A central extension, if present, is allowed by virtue of the freedom up to complex phases in quantum mechanics and they often play a crucial role. The consistency condition (5.127) is necessary for $\{\hat{T}_a\}$ to satisfy the Jacobi identity, if (5.129) holds then the central extension may be removed by the redefinition $\hat{T}_a \to \hat{T}_a + c_a\, 1$.

As shown subsequently non trivial central extensions are not present for semi-simple Lie algebras, it necessary for there to be an abelian subalgebra. A simple example arises for the Lie algebra $\mathfrak{iso}(2)$, given in (4.139), which has a central extension

$$[J_3, E_1] = iE_2\,, \qquad [J_3, E_2] = -iE_1\,, \qquad [E_1, E_2] = ic\, 1\,. \quad (5.134)$$

[16] Alternatively, using the left invariant one forms in (5.48) and defining $h = \frac{1}{2} h_{ab}\, \omega^a \wedge \omega^b$, then (5.127) is equivalent, by virtue of (5.49), to $\mathrm{d}h = 0$, so that h is closed, while the trivial solution (5.129) may be identified with $h = -\mathrm{d}c$, corresponding to h being exact, for $c = c_a \omega^a$. Thus projective representations depend on the cohomology classes of closed, modulo exact, two forms on $\mathcal{M}_G$.

5.6.1 Galilean Group

As an illustration of the significance of central extensions we consider the *Galilean Group*. Acting on space-time coordinated $\mathbf{x}, t$ this is defined by the transformations involving rotations, translations and velocity boosts

$$\mathbf{x}' = R\mathbf{x} + \mathbf{a} + \mathbf{v}t\,, \qquad t' = t + b\,, \tag{5.135}$$

where R is a rotation belonging to $SO(3)$. If we consider a limit of the Poincaré Lie algebra, with generators $\mathbf{J}, \mathbf{K}, \mathbf{P}, H$, by letting $\mathbf{K} \to c\mathbf{K}$, $H \to cM + c^{-1}H$ and take the limit $c \to \infty$ then the commutation relations from (4.42), (4.43) and (4.107), (4.108) become

$$\begin{aligned} &[J_i, J_j] = i\varepsilon_{ijk} J_k\,, \qquad [J_i, K_j] = i\varepsilon_{ijk} K_k\,, \qquad [K_i, K_j] = 0\,, \qquad [K_i, H] = iP_i\,, \\ &[K_i, P_j] = i\delta_{ij}\, M\,, \qquad [\mathbf{J}, M] = [\mathbf{K}, M] = [\mathbf{P}, M] = [H, M] = 0\,. \end{aligned} \tag{5.136}$$

When the Lie algebra is calculated just from the transformations in (5.135) the terms involving M are absent, the terms involving M are a central extension.

If we consider the just the subgroup formed by boosts and spatial translations then writing the associated unitary operators as

$$U[\mathbf{v}, \mathbf{a}] = e^{-i\mathbf{a}\cdot\mathbf{P}}\, e^{i\mathbf{v}\cdot\mathbf{K}}\,, \tag{5.137}$$

then a straightforward calculation shows that

$$U[\mathbf{v}', \mathbf{a}']\, U[\mathbf{v}, \mathbf{a}] = e^{iM\,\mathbf{v}'\cdot\mathbf{a}}\, U[\mathbf{v}' + \mathbf{v}, \mathbf{a}' + \mathbf{a}]\,. \tag{5.138}$$

For comparison with the preceding general discussion we should take $T_a \to (\boldsymbol{\nabla}_{\mathbf{v}}, \boldsymbol{\nabla}_{\mathbf{a}})$ and $\hat{T}_a \to (-\mathbf{K}, \mathbf{P})$. From (5.138) then $\gamma_a \to M(\mathbf{0}, \mathbf{v})$ and from (5.125) $h_{ab} \to M\left(\begin{smallmatrix} 0 & I_3 \\ -I_3 & 0 \end{smallmatrix}\right)$.

For representations of the Galilean group in quantum mechanics the central extension plays an essential role. Using (5.136)

$$e^{-i\mathbf{v}\cdot\mathbf{K}}\mathbf{P}e^{i\mathbf{v}\cdot\mathbf{K}} = \mathbf{P} + M\mathbf{v}\,, \qquad e^{i\mathbf{v}\cdot\mathbf{K}}He^{i\mathbf{v}\cdot\mathbf{K}} = H + \mathbf{P}\cdot\mathbf{v} + \tfrac{1}{2}M\mathbf{v}^2\,. \tag{5.139}$$

In a similar fashion to the Poincaré group we may define irreducible representations in terms of a basis for a space $\mathcal{V}_M$ obtained from a vector $|\mathbf{0}\rangle$, such that $\mathbf{P}|\mathbf{0}\rangle = \mathbf{0}$, by

$$|\mathbf{p}\rangle = e^{i\mathbf{v}\cdot\mathbf{K}}|\mathbf{0}\rangle\,, \qquad \mathbf{p} = M\mathbf{v}\,, \tag{5.140}$$

so that as a consequence of (5.140)

$$\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle\,, \qquad H|\mathbf{p}\rangle = \big(E_0 + \tfrac{\mathbf{p}^2}{2M}\big)|\mathbf{p}\rangle\,. \tag{5.141}$$

Clearly $\mathcal{V}_M$ corresponds to states of a nonrelativistic particle of mass M. The representation can easily be extended to include spin by requiring that $|\mathbf{0}\rangle$ belong to an irreducible representation of the rotation group.

5.7 Integration over a Lie Group, Compactness

For a discrete finite group $G = \{g_i\}$ then an essential consequence of the group axioms is that, for any function f on G, the sum $\sum_i f(g_i) = \sum_i f(gg_i)$ is invariant for any arbitrary $g \in G$. This result played a vital role in the proof of results about representations such as Schur's lemmas and the equivalence of any representation to a unitary representation. Here we describe how this may be extended to Lie groups where, since the group elements depend on continuously varying parameters, the discrete sum is replaced by a correspondingly invariant integration.

If we consider first the simplest case of $U(1)$, with elements as in (1.52) depending on an angle θ then a general function f on $U(1)$ is just a periodic function of θ, $f(\theta+2\pi) = f(\theta)$. Since the product rule for this abelian group is $e^{i\theta'}e^{i\theta} = e^{i(\theta'+\theta)}$ then, for periodic f,

$$\int_0^{2\pi} \mathrm{d}\theta \; f(\theta) = \int_0^{2\pi} \mathrm{d}\theta \; f(\theta'+\theta) \,. \tag{5.142}$$

provides the required invariant integration over $U(1)$. For the covering group $\mathbb{R}$, formed by real numbers under addition, the integration has to be extended to the whole real line.

For a general Lie group G then, with notation as in (5.1) and (5.2), we require an integration measure over the associated n-dimensional manifold $\mathcal{M}_G$ such that

$$\int_G \mathrm{d}\rho(b) \; f(g(b)) = \int_G \mathrm{d}\rho(b) \; f(g(c)) \quad \text{for} \quad g(c) = g(a)g(b) \,, \tag{5.143}$$

where $\mathrm{d}\rho(b) = \mathrm{d}^n b \, \rho(b)$. To determine $\rho(b)$ it suffices just to calculate the Jacobian J for the change of variables $b \to c(b)$, with fixed a, giving for the associated change of the n-dimensional integration volume elements

$$\mathrm{d}^n c = |J| \, \mathrm{d}^n b \,, \qquad J = \det \left[\frac{\partial c^r}{\partial b^s} \right] , \tag{5.144}$$

and then require, to satisfy (5.143),

$$\mathrm{d}\rho(b) = \mathrm{d}\rho(c) \quad \Rightarrow \quad \rho(b) = |J| \, \rho(c) \,. \tag{5.145}$$

For a Lie group the fundamental result (5.31), with (5.30), ensures that

$$J = \det \big[\lambda(b)\big] \; \det \big[\mu(c)\big] = \frac{\det \big[\mu(c)\big]}{\det \big[\mu(b)\big]} \,. \tag{5.146}$$

Comparing (5.144) and (5.146) with (5.145) show that the invariant integration measure over a general Lie group G is obtained by taking

$$\mathrm{d}\rho(b) = \frac{C}{\big| \det \big[\mu(b)\big] \big|} \, \mathrm{d}^n b \,. \tag{5.147}$$

for some convenient constant C. The normalisation of the measure is dictated by the form near the identity since for $b \approx 0$ then $\mathrm{d}\rho(b) \approx C \, \mathrm{d}^n b$.

A Lie group G is *compact* if the group volume is finite,

$$\int_G \mathrm{d}\rho(b) = |G| < \infty \,, \tag{5.148}$$

otherwise it is *non compact.* By rescaling $\rho(b)$ we may take $|G| = 1$. For a compact Lie group many of the essential results for finite groups remain valid, in particular all representations are equivalent to unitary representations, and correspondingly the matrices representing the Lie algebra can be chosen as anti-hermitian or hermitian, according to convention. Amongst matrix groups $SU(n), SO(n)$ are compact while $SU(n,m), SO(n,m)$, for $n,m > 0$, are non compact.

5.7.1 $SU(2)$ Example

For $SU(2)$ with the parameterisation in (5.61) the corresponding 3×3 matrix $[\mu_{ji}(\mathbf{u})]$ was computed in (5.63). It is not difficult to see that the eigenvalues are $u_0, u_0 \pm i|\mathbf{u}|$ so that in this case, since $u_0{}^2 + \mathbf{u}^2 = 1$,

$$\det[\mu_{ji}(\mathbf{u})] = u_0 \,. \tag{5.149}$$

Hence (5.147) requires

$$\mathrm{d}\rho(\mathbf{u}) = \frac{1}{|u_0|}\,\mathrm{d}^3 u \,, \qquad -1 \le u_0 \le 1 \,, \quad |\mathbf{u}| \le 1 \,. \tag{5.150}$$

where range of $u_0, \mathbf{u}$ is determined in order to cover $SU(2)$ matrices in (5.8). For the parameterisation in terms of $\theta, \mathbf{n}$, $\mathbf{n}^2 = 1$, as given by (2.28)

$$u_0 = \cos\tfrac{1}{2}\theta \,, \quad \mathbf{u} = -\sin\tfrac{1}{2}\theta \; \mathbf{n} \,, \qquad \mathrm{d}^3 u = |\mathbf{u}|^2 \mathrm{d}|\mathbf{u}|\,\mathrm{d}^2 n \,, \tag{5.151}$$

so that

$$\mathrm{d}\rho(\theta,\mathbf{n}) = \tfrac{1}{2}\sin^2\tfrac{1}{2}\theta\,\mathrm{d}\theta\,\mathrm{d}^2 n \,, \qquad 0 \le \theta \le 2\pi \,. \tag{5.152}$$

Since $\int_{S^2}\mathrm{d}^2 n = 4\pi$ the group volume is easily found

$$\int_{SU(2)} \mathrm{d}\rho(\theta,\mathbf{n}) = 2\pi^2 \,. \tag{5.153}$$

An alternative common parameterisation for $SU(2)$ in terms of Euler angles ϕ, θ, ψ is obtained by expressing a general $SU(2)$ matrix in the form

$$A = e^{-i\frac{1}{2}\phi\sigma_3} e^{-i\frac{1}{2}\theta\sigma_2} e^{-i\frac{1}{2}\psi\sigma_3} \,, \qquad 0 \le \phi \le 2\pi \,, \; 0 \le \theta \le \pi \,, \; 0 \le \psi \le 4\pi \,, \tag{5.154}$$

where the ranges are dictated by the need to cover all $SU(2)$ matrices. In terms of $u_0, \mathbf{u}$ as in (5.8),

$$\begin{aligned} u_0 &= \cos\tfrac{1}{2}\theta \, \cos\tfrac{1}{2}(\phi+\psi) \,, \qquad & u_3 &= -\cos\tfrac{1}{2}\theta \, \sin\tfrac{1}{2}(\phi+\psi) \,, \\ u_1 &= \sin\tfrac{1}{2}\theta \, \sin\tfrac{1}{2}(\phi-\psi) \,, \qquad & u_2 &= \sin\tfrac{1}{2}\theta \, \cos\tfrac{1}{2}(\phi-\psi) \,. \end{aligned} \tag{5.155}$$

Using $\mathrm{d}u_1 \wedge \mathrm{d}u_2 = -\frac{1}{8}\sin\theta\, \mathrm{d}\theta \wedge \mathrm{d}(\phi-\psi)$ and $\mathrm{d}u_1 \wedge \mathrm{d}u_2 \wedge \mathrm{d}u_3 = \frac{1}{8}\sin\theta\, u_0\, \mathrm{d}\theta \wedge \mathrm{d}\phi \wedge \mathrm{d}\psi$ then

$$\mathrm{d}\rho(\phi,\theta,\psi) = \tfrac{1}{8}\sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi\, \mathrm{d}\psi\,. \tag{5.156}$$

For $SO(3)$, since $SU(2)$ is a double cover, the group volume is halved. In terms of the parameterisation (θ, n) used in (5.152) we should take $0 \le \theta \le \pi$ or in terms of the Euler angles instead of (5.154) $0 \le \psi \le 2\pi$.

For compact Lie groups the orthogonality relations for representations (1.39) or characters (1.42) remain valid if the summation is replaced by invariant integration over the group and $|G|$ by the group volume as in (5.148). For $SU(2)$ the characters are given in (2.81) and using (5.152) and integrating over $n \in S^2$ we then obtain

$$\int_0^{2\pi} \mathrm{d}\theta\, \sin^2\tfrac{1}{2}\theta\; \chi_j(\theta)\, \chi_{j'}(\theta) = \pi\, \delta_{jj'}\,. \tag{5.157}$$

This may be easily verified directly using the explicit formula for χ_j in (2.81). For $SO(3)$, when j, j' are integral, the integration range may be reduced to $[0,\pi]$ with the coefficient on the right hand side halved.

5.7.2 Non Compact $Sl(2,\mathbb{R})$ Example

As an illustration of a non compact Lie group, we consider $Sl(2,\mathbb{R})$ consisting of real 2×2 matrices with determinant 1. With the Pauli matrices in (2.11) a general real 2×2 matrix may be expressed as

$$A = v_0 + v_1\sigma_1 + v_2\, i\sigma_2 + v_3\sigma_3\,, \tag{5.158}$$

where, for $A \in Sl(2,\mathbb{R})$, $v_0, \mathbf{v}$ are real and we must further impose

$$\det A = {v_0}^2 + {v_2}^2 - {v_1}^2 - {v_3}^2 = 1\,. \tag{5.159}$$

If we choose $\mathbf{v} = (v_1, v_2, v_3)$ as independent parameters, so that we may write $A(\mathbf{v})$, then for a infinitesimal $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ under matrix multiplication

$$A(\mathbf{v})\, A(\boldsymbol{\theta}) = A(\mathbf{v} + \mathrm{d}\mathbf{v})\,, \tag{5.160}$$

where, using the multiplication rules (2.12),

$$\begin{pmatrix} \mathrm{d}v_1 & \mathrm{d}v_2 & \mathrm{d}v_3 \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 \end{pmatrix} \begin{pmatrix} v_0 & v_3 & v_2 \\ v_3 & v_0 & -v_1 \\ -v_2 & -v_1 & v_0 \end{pmatrix}. \tag{5.161}$$

This defines the matrix $\mu(\mathbf{v})$, as in (5.27), for $Sl(2,\mathbb{R})$ with the parameter choice in (5.158). It is easy to calculate, with (5.159),

$$\det \mu(\mathbf{v}) = v_0\,, \tag{5.162}$$

so that the invariant integration measure becomes

$$\mathrm{d}\rho(\mathbf{v}) = \frac{1}{|v_0|}\, \mathrm{d}^3 v\,. \tag{5.163}$$

Unlike the case for $SU(2)$ the parameters $\mathbf{v}$ have an infinite range so that the group volume diverges.

For an alternative parameterisation we may take

$$v_0 = \cosh\alpha \, \cos\beta \, , \; v_2 = \cosh\alpha \, \sin\beta \, , \; v_1 = \sinh\alpha \, \cos\gamma \, , \; v_3 = \sinh\alpha \, \sin\gamma \, , \\ \alpha \geq 0 \, , \qquad 0 \leq \beta, \gamma \leq 2\pi \, . \tag{5.164}$$

In this case the $Sl(2,\mathbb{R})$ integration measure becomes

$$\mathrm{d}\rho(\alpha,\beta,\gamma) = \tfrac{1}{2} \sinh 2\alpha \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma \, , \tag{5.165}$$

which clearly demonstrates the diverging form of the α integration. For $\beta, \gamma = 0$ the $Sl(2,\mathbb{R})$ matrix given by (5.164) reduces to one for $SO(1,1)$ as in (1.59).

The group $Sl(2,\mathbb{R})$ is related to a pseudo-orthogonal group in a similar fashion to $SU(2)$ and $SO(3)$. For a basis of real traceless 2×2 matrices

$$\hat{\boldsymbol{\sigma}} = (\sigma_1, i\sigma_2, \sigma_3) \, , \tag{5.166}$$

then we may define, for arbitrary real 3 vectors $\mathbf{x}$, a linear transformation $\mathbf{x} \to \mathbf{x}'$ by

$$\hat{\boldsymbol{\sigma}} \cdot \mathbf{x}' = A \, \hat{\boldsymbol{\sigma}} \cdot \mathbf{x} A^{-1} \, , \qquad A \in Sl(2,\mathbb{R}) \, , \tag{5.167}$$

such that the quadratic form

$$\det \hat{\boldsymbol{\sigma}} \cdot \mathbf{x} = x_1{}^2 + x_3{}^2 - x_2{}^2 \, , \tag{5.168}$$

is invariant. This then demonstrates that $SO(2,1) \simeq Sl(2,\mathbb{R})/\mathbb{Z}_2$.

Additionally $Sl(2,\mathbb{R}) \simeq SU(1,1)$. For any $B \in SU(1,1)$ we must have

$$B \sigma_3 B^\dagger = \sigma_3 \, , \qquad \det B = 1 \, . \tag{5.169}$$

Writing

$$B = w_0 + w_1 \sigma_1 + w_2 \sigma_2 - w_3 \, i\sigma_3 \, , \tag{5.170}$$

then for $w_0, \mathbf{w}$ real $\sigma_3 B^\dagger = B^{-1}\sigma_3$ so long as

$$\det B = w_0{}^2 + w_3{}^2 - w_1{}^2 - w_3{}^2 = 1 \, . \tag{5.171}$$

Hence for any $A(\mathbf{v}) \in Sl(2,\mathbb{R})$ it is clear that $e^{i\frac{1}{4}\pi\sigma_1} A(\mathbf{v}) e^{-i\frac{1}{4}\pi\sigma_1} = B(\mathbf{w}) \in SU(1,1)$, with $\mathbf{w} = (v_1, v_3, v_2)$, showing the isomorphism between these two non compact Lie groups.

5.8 Adjoint Representation and its Corollaries

A Lie algebra $\mathfrak{g}$ is just a vector space with also a bilinear commutator, $[\,,\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, subject only to the requirement that the commutator is antisymmetric and satisfies the Jacobi identity. The vector space defines the representation space for the *adjoint representation* which plays an absolutely fundamental role in the analysis of Lie algebras.

For any $X, Y \in \mathfrak{g}$ then

$$Y \underset{X}{\rightarrow} [X,Y] = X^{\rm ad}\, Y\,, \tag{5.172}$$

defines the linear map$X^{\rm ad} : \mathfrak{g} \to \mathfrak{g}$. There is also a corresponding adjoint representation for the associated Lie group G. For any $X \in \mathfrak{g}$ the associated one parameter group is given by $\exp(sX) \in G$ and then the adjoint representation $D^{\rm ad}$ is defined by

$$Y \xrightarrow[\exp(X)]{} D^{\rm ad}\big(\exp(sX)\big)Y = e^{sX^{\rm ad}}Y = \sum_{n=0}^{\infty} \frac{s^n}{n!} \underbrace{\big[X, \dots \big[X}_{n}, Y\big] \dots\big]\,, \tag{5.173}$$

with similar notation to (5.96). To verify that (5.172) provides a representation of the Lie algebra the Jacobi identity is essential since from

$$Z^{\rm ad} X^{\rm ad}\, Y = [Z, [X,Y]]\,, \tag{5.174}$$

we obtain for the adjoint commutator, using (5.17),

$$\big[Z^{\rm ad}, X^{\rm ad}\big]\, Y = [Z,[X,Y]] - [X,[Z,Y]] = [[Z,X],Y] = [Z,Y]^{\rm ad}\, Y\,, \tag{5.175}$$

and hence in general

$$\big[Z^{\rm ad}, X^{\rm ad}\big] = [Z,Y]^{\rm ad}\,. \tag{5.176}$$

Explicit adjoint representation matrices are obtained by choosing a basis for $\mathfrak{g}$, $\{T_a\}$ so that for any $Y \in \mathfrak{g}$ then $Y = T_a Y^a$ and (5.172) becomes

$$X^{\rm ad}\, Y = T_c\, (X^{\rm ad})^c{}_b Y^b\,. \tag{5.177}$$

For the generators T_a the corresponding adjoint representation matrices are then given by

$$[T_a, T_b] = T_c\, (T_a{}^{\rm ad})^c{}_b \qquad \Rightarrow \qquad (T_a{}^{\rm ad})^c{}_b = f^c{}_{ab}\,, \tag{5.178}$$

using (5.41). The commutator

$$\big[T_a{}^{\rm ad}, T_b^{\rm ad}\big] = f^c{}_{ab}\, T_c^{\rm ad}\,, \tag{5.179}$$

is directly equivalent to (5.42). The group representation matrices $D^{\rm ad}(\exp X) = e^{X^{\rm ad}}$, with $X^{\rm ad} = T_a{}^{\rm ad} X^a$, are then obtained using the matrix exponential. Close to the identity, in accord with (5.67),

$$D^{\rm ad}(\exp X) = I + X^{\rm ad} + {\rm O}(X^2)\,. \tag{5.180}$$

If the Lie algebra is abelian then clearly $X^{\rm ad} = 0$ for all X so the adjoint representation is trivial.

For $\mathfrak{su}(2)$

$$[T_i, T_j] = i\varepsilon_{ijk}\, T_k \qquad \Rightarrow \qquad (T_i^{\rm ad})_{jk} = -i\varepsilon_{ijk}\,, \tag{5.181}$$

where $\mathbf{T}^{\rm ad}$ are three 3×3 hermitian matrices. If $\mathbf{n}$ is a unit vector $(\mathbf{n}\cdot\mathbf{T}^{\rm ad})^2 = I - n\, n^T$ from which we may deduce that $\mathbf{n}\cdot\mathbf{T}^{\rm ad}$ has eigenvalues $\pm 1, 0$ so that this is the spin 1 representation. For the the Lie algebra $\mathfrak{iso}(2)$, as given in (4.139), we have

$$E_1^{\rm ad} = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_2^{\rm ad} = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad J_3^{\rm ad} = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.182}$$

5.8.1 Killing Form

The *Killing*[17] *form*, although apparently due to Cartan, provides a natural symmetric bilinear form, analogous to a metric, for the Lie algebra $\mathfrak{g}$. It is defined using the trace, over the vector space $\mathfrak{g}$, of the adjoint representation matrices by

$$\kappa(X,Y) = \mathrm{tr}\big(X^{\mathrm{ad}}Y^{\mathrm{ad}}\big) \qquad \text{for all} \quad X,Y \in \mathfrak{g}\,, \tag{5.183}$$

or in terms of a basis as in (5.178)

$$\kappa_{ab} = \kappa(T_a,T_b) = f^c{}_{ad}\, f^d{}_{bc}\,, \tag{5.184}$$

so that $\kappa(X,Y) = \kappa_{ab}X^aY^b$. Clearly it is symmetric $\kappa_{ab} = \kappa_{ba}$.

The importance of the Killing form arises from the crucial invariance condition

$$\kappa([Z,X],Y) + \kappa(X,[Z,Y]) = 0\,. \tag{5.185}$$

The verification of this is simple since, from (5.176),

$$\kappa([Z,X],Y) = \mathrm{tr}\big([Z,X]^{\mathrm{ad}}Y^{\mathrm{ad}}\big) = \mathrm{tr}\big([Z^{\mathrm{ad}},X^{\mathrm{ad}}]\,Y^{\mathrm{ad}}\big)\,, \tag{5.186}$$

and then (5.185) follows from $\mathrm{tr}\big([Z^{\mathrm{ad}},X^{\mathrm{ad}}]\,Y^{\mathrm{ad}}\big) + \mathrm{tr}\big(X^{\mathrm{ad}}\,[Z^{\mathrm{ad}},Y^{\mathrm{ad}}]\big) = 0$, using cyclic symmetry of the matrix trace. The result (5.185) also shows that the Killing form is invariant under the action of the corresponding Lie group G since

$$\kappa\big(e^{sZ^{\mathrm{ad}}}X,\, e^{sZ^{\mathrm{ad}}}Y\big) = \kappa\big(X,Y\big)\,, \tag{5.187}$$

which follows from (5.173) and differentiating with respect to s and then using (5.185).

Alternatively (5.185) may be expressed in terms of components using

$$\kappa\big([T_c,T_a],T_b\big) = f^d{}_{ca}\kappa(T_d,T_b) = f^d{}_{ca}\,\kappa_{db} \equiv f_{cab}\,, \tag{5.188}$$

in a form expressing κ_{ab} as an invariant tensor for the adjoint representation

$$\kappa_{db}\, f^d{}_{ca} + \kappa_{ad}\, f^d{}_{cb} = 0 \qquad \Leftrightarrow \qquad f_{cab} + f_{cba} = 0\,. \tag{5.189}$$

Since, from (5.39), $f_{cab} + f_{cba} = 0$ this implies

$$f_{abc} = f_{[abc]}\,. \tag{5.190}$$

If the Lie algebra $\mathfrak{g}$ contains an invariant subalgebra $\mathfrak{h}$ then in an appropriate basis we may write

$$T_a = (T_i,T_r)\,, \quad T_i \in \mathfrak{h} \qquad [T_i,T_j] = f^k{}_{ij}T_k\,, \quad [T_r,T_i] = f^j{}_{ri}T_j\,, \tag{5.191}$$

so that the Killing form restricted to $\mathfrak{h}$ is just

$$\kappa_{ij} = f^k{}_{il}\, f^l{}_{jk} = \mathrm{tr}_{\mathfrak{h}}(T_i^{\mathrm{ad}}T_j^{\mathrm{ad}})\,. \tag{5.192}$$

[17] Wilhelm Karl Joseph Killing, 1847-1923, German.

The crucial property of the Killing form is the invariance condition (5.185). If g_{ab} also defines an invariant bilinear form on the Lie algebra, as in (5.185), so that

$$g_{ab}\big([Z,X]^a Y^b + X^a [Z,Y]^b\big) = 0\,, \tag{5.193}$$

then, for any solution λ_i of $\det[\kappa_{ab} - \lambda g_{ab}] = 0$, $\mathfrak{h}_i = \{X_i : (\kappa_{ab} - \lambda_i g_{ab})X_i^b = 0\}$ forms, by virtue of the invariance condition (5.193), an invariant subalgebra $\mathfrak{h}_i \subset \mathfrak{g}$. Restricted to $\mathfrak{h}_i$ the Killing form κ_{ab} and g_{ab} are proportional. For a simple Lie algebra, when there are no invariant subalgebras, the Killing form is essentially unique.

For a compact group the adjoint representation $D^{\rm ad}$ may be chosen to be unitary so that in (5.180) the adjoint Lie algebra generators are anti-hermitian, as in (5.68),

$$X^{\rm ad\dagger} = -X^{\rm ad}\,. \tag{5.194}$$

In this case

$$\kappa(X,X) \le 0\,, \qquad \kappa(X,X) = 0 \quad \Leftrightarrow \quad X^{\rm ad} = 0\,. \tag{5.195}$$

For $\mathfrak{su}(2)$ using (5.181)

$$\kappa_{ij} = {\rm tr}(T_i^{\rm ad}\, T_i^{\rm ad}) = i^2 \varepsilon_{ikl}\, \varepsilon_{jlk} = 2\, \delta_{ij}\,. \tag{5.196}$$

However for $\mathfrak{iso}(2)$ then, if $T_a = (E_1, E_2, J_3)$, $a = 1,2,3$, it is easy to see from (5.182)

$$[\kappa_{ab}] = 2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,. \tag{5.197}$$

5.8.2 Conditions for Non Degenerate Killing Form

For the Killing form to play the role of a metric on the Lie algebra then it should be non-degenerate, which requires that if $\kappa(Y,X) = 0$ for all $Y \in \mathfrak{g}$ then $X = 0$ or more simply $\det[\kappa_{ab}] \neq 0$ so that $\kappa_{ab} Y^b = 0$ has no non trivial solution. An essential theorem due to Cartan gives the necessary and sufficient conditions for this to be true. Using the definition of a semi-simple Lie algebra given in 5.2 we have;

Theorem The Killing form is non-degenerate if and only if the Lie algebra is semi-simple.

To demonstrate that if the Lie algebra is not semi-simple the Killing form is degenerate is straightforward. Assume there is an invariant abelian subalgebra $\mathfrak{h}$ with a basis $\{T_i\}$ so that

$$T_a = (T_i, T_r) \qquad \Rightarrow \qquad [T_i, T_j] = 0\,, \quad [T_r, T_i] = f^j{}_{ri} T_j\,. \tag{5.198}$$

Then from (5.184)

$$\kappa_{ai} = f^c{}_{ad}\, f^d{}_{ic} = f^r{}_{aj}\, f^j{}_{ir} = 0\,, \qquad \text{since} \quad f^r{}_{sj} = f^r{}_{kj} = 0\,, \tag{5.199}$$

which is equivalent to $\kappa(Y,X) = 0$ for $X \in \mathfrak{h}$ and all $Y \in \mathfrak{g}$. The converse is less trivial. For a Lie algebra $\mathfrak{g}$, if $\det[\kappa_{ab}] = 0$ then $\mathfrak{h} = \{X : \kappa(Y,X) = 0, \text{for all } Y \in \mathfrak{g}\}$ forms a non trivial invariant subalgebra, since $\kappa(Y,[Z,X]) = -\kappa([Z,Y],X) = 0$, for any $Z,Y \in \mathfrak{g}$, $X \in \mathfrak{h}$. Thus $\mathfrak{g}$ is not simple. The proof that $\mathfrak{g}$ is not semi-simple then consists in showing that $\mathfrak{h}$ is solvable, so that, with the definition in (5.52), $\mathfrak{h}^{(n)}$ is abelian for some n. The alternative would require $\mathfrak{h}^{(n)} = \mathfrak{h}^{(n+1)}$, for some n, but this is incompatible with $\kappa(X,Y) = 0$ for all $X,Y \in \mathfrak{h}$.

The results (5.196) and (5.197) illustrate that $\mathfrak{su}(2)$ is semi-simple, whereas $\mathfrak{iso}(2)$ is not, it contains an invariant abelian subalgebra.

For a compact Lie group G the result that a degenerate Killing form for a Lie algebra $\mathfrak{g}$ implies the presence of an abelian invariant subalgebra follows directly from (5.195) since if $X^{\mathrm{ad}} = 0$, X commutes with all elements in $\mathfrak{g}$. For the compact case the Lie algebra can be decomposed into a semi-simple part and an abelian part so that the group has the form

$$G \simeq G_{\mathrm{semi-simple}} \otimes U(1) \otimes \cdots \otimes U(1)/F\,, \tag{5.200}$$

with a $U(1)$ factor for each independent Lie algebra element with $X^{\mathrm{ad}} = 0$ and where F is some finite abelian group belonging to the centre of G.

5.8.3 Decomposition of Semi-simple Lie Algebras

If a semi-simple Lie algebra $\mathfrak{g}$ contains an invariant subalgebra $\mathfrak{h}$ then the adjoint representation is reducible. However it may be decomposed into a direct sum of simple Lie algebras for each of which the adjoint representation is irreducible. To verify this let

$$\mathfrak{h}_\perp = \{X : \kappa(X,Y) = 0,\, Y \in \mathfrak{h}\}\,. \tag{5.201}$$

Then $\mathfrak{h}_\perp$ is also an invariant subalgebra since, for any $X \in \mathfrak{h}_\perp$ and $Z \in \mathfrak{g}, Y \in \mathfrak{h}$, $\kappa([Z,X],Y) = -\kappa(X,[Z,Y]) = 0$. Furthermore $\mathfrak{h}_\perp \cap \mathfrak{h} = 0$ since otherwise, by the definition of $\mathfrak{h}_\perp$ in (5.201), there would be a $X \in \mathfrak{h}_\perp$ and also $X \in \mathfrak{h}$ so that $\kappa(X,Z) = 0$ for all $Z \in \mathfrak{g}$ which contradicts the Killing form being non-degenerate. Hence

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}_\perp\,. \tag{5.202}$$

This decomposition may be continued to give until there are no remaining invariant spaces

$$\mathfrak{g} = \bigoplus_i \mathfrak{g}_i\,, \qquad \mathfrak{g}_i \ \text{ simple}\,. \tag{5.203}$$

For the Lie algebra there is then a basis $\{T_a^{(i)}\}$, such that for each individual i this represents a basis for $\mathfrak{g}_i$, $a = 1, \dots \dim \mathfrak{g}_i$, and with the generators for $\mathfrak{g}_i, \mathfrak{g}_j$, $i \neq j$ commuting as in (5.53) and $\kappa(T_a^{(i)}, T_b^{(j)}) = 0$, $i \neq j$. For any $X,Y \in \mathfrak{g}$ then the Killing form becomes a sum

$$X = \sum_i X_i\,, \quad Y = \sum_i Y_i\,, \qquad \kappa(X,Y) = \sum_i \mathrm{tr}_{\mathfrak{g}_i}\big(X_i^{\mathrm{ad}} Y_i^{\mathrm{ad}}\big)\,, \tag{5.204}$$

The corresponding decomposition for the associated Lie group becomes $G = \otimes_i G_i$.

With this decomposition the study of semi-simple Lie algebras is then reduced to just simple Lie algebras.

5.8.4 Casimir Operators and Central Extensions

For semi-simple Lie algebras we may easily construct a quadratic Casimir operator for any representation and also show that there are no non trivial central extensions.

The restriction to semi-simple Lie algebras, $\det[\kappa_{ab}] \neq 0$, ensures that the Killing form $\kappa = [\kappa_{ab}]$ has an inverse $\kappa^{-1} = [\kappa^{ab}]$, so that $\kappa_{ac}\,\kappa^{cb} = \delta_a{}^b$, and we may then use κ^{ab} and κ_{ab} to raise and lower Lie algebra indices, just as with a metric. The invariance condition (5.189) becomes $\kappa\, T_a^{\rm ad} + T_a^{{\rm ad}T}\kappa = 0$ so that from $[T_a^{\rm ad}, \kappa^{-1}\kappa] = 0$ we obtain $T_a^{\rm ad}\kappa^{-1} + \kappa^{-1}T_a^{{\rm ad}T} = 0$ or

$$f^b{}_{ad}\,\kappa^{dc} + f^c{}_{ad}\,\kappa^{bd} = 0\,, \tag{5.205}$$

showing that κ^{ab} is also an invariant tensor. Hence, for any representation of the Lie algebra in terms of $\{t_a\}$ satisfying (5.60), then

$$[t_a, \kappa^{bc}\, t_b t_c] = \kappa^{bc}\big(f^d{}_{ab}\, t_d t_c + f^d{}_{ac}\, t_b t_d\big) = \big(\kappa^{be} f^d{}_{ab} + \kappa^{dc} f^e{}_{ac}\big)\, t_d t_e = 0\,. \tag{5.206}$$

In consequence $\kappa^{ab}\, t_a t_b$ is a *quadratic Casimir operator.*

To discuss central extensions we rewrite the fundamental consistency condition (5.127) in the form

$$h_{ae}\, f^e{}_{cd} = -h_{de}\, f^e{}_{ac} - h_{ce}\, f^e{}_{da}\,. \tag{5.207}$$

Then using (5.205)

$$h_{ae}\, f^e{}_{cd}\, f^c{}_{bg}\, \kappa^{gd} = -h_{ae}\, f^g{}_{cd}\, f^c{}_{bg}\, \kappa^{de} = h_{ae}\, \kappa_{db}\, \kappa^{ed} = h_{ab}\,, \tag{5.208}$$

and also, with (5.205) again,

$$\begin{aligned}\big(h_{de}\, f^e{}_{ac} + h_{ce}\, f^e{}_{da}\big) f^c{}_{bg}\, \kappa^{gd} &= \big(h_{de}\, f^e{}_{ac}\, f^c{}_{bg} + h_{ec}\, f^c{}_{bg}\, f^e{}_{ad}\big)\kappa^{gd} \\ &= h_{de}\, f^e{}_{ac}\, f^c{}_{bg}\, \kappa^{gd} - h_{ec}\, f^c{}_{bg}\, f^g{}_{ad}\, \kappa^{de}\end{aligned} \tag{5.209}$$

we may obtain from (5.207), re-expressing (5.209) as a matrix trace,

$$h_{ab} = -{\rm tr}\big(h\,[T_a^{\rm ad}, T_b^{\rm ad}]\kappa^{-1}\big) = -{\rm tr}\big(h\, T_c^{\rm ad}\kappa^{-1}\big) f^c{}_{ab}\,. \tag{5.210}$$

Hence h_{ab} is of the form given in (5.129) which demonstrates that for sem-simple Lie algebras there are no non trivial central extensions. Central extensions therefore arise only when are invariant abelian subalgebras.

5.9 Bases for Lie Algebras for Matrix Groups

Here we obtain the Lie algebras $\mathfrak{g}$ corresponding to the various continuous matrix groups G described in section 1.5 by considering matrices close to the identity

$$M = I + X + {\rm O}(X^2)\,, \tag{5.211}$$

with suitable conditions on X depending on the particular group.

For $\mathfrak{u}(n)$, X is a complex $n\times n$ matrix satisfying $X^\dagger = -X$ and for $\mathfrak{su}(n)$, also $\mathrm{tr}(X) = 0$. It is convenient to consider first a basis formed by the n^2, $n \times n$, matrices $\{R^i{}_j\}$, where $R^i{}_j$ has 1 in the i'th row and j'th column and is otherwise zero,

$$R^i{}_j = \begin{matrix} \\ \\ i \\ \\ \\ \\ \end{matrix}\!\!\begin{pmatrix} 0 & 0 & \dots & \overset{j}{0} & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & & & 1 & & 0 \\ \vdots & & & & & \vdots \\ 0 & & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix}, \quad i,j = 1,\dots,n\,, \qquad (R^i{}_j)^\dagger = R^j{}_i\,. \tag{5.212}$$

These matrices satisfy

$$\big[R^i{}_j, R^k{}_l\big] = \delta^k{}_j\, R^i{}_l - \delta^i{}_l\, R^k{}_j\,, \tag{5.213}$$

and

$$\mathrm{tr}\big(R^i{}_j\, R^k{}_l\big) = \delta^k{}_j\, \delta^i{}_l\,. \tag{5.214}$$

In general $X = R^i{}_j X^j{}_i \in \mathfrak{gl}(n)$ for arbitrary $X^j{}_i$ so that $\{R^i{}_j\}$ form a basis for $\mathfrak{gl}(n)$. If $\sum_j X^j{}_j = 0$ then $X \in \mathfrak{sl}(n)$ while if $(X^j{}_i)^* = -X^i{}_j$ then $X = -X^\dagger \in \mathfrak{u}(n)$. For the associated adjoint matrices

$$[X, R^i{}_j] = X^i{}_k\, R^k{}_j - T^i{}_k\, X^k{}_j \quad \Rightarrow \quad (X^{\mathrm{ad}})^l{}_{k,}{}^i{}_j = X^i{}_k\, \delta^l{}_j - X^l{}_j\, \delta^i{}_k\,. \tag{5.215}$$

Hence, for $X = R^i{}_j X^j{}_i, Y = R^i{}_j Y^j{}_i$,

$$\kappa(X,Y) = \mathrm{tr}\big(X^{\mathrm{ad}} Y^{\mathrm{ad}}\big) = 2\big(n \textstyle\sum_{i,j} X^j{}_i\, Y^i{}_j - \sum_i X^i{}_i \sum_j Y^j{}_j\big)\,. \tag{5.216}$$

Restricting to $\mathfrak{u}(n)$

$$\kappa(X,X) = -2n \textstyle\sum_{i,j} |\hat{X}^j{}_i|^2\,, \qquad \hat{X}^j{}_i = X^j{}_i - \frac{1}{n}\, \delta^j{}_i \sum_k X^k{}_k\,. \tag{5.217}$$

Clearly $\kappa(X,X) = 0$ for $X \propto I$ reflecting that $\mathfrak{u}(n)$ contains an invariant abelian subalgebra. For $\mathfrak{su}(n)$, when $\sum_k X^k{}_k = 0$ and hence $\mathrm{tr}(X) = 0$, then $\kappa(X,X) = 2n\, \mathrm{tr}(X^2) < 0$.

For $\mathfrak{o}(n)$ or $\mathfrak{so}(n)$ then in (5.211) we must require $X^T = -X$ so that $\mathrm{tr}(X) = 0$. A basis for $n \times n$ antisymmetric matrices is given by the $\frac{1}{2}n(n-1)$ matrices $\{S_{ij} : i < j\}$ where

$$S_{ij} = -S_{ji} = \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix}\!\!\begin{pmatrix} 0 & \dots & \overset{i}{0} & \dots & \overset{j}{0} & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & & 0 & & 1 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & & -1 & & 0 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}, \qquad i \neq j = 1,\dots,n\,. \tag{5.218}$$

These satisfy

$$\big[S_{ij}, S_{kl}\big] = \delta_{jk}\, S_{il} - \delta_{ik}\, S_{jl} - \delta_{jl}\, S_{ik} + \delta_{il}\, S_{jk}\,, \tag{5.219}$$

and

$$\mathrm{tr}\big(S_{ij}\, S_{kl}\big) = 2\big(\delta_{il}\,\delta_{jk} - \delta_{ik}\,\delta_{jl}\big)\,. \tag{5.220}$$

For arbitrary $X \in \mathfrak{so}(2n)$ then $X = \frac{1}{2} X_{ij} S_{ij}$, where $X_{ij} = -X_{ji}$ is real. From (5.219)

$$[X, S_{ij}] = X_{ki}\, S_{kj} - X_{kj}\, S_{ki} \quad \Rightarrow \quad X^{\mathrm{ad}}_{kl,ij} = X_{ki}\,\delta_{lj} - X_{kj}\,\delta_{li} - X_{li}\,\delta_{kj} + X_{lj}\,\delta_{ki}\,, \tag{5.221}$$

and hence

$$\kappa(X,Y) = \tfrac{1}{4} X^{\mathrm{ad}}_{kl,ij}\, Y^{\mathrm{ad}}_{ij,kl} = -(n-2)\, X_{ij} Y_{ij}\,. \tag{5.222}$$

The matrices (5.218) are the generators for the vector representation of $SO(n)$ which is of course real, as described later there are also complex representations involving spinors.

For $\mathfrak{sp}(2n,\mathbb{R})$ or $\mathfrak{sp}(2n,\mathbb{C})$ the condition (1.53) translates into

$$XJ + JX^T = 0 \qquad \Rightarrow \qquad JX = (JX)^T\,, \tag{5.223}$$

where J is the standard antisymmetric matrix given in (1.54). It can be represented by

$$J_{ij} = -(-1)^i\,\delta_{ij'}\,, \qquad j' = j - (-1)^j\,. \tag{5.224}$$

A basis for $\mathfrak{sp}(2n,\mathbb{R})$, or $\mathfrak{sp}(2n,\mathbb{C})$, is provided by the $2n \times 2n$ matrices satisfying (5.223)

$$T_{ij} = J_{jk}\, T_{kl}\, J_{li} = \begin{array}{c} \\ \\ \\ i \\ \\ j' \\ \\ \\ \end{array}\!\!\begin{pmatrix} 0 & \dots & \overset{i'}{0} & \dots & \overset{j}{0} & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & & 0 & & 1 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & & -(-1)^{i+j} & & 0 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}, \qquad i,j = 1,\dots,2n\,. \tag{5.225}$$

An independent basis is given by $\{T_{ij},\ 1 \le i < j \le 2n; T_{2i-1\,2i-1},\ T_{2i\,2i-1},\ 1 \le i \le n\}$. The matrices $\{T_{ij}\}$ satisfy

$$\big[T_{ij}, T_{kl}\big] = \delta_{jk}\, T_{il} - \delta_{il}\, T_{kj} - J_{ik}\, J_{jm} T_{ml} - J_{jl}\, T_{im} J_{mk}\,, \tag{5.226}$$

and also

$$\mathrm{tr}\big(T_{ij}\, T_{kl}\big) = 2\big(\delta_{il}\,\delta_{jk} - J_{ik}\, J_{jl}\big)\,. \tag{5.227}$$

For any $X \in \mathfrak{sp}(2n,\mathbb{R})$, or $\mathfrak{sp}(2n,\mathbb{C})$, then $X = \frac{1}{2} X_{ij} T_{ij}$, with $X_{ij} = J_{jk} X_{kl} J_{li}$ real or complex. Using (5.226)

$$[X, T_{ij}] = X_{ki}\, T_{kj} - X_{jk}\, T_{ik}\,, \tag{5.228}$$

so that

$$X^{\mathrm{ad}}_{kl,ij} = X_{ki}\,\delta_{lj} - X_{jl}\,\delta_{ki} - X_{km} J_{mj}\, J_{li} - J_{im} X_{ml}\, J_{kj}\,, \tag{5.229}$$

and hence

$$\kappa(X,Y) = \tfrac{1}{4} X^{\rm ad}_{kl,ij}\, Y^{\rm ad}_{ij,kl} = 2(n+1)\, X_{ij} Y_{ji} \,. \tag{5.230}$$

For the corresponding compact group $Sp(n) = Sp(2n,\mathbb{C}) \cap SU(2n)$ we impose, as well as (5.223), for the corresponding Lie algebra

$$X^\dagger = -X \qquad \text{or} \qquad X_{ij} = -X_{ji}{}^* \,. \tag{5.231}$$

Then (5.230) gives

$$\kappa(X,X) = -2(n+1) \textstyle\sum_{i,j} |X_{ij}|^2 \,. \tag{5.232}$$

From (5.223)

$$JXJ^{-1} = -X^T \,, \tag{5.233}$$

so that, following the discussion in section 5.3.2, the fundamental representation of compact $Sp(n)$ is pseudo-real.

5.10 Orthogonal and Spin Groups

The relation $SO(3)$ and $SU(2)$, which is described in section 2.2, and also the introduction of spinorial representations, described in section 2.10, may be extended to higher orthogonal groups. In the discussion for $SO(3)$ and $SU(2)$ an essential role was played by the Pauli matrices. For $SO(n)$ we introduce similarly gamma matrices, γ_i, $i = 1, \dots, n$ satisfying the *Clifford*[18] *algebra*,

$$\gamma_i \gamma_j + \gamma_j \gamma_i = 2\delta_{ij}\, I \,, \qquad \gamma_i^\dagger = \gamma_i \,. \tag{5.234}$$

The algebra may be extended to pseudo-orthogonal groups such as the Lorentz group, which involve a metric g_{ij} as in (1.65), by taking $\delta_{ij} \to g_{ij}$ on the right hand side of (5.234). To obtain explicit gamma matrices for $SO(n,m)$ it is sufficient for each j with $g_{jj} = -1$ just to let $\gamma_j \to i\gamma_j$ for the corresponding $SO(n+m)$ gamma matrices. For the non compact group the gamma matrices are not all hermitian. (For g_{ij} as in (1.65) then if $A = \gamma_1 \dots \gamma_n$ then $A\gamma_i A^{-1} = -(-1)^n \gamma_i{}^\dagger$.)

The representations of the Clifford algebra (5.234), acting on a representation space $\mathcal{S}$, are irreducible if $\mathcal{S}$ has no invariant subspaces under the action of arbitrary products of γ_i's. As will become apparent there is essentially one irreducible representation for even n and two, related by a change of sign, for odd n. If $\{\gamma'_i\}$, like $\{\gamma_i\}$, are matrices forming an irreducible representation of (5.234) then $\gamma'_i = A\gamma_i A^{-1}$, or possibly $\gamma'_i = -A\gamma_i A^{-1}$ for n odd, for some A. As a consequence of (5.234)

$$(\gamma \cdot x)^2 = x^2\, I, \qquad x \in \mathbb{R}^n \,. \tag{5.235}$$

To show the connection with $SO(n)$ we first define

$$s_{ij} = \tfrac{1}{2}\, \gamma_{[i}\, \gamma_{j]} = -s_{ij}{}^\dagger \,. \tag{5.236}$$

[18] William Kingdon Clifford, 1845-1879, English, second wrangler 1867.

Using just (5.234) it is easy to obtain

$$[s_{ij}, \gamma_k] = \delta_{jk}\,\gamma_i - \delta_{ik}\,\gamma_j\,, \tag{5.237}$$

and hence

$$\big[s_{ij}, s_{kl}\big] = \delta_{jk}\,s_{il} - \delta_{ik}\,s_{jl} - \delta_{jl}\,s_{ik} + \delta_{il}\,s_{jk}\,. \tag{5.238}$$

This is identical with (5.219), the Lie algebra $\mathfrak{so}(n)$. Moreover for finite transformations, which involve the matrix exponential of $\frac{1}{2}\,\omega_{ij}s_{ij}$, $\omega_{ij} = -\omega_{ji}$,

$$e^{-\frac{1}{2}\,\omega_{ij}s_{ij}}\,\gamma\cdot x\;e^{\frac{1}{2}\,\omega_{ij}s_{ij}} = \gamma\cdot x'\,, \qquad x' = Rx\,, \quad R = e^{-\frac{1}{2}\,\omega_{ij}S_{ij}} \in SO(n)\,, \tag{5.239}$$

with $S_{ij} \in \mathfrak{so}(n)$ as in (5.218). It is easy to see that $x'^2 = x^2$, as required for rotations, as a consequence of (5.235). To show the converse we note that $\gamma'_i = \gamma_j R_{ji}$ also satisfies (5.234) for $[R_{ji}] \in O(n)$ so that $\gamma'_i = A(R)\gamma_i A(R)^{-1}$ where $A(R) = e^{-\frac{1}{2}\,\omega_{ij}s_{ij}}$ for R continuously connected to the identity.

The exponentials of the spin matrices form the group

$$\mathrm{Spin}(n) = \big\{e^{-\frac{1}{2}\,\omega_{ij}s_{ij}} : \omega_{ij} = -\omega_{ji} \in \mathbb{R}\big\}\,. \tag{5.240}$$

Clearly $\mathrm{Spin}(n)$ and $SO(n)$ have the same Lie algebra. For $n = 3$ we may let $\gamma_i \to \sigma_i$ and $s_{ij} = \frac{1}{2} i\varepsilon_{ijk}\sigma_k$ so that $\mathrm{Spin}(3) \simeq SU(2)$. In general, since $\pm I \in \mathrm{Spin}(n)$ are mapped to $I \in SO(n)$, we have $SO(n) \simeq \mathrm{Spin}(n)/\mathbb{Z}_2$.

Unlike $SO(n)$, $\mathrm{Spin}(n)$ is simply connected and is the covering group for $SO(n)$. For further analysis we define

$$\Gamma = \gamma_1\gamma_2\dots\gamma_n = (-1)^{\frac{1}{2}n(n-1)}\,\Gamma^\dagger\,, \qquad \Gamma^\dagger = \gamma_n\gamma_{n-1}\dots\gamma_1\,, \tag{5.241}$$

so that

$$\Gamma^2 = (-1)^{\frac{1}{2}n(n-1)}\,I\,. \tag{5.242}$$

Directly from (5.234)

$$[\Gamma, \gamma_i] = 0\,, \quad n \text{ odd}\,, \quad \Gamma\,\gamma_i + \gamma_i\,\Gamma = 0\,, \quad n \text{ even}\,, \qquad i = 1, \dots n\,. \tag{5.243}$$

Using, similarly to (2.28),

$$e^{\alpha\, s_{ij}} = \cos\tfrac{1}{2}\alpha\, I + \sin\tfrac{1}{2}\alpha\, 2s_{ij}\,, \tag{5.244}$$

then

$$e^{\pi \sum_{i=1}^m s_{2i-1\,2i}} = \Gamma\,, \qquad e^{-\pi \sum_{i=1}^m s_{2i-1\,2i}} = (-1)^m\Gamma\,, \qquad \text{for } \; n = 2m \; \text{ even}\,. \tag{5.245}$$

This allows the identification of the centres of the spin groups

$$\begin{aligned} \mathcal{Z}(\mathrm{Spin}(n)) &= \{I, -I, \Gamma, -\Gamma\} \simeq \begin{cases} \mathbb{Z}_2 \otimes \mathbb{Z}_2\,, & n = 4m\,, \\ \mathbb{Z}_4\,, & n = 4m+2\,, \end{cases} \\ \mathcal{Z}(\mathrm{Spin}(n)) &= \{I, -I\} \simeq \mathbb{Z}_2\,, \qquad\qquad\qquad n = 2m+1\,. \end{aligned} \tag{5.246}$$

Spinors for general rotational groups are defined as belonging to the fundamental representation space $\mathcal{S}$ for $\mathrm{Spin}(n)$, so they form projective representations, up to a sign, of $SO(n)$.

5.10.1 Products and Traces of Gamma Matrices

For products of gamma matrices if the same gamma matrix γ_i appears twice in the product then, since it anti-commutes with all other gamma matrices, as a consequence of (5.234), and also $\gamma_i^2 = I$, it may be removed from the product, leaving the remaining matrices unchanged apart from a possible change of sign. Linearly independent matrices are obtained by considering products of different gamma matrices. Accordingly we define, for $i_, \dots i_r$ all different indices,

$$\Gamma_{i_1\dots i_r} = \gamma_{[i_1} \dots \gamma_{i_r]} = (-1)^{\frac{1}{2}r(r-1)}\, {\Gamma_{i_1\dots i_r}}^\dagger \,, \quad {\Gamma_{i_1\dots i_r}}^\dagger = \Gamma_{i_r\dots i_1} \,, \qquad r = 1, \dots n \,, \tag{5.247}$$

where ${\Gamma_{i_1\dots i_r}}^2 = (-1)^{\frac{1}{2}r(r-1)}\, I$. From the definition (5.241)

$$\Gamma_{i_1\dots i_n} = \varepsilon_{i_1\dots i_n}\, \Gamma \,. \tag{5.248}$$

We also have the relations

$$\Gamma_{i_1\dots i_r} = (-1)^{\frac{1}{2}n(n-1)+\frac{1}{2}r(r-1)}\, \tfrac{1}{s!}\, \varepsilon_{i_1\dots i_r\, j_1\dots j_s} \Gamma_{j_1\dots j_s} \Gamma \,, \qquad r+s=n \,. \tag{5.249}$$

An independent basis for these products is given by $\mathcal{C}_r = \{\Gamma_{i_1\dots i_r} : i_1 < i_2 < \dots < i_n\}$, with $\dim \mathcal{C}_r = \binom{n}{r}$, $\mathcal{C}_n = \{\Gamma\}$. It is easy to see that $\mathcal{C}^{(n)} = \{\pm I, , \pm \mathcal{C}_1, \dots, \pm \mathcal{C}_{n-1}, \pm\Gamma\}$ is closed under multiplication and therefore forms a finite matrix group, with $\dim \mathcal{C}^{(n)} = 2\sum_{r=0}^n \binom{n}{r} = 2^{n+1}$. The matrices $\{I, \mathcal{C}_1, \dots, \mathcal{C}_n\}$ may also be regarded as the basis vectors for a 2^n-dimensional vector space which is also a group under multiplication, and so this forms a field.

When n is odd then from (5.243) Γ commutes with all elements in $\mathcal{C}^{(n)}$ and so for an irreducible representation we must have $\Gamma \propto I$. Taking into account (5.242)

$$\Gamma = \begin{cases} \pm I \,, & n = 4m+1 \\ \pm i\, I \,, & n = 4m+3 \end{cases} . \tag{5.250}$$

The $\pm$ signs correspond to inequivalent representations, linked by taking $\gamma_i \to -\gamma_i$. For an independent basis then, as a consequence of (5.249), the products of gamma matrices are no longer independent if $r > \frac{1}{2}n$, so that $\mathcal{C}^{(4m+1)} = \{\pm I, \pm\mathcal{C}_1, \dots, \pm\mathcal{C}_{2m}\}$ or, since the products may involve i from (5.250), $\mathcal{C}^{(4m+3)} = \{\pm I, \pm iI, \pm\mathcal{C}_1, \pm i\mathcal{C}_1, \dots, \pm i\mathcal{C}_{2m}\}$.

For n even $\mathcal{C}^{(n)}$ does not contain any elements commuting with all γ_i but

$$[\Gamma, s_{ij}] = 0 \,. \tag{5.251}$$

Hence we may decompose the representation space $\mathcal{S} = \mathcal{S}_+ \oplus \mathcal{S}_-$, such that $\Gamma\mathcal{S}_\pm = \mathcal{S}_\pm$ and, since γ_i anti-commutes with Γ, $\gamma_i\, \mathcal{S}_\pm = \mathcal{S}_\mp$. Hence there is a corresponding decomposition of the gamma matrices with Γ diagonal and where, using (5.242),

$$\Gamma = \begin{cases} \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right), & n = 4m \,, \\ i\left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right), & n = 4m+2 \,, \end{cases} \qquad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ \bar\sigma_i & 0 \end{pmatrix} , \qquad s_{ij} = \begin{pmatrix} s_{+ij} & 0 \\ 0 & s_{-ij} \end{pmatrix} . \tag{5.252}$$

Clearly $\bar\sigma_i = \sigma_i{}^\dagger$ and $s_{+ij} = \frac{1}{2}\,\sigma_{[i}\,\bar\sigma_{j]}$, $s_{-ij} = \frac{1}{2}\,\bar\sigma_{[i}\,\sigma_{j]}$ and just as in (5.236) we have

$$s_{\pm ij}{}^\dagger = -s_{\pm ij}\,. \tag{5.253}$$

With the decomposition in (5.252) the Clifford algebra (5.234) is equivalent to

$$\sigma_i\,\bar\sigma_j + \sigma_j\,\bar\sigma_i = 2\delta_{ij}\,I\,, \qquad \bar\sigma_i\,\sigma_j + \bar\sigma_j\,\sigma_i = 2\delta_{ij}\,I\,. \tag{5.254}$$

For traces of gamma matrices and their products we first note that from (5.234)

$$\mathrm{tr}\big(\gamma_j(\gamma_i\gamma_j + \gamma_j\gamma_i)\big) = 2\,\mathrm{tr}(\gamma_j\gamma_j\gamma_i) = 2\,\mathrm{tr}(\gamma_i) = 0\,, \quad j \neq i\,, \quad \text{no sum on } j\,. \tag{5.255}$$

We may similarly use $\gamma_j\,\Gamma_{i_1\dots i_r} + \Gamma_{i_1\dots i_r}\gamma_j = 0$, when r is odd and for $j \neq i_1,\dots,i_r$, or $\gamma_j\,\Gamma_{j\,i_2\dots i_r} + \Gamma_{j\,i_2\dots i_r}\gamma_j = 0$, when r is even and with no sum on j, to show that

$$\mathrm{tr}(\Gamma_{i_1\dots i_r}) = 0\,, \qquad \text{except when } r = n,\ n \text{ odd}\,. \tag{5.256}$$

Hence in general, for $r,s = 0,\dots,n$ for n even, or with $r,s < \frac{1}{2}n$ for n odd,

$$\mathrm{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{j_1\dots j_s}\big) = 0\,, \qquad r \neq s\,, \tag{5.257}$$

and

$$\mathrm{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{j_1\dots j_r}\big) = \begin{cases} \pm\mathrm{tr}(I) & \text{if } (j_1,\dots,j_r) \text{ is an even/odd permutation of } (i_r,\dots,i_1)\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{5.258}$$

In general these products of gamma matrices form a complete set so that for any $d_n \times d_n$ matrix A, where $d_n = 2^{\frac{1}{2}n}$ for n even and $d_n = 2^{\frac{1}{2}(n-1)}$ for n odd,

$$d_n\,A = I\,\mathrm{tr}(A) + \sum_r \tfrac{1}{r!}\,\Gamma_{i_1\dots i_r}\,\mathrm{tr}\big(\Gamma_{i_r\dots i_1}\,A\big)\,, \tag{5.259}$$

with $r = 1,\dots,n$ for n even, $r = 1,\dots,\frac{1}{2}(n-1)$ for n odd.

5.10.2 Construction of Representations of the Clifford Algebra

For $n = 2m$ an easy way to construct the γ-matrices satisfying the Clifford algebra (5.234) explicitly is to define

$$a_r = \tfrac{1}{2}(\gamma_{2r-1} + i\gamma_{2r})\,, \qquad a_r{}^\dagger = \tfrac{1}{2}(\gamma_{2r-1} - i\gamma_{2r})\,, \qquad r = 1,\dots,m\,. \tag{5.260}$$

Then (5.234) becomes

$$a_r\,a_s + a_s\,a_r = 0\,, \qquad a_r\,a_s{}^\dagger + a_s\,a_r{}^\dagger = \delta_{rs}\,I\,, \tag{5.261}$$

which is just the algebra for m fermionic creation and annihilation operators, the femionic analogue of the usual bosonic harmonic oscillator operators. The construction of the essentially unique representation space $\mathcal{S}$ for such operators is standard, there is a vacuum state annihilated by all the a_r's and all other states in the space are obtained by acting on the vacuum state with linear combinations of products of ${a_r}^\dagger$'s. In general, since ${a_r}^{\dagger\,2} = 0$, a basis is formed by restricting to products of the form $\prod_{r=1}^m ({a_r}^\dagger)^{s_r}$ with $s_r = 0,1$ for each r. There are then 2^m independent basis vectors, giving $\dim S = 2^m$. For $m = 1$ then we may take, with the 'vacuum state' represented by $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$,

$$a = \sigma_+ = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right), \qquad a^\dagger = \sigma_- = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right), \qquad \gamma_1\gamma_2 = i\,\sigma_3 = i\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right). \tag{5.262}$$

The general case is obtained using tensor products

$$a_r = \underbrace{I\otimes\cdots\otimes I}_{r-1}\otimes\sigma_+\otimes\underbrace{\sigma_3\otimes\cdots\otimes\sigma_3}_{m-r}, \qquad {a_r}^\dagger = \underbrace{I\otimes\cdots\otimes I}_{r-1}\otimes\sigma_-\otimes\underbrace{\sigma_3\otimes\cdots\otimes\sigma_3}_{m-r}. \tag{5.263}$$

The σ_3's appearing in the tensor products follow from the requirement that a_r, a_s, and ${a_r}^\dagger, {a_s}^\dagger$, anti-commute for $r \neq s$. With (5.263) $\gamma_{2r-1}\gamma_{2r} = i\,I\otimes\cdots\otimes I\otimes\sigma_3\otimes I\cdots\otimes I$ so that

$$\Gamma = i^m\,\underbrace{\sigma_3\otimes\sigma_3\otimes\cdots\otimes\sigma_3}_{m}. \tag{5.264}$$

These results are equivalent to defining the gamma matrices for increasing n, where ${\gamma_i}^{(n)}{\gamma_j}^{(n)} + {\gamma_j}^{(n)}{\gamma_i}^{(n)} = 2\delta_{ij}\,I^{(n)}$, recursively in terms of the Pauli matrices by

$$\begin{aligned} &{\gamma_i}^{(2m+2)} = {\gamma_i}^{(2m)}\otimes\sigma_3\,, \qquad i = 1,\dots,2m\,,\\ &\gamma_{2m+1}^{(2m+2)} = I^{(2m)}\otimes\sigma_1\,, \qquad \gamma_{2m+2}^{(2m+2)} = I^{(2m)}\otimes\sigma_2\,,\\ &\Gamma^{(2m+2)} = i\,\Gamma^{(2m)}\otimes\sigma_3\,. \end{aligned} \tag{5.265}$$

Note that we may take ${\gamma_i}^{(2)} = \sigma_i$, $i = 1,2$ with $\Gamma^{(2)} = i\,\sigma_3$. For odd n the gamma matrices may be defined in terms of those for $n-1$ by

$${\gamma_i}^{(2m+1)} = {\gamma_i}^{(2m)}, \;\; i = 1,\dots,2m\,, \quad \gamma_{2m+1}^{(2m+1)} = c_m\,\Gamma^{(2m)}, \;\; c_m = \begin{cases}\pm 1\,, & m \text{ even}\,,\\ \pm i\,, & m \text{ odd}\,,\end{cases} \tag{5.266}$$

where the $\pm$ signs correspond to inequivalent representations. Thus ${\gamma_i}^{(3)} = (\sigma_1, \sigma_2, \mp\sigma_3)$.

5.10.3 Conjugation Matrix for Gamma Matrices

It is easy to see that γ_i^T also obeys the Clifford algebra in (5.234) so that for an irreducible representation we must have

$$\begin{aligned} &C\gamma_i\,C^{-1} = -\,\gamma_i^T \qquad \Rightarrow \qquad C\,\Gamma\,C^{-1} = (-1)^{\frac12 n(n+1)}\,\Gamma^T\\ \text{or}\quad &C\gamma_i\,C^{-1} = \gamma_i^T \qquad \Rightarrow \qquad C\,\Gamma C^{-1} = (-1)^{\frac12 n(n-1)}\,\Gamma^T\,. \end{aligned} \tag{5.267}$$

When n is even then, by taking $C \to C\Gamma$, the two cases are equivalent. When n is odd, and we require (5.250), then for $n = 4m+1$, C must satisfy $C\gamma_i\,C^{-1} = \gamma_i^T$, for $n = 4m+3$, then $C\gamma_i\,C^{-1} = -\gamma_i^T$. In either case for the spin matrices in (5.236)

$$C\,s_{ij}\,C^{-1} = -{s_{ij}}^T\,, \tag{5.268}$$

so that for the matrices defining Spin(n)

$$e^{-\frac{1}{2}\,\omega_{ij}s_{ij}}\, C\, \big(e^{-\frac{1}{2}\,\omega_{ij}s_{ij}}\big)^T = C\,. \qquad (5.269)$$

With the recursive construction of the gamma matrices $\gamma_i^{(n)}$ in (5.265) we may also construct in a similar fashion $C^{(n)}$ iteratively since, using (5.75),

$$\begin{aligned} & C^{(n)}\,\gamma_i^{(n)}\,C^{(n)\,-1} = \gamma_i^{(n)\,T} \\ & \quad\Rightarrow \quad C^{(n+2)} = C^{(n)} \otimes\, i\sigma_2 \quad \text{ensures} \quad C^{(n+2)}\,\gamma_i^{(n+2)}\,C^{(n+2)\,-1} = -\gamma_i^{(n+2)\,T}\,, \end{aligned} \qquad (5.270)$$

and, using $\sigma_1\sigma_i\sigma_1 = \sigma_i{}^T$, $i=1,2$, $\sigma_1\sigma_3\sigma_1 = -\sigma_3{}^T$,

$$\begin{aligned} & C^{(n)}\,\gamma_i^{(n)}\,C^{(n)\,-1} = -\gamma_i^{(n)\,T} \\ & \quad\Rightarrow \quad C^{(n+2)} = C^{(n)} \otimes\, \sigma_1 \quad \text{ensures} \quad C^{(n+2)}\,\gamma_i^{(n+2)}\,C^{(n+2)\,-1} = \gamma_i^{(n+2)\,T}\,. \end{aligned} \qquad (5.271)$$

Starting from $n=0$, or $n=2$, this construction gives (note that $(X\otimes Y)^T = X^T \otimes Y^T$),

$$\begin{array}{llll} C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \Gamma^T\,, & C = C^T\,, & n=8k\,, \\ C\,\gamma_i\,C^{-1} = \,-\,\gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \,-\,\Gamma^T\,, & C = \,-\,C^T\,, & n=8k+2\,, \\ C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \Gamma^T\,, & C = \,-\,C^T\,, & n=8k+4\,, \\ C\,\gamma_i\,C^{-1} = \,-\,\gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \,-\,\Gamma^T\,, & C = C^T\,, & n=8k+6\,. \end{array} \qquad (5.272)$$

In each case we have $Cs_{ij}C^{-1} = -s_{ij}{}^T$. Starting from (5.272) and with the construction in (5.266) for odd n,

$$\begin{array}{lll} C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C = C^T\,, & n=8k+1\,, \\ C\,\gamma_i\,C^{-1} = \,-\,\gamma_i{}^T\,, & C = \,-\,C^T\,, & n=8k+3\,, \\ C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C = \,-\,C^T\,, & n=8k+5\,, \\ C\,\gamma_i\,C^{-1} = \,-\,\gamma_i{}^T\,, & C = C^T\,, & n=8k+7\,. \end{array} \qquad (5.273)$$

The definition of C for $n=2m+1$ remains the same as in (5.272) for $n=2m$ since in each n odd case we have $C\gamma_1\gamma_2\ldots\gamma_n C^{-1} = (\gamma_1\gamma_2\ldots\gamma_n)^T$.

If we consider a basis in which Γ is diagonal, as in (5.252), then for $n=8k, 8k+4$ $[C,\Gamma]=0$, so that C is block diagonal, while for $n=8k+2, 8k+6$ $C\Gamma+\Gamma C=0$, so that we may take C to have a block off diagonal form. By considering the freedom under $C\to S^TCS$ with $S\Gamma S^{-1}=\Gamma$ we may choose with the basis in (5.252),

$$\begin{array}{llll} C = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\,, & \bar\sigma_i = \sigma_i^T, & s_{\pm ij} = \,-\,s_{\pm ij}{}^T\,, & n=8k\,, \\ C = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\,, & \sigma_i = \sigma_i^T,\ \bar\sigma_i = \bar\sigma_i^T\,, & s_{\pm ij} = \,-\,s_{\mp ij}{}^T, & n=8k+2\,, \\ C = \begin{pmatrix} J & 0 \\ 0 & J \end{pmatrix}\,, & J = \,-\,J^T,\ J\bar\sigma_i = -(J\sigma_i)^T\,, & Js_{\pm ij} = (Js_{\pm ij})^T, & n=8k+4\,, \\ C = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\,, & \sigma_i = \,-\,\sigma_i^T,\ \bar\sigma_i = -\bar\sigma_i^T\,, & s_{\pm ij} = \,-\,s_{\mp ij}{}^T, & n=8k+6\,. \end{array} \qquad (5.274)$$

Here the antisymmetric matrix J can be taken to be of the standard form as in (1.54). For $n = 8k$ the matrices are real.

Since the generators of the two fundamental spinor representations satisfy (5.253) then as a consequence of the discussion in section 5.3.2 we have for these representations of $\mathrm{Spin}(n)$, for n even, from (5.274)

$$\begin{aligned} &\mathrm{Spin}(8k): \ \text{real}\,, \qquad \mathrm{Spin}(8k+4): \ \text{pseudo-real}\,, \\ &\mathrm{Spin}(8k+2)\,,\ \mathrm{Spin}(8k+6): \ \text{complex}\,. \end{aligned} \tag{5.275}$$

Furthermore for n odd the single spinor representation, from (5.273), satisfies

$$\mathrm{Spin}(8k+1)\,, \mathrm{Spin}(8k+7): \ \text{real}\,, \qquad \mathrm{Spin}(8k+3)\,, \mathrm{Spin}(8k+5): \ \text{pseudo-real}\,. \tag{5.276}$$

5.10.4 Special Cases

When $n = 2$ we may take

$$\sigma_i = (1, -i)\,, \qquad \bar\sigma_i = (1, i)\,, \tag{5.277}$$

while for $n = 4$ we may express $\sigma_i, \bar\sigma_i$ in terms of unit quaternions

$$\sigma_i = (1, -i, -j, -k)\,, \qquad \bar\sigma_i = (1, i, j, k)\,. \tag{5.278}$$

For low n results for γ-matrices may be used to identify $\mathrm{Spin}(n)$ with other groups. Thus

$$\begin{aligned} &\mathrm{Spin}(3) \simeq SU(2)\,, \qquad &&\mathrm{Spin}(4) \simeq SU(2) \otimes SU(2)\,, \\ &\mathrm{Spin}(5) \simeq Sp(2)\,, \qquad &&\mathrm{Spin}(6) \simeq SU(4)\,. \end{aligned} \tag{5.279}$$

For $n = 3$ it is evident directly that $e^{-\frac{1}{2}\omega_{ij} s_{ij}} \in SU(2)$. For $n = 4$ as a consequence of (5.249), with the decomposition in (5.252), we have

$$s_{\pm ij} = \pm\tfrac{1}{2}\, \varepsilon_{ijkl}\, s_{\pm kl}\,, \tag{5.280}$$

so that $e^{-\frac{1}{2}\omega_{ij} s_{ij}} = e^{-\frac{1}{2}\omega_{+ij} s_{+ij}} \otimes e^{-\frac{1}{2}\omega_{-ij} s_{-ij}}$ factorises a 4×4 $\mathrm{Spin}(4)$ matrix into a product of two independent $SU(2)$ matrices as $\omega_{\pm ij} = \frac{1}{2}\,\omega_{ij} \pm \frac{1}{4}\,\varepsilon_{ijkl}\,\omega_{kl}$ are independent. For $n = 5$ then the 4×4 matrix $e^{-\frac{1}{2}\omega_{ij} s_{ij}} \in SU(4) \cap Sp(4, \mathbb{C})$, using (5.269) with $C^T = -C$. In this case there are 10 independent s_{ij} which matches with the dimension of the compact $Sp(2)$. For $n = 6$, $e^{-\frac{1}{2}\omega_{+ij} s_{+ij}} \in SU(4)$ with the 15 independent 4×4 matrices s_{+ij} matching the dimension of $SU(4)$. Note also that, from (5.246), $\mathcal{Z}(\mathrm{Spin}(6)) \simeq \mathbb{Z}_4 \simeq \mathcal{Z}(SU(4))$. Using (5.274) with (5.252), the transformation (5.239) can be rewritten just in terms of the $SU(4)$ matrix

$$e^{-\frac{1}{2}\omega_{ij} s_{+ij}}\, \sigma \cdot x\, \big(e^{-\frac{1}{2}\omega_{ij} s_{+ij}}\big)^T = \sigma \cdot x'\,, \tag{5.281}$$

which is analogous to (2.19). The result that the transformation $x \to x'$ satisfies $x^2 = x'^2$ also follows in a similar fashion to (2.21), but in this case using the Pfaffian (1.55) instead of the determinant since we require $\mathrm{Pf}(\sigma \cdot x) = x^2$ (from $\sigma \cdot x\, \bar\sigma \cdot x = x^2\, I$ then, with $n = 6$, $\det(\sigma \cdot x) = (x^2)^2$).

6 $SU(3)$ and its Representations

$SU(3)$ is an obvious generalisation of $SU(2)$ although that was not the perception in the 1950's when many physicists were searching for a higher symmetry group, beyond $SU(2)$ and isospin, to accommodate and classify the increasing numbers of resonances found in particle accelerators with beams of a few GeV. Although the discovery of the relevance of $SU(3)$ as a hadronic symmetry group was a fundamental breakthrough, leading to the realisation that quarks are fundamental constituents, it now appears that $SU(3)$ symmetry is just an almost accidental consequence of the fact that the three lightest quarks have a mass which is significantly less than the typical hadronic mass scale.

Understanding $SU(2)$ and its representations is an essential first step before discussing general simple Lie groups. Extending to $SU(3)$ introduces many of the techniques which are needed for the general case in a situation where the algebra is still basically simple and undue mathematical sophistication is not required. For general $SU(N)$ the Lie algebra is given, for the associated chosen basis, by (5.213) where, since the corresponding matrices in (5.212) are not anti-hermitian, we are regarding the Lie algebra as a complex vector space. To set the scene for $SU(3)$ we reconsider first $SU(2)$.

6.1 Recap of $\mathfrak{su}(2)$

For the basic generators of $\mathfrak{su}(2)$ we define in terms of 2×2 matrices as in (5.212)

$$e_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \qquad e_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} , \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{6.1}$$

which satisfy the Lie algebra

$$[e_+, e_-] = h \, , \qquad [h, e_\pm] = \pm 2 \, e_\pm \, . \tag{6.2}$$

These matrices satisfy

$${e_+}^\dagger = e_- \, , \qquad h^\dagger = h \, . \tag{6.3}$$

Under interchange of the rows and columns

$$b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \Rightarrow \qquad b\big\{e_+, e_-, h\big\}b^{-1} = \big\{e_-, e_+, -h\big\} \, . \tag{6.4}$$

Clearly $b^2 = I$ and $\big\{e_+, e_-, h\big\}$, $\big\{e_-, e_+, -h\big\}$ must satisfy the same commutation relations as in (6.2) so b generates an automorphism.

For representations of the $\mathfrak{su}(2)$ Lie algebra then we require operators

$$\mathfrak{l} = \{E_+, E_-, H\} \, , \qquad [E_+, E_-] = H \, , \quad [H, E_\pm] = \pm 2 \, E_\pm \, . \tag{6.5}$$

It is easy to see that the commutation relations are identical with (2.41a) and (2.41a), and also the hermeticity conditions with (2.42), by taking $J_\pm \to E_\pm$, $2J_3 \to H$. Indeed the representation matrices in (6.1) then correspond exactly with (2.77).

An important role in the general theory of Lie groups is played by the automorphism symmetries of a privileged basis for the Lie algebra which define the *Weyl*[19] *group.* For $\mathfrak{su}(2)$ the relevant basis is given by (6.5) and then from (6.4) there is just one non trivial automorphism

$$\mathfrak{l} \underset{b}{\rightarrow} \mathfrak{l}_R = \{E_-, E_+, -H\}\,. \tag{6.6}$$

Since $b^2 = I$ the Weyl group for $\mathfrak{su}(2)$, $W(\mathfrak{su}(2)) \simeq \mathbb{Z}_2$.

For representations we require a finite dimensional representation space on which there are operators $E_\pm, H$ which obey the commutation relations (6.5) and subsequently require there is a scalar product so that the operators satisfy the hermeticity conditions in (6.3). A basis for a representation space for $\mathfrak{su}(2)$ is given by $\{|r\rangle\}$ where

$$H|r\rangle = r|r\rangle\,. \tag{6.7}$$

The eigenvalue r is termed the *weight.* It is easy to see from (6.5) that

$$E_\pm|r\rangle \propto |r \pm 2\rangle \qquad \text{unless} \quad E_+|r\rangle = 0 \quad \text{or} \quad E_-|r\rangle = 0\,. \tag{6.8}$$

We consider representations where there is a *highest weight*, $r_{\max} = n$, and hence a highest weight vector $|n\rangle_{\rm hw}$ satisfying

$$E_+|n\rangle_{\rm hw} = 0\,. \tag{6.9}$$

The representation space V_n is then spanned by

$$\big\{ {E_-}^r|n\rangle_{\rm hw} : r = 0, 1, \dots \big\}\,. \tag{6.10}$$

On this basis

$$H {E_-}^r|n\rangle_{\rm hw} = (n - 2r)\, {E_-}^r|n\rangle_{\rm hw}\,, \tag{6.11}$$

and using

$$\big[E_+, {E_-}^r\big] = \sum_{s=0}^{r-1} {E_-}^{r-s-1}[E_+, E_-]{E_-}^s = {E_-}^{r-1} \sum_{s=0}^{r-1} \big(H - 2s\big) = {E_-}^{r-1}\, r(H - r + 1)\,, \tag{6.12}$$

then from (6.9),

$$E_+\, {E_-}^r|n\rangle_{\rm hw} = r(n - r + 1)\, {E_-}^{r-1}|n\rangle_{\rm hw}\,. \tag{6.13}$$

(6.11) and (6.13) ensure that the commutation relations (6.5) are realised on V_n.

If $n \in \mathbb{N}_0$, or $n = 0, 1, 2, \dots$, then from (6.13)

$$|{-n-2}\rangle_{\rm hw} = {E_-}^{n+1}|n\rangle_{\rm hw} \in V_n\,, \tag{6.14}$$

is also a highest weight vector, satisfying (6.9). From $|{-n-2}\rangle_{\rm hw}$ we may construct, just as in (6.10), a basis for an associated invariant subspace

$$V_{-n-2} \subset V_n\,. \tag{6.15}$$

[19] Hermann Klaus Hugo Weyl, 1885-1955, German.

Hence the representation space defined by the basis V_n is therefore reducible under the action of $\mathfrak{su}(2)$. An irreducible representation is obtained by restricting to the finite dimensional quotient space

$$\mathcal{V}_n = V_n / V_{-n-2} \,. \tag{6.16}$$

In general for a vector space V with a subspace U the quotient V/U is defined by

$$V/U = \big\{ \, |v\rangle/\!\sim \, : |v\rangle \sim |v'\rangle \text{ if } |v\rangle - |v'\rangle \in U \big\} \,. \tag{6.17}$$

It is easy to verify that V/U is a vector space and, if V, U are finite-dimensional, $\dim(V/U) = \dim V - \dim U$. If X is a linear operator acting on V then

$$U \underset{X}{\rightarrow} U \quad \Rightarrow \quad \{X|v\rangle/\sim\} = \{X|v'\rangle/\sim\} \text{ if } |v\rangle \sim |v'\rangle \quad \Rightarrow \quad X : V/U \rightarrow V/U \,. \tag{6.18}$$

Thus, if $U \subset V$ is an invariant subspace under X, then X has a well defined action on V/U. Furthermore for traces

$$\mathrm{tr}_{V/U}(X) = \mathrm{tr}_V(X) - \mathrm{tr}_U(X) \,. \tag{6.19}$$

Since V_{-n-2} is an invariant subspace under the action of the $\mathfrak{su}(2)$ Lie algebra generators we may then define $E_\pm, H$ to act linearly on the quotient $\mathcal{V}_n$ given by (6.16). On $\mathcal{V}_n$ this ensures

$$E_-{}^{n+1} |n\rangle_{\rm hw} = 0 \,, \tag{6.20}$$

so that there is a finite basis $\{E_-{}^r |n\rangle_{\rm hw} : r = 0, \dots, n\}$. In terms of the angular momentum representations constructed in section 2, $n = 2j$. The space $\mathcal{V}_n$ may equally be constructed from a lowest weight state $|-n\rangle$ satisfying $H|-n\rangle = -n\,|-n\rangle$, $E_-|-n\rangle = 0$, in accord with the automorphism symmetry (6.4) of the $\mathfrak{su}(2)$ Lie algebra.

If we define a formal trace over all vectors belonging to V_n then

$$C_n(t) = \widetilde{\mathrm{tr}}_{V_n}\big(t^H\big) = \sum_{r=0}^{\infty} t^{n-2r} = \frac{t^{n+2}}{t^2-1} \,, \tag{6.21}$$

where convergence of the sum requires $|t| > 1$. Then for the irreducible representation defined on the quotient $\mathcal{V}_n$, by virtue of (6.19), the character is

$$\chi_n(t) = \mathrm{tr}_{\mathcal{V}_n}\big(t^H\big) = C_n(t) - C_{-n-2}(t) = \frac{t^{n+2} - t^{-n}}{t^2 - 1} = \frac{t^{n+1} - t^{-n-1}}{t - t^{-1}} \,. \tag{6.22}$$

This is just the same as (2.81) with $t \rightarrow e^{i\frac{1}{2}\theta}$ and $n \rightarrow 2j$. It is easy to see that

$$\chi_n(1) = \dim \mathcal{V}_n = n + 1 \,. \tag{6.23}$$

Although the irreducible representation of $\mathfrak{su}(2)$ are labelled by $n \in \mathbb{N}_0$ the characters may be extended to any integer n with the property

$$\chi_n(t) = -\chi_{-n-2}(t) \,, \tag{6.24}$$

as follows directly from (6.22). Clearly $\chi_{-1}(t) = 0$.

The $\mathfrak{su}(2)$ Casimir operator in this basis

$$C = E_+E_- + E_-E_+ + \tfrac{1}{2}H^2 = 2E_-E_+ + \tfrac{1}{2}H^2 + H\,, \tag{6.25}$$

and it is easy to see that

$$C|n\rangle_{\rm hw} = c_n|n\rangle_{\rm hw} \qquad \text{for} \qquad c_n = \tfrac{1}{2}n(n+2)\,. \tag{6.26}$$

Note that $c_{-n-2} = c_n$ as required from (6.14) as all vectors belonging to V_n must have the same eigenvalue for C.

6.2 A $\mathfrak{su}(3)$ Lie algebra basis and its automorphisms

We consider a basis for the $\mathfrak{su}(3)$ Lie algebra in terms of 3×3 matrices as in (5.212). Thus we define

$$e_{1+} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad e_{2+} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad e_{3+} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6.27}$$

and their conjugates

$$e_{i-} = {e_{i+}}^\dagger\,, \quad i = 1,2,3\,, \tag{6.28}$$

together with the hermitian traceless diagonal matrices

$$h_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad h_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{6.29}$$

The commutator algebra satisfied by $\{e_{1\pm}, e_{2\pm}, e_{3\pm}, h_1, h_2\}$ is invariant under simultaneous permutations of the rows and columns of each matrix. For b corresponding to the permutation $(1\,2)$ and a to the cyclic permutation $(1\,2\,3)$

$$b = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad a = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{6.30}$$

then

$$\begin{aligned} b\{h_1, h_2\}b^{-1} &= \{-h_1, h_1 + h_2\}\,, \qquad & b\{e_{1\pm}, e_{2\pm}, e_{3\pm}\}b^{-1} &= \{e_{1\mp}, e_{3\pm}, e_{2\pm}\}\,, \\ a\{h_1, h_2\}a^{-1} &= \{h_2, -h_1 - h_2\}\,, \qquad & a\{e_{1\pm}, e_{2\pm}, e_{3\pm}\}a^{-1} &= \{e_{2\pm}, e_{3\mp}, e_{1\mp}\}\,. \end{aligned} \tag{6.31}$$

The matrices in (6.30) satisfy

$$b^2 = I\,, \qquad a^3 = I\,, \qquad ab = ba^2\,, \tag{6.32}$$

so that they generate the permutation group $S_3 = \{e, a, a^2, b, ab, a^2b\}$.

For representations of $\mathfrak{su}(3)$ it is then sufficient to require operators

$$\{E_{1\pm}, E_{2\pm}, E_{3\pm}, H_1, H_2\} \;\to\; \big[\hat{R}^i{}_j\big] = \begin{pmatrix} \frac{1}{3}(2H_1+H_2) & E_{1+} & E_{3+} \\ E_{1-} & \frac{1}{3}(-H_1+H_2) & E_{2+} \\ E_{3-} & E_{2-} & -\frac{1}{3}(H_1+2H_2) \end{pmatrix}, \tag{6.33}$$

acting on a vector space, and satisfying the same commutation relations as the corresponding matrices $\{e_{1\pm}, e_{2\pm}, e_{3\pm}, h_1, h_2\}$. The commutation relations may be summarised in terms of $\hat{R}^i{}_j$ by

$$\big[\hat{R}^i{}_j,\, \hat{R}^k{}_l\big] = \delta^k{}_j\, \hat{R}^i{}_l - \delta^i{}_l\, \hat{R}^k{}_j\,, \tag{6.34}$$

since, for X, Y appropriate matrices, (6.34) requires

$$\big[\mathrm{tr}(X\hat{R}), \mathrm{tr}(Y\hat{R})\big] = \mathrm{tr}\big([X,Y]\hat{R}\big)\,, \tag{6.35}$$

and with the definitions (6.27) and (6.29) we have, from (6.33), $\mathrm{tr}(e_{i\pm}\hat{R}) = E_{i\pm}$, $i = 1,2,3$ and $\mathrm{tr}(h_i\hat{R}) = H_i$, $i = 1,2$.

Just as with $\mathfrak{su}(2)$ the possible irreducible representation spaces may be determined algebraically from the commutation relations of the operators in the privileged basis given in (6.33). Crucially there are two commuting generators H_1, H_2 so that

$$[H_1, H_2] = 0\,. \tag{6.36}$$

For E_{i+} the commutation relations are

$$[E_{1+}, E_{2+}] = E_{3+}\,, \qquad [E_{1+}, E_{3+}] = [E_{2+}, E_{3+}] = 0\,. \tag{6.37}$$

while under commutation with H_1, H_2

$$\begin{aligned} \big[H_1, \{E_{1\pm}, E_{2\pm}, E_{3\pm}\}\big] &= \pm\{2E_{1\pm}, -E_{2\pm}, E_{3\pm}\}\,, \\ \big[H_2, \{E_{1\pm}, E_{2\pm}, E_{3\pm}\}\big] &= \pm\{-E_{1\pm}, 2E_{2\pm}, E_{3\pm}\}\,. \end{aligned} \tag{6.38}$$

The remaining commutators involving $E_{i\pm}$ are

$$\begin{aligned} &[E_{1+}, E_{1-}] = H_1\,, \qquad [E_{1+}, E_{2-}] = 0\,, \qquad [E_{2+}, E_{2-}] = H_2\,, \\ &[E_{3+}, E_{1-}] = -E_{2+}\,, \quad [E_{3+}, E_{2-}] = E_{1+}\,, \quad [E_{3+}, E_{3-}] = H_1 + H_2\,, \end{aligned} \tag{6.39}$$

together with those obtained by conjugation, $[X,Y]^\dagger = -[X^\dagger, Y^\dagger]$, where $E_{i\pm}{}^\dagger = E_{i\mp}$ and $H_i^\dagger = H_i$.

The $\mathfrak{su}(3)$ Lie algebra basis in (6.33) can be decomposed into three $\mathfrak{su}(2)$ Lie algebras,

$$\mathfrak{l}_1 = \{E_{1+}, E_{1-}, H_1\}\,, \qquad \mathfrak{l}_2 = \{E_{2+}, E_{2-}, H_2\}\,, \qquad \mathfrak{l}_3 = \{E_{3+}, E_{3-}, H_1 + H_2\}\,, \tag{6.40}$$

where each $\mathfrak{l}_i$ satisfies (6.5). From (6.31) the automorphism symmetries of the privileged basis in (6.33) are generated by

$$\mathfrak{l}_1 \underset{b}{\to} \mathfrak{l}_{1R}\,, \quad \mathfrak{l}_2 \underset{b}{\to} \mathfrak{l}_3\,, \quad \mathfrak{l}_3 \underset{b}{\to} \mathfrak{l}_2\,, \qquad \mathfrak{l}_1 \underset{a}{\to} \mathfrak{l}_2\,, \quad \mathfrak{l}_2 \underset{a}{\to} \mathfrak{l}_{3R}\,, \quad \mathfrak{l}_3 \underset{a}{\to} \mathfrak{l}_{1R}\,, \tag{6.41}$$

with the reflected $\mathfrak{su}(2)$ Lie algebra defined by (6.6). The corresponding Weyl group, defined in terms of transformations a, b satisfying (6.32), $W(\mathfrak{su}(3)) \simeq S_3$.

If we define

$$H_\perp = \tfrac{1}{\sqrt{3}}(H_1 + 2H_2)\,, \tag{6.42}$$

then the automorphism symmetries become

$$\big(H_1, H_\perp\big) \underset{b}{\to} \big(-H_1, H_\perp\big)\,, \qquad \big(H_1, H_\perp\big) \underset{a}{\to} \big(-\tfrac{1}{2}H_1 + \tfrac{\sqrt{3}}{2}H_\perp, -\tfrac{\sqrt{3}}{2}H_1 - \tfrac{1}{2}H_\perp\big)\,. \tag{6.43}$$

Regarding $H_1, H_\perp$ as corresponding to Cartesian x, y coordinates then b represents a reflection in the y-axis and a a rotation through $2\pi/3$.

6.3 Highest Weight Representations for $\mathfrak{su}(3)$

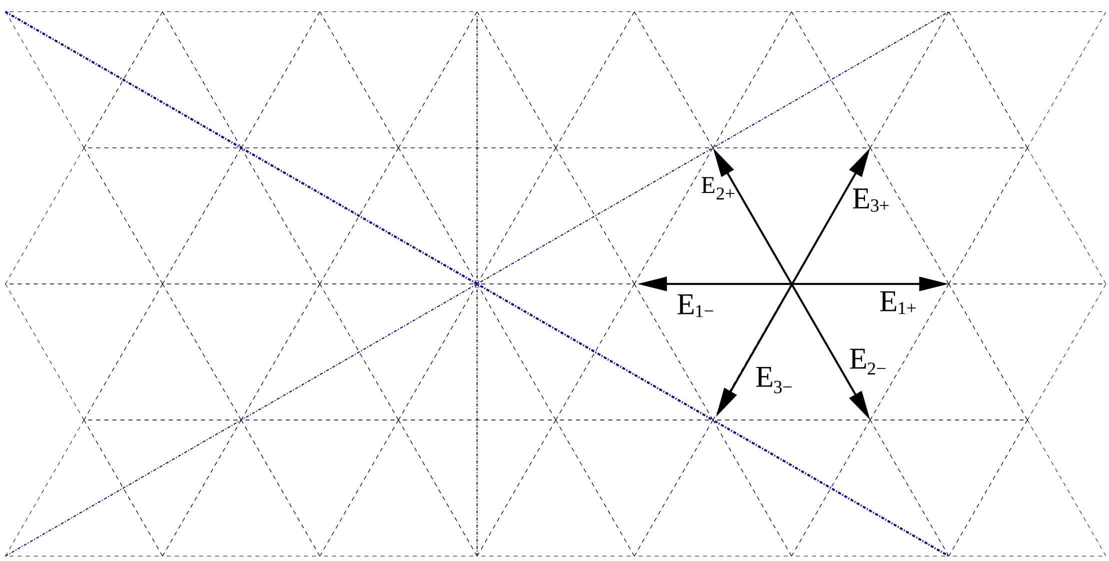
H_1, H_2 commute, (6.36), and a standard basis for the representation space for $\mathfrak{su}(3)$ is given by their simultaneous eigenvectors $|r_1, r_2\rangle$ where

$$H_1|r_1, r_2\rangle = r_1|r_1, r_2\rangle\,, \qquad H_2|r_1, r_2\rangle = r_2|r_1, r_2\rangle\,. \tag{6.44}$$

As a consequence of (6.38) we must then have

$$\begin{aligned} E_{1\pm}|r_1, r_2\rangle &\propto |r_1 \pm 2, r_2 \mp 1\rangle\,, \\ E_{2\pm}|r_1, r_2\rangle &\propto |r_1 \mp 1, r_2 \pm 2\rangle\,, \\ E_{3\pm}|r_1, r_2\rangle &\propto |r_1 \pm 1, r_2 \pm 1\rangle\,, \end{aligned} \tag{6.45}$$

unless E_{i+} and/or E_{i-} annihilate $|r_1, r_2\rangle$ for one or more individual i. The set of values $[r_1, r_2]$, linked by (6.45), are the weights of the representation. The may be plotted on a triangular lattice with r_1 along the x-axis and $\frac{1}{\sqrt{3}}(r_1 + 2r_2)$ along the y-axis.



For any element $\sigma \in W(\mathfrak{su}(3))$ there is an associated action on the weights for $\mathfrak{su}(3)$, $\sigma[r_1, r_2]$, such that

$$H_i \underset{\sigma}{\rightarrow} H'_i\,, \quad H'_i|r_1,r_2\rangle = r'_i|r_1,r_2\rangle\,, \ i=1,2 \quad \Rightarrow \quad [r'_1, r'_2] = \sigma[r_1,r_2]\,. \tag{6.46}$$

From (6.41) this is given by

$$\begin{aligned} &b[r_1,r_2] = [-r_1, r_1+r_2]\,, \quad ab[r_1,r_2] = [r_1+r_2, -r_2]\,, \quad a^2b[r_1,r_2] = [-r_2,-r_1]\,, \\ &a[r_1,r_2] = [r_2, -r_1-r_2]\,, \qquad a^2[r_1,r_2] = [-r_1-r_2, r_1]\,. \end{aligned} \tag{6.47}$$

As will become apparent the set of weights for any representation is invariant under the action of the Weyl group, thus $\mathfrak{su}(3)$ weight diagrams are invariant under rotations by $2\pi/3$ and reflections in the y-axis.

For a highest weight representation there is a unique vector $|n_1,n_2\rangle_{\rm hw}$, such that for all other weights $r_1+r_2 < n_1+n_2$. $[n_1,n_2]$ is the highest weight and we must then have

$$E_{1+}|n_1,n_2\rangle_{\rm hw} = E_{2+}|n_1,n_2\rangle_{\rm hw} = 0 \quad \Rightarrow \quad E_{3+}|n_1,n_2\rangle_{\rm hw} = 0\,. \tag{6.48}$$

The corresponding representation space $V_{[n_1,n_2]}$ is formed by the action of arbitrary products of the lowering operators E_{i-}, $i=1,2,3$ on the highest weight vector. For a basis for $V_{[n_1,n_2]}$ we may choose

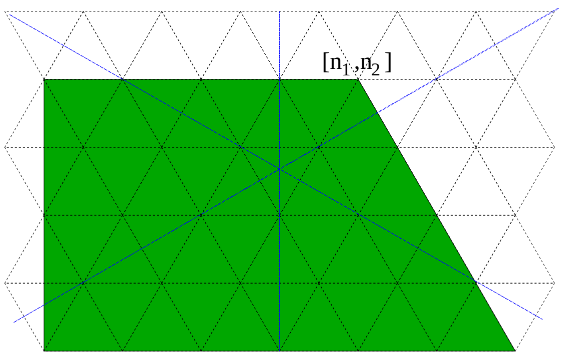
$$\big\{ E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1,n_2\rangle_{\rm hw} : r,s,t = 0,1,\dots \big\}\,, \tag{6.49}$$

where the ordering of E_{1-}, E_{2-}, E_{3-} in (6.49) reflects an arbitrary choice, any polynomial in E_{1-}, E_{2-}, E_{3-} acting on $|n_1,n_2\rangle$ may be expressed uniquely in terms of the basis (6.49) using the commutation relations given by the conjugate of (6.37).

For these basis vectors

$$\begin{aligned} H_1\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1,n_2\rangle_{\rm hw} &= (n_1 - 2r + s - t)\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1,n_2\rangle_{\rm hw}\,, \\ H_2\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1,n_2\rangle_{\rm hw} &= (n_2 + r - 2s - t)\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1,n_2\rangle_{\rm hw}\,, \end{aligned} \tag{6.50}$$

so that the weights of vectors belonging to $V_{[n_1,n_2]}$ are those belonging to a $2\pi/3$ segment in the weight diagram with vertex at $[n_1,n_2]$, as shown by the shaded region in the figure below.



It is clear that the basis (6.49) for $V_{[n_1,n_2]}$ requires that in general the allowed weights are degenerate, *i.e.* there are multiple vectors for each allowed weight in the representation space $V_{[n_1,n_2]}$ except on the boundary. For a particular weight $[r_1, r_2]$, (6.49) gives the $k+1$ or $l+1$, depending on which is the less, independent vectors,

$$E_{3-}{}^{t} E_{2-}{}^{l-t} E_{1-}{}^{k-t} |n_1, n_2\rangle_{\rm hw} \, , \qquad 0 \le t \le k, l \, , \tag{6.51}$$

where

$$k = \tfrac{1}{3}\big(2n_1 + n_2 - 2r_1 - r_2\big) \, , \qquad l = \tfrac{1}{3}\big(n_1 + 2n_2 - r_1 - 2r_2\big) \, . \tag{6.52}$$

The representation of $\mathfrak{su}(3)$ is determined then in terms of the action of $E_{i\pm}$ on the basis (6.49). For the lowering operators it is easy to see that

$$\begin{aligned} E_{3-}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} &= E_{3-}{}^{t+1} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \, , \\ E_{2-}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} &= E_{3-}{}^{t} E_{2-}{}^{s+1} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \, , \\ E_{1-}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} &= E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r+1} |n_1, n_2\rangle_{\rm hw} \\ &\qquad - s\, E_{3-}{}^{t+1} E_{2-}{}^{s-1} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \, , \end{aligned} \tag{6.53}$$

using $[E_{1-}, E_{2-}{}^{s}] = -s\, E_{3-} E_{2-}{}^{s-1}$.

The action of E_{i+} on the basis (6.49) may then be determined by using the basic commutation relations (6.39), with (6.38) and (6.37), and then applying (6.48). Just as in (6.12) we may obtain

$$\big[E_{1+}, E_{1-}{}^{r}\big] = E_{1-}{}^{r-1}\, r(H_1 - r + 1) \, , \quad \big[E_{1+}, E_{2-}{}^{s}\big] = 0 \, , \quad \big[E_{1+}, E_{3-}{}^{t}\big] = -t\, E_{3-}{}^{t-1} E_{2-} \, , \tag{6.54}$$

so that

$$\begin{aligned} & E_{1+}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \\ & \quad = r(n_1 - r + 1)\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r-1} |n_1, n_2\rangle_{\rm hw} - t\, E_{3-}{}^{t-1} E_{2-}{}^{s+1} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \, . \end{aligned} \tag{6.55}$$

Similarly

$$\begin{aligned} \big[E_{2+}, E_{3-}{}^{t}\big] &= t\, E_{3-}{}^{t-1} E_{1-} \, , & \big[E_{1-}, E_{2-}{}^{s}\big] &= -s\, E_{3-} E_{2-}{}^{s-1} \, , \\ \big[E_{2+}, E_{2-}{}^{s}\big] &= E_{2-}{}^{s-1}\, s(H_2 - s + 1) \, , & \big[E_{2+}, E_{1-}{}^{r}\big] &= 0 \, , \end{aligned} \tag{6.56}$$

which leads to

$$\begin{aligned} & E_{2+}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \\ & = s(n_2 + r - s - t + 1)\, E_{3-}{}^{t} E_{2-}{}^{s-1} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} + t\, E_{3-}{}^{t-1} E_{2-}{}^{s} E_{1-}{}^{r+1} |n_1, n_2\rangle_{\rm hw} \, . \end{aligned} \tag{6.57}$$

Furthermore

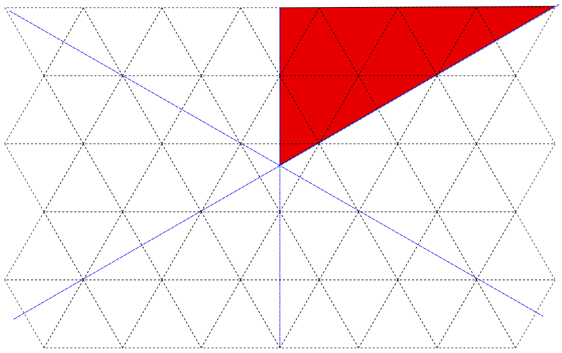
$$\begin{aligned} E_{3+}\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} &= [E_{1+}, E_{2+}]\, E_{3-}{}^{t} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \\ &= t(n_1 + n_2 - r - s - t + 1)\, E_{3-}{}^{t-1} E_{2-}{}^{s} E_{1-}{}^{r} |n_1, n_2\rangle_{\rm hw} \\ &\quad + rs(n_1 - r + 1)\, E_{3-}{}^{t} E_{2-}{}^{s-1} E_{1-}{}^{r-1} |n_1, n_2\rangle_{\rm hw} \, . \end{aligned} \tag{6.58}$$

The results (6.50), (6.53) with (6.55), (6.57) and (6.58) demonstrate how $V_{[n_1,n_2]}$ forms a representation space for $\mathfrak{su}(3)$.

Defining now

$$\mathcal{W} = \{[m,n] : m,n \in \mathbb{N}_0\}\,, \tag{6.59}$$

which corresponds to the sector of a weight diagram shown below,



then, if $[n_1,n_2] \in \mathcal{W}$, $V_{[n_1,n_2]}$ contains further highest weight vectors, satisfying (6.48), which may be used to construct invariant subspaces. Directly from (6.55) and (6.57) it is easy to see that

$$\begin{aligned}
|-n_1-2, n_1+n_2+1\rangle_{\rm hw} &= E_{1-}{}^{n_1+1}|n_1,n_2\rangle_{\rm hw}\,,\\
|n_1+n_2+1, -n_2-2\rangle_{\rm hw} &= E_{2-}{}^{n_2+1}|n_1,n_2\rangle_{\rm hw}\,,\\
|n_2, -n_1-n_2-3\rangle_{\rm hw} &= E_{2-}{}^{n_2+n_1+2}|-n_1-2, n_1+n_2+1\rangle_{\rm hw}\,,
\end{aligned} \tag{6.60}$$

satisfy the necessary conditions (6.48). In general, a linear combination of the vectors in (6.51)

$$|r_1,r_2\rangle = \sum_{0\le t\le k,l} a_t\, E_{3-}{}^{t} E_{2-}{}^{l-t} E_{1-}{}^{k-t}|n_1,n_2\rangle_{\rm hw}\,, \tag{6.61}$$

satisfies the highest weight conditions (6.48), by virtue of (6.55) and (6.57), only if

$$\begin{aligned}
(k-t)(n_1-k+1+t)\,a_t - (t+1)\,a_{t+1} &= 0\,,\\
(l-t)(n_2+k-l+1-t)\,a_t + (t+1)\,a_{t+1} &= 0\,,
\end{aligned} \tag{6.62}$$

for all relevant t. This requires

$$(k-t)(n_1-k+1+t) = -(l-t)(n_2+k-l+1-t)\,, \tag{6.63}$$

which has two solutions

$$k = n_1+n_2+2 \quad \Rightarrow \quad l = n_2+1 \ \text{ or } \ l = n_1+n_2+2\,. \tag{6.64}$$

It is then possible to construct two further highest weight vectors $|-n_1-n_2-3, n_1\rangle_{\rm hw}$ and $|-n_2-2, -n_1-2\rangle_{\rm hw}$. Just as in (6.60) we may write

$$|-n_1-n_2-3, n_1\rangle_{\rm hw} = E_{1-}{}^{n_2+n_1+2}|n_1+n_2+1, -n_2-2\rangle_{\rm hw}\,, \tag{6.65}$$

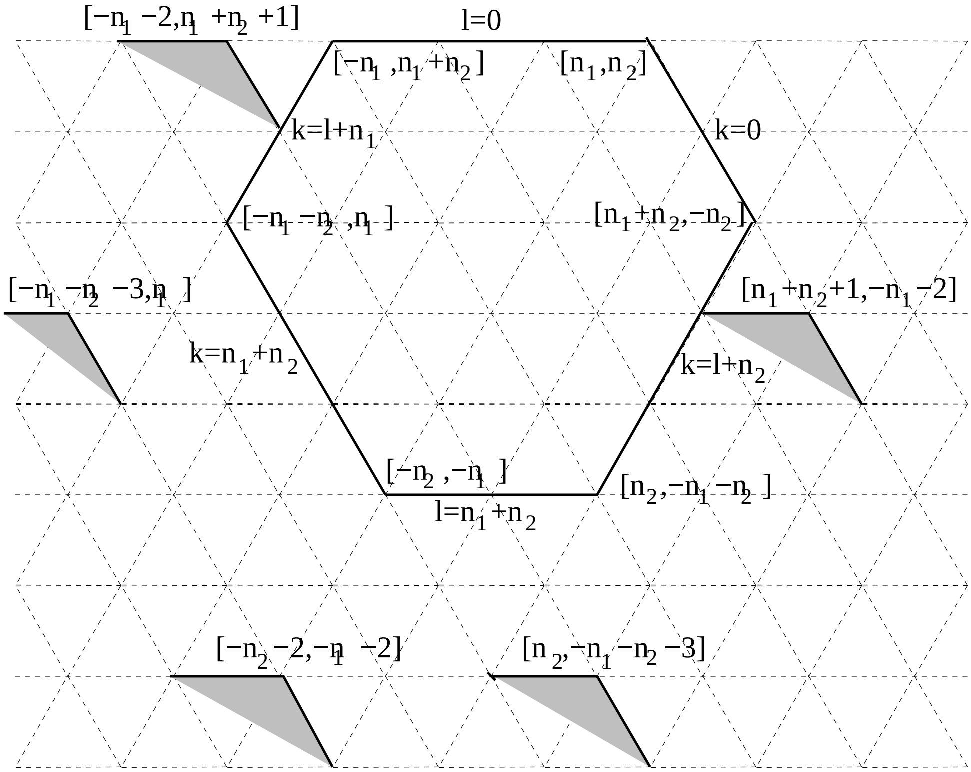
and furthermore[20]

$$|{-n_2-2}, -n_1-2\rangle_{\rm hw} = E_{1-}{}^{n_2+1}|n_2, -n_1-n_2-3\rangle_{\rm hw} = E_{2-}{}^{n_1+1}|{-n_1-n_2-3}, n_1\rangle_{\rm hw}\, . \qquad (6.66)$$

It is not difficult to see that for each of highest weight vectors given in (6.60), (6.65) and (6.66), $|n'_1, |n'_2\rangle$, there are associated invariant, under the action of $\mathfrak{su}(3)$, subspaces $V_{n'_1,|n'_2}$ contained in $V_{[n_1,n_2]}$. In particular

$$\begin{aligned}
&V_{[-n_1-2,\,n_1+n_2+1]}\, ,\ V_{[n_1+n_2+1,-n_2-2]} \subset V_{[n_1,n_2]}\, ,\\
&V_{[n_2,-n_1-n_2-3]} \subset V_{[-n_1-2,\,n_1+n_2+1]}\, ,\ V_{[-n_1-n_2-3,\,n_1]} \subset V_{[n_1+n_2+1,-n_2-2]}\, ,\\
&V_{[-n_2-2,-n_1-2]} \subset V_{[n_2,-n_1-n_2-3]} \cap V_{[-n_1-n_2-3,\,n_1]}\, . \qquad (6.67)
\end{aligned}$$

The highest weight vectors which are present are illustrated on the weight diagram below, with the shaded regions indicating where the associated invariant subspaces are present.



The reduction to an irreducible representation space becomes less trivial than that given by (6.16) for $\mathfrak{su}(2)$ due to this nested structure of invariant subspaces. Using the same definition of the quotient of a vector space by a subspace as in (6.16) we may define

$$\begin{aligned}
\mathcal{V}^{(2)}_{[n_1,n_2]} &= \big(V_{[n_2,-n_1-n_2-3]} \oplus V_{[-n_1-n_2-3,\,n_1]}\big)\big/\, V_{[-n_2-2,-n_1-2]}\, ,\\
\mathcal{V}^{(1)}_{[n_1,n_2]} &= \big(V_{[-n_1-2,\,n_1+n_2+1]} \oplus V_{[n_1+n_2+1,-n_2-2]}\big)\big/\, \mathcal{V}^{(2)}_{[n_1,n_2]}\, ,\\
\mathcal{V}_{[n_1,n_2]} &= V_{[n_1,n_2]}\big/\, \mathcal{V}^{(1)}_{[n_1,n_2]}\, . \qquad (6.68)
\end{aligned}$$

[20]To show this use $E_{1-}{}^{r}E_{2-}{}^{s} = \sum_{t=0}^{r}(-1)^t\binom{r}{t}\frac{s!}{(s-t)!}\,E_{3-}{}^{t}E_{2-}{}^{s-t}E_{1-}{}^{r-t}$.

In $\mathcal{V}_{[n_1,n_2]}$ there then are no highest weight vectors other than $|n_1,n_2\rangle_{\rm hw}$ so invariant subspaces are absent and $\mathcal{V}_{[n_1,n_2]}$ is a representation space for an irreducible representation of $\mathfrak{su}(3)$. Although it remains to be demonstrated the representation space is then finite-dimensional and the corresponding weight diagram has vertices with weights

$$[n_1,n_2]\,, \quad [-n_1,n_1{+}n_2]\,, \quad [n_1{+}n_2,-n_2]\,, \quad [-n_1-n_2,n_1]\,, \quad [n_2,-n_1-n_2]\,, \quad [-n_2,-n_1]\,, \tag{6.69}$$

which are related by the transformations of the Weyl group as in (6.47).

6.3.1 Analysis of the Weight Diagram

To show how (6.68) leads to a finite-dimensional representation we consider how it applies to for the vectors corresponding to particular individual weights $[r_1,r_2]$. Accordingly we consider restrictions of the highest weight spaces. For $V_{[n_1,n_2]}$ the relevant subspace is formed by the basis in (6.51)

$$V_{[n_1,n_2]}^{(k,l)} = \big\{\textstyle\sum_{0\le t\le k,l} a_t \; E_{3-}{}^t E_{2-}{}^{l-t} E_{1-}{}^{k-t}|n_1,n_2\rangle_{\rm hw}\big\}\,, \tag{6.70}$$

where k,l are determined as in (6.52). Clearly these subspaces are finite-dimensional with

$$\dim V_{[n_1,n_2]}^{(k,l)} = \begin{cases} k+1\,, & k\le l\,, \\ l+1\,, & l \le k\,. \end{cases} \tag{6.71}$$

In a similar fashion we may define

$$\begin{aligned} &V_{[-n_1-2,n_1+n_2+1]}^{(k-n_1-1,l)}\,, \qquad V_{[n_1+n_2+1,-n_2-2]}^{(k,l-n_2-1)}\,, \\ &V_{[n_2,-n_1-n_2-3]}^{(k-n_1-1,l-n_1-n_2-2)}\,, \qquad V_{[-n_1-n_2-3,n_1]}^{(k-n_1-n_2-2,l-n_2-1)}\,, \qquad V_{[-n_2-2,-n_1-2]}^{(k-n_1-n_2-2,l-n_1-n_2-2)}\,, \end{aligned} \tag{6.72}$$

which form nested subspaces, just as in (6.67), and whose dimensions are given by the obvious extension of (6.71).

To illustrate how the construction of the representation space $\mathcal{V}_{[n_1,n_2]}$ in terms of quotient spaces leads to cancellations outside a finite region of the weight diagram we describe how this is effected in particular regions of the weight diagram by showing that the dimensions of the quotient spaces outside the finite region of the weight diagram specified by vertices in (6.69) are zero and also that on the boundary the dimension is one. For $k\le n_1$, $l\le n_2$ there are no cancellations for $V_{[n_1,n_2]}^{(k,l)}$. Taking into account the contributions from $V_{[-n_1-2,n_1+n_2+1]}^{(k-n_1-1,l)}$ and $V_{[n_1+n_2+1,-n_2-2]}^{(k,l-n_2-1)}$ gives

$$\dim V_{[n_1,n_2]}^{(k,l)} - \dim V_{[-n_1-2,n_1+n_2+1]}^{(k-n_1-1,l)} = \begin{cases} 0 & \text{if} \quad k\ge l+n_1+1,\, l\ge 0\,, \\ 1 & \text{if} \quad k = l+n_1,\, l\ge 0\,, \end{cases} \tag{6.73}$$

and

$$\dim V_{[n_1,n_2]}^{(k,l)} - \dim V_{[n_1+n_2+1,-n_2-2]}^{(k,l-n_2-1)} = \begin{cases} 0 & \text{if} \quad l\ge k+n_2+1,\, k\ge 0\,, \\ 1 & \text{if} \quad l = k+n_2,\, k\ge 0\,. \end{cases} \tag{6.74}$$

Furthermore

$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} \\
&= \begin{cases} l+1-n_2-(l-n_2)=1\,, & k=n_1+n_2,\ n_2\le l\le n_1+n_2\,, \\ k+1-(k-n_1)-n_1=1\,, & l=n_1+n_2,\ n_1\le k\le n_1+n_2\,. \end{cases}
\end{aligned} \tag{6.75}$$

The remaining contributions, when present, give rise to a complete cancellation so that the representation space given by (6.68) is finite dimensional. When $l \ge n_2$, $k \ge n_1+n_2+1$,

$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} + \dim V^{(k-n_1-n_2-2,l-n_2-1)}_{[-n_1-n_2-3,n_1]} \\
&= \begin{cases} (l+1)-(l+1)-(l-n_2)+(l-n_2)\,, & k\ge l+n_1+1 \\ (l+1)-(k-n_1)-(l-n_2)+(k-n_1-n_2-1)\,, & l\le k\le l+n_1+1 \end{cases} \\
&= 0\,,
\end{aligned} \tag{6.76}$$

and for $k, l \ge n_1+n_2+1$ in an analogous fashion

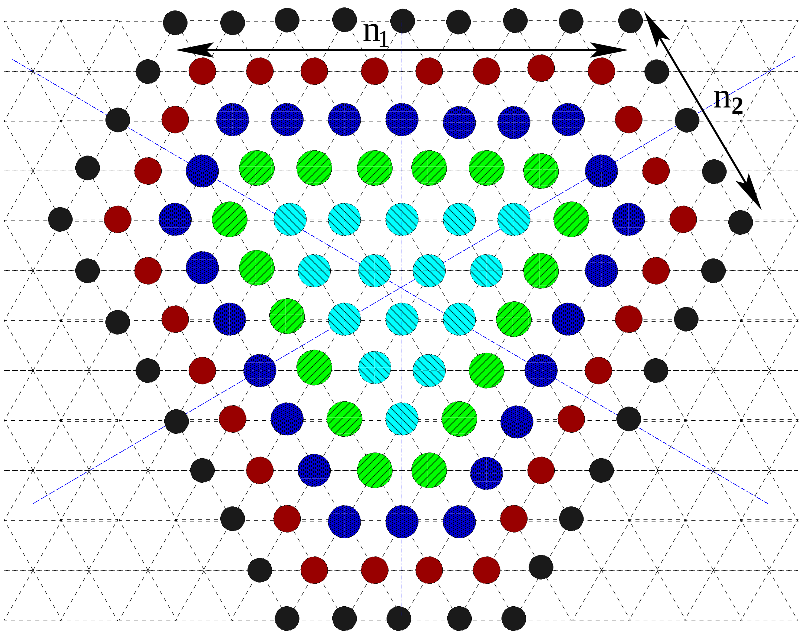
$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} + \dim V^{(k-n_1-1,l-n_1-n_2-2)}_{[n_2,-n_1-n_2-3]} \\
&+ \dim V^{(k-n_1-n_2-2,l-n_2-1)}_{[-n_1-n_2-3,n_1]} - \dim V^{(k-n_1-n_2-2,l-n_1-n_2-2)}_{[-n_2-2,-n_1-2]} = 0\,.
\end{aligned} \tag{6.77}$$

For the finite representation space $\mathcal{V}_{[n_1,n_2]}$ then at each vertex of the weight diagram as in (6.69) there are associated vectors which satisfy analogous conditions to (6.48), in particular

$$\begin{aligned}
&(E_{1-}, E_{3+})\,|{-n_1}, n_1+n_2\rangle = 0\,, \qquad (E_{2-}, E_{3+})\,|n_1+n_2, -n_2\rangle\,, \\
&(E_{2+}, E_{3-})\,|{-n_1}-n_2, n_1\rangle = 0\,, \qquad (E_{1+}, E_{3-})\,|n_2, -n_1-n_2\rangle = 0\,, \\
&\quad (E_{1-}, E_{2-})\,|{-n_2}, -n_1\rangle = 0\,.
\end{aligned} \tag{6.78}$$

Each vector may be use to construct the representation space by acting on it with appropriate lowering operators. In this fashion $\mathcal{V}_{[n_1,n_2]}$ may be shown to be invariant under $W(\mathfrak{su}(3))$.

A generic weight diagram has the structure shown below. The multiplicity for each weight is the same on each layer. For $n_1 \ge n_2$ there are n_2+1 six-sided layers and then the layers become triangular. For the six-sided layers the multiplicity increases by one as one moves from the outside to the inside, the triangular layers all have multiplicity n_2+1. In the diagram different colours have the same multiplicity.



6.3.2 $SU(3)$ Characters

A much more straightforward procedure for showing how finite dimensional representations of $SU(3)$ are formed is to construct their characters following the approach described for $SU(2)$ based on (6.21) and (6.22). For the highest weight representation space $V_{[n_1,n_2]}$ we then define in terms of the basis (6.49)

$$\begin{aligned} C_{[n_1,n_2]}(t_1,t_2) = \widetilde{\mathrm{tr}}_{V_{[n_1,n_2]}}\big({t_1}^{H_1}\,{t_2}^{H_2}\big) &= \sum_{r,s,t\geq 0} {t_1}^{n_1-2r+s-t}\,{t_2}^{n_2+r-2s-t} \\ &= {t_1}^{n_1}\,{t_2}^{n_2} \sum_{r,s,t\geq 0} \big(t_2/{t_1}^2\big)^r \big(t_1/{t_2}^2\big)^s \big(1/t_1t_2\big)^t \,. \end{aligned} \tag{6.79}$$

For a succinct final expression it is more convenient to use the variables

$$u = (u_1,u_2,u_3)\,, \quad u_1 = t_1\,, \quad u_3 = 1/t_2\,, \quad u_1u_2u_3 = 1\,, \tag{6.80}$$

so that $t_2/t_1^2 = u_2/u_1$, $t_1/t_2^2 = u_3/u_2$, $1/t_1t_2 = u_3/u_1$ and convergence of the sum requires $u_1 > u_2 > u_3$. Then

$$C_{[n_1,n_2]}(u) = \frac{{u_1}^{n_1+n_2+2}\,{u_2}^{n_2+1}}{(u_1-u_2)(u_2-u_3)(u_1-u_3)}\,. \tag{6.81}$$

Following (6.68) the character for the irreducible representation of $\mathfrak{su}(3)$ obtained from the highest weight vector $|n_1, n_2\rangle_{\rm hw}$ is then

$$\begin{aligned}\chi_{[n_1,n_2]}(u) &= C_{[n_1,n_2]}(u) - C_{[-n_1-2,n_1+n_2+1]}(u) - C_{[n_1+n_2+1,-n_2-2]}(u) \\ &\quad + C_{[-n_1-n_2-3,n_1]}(u) + C_{[n_2,-n_1-n_2-3]}(u) - C_{[-n_2-2,-n_1-2]}(u) \\ &= \frac{1}{(u_1-u_2)(u_2-u_3)(u_1-u_3)} \\ &\qquad \times \big({u_1}^{n_1+n_2+2}\, {u_2}^{n_2+1} - {u_2}^{n_1+n_2+2}\, {u_1}^{n_2+1} - {u_1}^{n_1+n_2+2}\, {u_3}^{n_2+1} \\ &\qquad\quad + {u_2}^{n_1+n_2+2}\, {u_3}^{n_2+1} - {u_3}^{n_1+n_2+2}\, {u_2}^{n_2+1} + {u_3}^{n_1+n_2+2}\, {u_1}^{n_2+1} \big) \,. \end{aligned} \tag{6.82}$$

It is easy to see that both the numerator and the denominator are completely antisymmetric so that $\chi_{[n_1,n_2]}(u)$ is a symmetric function of u_1, u_2, u_3, the $S_3 \simeq W(\mathfrak{su}(3))$.

If we consider a particular restriction we get

$$\chi_{[n_1,n_2]}(q, 1, q^{-1}) = \frac{1-q^{n_1+1}}{1-q}\, \frac{1-q^{n_2+1}}{1-q}\, \frac{1-q^{-n_1-n_2-2}}{1-q^{-2}} \,, \tag{6.83}$$

and hence it is then easy to calculate

$$\dim \mathcal{V}_{[n_1,n_2]} = \chi_{[n_1,n_2]}(1,1,1) = \tfrac{1}{2}(n_1+1)(n_2+1)(n_1+n_2+2) \,. \tag{6.84}$$

The relation of characters to the Weyl group is made evident by defining, for any element $\sigma \in W(\mathfrak{su}(3))$, a transformation on the weights such that

$$[r_1, r_2]^\sigma = \sigma[r_1+1, r_2+1] - [1,1] \,. \tag{6.85}$$

Directly from (6.47) we easily obtain

$$\begin{aligned} [r_1,r_2]^b &= [-r_1-2, r_1+r_2+1] \,, & [r_1,r_2]^{ab} &= [r_1+r_2+1, -r_2-2] \,, \\ [r_1,r_2]^a &= [r_2, -r_1-r_2-3] \,, & [r_1,r_2]^{a^2} &= [-r_1-r_2-3, r_1] \,, \\ [r_1,r_2]^{a^2 b} &= [-r_2-2, -r_1-2] \,. \end{aligned} \tag{6.86}$$

Clearly $[n_1, n_2]^\sigma$ generates the weights for the highest weight vectors contained in $\mathcal{V}_{[n_1,n_2]}$, as shown in (6.60), (6.65) and (6.66). Thus (6.82) may be written more concisely as

$$\chi_{[n_1,n_2]}(u) = \sum_{\sigma \in S_3} P_\sigma\, C_{[n_1,n_2]^\sigma}(u) = \sum_{\sigma \in S_3} C_{[n_1,n_2]}(\sigma u) \,, \tag{6.87}$$

with, for $\sigma \in S_3$,

$$P_\sigma = \begin{cases} -1 \,, & \sigma \text{ odd permutation} \,, \\ 1 \,, & \sigma \text{ even permutation} \,, \end{cases} \tag{6.88}$$

and where σu denotes the corresponding permutation, so that $b(u_1,u_2,u_3) = (u_2,u_1,u_3)$, $a(u_1,u_2,u_3) = (u_2,u_3,u_1)$. The definition of $\chi_{[n_1,n_2]}(u)$ extends to any $[n_1, n_2]$ by taking

$$\chi_{[n_1,n_2]^\sigma}(u) = P_\sigma \chi_{[n_1,n_2]}(u) \,. \tag{6.89}$$

Since $[-1,r]^b = [-1,r]$, $[r,-1]^{ab} = [r,-1]$ and $[r,-r-2]^{a^2b} = [r,-r-2]$ we must then have

$$\chi_{[-1,r]}(u) = \chi_{[r,-1]}(u) = \chi_{[r,-r-2]}(u) = 0\,. \tag{6.90}$$

This shows the necessity of the three factors in the dimension formula (6.84). It is important to note that for any $[n_1, n_2]$

$$n_1, n_2 \neq -1,\ n_1 + n_2 \neq -2\,, \qquad [n_1,n_2]^\sigma \in \mathcal{W} \quad \text{for a unique} \ \ \sigma \in S_3\,, \tag{6.91}$$

where $\mathcal{W}$ is defined in (6.59).

6.3.3 Casimir operator

For the basis in (6.33) the $\mathfrak{su}(3)$ quadratic Casimir operator is given by

$$\begin{aligned} C = \hat{R}^i{}_j\, \hat{R}^j{}_i &= \textstyle\sum_{i=1}^3 \big(E_{i+}E_{i-} + E_{i-}E_{i+}\big) + \tfrac{2}{3}\big({H_1}^2 + {H_2}^2 + H_1 H_2\big) \\ &= \textstyle\sum_{i=1}^3 E_{i-}E_{i+} + \tfrac{2}{3}\big({H_1}^2 + {H_2}^2 + H_1 H_2\big) + 2\big(H_1 + H_2\big)\,. \end{aligned} \tag{6.92}$$

Acting on a highest weight vector

$$C|n_1,n_2\rangle_{\rm hw} = c_{[n_1,n_2]}|n_1,n_2\rangle_{\rm hw}\,, \tag{6.93}$$

where, from the explicit form in (6.92),

$$c_{[n_1,n_2]} = \tfrac{2}{3}({n_1}^2 + {n_2}^2 + n_1 n_2) + 2(n_1 + n_2)\,. \tag{6.94}$$

It is an important check that $c_{[n_1,n_2]^\sigma} = c_{[n_1,n_2]}$ as required since C has the same eigenvalue $c_{[n_1,n_2]}$ for all vectors belonging to $V_{[n_1,n_2]}$.

6.3.4 Particular $SU(3)$ Representations

We describe here how the general results for constructing a finite dimensional $\mathfrak{su}(3)$ irreducible representation spaces $\mathcal{V}_{[n_1,n_2]}$ apply in some simple cases which are later of physical relevance. The general construction in (6.68) ensures that the resulting weight diagram is finite but in many cases the results can be obtained quite simply by considering the $\mathfrak{su}(2)$ subalgebras in (6.40) and then using results for $\mathfrak{su}(2)$ representations.

The trivial singlet representation of course arises for $n_1 = n_2 = 0$ when there is unique vector $|0,0\rangle$ annihilated by $E_{i\pm}$ and H_i.

A particularly simple class of representations arises when $n_2 = 0$. In this case applying the $\mathfrak{su}(2)$ representation condition (6.20) the highest weight vector must satisfy

$${E_{1-}}^{n_1+1}|n_1,0\rangle_{\rm hw} = 0\,, \qquad E_{2-}|n_1,0\rangle_{\rm hw} = 0\,. \tag{6.95}$$

Furthermore, using $[E_{3+}, {E_{1-}}^r] = -r{E_{1-}}^{r-1}E_{2+}$,

$$E_{3+}\,{E_{1-}}^r|n_1,0\rangle_{\rm hw} = 0\,, \qquad (H_1 + H_2)\,{E_{1-}}^r|n_1,0\rangle_{\rm hw} = (n_1 - r)\,{E_{1-}}^r|n_1,0\rangle_{\rm hw}\,, \tag{6.96}$$

so that $E_{1-}{}^{r}|n_1,0\rangle_{\rm hw}$ is a $\mathfrak{su}(2)_{\rm i_3}$ highest weight vector so that from (6.20) again

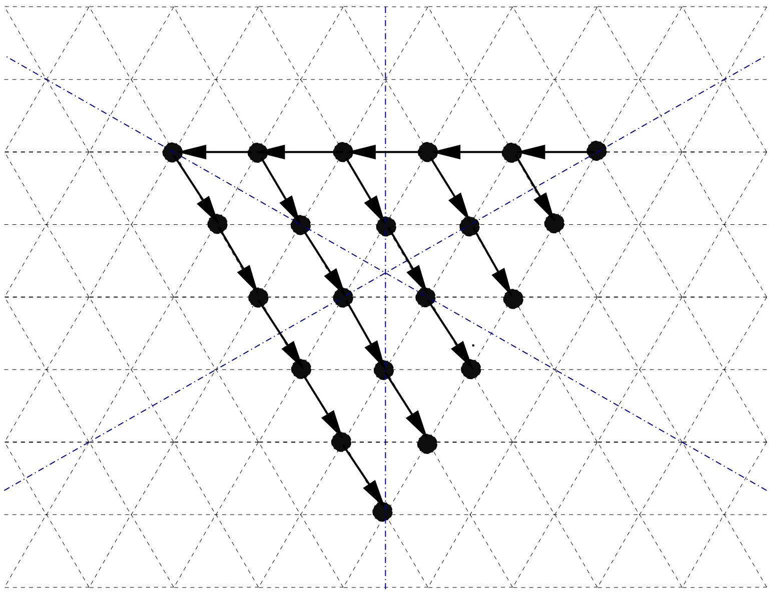
$$E_{3-}{}^{n_1-r+1}\, E_{1-}{}^{r}|n_1,0\rangle_{\rm hw} = 0\,. \tag{6.97}$$

Hence a finite dimensional basis for $\mathcal{V}_{[n_1,0]}$ is given by

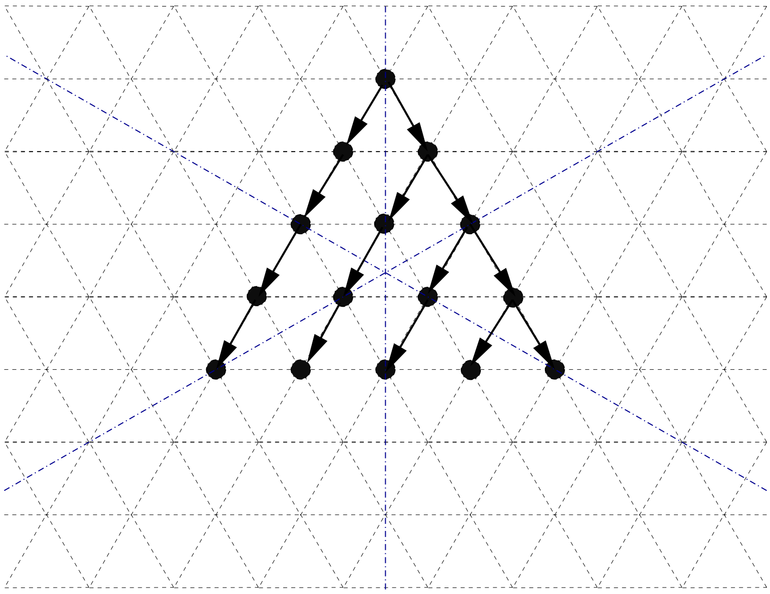
$$E_{3-}{}^{t}E_{1-}{}^{r}|n_1,0\rangle_{\rm hw}\,, \qquad t=0,\dots n_1-r\,,\ r=0,\dots,n_1\,, \tag{6.98}$$

where there is a unique vector for each weight $[n_1-2r-t, r-t]$, which therefore has multiplicity one. It is easy to check that this is in accord with the dimension of this representation $\dim\mathcal{V}_{[n_1,0]} = \frac{1}{2}(n_1+1)(n_1+2)$.

These representations have triangular weight diagrams as shown below.



A corresponding case arises when $n_1=0$ and the roles of E_{1-} and E_{2-} are interchanged. In this case the basis vectors for $\mathcal{V}_{[0,n_2]}$ are just $E_{3-}{}^{t}E_{2-}{}^{s}|0,n_2\rangle_{\rm hw}$ for $t=0,\dots n_2-s$, $s=0,\dots,n_2$ and the weight diagram is also triangular.



In general the weight diagrams for $\mathcal{V}_{[n_2,n_1]}$ may be obtained from that for $\mathcal{V}_{[n_1,n_2]}$ by rotation by π, these two representations are conjugate to each other.

The next simplest example arises for $n_1 = n_2 = 1$. The $\mathfrak{su}(2)$ conditions (6.20) for the highest weight state require

$$E_{1-}{}^2|1,1\rangle_{\rm hw} = E_{2-}{}^2|1,1\rangle_{\rm hw} = E_{3-}{}^3|1,1\rangle_{\rm hw} = 0\,. \tag{6.99}$$

Since $E_{1-}|1,1\rangle_{\rm hw}$ is a highest weight vector for $\mathfrak{su}(2)_{\rm i_2}$ and, together with $E_{2-}|1,1\rangle_{\rm hw}$, is also a $\mathfrak{su}(2)_{\rm i_3}$ highest weight vector then the weights and associated vectors obtained from $|1,1\rangle_{\rm hw}$ in terms of the basis (6.49) are then restricted to just

$$\begin{aligned}
&[-1,2]:\ E_{1-}|1,1\rangle_{\rm hw}\,, \quad [2,-1]:\ E_{2-}|1,1\rangle_{\rm hw}\,, \quad [0,0]:\ E_{3-}|1,1\rangle_{\rm hw}\,,\ E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,,\\
&[-2,1]:\ E_{3-}E_{1-}|1,1\rangle_{\rm hw}\,, \qquad [1,0]:\ E_{3-}E_{2-}|1,1\rangle_{\rm hw}\,,\ E_{2-}{}^2E_{1-}|1,1\rangle_{\rm hw}\,,\\
&[-1,-1]:E_{3-}{}^2|1,1\rangle_{\rm hw}\,,\ E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,.
\end{aligned} \tag{6.100}$$

However (6.99) requires further relations since

$$E_{2-}{}^2E_{1-}|1,1\rangle_{\rm hw} = \big(E_{2-}E_{1-}E_{2-} + E_{3-}E_{2-}\big)|1,1\rangle_{\rm hw} = 2\,E_{3-}E_{2-}|1,1\rangle_{\rm hw}\,, \tag{6.101}$$

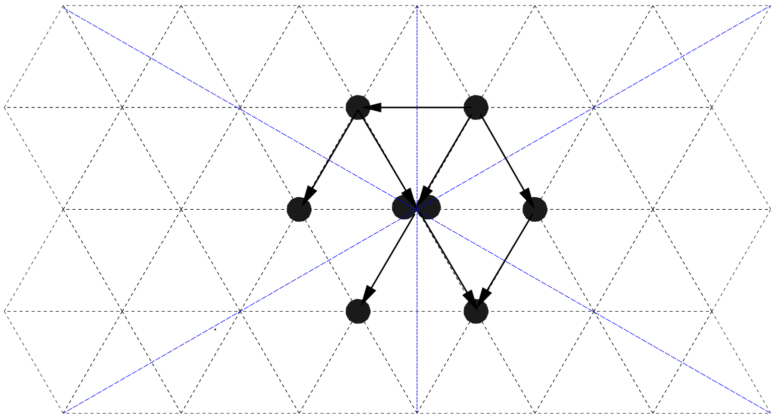
which then entails

$$\begin{aligned}
E_{1-}E_{2-}{}^2E_{1-}|1,1\rangle_{\rm hw} &= -2\,E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw}\\
&= 2\,E_{1-}E_{3-}E_{2-}|1,1\rangle_{\rm hw} = 2\big(E_{3-}E_{2-}E_{1-} - E_{3-}{}^2\big)|1,1\rangle_{\rm hw}\,,
\end{aligned} \tag{6.102}$$

so that furthermore

$$E_{3-}{}^2|1,1\rangle_{\rm hw} = 2\,E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,. \tag{6.103}$$

All weights therefore have multiplicity one except for $[0,0]$ which has multiplicity two. The overall dimension is then 8 and $\mathcal{V}_{[1,1]}$ corresponds to the $SU(3)$ adjoint representation. The associated weight diagram is just a regular hexagon, invariant under the dihedral group $D_3 \simeq S_3$, with the additional symmetry under rotation by π since this representation is self-conjugate.



6.4 $SU(3)$ Tensor Representations

Just as with the rotational group $SO(3)$, and also with $SU(2)$, representations may be defined in terms of tensors. The representation space for a rank r tensor is defined by the direct product of r copies of a fundamental representation space, formed by 3-vectors for $SO(3)$ and 2-spinors for $SU(2)$, and so belongs to the r-fold direct product of the fundamental representation. Such tensorial representations are reducible for any $r \geq 2$ with reducibility related to the existence of invariant tensors. Contraction of a tensor with an invariant tensor may lead to a tensor of lower rank so that these form an invariant subspace under the action of the group. Tensor representations become irreducible once conditions have been imposed to ensure all relevant contractions with invariant tensors are zero.

For $SU(N)$ it is necessary to consider both the N-dimensional fundamental representation and its conjugate, $SU(2)$ is a special case where these are equivalent. When $N = 3$ we then consider a complex 3-vector q^i and its conjugate $\bar{q}_i = (q^i)^*$, $i = 1, 2, 3$, belonging to the vector space $\mathcal{S}$ and its conjugate $\bar{\mathcal{S}}$, and which transform as

$$q^i \to A^i{}_j \, q^j \, , \qquad \bar{q}_i \to \bar{q}_j \, (A^{-1})^i{}_j \, , \qquad \big[A^j{}_i\big] \in SU(3) \, . \tag{6.104}$$

A (r, s)-tensor $T^{i_1 \dots i_r}_{j_1 \dots j_s}$ is then one which belongs to $\mathcal{S}(\otimes \mathcal{S})^{r-1}(\otimes \bar{\mathcal{S}})^s$ and which transforms as

$$T^{i_1 \dots i_r}_{j_1 \dots j_s} \to A^{i_1}{}_{k_1} \dots A^{i_r}{}_{k_r} \, T^{k_1 \dots k_r}_{l_1 \dots l_s} \, (A^{-1})^{l_1}{}_{j_1} \dots (A^{-1})^{l_s}{}_{j_s} \, . \tag{6.105}$$

The conjugate of a (r, s)-tensor is a (s, r)-tensor

$$\bar{T}^{j_1 \dots j_s}_{i_1 \dots i_r} = \big(T^{i_1 \dots i_r}_{j_1 \dots j_s}\big)^* \, . \tag{6.106}$$

The invariant tensors are a natural extension of those for $SU(2)$, as exhibited in (2.156) and (2.157). Thus there are the 3-index antisymmetric ε-symbols, forming $(3, 0)$ and $(0, 3)$-tensors, and the Kronecker δ, which is a $(1, 1)$-tensor,

$$\varepsilon^{ijk} \, , \qquad \varepsilon_{ijk} \, , \qquad \delta^i_j \, . \tag{6.107}$$

That ε^{ijk} and ε_{ijk} are invariant tensors is a consequence of the transformation matrix A satisfying $\det A = 1$. The transformation rules (6.105) guarantee that the contraction of an upper and lower index maintains the tensorial transformation properties. In consequence from a tensor $T^{i_1\dots i_r}_{j_1\dots j_s}$ then contracting with $\varepsilon^{ij_m j_n}$ or $\varepsilon_{j i_m i_n}$, for some arbitrary pair of indices, generates a $(r+1, s-2)$ or a $(r-2, s+1)$-tensor. Similarly using $\delta^i{}_j$ we may form a $(r-1, s-1)$-tensor. Thus the vector space of arbitrary (r,s)-tensors contains invariant subspaces, except for the fundamental $(1,0)$ or $(0,1)$ tensors or the trivial $(0,0)$ singlet. Just as for $SO(3)$ or $SU(2)$ we may form an irreducible representation space by requiring all such contractions give zero, so we restrict to (r,s)-tensors with all upper and lower indices totally symmetric, and also traceless on contraction of any upper and lower index,

$$S^{i_1\dots i_r}_{j_1\dots j_s} = S^{(i_1\dots i_r)}_{(j_1\dots j_s)}\,, \qquad S^{i_1\dots i_{r-1}\, i}_{j_1\dots j_{s-1}\, i} = 0\,. \tag{6.108}$$

The vector space formed by such symmetrised traceless tensors forms an irreducible $SU(3)$ representation space $\mathcal{V}_{[r,s]}$. To determine its dimension we may use the result in (2.124) for the dimension of the space of symmetric tensors, with indices taking three values, for $n = r, s$ and then take account of the trace conditions by subtracting the results for $n = r-1, s-1$. This gives

$$\begin{aligned} \dim \mathcal{V}_{[r,s]} &= \tfrac{1}{2}(r+1)(r+2)\,\tfrac{1}{2}(s+1)(s+2) - \tfrac{1}{2}r(r+1)\,\tfrac{1}{2}s(s+1) \\ &= \tfrac{1}{2}(r+1)(s+1)(r+s+2)\,. \end{aligned} \tag{6.109}$$

This is of course identical to (6.84). The irreducible representation space constructed in terms of (r,s)-tensors is isomorphic with the finite dimensional irreducible space constructed previously by analysis of the Lie algebra commutation relations.

6.4.1 $\mathfrak{su}(3)$ Lie algebra again

For many applications involving $SU(3)$ symmetry it is commonplace in physics papers to use a basis of hermitian traceless 3×3 matrices, forming a basis for the $\mathfrak{su}(3)$ Lie algebra, which are a natural generalisation of the Pauli matrices in (2.11), the Gell-Mann λ-matrices λ_a, $a = 1,\dots,8$,

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{6.110}$$

These satisfy

$$\mathrm{tr}(\lambda_a \lambda_b) = 2\,\delta_{ab}\,, \tag{6.111}$$

and

$$[\lambda_a, \lambda_b] = 2i\, f_{abc}\lambda_c\,, \tag{6.112}$$

for totally antisymmetric structure constants, f_{abc}. In terms of the matrices defined in (6.27) and (6.29) it is easy to see that $e_{1+} = \frac{1}{2}(\lambda_1 + i\lambda_2)$, $e_{2+} = \frac{1}{2}(\lambda_6 + i\lambda_7)$, $e_{3+} = \frac{1}{2}(\lambda_4 + i\lambda_5)$ and also $\lambda_3 = h_1$, $\lambda_8 = \frac{1}{\sqrt{3}}(h_1 + 2h_2)$.

The relation between $SU(3)$ matrices and the λ-matrices is in many similar to that for $SU(2)$ and the Pauli matrices, for an infinitesimal transformation the relation remains just as in (2.28). (2.15) needs only straightforward modification while instead of (2.12) we now have

$$\lambda_a \lambda_b = \tfrac{2}{3} I + d_{abc}\, \lambda_c + i f_{abc}\, \lambda_c \,, \qquad (6.113)$$

with d_{abc} totally symmetric and satisfying $d_{abb} = 0$.

6.5 $SU(3)$ and Physics

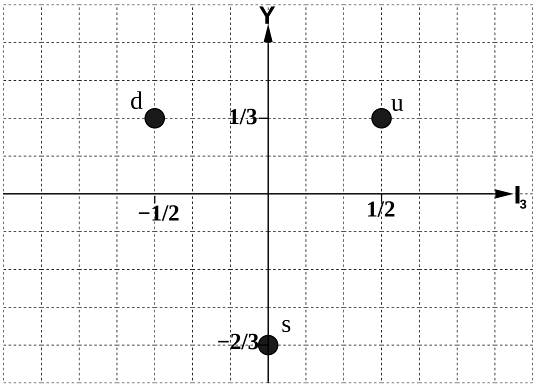
Besides its virtues in terms of understanding more general Lie groups a major motivation in studying $SU(3)$ is in terms of its role in physics. Historically $SU(3)$ was introduced, as a generalisation of the isospin $SU(2)_I$, to be an approximate symmetry group for strong interactions, in current terminology a flavour symmetry group, and the group in this context is often denoted as $SU(3)_F$. Unlike isospin, which was hypothesised to be an exact symmetry for strong interactions, neglecting electromagnetic interactions, $SU(3)_F$ is intrinsically approximate. The main evidence is the classification of particles with the same spin, parity into multiplets corresponding to $SU(3)$ representations. For the experimentally observed $SU(3)_F$ particle multiplets, unlike for isospin multiplets, the masses are significantly different.

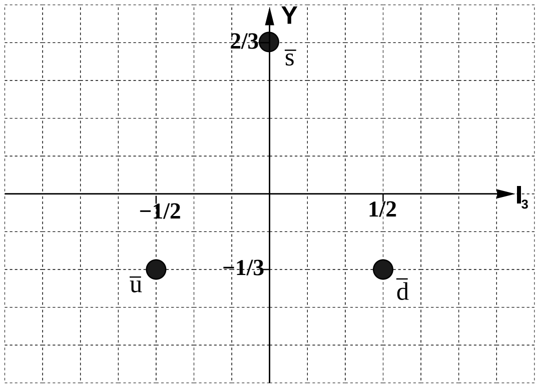
For $SU(3)_F$ the two commuting generators are identified with I_3, belonging to $SU(2)_I$, and also the hypercharge Y, where $[I_i, Y] = 0$ so that Y takes the same value for any isospin multiplet. Y is related to strangeness S, a quantum number invented to explain why the newly discovered, in the 1940's, so-called strange particles were only produced in pairs, the precise relation is $Y = B + S$, with B the baryon number. For any multiplet we must have $\mathrm{tr}(I_3) = \mathrm{tr}(Y) = 0$. Expressed in terms of the $\mathfrak{su}(3)$ operators H_1, H_2, $I_3 = H_1$, $Y = \frac{1}{3}(H_1 + 2H_2)$. For $SU(3)_F$ multiplets the electric charge is determined by $Q = I_3 + \frac{1}{2}Y$ and so must be always conserved, but Y is not conserved by weak interactions which are responsible for the decay of strange particles into non-strange particles.

For $SU(3)_F$ symmetry of strong interactions to be realised there must be 8 operators satisfying the $\mathfrak{su}(3)$ Lie algebra. If the same basis as for the λ-matrices in (6.110) is adopted then these are F_a, $a = 1, \dots, 8$, where F_a are hermitian, and

$$[F_a, F_b] = i f_{abc} F_c \,, \qquad F_i = I_i \,, \ \ i = 1,2,3 \,, \qquad F_8 = \tfrac{1}{\sqrt{3}} Y \,. \qquad (6.114)$$

From a more modern perspective $SU(3)_F$ is understood to be a consequence of the fact that low mass hadrons are composed of the three light quarks $q = (u, d, s)$ and their anti-particles $\bar{q} = (\bar{u}, \bar{d}, \bar{s})$, corresponding to three quark flavours. These belong respectively to the fundamental $[1,0]$ and $[0,1]$ representations, more often denoted by $\mathbf{3}$ and $\mathbf{3}^*$. On a weight diagram these are the simplest triangular representations. With axes labelled by I_3, Y these are

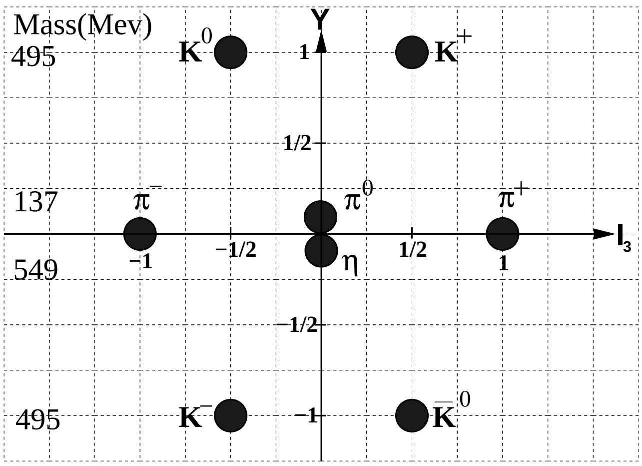




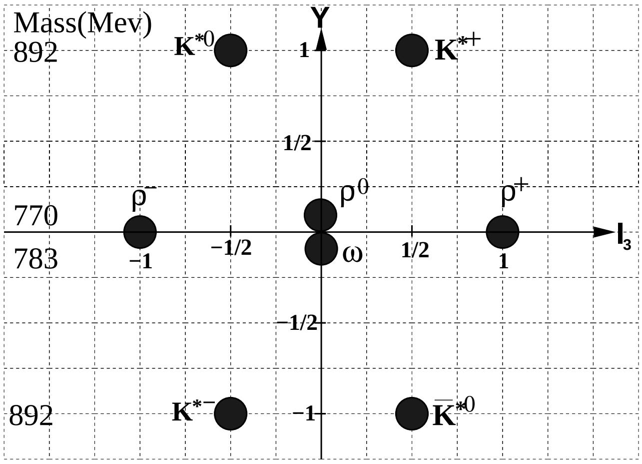
The charges of quarks are dictated by the requirement $Q = I_3 + \frac{1}{2}Y$ and so for q are fractional, $\frac{2}{3}$ and $-\frac{1}{3}$, while for $\bar{q}$ they are the opposite sign. We may further interpret the quantum numbers in terms of the numbers of particular quarks minus their anti-quarks, hence $I_3 = N_u - N_{\bar{u}} - N_d + N_{\bar{d}}$ and $S = -N_s + N_{\bar{s}}$, where each q has baryon number $B = \frac{1}{3}$ and each $\bar{q}$, $B = -\frac{1}{3}$,

As is well known isolated quarks are not observed, they are present as constituents of the experimentally observed mesons, which are generally $q\bar{q}$ composites, or baryons, whose quantum numbers are consistent with a qqq structure. The associated representations have zero triality, elements belonging to the centre $\mathcal{Z}(SU(3))$ act trivially, or equivalently the observed representations correspond to the group $SU(3)/\mathbb{Z}_3$.

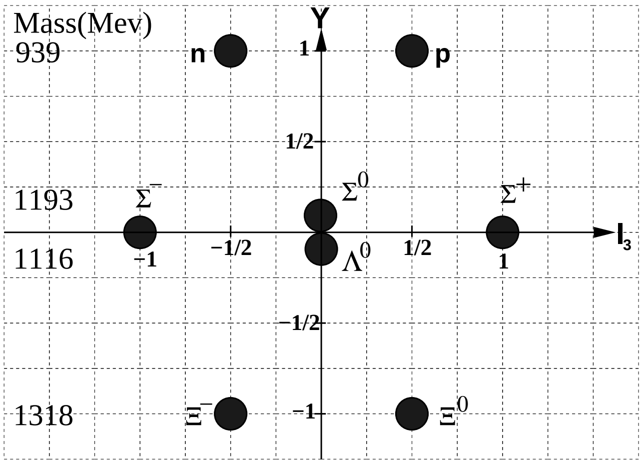
For the mesons we have self-conjugate octets belonging to the $[1,1]$, or $\mathbf{8}$, $SU(3)$ representations. The weight diagram for the lightest spin-0 negative parity mesons is



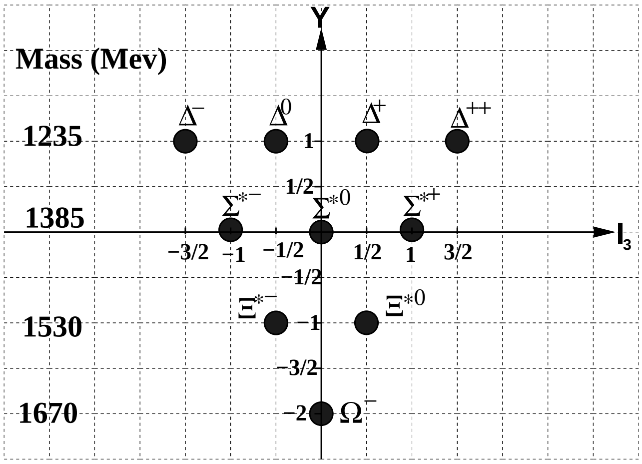
Here the kaons K^+, K^0 and $\bar{K}^0, K^-$ are $I = \frac{1}{2}$ strange particles with $S = 1$ and $S = -1$. A similar pattern emerges for the next lightest spin one negative parity mesons.



The lightest multiplet of spin-$\frac{1}{2}$ baryons is also an octet, with a similar weight diagram, the same set of I_3, Y although of course different particle assignments.



The novelty for baryons is that there are also decuplets, corresponding to the $[3,0]$ and $[0,3]$ representations, or labelled by their dimensionality $\mathbf{10}$ and $\mathbf{10}^*$. The next lightest spin-$\frac{3}{2}$ baryons and their anti-particles belong to decuplets.



Except for the Ω^- the particles in the decuplet are resonances, found as peaks in the invariant mass distribution for various cross sections. Since $m_\Xi + m_K > m_{\Omega^-}$ the Ω^- can decay only via weak interactions and its lifetime is long enough to leave an observable track.

6.5.1 $SU(3)_F$ Symmetry Breaking

Assuming quark masses are not equal there are no exact flavour symmetries in strong interactions, or equivalently QCD, save for a $U(1)$ for each quark. Even isospin symmetry is not exact since $m_u \neq m_d$. Restricting to the three light $q = (u, d, s)$ quarks the relevant QCD mass term may be written as

$$\begin{aligned}\mathcal{L}_m =& - m_u\, \bar{u}u - m_d\, \bar{d}d - m_s\, \bar{s}s \\ =& - \bar{m}\, \bar{q}q - \tfrac{1}{2}(m_u - m_d)\, \bar{q}\lambda_3 q - \tfrac{1}{2\sqrt{3}}(m_u + m_d - 2m_s)\, \bar{q}\lambda_8 q\,, \end{aligned} \tag{6.115}$$

for $\bar{m} = \frac{1}{3}(m_u + m_d + m_s)$. If the difference between m_u, m_d is neglected then the strong interaction Hamiltonian must be of the form

$$H = H_0 + T_8\,, \tag{6.116}$$

where H_0 is a $SU(3)$ singlet and T_8 is part of an octet of operators $\{T_a\}$ so that, with the $SU(3)$ operators $\{F_a\}$ as in (6.114), we have the commutation relations $[F_a, H_0] = 0$ and $[F_a, T_b] = if_{abc}T_c$. The Hamiltonian in (6.116) is invariant under isospin symmetry since $[I_i, T_8] = 0$.

In any $SU(3)$ multiplet the particle states may be labelled $|II_3, Y\rangle$ for various isospins I and hypercharges Y, depending on the particular representation. For $I_3 = -I, -I+1, \ldots, I$ the vectors $|II_3, Y\rangle$ form a standard basis under $SU(2)_I$. With isospin symmetry the particle masses are independent of I_3 and to first order in $SU(3)$ symmetry breaking

$$m_{I,Y} = m_0 + \langle II_3, Y|T_8|II_3, Y\rangle\,. \tag{6.117}$$

It remains to determine a general expression for $\langle II_3, Y|T_8|II_3, Y\rangle$, which is essentially equivalent to finding the extension of the Wigner-Eckart theorem, described in section 2.9, to $SU(3)$.

Instead of finding results for $SU(3)$ Clebsch-Gordan coefficients the necessary calculation may be accomplished, in this particular case, with less effort. It is necessary to recognise that the crux of the Wigner-Eckart theorem is that, as far as the I, Y dependence is concerned, $\langle II_3, Y|T_8|II_3, Y\rangle$ is determined just by the $SU(3)$ transformation properties of T_8. Hence, apart from overall undetermined constants, T_8 may be replaced by any other operator with the same transformation properties. For convenience we revert to a tensor basis for the octet $T_a \to T^i{}_j$, $T^i{}_i = 0$, and then with $F_a \to \hat{R}^i{}_j$ as in (6.33),

$$\big[\hat{R}^i{}_j, T^k{}_l\big] = \delta^k{}_j\, T^i{}_l - \delta^i{}_l\, T^k{}_j\,, \qquad T_8 = \tfrac{1}{3}\big(T^1{}_1 + T^2{}_2 - 2T^3{}_3\big)\,. \tag{6.118}$$

This ensures that $T^i{}_j$ is a traceless $(1,1)$ irreducible tensor operator. Any such tensor operator constructed in terms of $\hat{R}^i{}_j$ has the same $SU(3)$ transformation properties. The simplest case is if $T^i{}_j = \hat{R}^i{}_j$ when (6.118) requires

$$T_8 = \tfrac{1}{3}\big(H_1 + 2H_2\big) = Y\,, \tag{6.119}$$

with Y the hypercharge operator. An further independent $(1,1)$ operator is also given by the quadratic expression $T^i{}_j = \frac{1}{2}(\hat{R}^i{}_k\hat{R}^k{}_j + \hat{R}^k{}_j\hat{R}^i{}_k) - \frac{1}{3}\delta^i{}_j\, \hat{R}^k{}_l\hat{R}^l{}_k$ which then leads to

$$T_8 = \tfrac{1}{4}\big(\hat{R}^1{}_k\hat{R}^k{}_1 + \hat{R}^k{}_1\hat{R}^1{}_k + \hat{R}^2{}_k\hat{R}^k{}_2 + \hat{R}^k{}_2\hat{R}^2{}_k - \hat{R}^3{}_k\hat{R}^k{}_3 - \hat{R}^k{}_3\hat{R}^3{}_k\big) - \tfrac{1}{6}\,C\,, \tag{6.120}$$

where C is the $SU(3)$ Casimir operator defined in (6.92). Using (6.33) then

$$\begin{aligned} T_8 &= \tfrac{1}{2}\big(E_{1+}E_{1-} + E_{1-}E_{1+} + \tfrac{1}{2}H_1{}^2\big) - \tfrac{1}{36}\big(H_1 + 2H_2\big)^2 - \tfrac{1}{6}\,C \\ &= I_iI_i - \tfrac{1}{4}\,Y^2 - \tfrac{1}{6}\,C\,, \end{aligned} \tag{6.121}$$

with I_i the isospin operators and (6.25) has been used for the $SU(2)_I$ Casimir operator. For a 3×3 traceless matrix R, $R^3 - \frac{1}{3}I\,\mathrm{tr}(R^3) = \frac{1}{2}R\,\mathrm{tr}(R^2)$ so that there are no further independent cubic, or higher order, traceless $(1,1)$ tensor operators formed from $\hat{R}^i{}_j$.

The results of the Wigner-Eckart theorem imply that, to calculate $\langle II_3, Y|T_8|II_3, Y\rangle$, it is sufficient to replace T_8 by an arbitrary linear combination of (6.119) and (6.121). Absorbing an I, Y independent constant into m_0 and replacing the operators I_iI_i and Y by their eigenvalues this gives the first order mass formula

$$m_{IY} = m_0 + aY + b\big(I(I+1) - \tfrac{1}{4}Y^2\big)\,, \tag{6.122}$$

with a, b undetermined coefficients.

For the baryon octet (6.122) gives $2(m_N + m_\Xi) = 3m_\Sigma + m_\Lambda$, which is quite accurate. For the decuplet the second term is proportional to the first so that the masses are linear in Y, again in accord with experimental data. For mesons, for various reasons, the mass formula is applied to m^2, so that $4m_K^2 = 3m_\pi^2 + m_\eta^2$.

6.5.2 $SU(3)$ and Colour

The group $SU(3)$ plays a more fundamental role, other than a flavour symmetry group, as the gauge symmetry group of QCD. Each quark then belongs to the three dimensional fundamental, **3** or $[1,0]$, representation space for $SU(3)_{\rm colour}$ so that there is an additional colour index $r=1,2,3$ and hence, for each of the six different flavours of quarks $q = u,d,s,c,b,t$ in the standard model, we have q^r. The antiquarks belong to the conjugate, $\mathbf{3}^*$ or $[0,1]$, representation space, $\bar q_r$. The crucial assumption, yet to be fully demonstrated, is that QCD is a confining theory, the states in the physical quantum mechanical space are all colour singlets. No isolated quarks are then possible and this matches with the observed mesons and baryons since the simplest colour singlets are just

$$\bar q_{1r} q_2{}^r \, , \qquad \varepsilon_{rst}\, q_1{}^r q_2{}^s q_3{}^t \, . \tag{6.123}$$

Baryons are therefore totally antisymmetric in the colour indices. Fermi statistics then requires that they should be symmetric under interchange with respect to all other variables, spatial, spin and flavour. This provides non trivial constraints on the baryon spectrum which match with experiment. The additional colour degrees of freedom also play a role in various dynamical calculations, such as the total cross section for e^-e^+ scattering or $\pi^0 \to \gamma\gamma$ decay.

6.6 Tensor Products for $SU(3)$

Just as for angular momentum it is essential to be able to decompose tensor products of $SU(3)$ representations into irreducible components in applications of $SU(3)$ symmetry. Only states belonging to the same irreducible representation will have the same physical properties, except for dynamical accidents or a hidden addition symmetry.

For small dimensional representations it is simple to use the tensor formalism described in section 6.4 with irreducible representations characterised by symmetric traceless tensors as in (6.108). Thus for the product of two fundamental representations it is sufficient to express it in terms of its symmetric and antisymmetric parts

$$q_1{}^i\, q_2{}^j = S^{ij} + \varepsilon^{ijk} \bar q_k \, , \qquad S^{ij} = q_1{}^{(i} q_2{}^{j)} \, , \quad \bar q_k = \tfrac{1}{2}\varepsilon_{kij} q_1{}^i q_2{}^j \, . \tag{6.124}$$

while for the product of the fundamental and its conjugate it is only necessary to separate out the trace

$$\bar q_i q^j = M_i^j + \delta_i^j S \, , \qquad M_i^j = \bar q_i q^j - \tfrac{1}{3}\delta_i^j \, \bar q_k q^k \, , \quad S = \tfrac{1}{3}\, \bar q_i q^i \, . \tag{6.125}$$

These correspond respectively to

$$\mathbf{3}\otimes\mathbf{3} = \mathbf{6}\oplus\mathbf{3}^* \, , \qquad \mathbf{3}^*\otimes\mathbf{3} = \mathbf{8}\oplus\mathbf{1} \, . \tag{6.126}$$

For the product of three fundamental representations then the decomposition may be expressed in terms of an irreducible $(3,0)$ tensor, two independent $(1,1)$ tenors and a singlet

$$\begin{aligned} q_1{}^i\, q_2{}^j\, q_3{}^k &= D^{ijk} + \varepsilon^{ikl} B_l^j + \varepsilon^{jkl} B_l^i + \varepsilon^{ijl} B'^k_l + \varepsilon^{ijk} S \, , \\ D^{ijk} &= q_1{}^{(i}\, q_2{}^j\, q_3{}^{k)} \, , \qquad\qquad S = \tfrac{1}{6}\varepsilon_{ijk}\, q_1{}^i\, q_2{}^j\, q_3{}^k \, , \\ B_l^i &= \tfrac{1}{3}\varepsilon_{jkl}\, q_1{}^{(i}\, q_2{}^{j)}\, q_3{}^k \, , \quad B'^k_l = \tfrac{1}{2}\varepsilon_{ijl}\, q_1{}^i\, q_2{}^j\, q_3{}^k - \delta_l^k\, S \, . \end{aligned} \tag{6.127}$$

To verify that this is complete it is necessary to recognise, since the indices take only three values, that

$$\varepsilon^{ijl} B_l^k + \varepsilon^{kil} B_l^j + \varepsilon^{jkl} B_l^i = \varepsilon^{ijk} B_l^l = 0 \, , \tag{6.128}$$

for any B^i_j belonging to the **8** representation. (6.127) then corresponds to

$$\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} = \mathbf{10}\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{1}\,. \qquad (6.129)$$

These of course are the baryon representations for $SU(3)_F$.

In general it is only necessary to use the invariant tensors in (6.107) to reduce the tensor products to irreducible tensors. Thus for the product of two octets the irreducible tensors are constructed by forming first the symmetric $(2,2)$, $(3,0)$, $(0,3)$ tensors as well as two $(1,1)$ tensors and also a singlet by

$$B^i_j\,B'^k_l \to B^{(i}_{(j}B'^{k)}_{l)}\,,\ \varepsilon^{jl(m}B^i_jB'^{k)}_l\,,\ \ \varepsilon_{ik(m}B^i_jB'^k_{l)}\,,\ \ B^i_j\,B'^j_l\,,\ \ B^i_j\,B'^j_i\,,\ \ B^i_j\,B'^j_i\,. \qquad (6.130)$$

and then subtracting the required terms to cancel all traces formed by contracting upper and lower indices, as in (6.125). This gives the decomposition

$$\mathbf{8}\otimes\mathbf{8} = \mathbf{27}\oplus\mathbf{10}\oplus\mathbf{10}^*\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{1}\,. \qquad (6.131)$$

6.6.1 Systematic Discussion of Tensor Products

For tensor products of arbitrary representations there is a general procedure which is quite simple to apply in practice. The derivation of this is straightforward using characters to find an algorithm for the expansion of the product of two characters for highest weight irreducible representations as in (1.47). For $\mathfrak{su}(3)$, characters are given by (6.82). In general these have an expansion in terms of a sum over the weights in the associated weight diagram

$$\chi_{\underline{\Lambda}}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda},\underline{\lambda}}\, u_1{}^{r_1+r_2+2}u_2{}^{r_2+1}\,, \qquad \underline{\Lambda} = [n_1,n_2]\,, \quad \underline{\lambda} = [r_1,r_2]\,, \qquad (6.132)$$

where $n_{\underline{\Lambda},\underline{\lambda}}$ is then the multiplicity in the representation space $\mathcal{V}_{\underline{\Lambda}}$ for vectors with weight $\underline{\lambda}$. Due to the symmetry of the weight diagram under the Weyl group we have

$$n_{\underline{\Lambda},\underline{\lambda}} = n_{\underline{\Lambda},\sigma\underline{\lambda}}\,. \qquad (6.133)$$

Using (6.81) it is easy to see that

$$C_{\underline{\Lambda}}(u)\,\chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\,C_{\underline{\Lambda}+\underline{\lambda}}(u)\,, \qquad (6.134)$$

and since, for the weights $\{\underline{\lambda}\}$ corresponding to the representation with highest weight $\underline{\Lambda}$,

$$\{\underline{\lambda}\} = \{\sigma\underline{\lambda}\}\,, \qquad (\underline{\Lambda}+\underline{\lambda})^\sigma = \underline{\Lambda}^\sigma + \sigma\underline{\lambda}\,, \qquad (6.135)$$

then, with (6.133), we may use (6.87) to obtain

$$\chi_{\underline{\Lambda}}(u)\,\chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\,\chi_{\underline{\Lambda}+\underline{\lambda}}(u)\,. \qquad (6.136)$$

However in general $\underline{\Lambda} + \underline{\lambda} \notin \mathcal{W}$, as defined in (6.59). In this case (6.89) may be used to rewrite (6.136) as

$$\chi_{\underline{\Lambda}}(u)\, \chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\,\underline{\lambda}}\, P_\sigma\, \chi_{(\underline{\Lambda}+\underline{\lambda})^\sigma}(u)\,, \qquad (\underline{\Lambda} + \underline{\lambda})^\sigma \in \mathcal{W}\,, \tag{6.137}$$

dropping all terms where $\underline{\Lambda}+\underline{\lambda}$ satisfies any of the conditions in (6.90) ensuring $\chi_{\underline{\Lambda}+\underline{\lambda}}(u) = 0$, so that, by virtue of (6.91), σ in (6.137) is then unique. Since in (6.137) some terms may now contribute with a negative sign there are then cancellations although the final result is still a positive sum of characters.

The result (6.137) may be re-expressed in terms of the associated representation spaces. For a highest weight $\underline{\Lambda}$ the representation space $\mathcal{V}_{\underline{\Lambda}}$ has a decomposition into subspaces for each weight,

$$\mathcal{V}_{\underline{\Lambda}} = \bigoplus_{\underline{\lambda}} \mathcal{V}_{\underline{\Lambda}}^{(\underline{\lambda})}\,, \qquad \dim \mathcal{V}_{\underline{\Lambda}}^{(\underline{\lambda})} = n_{\underline{\Lambda},\,\underline{\lambda}}\,, \tag{6.138}$$

and then (6.137) is equivalent to

$$\mathcal{V}_{\underline{\Lambda}} \otimes \mathcal{V}_{\underline{\Lambda}'} \simeq \bigoplus_{\underline{\lambda}} n_{\underline{\Lambda}',\,\underline{\lambda}}\, P_\sigma\, \mathcal{V}_{(\underline{\Lambda}+\underline{\lambda})^\sigma}\,, \qquad (\underline{\Lambda} + \underline{\lambda})^\sigma \in \mathcal{W}\,. \tag{6.139}$$

This implies the corresponding decomposition for the associated representations.

As applications we may consider tensor products involving $\mathcal{V}_{[1,0]}$ which has the weight decomposition

$$\mathcal{V}_{[1,0]} \to [1,0]\,,\ [-1,1]\,,\ [0,-1]\,, \tag{6.140}$$

and then

$$\begin{aligned} \mathcal{V}_{[n_1,n_2]} \otimes \mathcal{V}_{[1,0]} &\simeq \mathcal{V}_{[n_1+1,n_2]} \oplus \mathcal{V}_{[n_1-1,n_2+1]} \oplus \mathcal{V}_{[n_1,n_2-1]} \\ &= \begin{cases} \mathcal{V}_{[1,n_2]} \oplus \mathcal{V}_{[0,n_2-1]}\,, & n_1 = 0\,, \\ \mathcal{V}_{[n_1+1,0]} \oplus \mathcal{V}_{[n_1-1,1]}\,, & n_2 = 0\,. \end{cases} \end{aligned} \tag{6.141}$$

It is easy to see that this is in accord with the results in (6.129). For an octet

$$\mathcal{V}_{[1,1]} \to [1,1]\,,\ [2,-1]\,,\ [-1,2]\,,\ [0,0]^2\,,\ [1,-2]\,,\ [-2,1]\,,\ [-1,-1]\,, \tag{6.142}$$

so that, for $n_1, n_2 \geq 2$,

$$\begin{aligned} \mathcal{V}_{[n_1,n_2]} \otimes \mathcal{V}_{[1,1]} \simeq{}& \mathcal{V}_{[n_1+1,n_2+1]} \oplus \mathcal{V}_{[n_1+2,n_2-1]} \oplus \mathcal{V}_{[n_1-1,n_2+2]} \oplus \mathcal{V}_{[n_1,n_2]} \\ & \oplus \mathcal{V}_{[n_1,n_2]} \oplus \mathcal{V}_{[n_1+1,n_2-2]} \oplus \mathcal{V}_{[n_1-2,n_2+1]} \oplus \mathcal{V}_{[n_1-1,n_2-1]}\,, \end{aligned} \tag{6.143}$$

with special cases

$$\mathcal{V}_{[1,1]} \otimes \mathcal{V}_{[1,1]} \simeq \mathcal{V}_{[2,2]} \oplus \mathcal{V}_{[3,0]} \oplus \mathcal{V}_{[0,3]} \oplus \mathcal{V}_{[1,1]} \oplus \mathcal{V}_{[1,1]} \oplus \mathcal{V}_{[0,0]}\,, \tag{6.144}$$

which is in accord with (6.131), and

$$\mathcal{V}_{[3,0]} \otimes \mathcal{V}_{[1,1]} \simeq \mathcal{V}_{[4,1]} \oplus \mathcal{V}_{[2,2]} \oplus \mathcal{V}_{[3,0]} \oplus \mathcal{V}_{[1,1]}\,, \tag{6.145}$$

using $\mathcal{V}_{[4,-2]} \simeq -\mathcal{V}_{[3,0]}$. Equivalently, labelling the representations by their dimensions

$$\mathbf{10} \otimes \mathbf{8} = \mathbf{35} \oplus \mathbf{27} \oplus \mathbf{10} \oplus \mathbf{8}\,. \tag{6.146}$$

7 Gauge Groups and Gauge Theories

Gauge theories are fundamental to our understanding of theoretical physics, many successful theories such as superconductivity and general relativity are best understood in terms of an appropriate gauge symmetry and its implementation. High energy particle physics is based on quantum gauge field theories. A *gauge theory* is essentially one where there are redundant degrees of freedom, which cannot in general be eliminated, at least without violating other symmetries that are present. The presence of such superfluous degrees of freedom requires a careful treatment when gauge theories are quantised and a quantum vector space for physical states is constructed. If the basic variables in a gauge theory are denoted by q then gauge transformations $q \to q^g$, for $g \in G$ for some group G, are dynamical symmetries which define an equivalence $q \sim q^g$. The objects of interest are then functions of q which are invariant under G, in a physical theory these are the physical observables. For a solution $q(t)$ of the dynamical equations of motion then a gauge symmetry requires that $q^{g(t)}(t)$ is also a solution for arbitrary continuously differentiable $g(t) \in G_t \simeq G$. For this to be feasible G must be a Lie group, group multiplication is defined by $g(t)g'(t) = gg'(t)$ and the full group of gauge transformations is then essentially $\mathcal{G} \simeq \otimes_t G_t$. A gauge theory in general requires the introduction of additional dynamical variables which form a connection, depending on t, on $\mathcal{M}_G$ and so belongs to the Lie algebra $\mathfrak{g}$.

For a relativistic gauge field theory there are vector gauge fields, with a Lorentz index $A_\mu(x)$, belonging to $\mathfrak{g}$. Denoting the set of all vector fields, functions of x and taking values in $\mathfrak{g}$, by $\mathcal{A}$, we can then write

$$A_\mu \in \mathcal{A} \,. \tag{7.1}$$

In a formal sense, the gauge group $\mathcal{G}$ is defined by

$$\mathcal{G} \simeq \bigotimes_x G_x \,, \tag{7.2}$$

i.e. an element of $\mathcal{G}$ is a map from space-time points to elements of the Lie group G (the definition of $\mathcal{G}$ becomes precise when space-time is approximated by a lattice). Gauge transformations act on the gauge fields so that

$$A_\mu(x) \underset{g(x)}{\longrightarrow} A_\mu{}^{g(x)}(x) \sim A_\mu(x) \,. \tag{7.3}$$

Gauge transformations $g(x)$ are then the redundant variables and the physical space is determined by the equivalence classes of gauge fields modulo gauge transformations or

$$\mathcal{A}/\mathcal{G} \,. \tag{7.4}$$

If $A_\mu(x)$ is subject to suitable boundary conditions as $|x| \to \infty$, or we restrict $x \in \mathcal{M}$ for some compact $\mathcal{M}$, then this is topologically non trivial.

The most significant examples of quantum gauge field theories are[21],

Theory:	QED	Weinberg-Salam model	QCD,
Gauge Group:	$U(1)$	$SU(2) \otimes U(1)$	$SU(3)$.

[21]Steven Weinberg, (1933-), American. Abdus Salam, (1926-1996), Pakistani. Nobel Prizes 1979.

Renormalisable gauge field theories are almost uniquely determined by specifying the gauge group and then the representation content of any additional fields.

7.1 Abelian Gauge Theories

The simplest example arises for $G = U(1)$, which is the gauge group for Maxwell[22] electromagnetism, although the relevant gauge symmetry was only appreciated by the 1920's and later. For $U(1)$ the group elements are complex numbers of modulus one, so they can be expressed as $e^{i\alpha}$, $0 \le \alpha < 2\pi$. For a gauge theory the group transformations depend on x so we can then write $e^{i\alpha(x)}$. The representations of $U(1)$ are specified by $q \in \mathbb{R}$, physically the charge, so that for a complex field $\phi(x)$ the group transformations are

$$\phi \underset{e^{i\alpha}}{\longrightarrow} e^{iq\alpha}\phi = \phi' \, . \tag{7.5}$$

If the field ϕ forms a non projective representation we must have

$$q \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\} \, . \tag{7.6}$$

In quantum mechanics this is not necessary but if the $U(1)$ is embedded in a semi-simple Lie group then, with a suitable convention, q can be chosen to satisfy (7.6). For $U(1)$ the multiplication of representations is trivial, the charges just add, and also under complex conjugation $q \to -q$. It is then easy to construct lagrangians $\mathcal{L}_\phi$ which are invariant under (7.5) for *global transformations*, where α is independent of x. Restricting to first derivatives this requires

$$\mathcal{L}_\phi(\phi, \partial_\mu \phi) = \mathcal{L}_\phi(\phi', \partial_\mu \phi') \, , \tag{7.7}$$

and an obvious solution, which defines a Lorentz invariant theory for complex scalars ϕ, is then

$$\mathcal{L}_\phi(\phi, \partial_\mu \phi) = \partial^\mu \phi^\star \phi_\mu \phi - V(\phi^* \phi) \, . \tag{7.8}$$

For *local transformations*, when the elements of the gauge group depend on x, the initial lagrangian is no longer invariant due to the presence of derivatives since

$$\partial_\mu \phi' = e^{iq\alpha}\big(\partial_\mu \phi + iq\partial_\mu \alpha \, \phi\big) \, , \tag{7.9}$$

and the $\partial_\mu \alpha$ terms fail to cancel. This is remedied by introducing a connection, or gauge field, A_μ and then defining a covariant derivative on ϕ by

$$D_\mu \phi = \partial_\mu \phi - iqA_\mu \phi \, . \tag{7.10}$$

If under a local $U(1)$ gauge transformation, as in (7.5), the gauge field transforms as

$$A_\mu \underset{e^{i\alpha}}{\longrightarrow} A_\mu + \partial_\mu \alpha = A'_\mu \, , \tag{7.11}$$

so that

$$D'_\mu \phi' = e^{i\alpha} D_\mu \phi \, , \tag{7.12}$$

[22]James Clerk Maxwell, 1831-79, Scottish, second wrangler 1854.

and then it is easy to see that, for any globally invariant lagrangian satisfying (7.7),

$$\mathcal{L}_\phi\big(\phi, D_\mu\phi\big) = \mathcal{L}_\phi\big(\phi', D'_\mu\phi'\big)\,. \tag{7.13}$$

It is important to note that for abelian gauge theories $A_\mu \sim A'_\mu$, which corresponds precisely to the freedom of polarisation vectors in (4.197) when Lorentz vector fields are used for massless particles with helicities ± 1.

The initial scalar field theory then includes the gauge field A_μ, as well as the scalar fields ϕ, both gauge dependent. For well defined dynamics the scalar lagrangian $\mathcal{L}_\phi$ must be extended to include an additional gauge invariant kinetic term for A_μ. In the abelian case it is easy to see that the curvature

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = F'_{\mu\nu}\,, \tag{7.14}$$

is gauge invariant, since $\partial_\mu\partial_\nu\alpha = \partial_\nu\partial_\mu\alpha$. In electromagnetism $F_{\mu\nu}$ decomposes in to the electric and magnetic fields and is related to the commutator of two covariant derivatives since

$$[D_\mu, D_\nu]\phi = -iqF_{\mu\nu}\phi\,. \tag{7.15}$$

The simplest Lorentz invariant, gauge invariant, lagrangian is then

$$\mathcal{L} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_\phi\big(\phi, D_\mu\phi\big)\,, \qquad \mathcal{L}_{\text{gauge}} = -\frac{1}{4e^2}\, F^{\mu\nu}F_{\mu\nu}\,, \tag{7.16}$$

with e an arbitrary parameter, unimportant classically. It is commonplace to rescale the fields so that

$$A_\mu \to eA_\mu\,, \qquad D_\mu\phi = \partial_\mu\phi - ieqA_\mu\phi\,, \tag{7.17}$$

so that e disappears from the gauge field term in (7.16). The dynamical equations of motion which flow from (7.16) are, for the gauge field,

$$\frac{1}{e^2}\,\partial^\mu F_{\mu\nu} = j_\nu = -\frac{\partial}{\partial A^\nu}\mathcal{L}_\phi\big(\phi, D_\mu\phi\big)\,, \tag{7.18}$$

which are of course Maxwell's equations for an electric current j_ν and e becomes the basic unit of electric charge. A necessary consistency condition is that the current is conserved $\partial^\nu j_\nu = 0$. In addition $F_{\mu\nu}$ satisfies an identity, essentially the Bianchi identity, which follows directly from its definition in (7.14),

$$\partial_\omega F_{\mu\nu} + \partial_\nu F_{\omega\mu} + \partial_\mu F_{\nu\omega} = 0\,. \tag{7.19}$$

In the language of forms, $A = A_\mu \mathrm{d}x^\mu$, $F = \frac{1}{2}F_{\mu\nu}\,\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = \mathrm{d}A$, this is equivalent to $\mathrm{d}F = \mathrm{d}^2 A = 0$.

7.2 Non Abelian Gauge Theories

In retrospect the generalisation of gauge theories to non abelian Lie groups is a natural step. A fully consistent non abelian gauge theory was first described in 1954, for the group

$SU(2)$, by Yang and Mills[23] so they are often referred to, for the particular gauge invariant lagrangian generalising the abelian lagrangian given in (7.16) and obtained below, as Yang-Mills theories. Nevertheless the same theory was also developed, but not published, by R. Shaw[24] (it appeared as an appendix in his Cambridge PhD thesis submitted in 1955 although this work was done in early 1954). Such theories were not appreciated at first since they appeared to contain unphysical massless particles, and also since understanding their quantisation was not immediate.

Following the same discussion as in the abelian case we first consider fields ϕ belonging to the representation space $\mathcal{V}$ for a Lie group G. Under a local group transformation then

$$\phi(x) \underset{g(x)}{\longrightarrow} g(x)\phi(x) = \phi'(x) \,, \tag{7.20}$$

for $g(x) \in \mathcal{R}$ for $\mathcal{R}$ an appropriate representation, acting on $\mathcal{V}$, of G. Manifestly derivatives fail to transform in the same simple homogeneous fashion since

$$\partial_\mu\phi(x) \underset{g(x)}{\longrightarrow} g(x)\big(\partial_\mu\phi(x) + g(x)^{-1}\partial_\mu g(x)\,\phi(x)\big) = \partial_\mu\phi'(x) \,, \tag{7.21}$$

where $g^{-1}\partial_\mu g$ belongs to the corresponding representation of the Lie algebra of G, $\mathfrak{g}$, which is assumed to have a basis $\{t_a\}$ satisfying the Lie algebra (5.60). As before to define a covariantly transforming derivative D_μ it is necessary to introduce a connection belonging to this Lie algebra representation which may be expanded over the basis matrices t_a,

$$A_\mu(x) = A^a{}_\mu(x)\, t_a \,, \tag{7.22}$$

and then

$$D_\mu\phi = (\partial_\mu + A_\mu)\phi \,. \tag{7.23}$$

Requiring

$$D'_\mu\phi' = gD_\mu\phi \,, \tag{7.24}$$

or

$$g^{-1}A'_\mu\, g + g^{-1}\partial_\mu g = A_\mu \,, \tag{7.25}$$

then the gauge field must transform under a gauge transformation as

$$A_\mu \underset{g}{\longrightarrow} A'_\mu = gA_\mu\, g^{-1} - \partial_\mu g\, g^{-1} = gA_\mu\, g^{-1} + g\partial_\mu g^{-1} \,. \tag{7.26}$$

Hence if $\mathcal{L}_\phi(\phi, \partial_\mu\phi)$ is invariant under global transformations $\phi \to g\phi$ then $\mathcal{L}_\phi(\phi, D_\mu\phi)$ is invariant under the corresponding local transformations, so long as A_μ also transforms as in (7.26).

It is also useful to note, since G is a Lie group, the associated infinitesimal transformations when

$$g = I + \lambda \,, \qquad \lambda = \lambda^a t_a \,. \tag{7.27}$$

[23]Chen-Ning Franklin Yang, 1922-, Chinese then American, Nobel prize 1957. Robert L. Mills, 1927-99, American.

[24]Ron Shaw, 1929-, English.

Then from (7.20) and (7.24), for arbitrary $\lambda^a(x)$,

$$\delta\phi = \lambda\phi\,, \qquad \delta D_\mu\phi = \lambda D_\mu\phi\,, \tag{7.28}$$

and from (7.26)

$$\delta A_\mu = [\lambda, A_\mu] - \partial_\mu\lambda \qquad \Rightarrow \qquad \delta A^a{}_\mu = -f^a{}_{bc}A^b{}_\mu\lambda^c - \partial_\mu\lambda^a\,. \tag{7.29}$$

The associated curvature is obtained from the commutator of two covariant derivatives, as in the abelian case in (7.15), which gives

$$[D_\mu, D_\nu]\phi = F_{\mu\nu}\phi\,, \qquad F_{\mu\nu} = F^a{}_{\mu\nu}t_a\,, \tag{7.30}$$

so that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\,, \tag{7.31}$$

or

$$F^a{}_{\mu\nu} = \partial_\mu A^a{}_\nu - \partial_\nu A^a{}_\mu + f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\,. \tag{7.32}$$

Unlike the abelian case, but more akin to general relativity, the curvature is no longer linear. The same result is expressible more elegantly using differential form notation by

$$F = \mathrm{d}A + A\wedge A\,, \qquad A = A_\mu\,\mathrm{d}x^\mu\,, \quad A\wedge A = \tfrac{1}{2}[A_\mu, A_\nu]\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\,. \tag{7.33}$$

For a gauge transformation as in (7.26)

$$F_{\mu\nu} \underset{g}{\longrightarrow} F'_{\mu\nu} = gF_{\mu\nu}\,g^{-1}\,, \tag{7.34}$$

or, infinitesimally,

$$\delta F_{\mu\nu} = [\lambda, F_{\mu\nu}] \qquad \Rightarrow \qquad \delta F^a{}_{\mu\nu} = -f^a{}_{bc}F^b{}_{\mu\nu}\lambda^c\,, \tag{7.35}$$

which are homogeneous.

As a consistency check we verify the result (7.35) for $\delta F^a{}_{\mu\nu}$ from the expression (7.32) using (7.29) for $\delta A^a{}_\mu$. First

$$\delta\big(\partial_\mu A^a{}_\nu - \partial_\nu A^a{}_\mu\big) = -f^a{}_{bc}\big(\partial_\mu A^b{}_\nu - \partial_\nu A^b{}_\mu\big)\lambda^c - f^a{}_{bc}\big(A^b{}_\nu\partial_\mu\lambda^c - A^b{}_\mu\partial_\nu\lambda^c\big)\,. \tag{7.36}$$

Then

$$\delta\big(f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\big)\big|_{\partial\lambda} = -f^a{}_{bc}\big(\partial_\mu\lambda^b A^c{}_\nu + A^b{}_\mu\partial_\nu\lambda^c\big)\,, \tag{7.37}$$

which cancels, using (5.39), the $\partial\lambda$ terms in (7.36). Furthermore

$$\begin{aligned}\delta\big(f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\big)\big|_\lambda &= -\,f^a{}_{bc}\big(f^b{}_{de}A^d{}_\mu\lambda^e\,A^c{}_\nu + A^b{}_\mu\,f^c{}_{de}A^d{}_\nu\lambda^e\big)\\ &= -\,\big(f^a{}_{fd}f^f{}_{be} + f^a{}_{cf}f^f{}_{be}\big)A^b{}_\mu A^d{}_\nu\lambda^e = -f^a{}_{fe}\,f^f{}_{bd}A^b{}_\mu A^d{}_\nu\,\lambda^e\,,\end{aligned} \tag{7.38}$$

by virtue of the Jacobi identity in the form (5.43). Combining (7.36), (7.37) and (7.38) demonstrates (7.35) once more.

The gauge fields ${A^a}_\mu$ are associated with the adjoint representation of the gauge group G. For any adjoint field $\Phi^a t_a$ then the corresponding covariant derivative is given by

$$D_\mu \Phi = \partial_\mu \Phi + [A_\mu, \Phi] \qquad \Rightarrow \qquad (D_\mu \Phi)^a = \partial_\mu \Phi^a + {f^a}_{bc}\, {A^b}_\mu \Phi^c \,. \tag{7.39}$$

This is in accord with the general form given by (7.23), with (7.22), using (5.178) for the adjoint representation generators. Note that (7.29) can be written as $\delta {A^a}_\mu = -(D_\mu \lambda)^a$ and for an arbitrary variation $\delta {A^a}_\mu$ from (7.32),

$$\delta {F^a}_{\mu\nu} = (D_\mu \delta A_\nu)^a - (D_\nu \delta A_\mu)^a \,. \tag{7.40}$$

From the identity

$$\big([D_\omega, [D_\mu, D_\nu]] + [D_\nu, [D_\omega, D_\mu]] + [D_\mu, [D_\nu, D_\omega]]\big)\phi = 0 \,, \tag{7.41}$$

for any representation, we have the non abelian Bianchi identity, generalising (7.19),

$$D_\omega F_{\mu\nu} + D_\nu F_{\omega\mu} + D_\mu F_{\nu\omega} = 0 \,, \tag{7.42}$$

where the adjoint covariant derivatives are as defined in (7.39). Alternatively with the notation in (7.33)

$$\mathrm{d}F + A \wedge F - F \wedge A = 0 \,. \tag{7.43}$$

To construct a lagrangian leading to dynamical equations of motion which are covariant under gauge transformations it is necessary to introduce a group invariant metric $g_{ab} = g_{ba}$, satisfying (5.193) or equivalently

$$g_{db} {f^d}_{ca} + g_{ad} {f^d}_{cb} = 0 \,, \tag{7.44}$$

which also implies, for finite group transformations g and with X, Y belonging to the associated Lie algebra,

$$g_{ab}\, (gXg^{-1})^a\, (gYg^{-1})^b = g_{ab}\, X^a Y^b \,. \tag{7.45}$$

If X, Y are then adjoint representation fields the definition of the adjoint covariant derivative in (7.39) gives

$$\partial_\mu \big(g_{ab}\, X^a Y^b\big) = g_{ab} \big((D_\mu X)^a Y^b + X^a (D_\mu Y)^b\big) \,, \tag{7.46}$$

in a similar fashion to covariant derivatives in general relativity.

The simplest gauge invariant lagrangian, extending the abelian result in (7.16), is then, as a result of the transformation properties (7.34) or (7.35), just the obvious extension of that proposed by Yang and Mills for $SU(2)$

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4}\, g_{ab} F^{a\mu\nu} {F^b}_{\mu\nu} \,. \tag{7.47}$$

It is essential that the metric be non degenerate $\det[g_{ab}] \neq 0$, and then using (7.40) requiring the action to be stationary gives the gauge covariant dynamical equations

$$(D^\mu F_{\mu\nu})^a = 0 \,. \tag{7.48}$$

These equations, as well as (7.42) and unlike the abelian case, are non linear. As described before a necessary consequence of gauge invariance is that if A_μ is a solution then so is any gauge transform as in (7.26) and hence the time evolution of A_μ is arbitrary up to this extent, only gauge equivalence classes, belonging to (7.4), have a well defined dynamics. If the associated quantum field theory is to have a space of quantum states with positive norm then it is also necessary that the metric g_{ab} should be positive definite. This requires that the gauge group G should be compact and restricted to the form exhibited in (5.200). Each $U(1)$ factor corresponds to a simple abelian gauge theory as described in 7.1. If there are no $U(1)$ factors G is semi-simple and g_{ab} is determined by the Killing form for each simple group factor. For G simple then by a choice of basis we may take

$$g_{ab} = \frac{1}{g^2}\,\delta_{ab}\,, \tag{7.49}$$

with g the gauge coupling. For G a product of simple groups then there is a separate coupling for each simple factor, unless additional symmetries are imposed.

If the condition that the metric g_{ab} be positive definite is relaxed then the gauge group G may be non compact, but there are also examples of non semi-simple Lie algebras with a non-degenerate invariant metric. The simplest example is given by the Lie algebra $\mathfrak{iso}(2)$ with a central extension, which is given in (5.134). Choosing $T_a = (E_1, E_2, J_3, 1)$ then it is straightforward to verify that

$$[g_{ab}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & c \\ 0 & 0 & c & 0 \end{pmatrix}, \qquad \beta \text{ arbitrary}\,, \tag{7.50}$$

is invariant. The Killing form only involves the matrix with the element proportional to β non zero. Since it is necessary that $c \neq 0$ for the metric to be non-degenerate the presence of the central charge in the Lie algebra is essential. For any β it is easy, since $\det[g_{ab}] = -c^2$, to see that $[g_{ab}]$ has one negative eigenvalue.

An illustration of the application of identities such as (7.46) is given by the conservation of the gauge invariant energy momentum tensor defined by

$$T^{\mu\nu} = g_{ab}\big(F^{a\mu\sigma}F^{b\nu}{}_\sigma - \tfrac{1}{4}g^{\mu\nu}F^{a\sigma\rho}F^b{}_{\sigma\rho}\big)\,. \tag{7.51}$$

Then

$$\begin{aligned} \partial_\mu T^\mu{}_\nu &= g_{ab}\big((D_\mu F^{\mu\sigma})^a F^b{}_{\nu\sigma} + F^{a\mu\sigma}(D_\mu F_{\sigma\nu})^b - \tfrac{1}{2}F^{a\sigma\rho}(D_\nu F_{\sigma\rho})^b\big) \\ &= g_{ab}(D_\mu F^{\mu\sigma})^a - \tfrac{1}{2}g_{ab}F^{a\sigma\rho}\big((D_\rho F_{\nu\sigma})^b - (D_\sigma F_{\nu\rho})^b + (D_\nu F_{\sigma\rho})^b\big) \\ &= g_{ab}(D_\mu F^{\mu\sigma})^a\,, \end{aligned} \tag{7.52}$$

using the Bianchi identity (7.42). Clearly this is conserved subject to the dynamical equation (7.48).

7.2.1 Chern-Simons Theory

The standard gauge invariant lagrangian is provided by (7.47). However in order to obtain a gauge invariant action, given by the integral over space-time of the lagrangian, it is only necessary that the lagrangian is invariant up to a total derivative. This allows for additional possibility for gauge field theories, with gauge group G a general Lie group, in three space-time dimensions, termed Chern-Simons[25] theories.

First we note that in four dimensions the Bianchi identity (7.42) may be alternatively be written using the four dimensional antisymmetric symbol as

$$\varepsilon^{\mu\nu\sigma\rho} D_\nu F_{\sigma\rho} = 0 \, . \tag{7.53}$$

Apart from (7.47) there is then another similar gauge invariant and Lorentz invariant

$$\frac{1}{4}\, \varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} \, , \tag{7.54}$$

which may be used as an additional term in the lagrangian. However the corresponding contribution to the action is odd under $\mathbf{x} \to -\mathbf{x}$ or $t \to -t$. Such a term does not alter the dynamical equations since its variation is a total derivative and thus the variation of the corresponding term in the action vanishes. To show this under arbitrary variations of the gauge field we use (7.40) and (7.53) to give

$$\delta \frac{1}{4}\, \varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} = \varepsilon^{\mu\nu\sigma\rho} g_{ab} (D_\mu \delta A_\nu)^a F^b{}_{\sigma\rho} = \partial_\mu \big(\varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu \, F^b{}_{\sigma\rho} \big) \, . \tag{7.55}$$

This allows us to write

$$\frac{1}{4}\, \varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} = \partial_\mu \omega^\mu \, , \tag{7.56}$$

where

$$\omega^\mu = \varepsilon^{\mu\nu\sigma\rho} g_{ab} \big(A^a{}_\nu \partial_\sigma A^b{}_\rho + \tfrac{1}{3} f^b{}_{cd} A^a{}_\nu A^c{}_\sigma A^d{}_\rho \big) \, , \tag{7.57}$$

since this has the variation

$$\begin{aligned} \delta\omega^\mu &= \varepsilon^{\mu\nu\sigma\rho} g_{ab} \big(\delta A^a{}_\nu \partial_\sigma A^b{}_\rho + A^a{}_\nu \partial_\sigma \delta A^b{}_\rho + f^b{}_{cd}\, \delta A^a{}_\nu \, A^c{}_\sigma A^d{}_\rho \big) \\ &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \partial_\sigma \big(A^a{}_\nu \delta A^b{}_\rho \big) + \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu \big(2\partial_\sigma A^b{}_\rho + f^b{}_{cd} A^c{}_\sigma A^d{}_\rho \big) \\ &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \partial_\nu \big(\delta A^a{}_\sigma A^b{}_\rho \big) + \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu F^b{}_{\sigma\rho} \, , \end{aligned} \tag{7.58}$$

using that $g_{ab} f^b{}_{cd}$ is totally antisymmetric as a consequence of (7.44). The result is then in agreement with (7.55).

If the variation is a gauge transformation so that

$$A^a{}_\mu \underset{g}{\to} A'^a{}_\mu \qquad \Rightarrow \qquad \omega^\mu \underset{g}{\to} \omega'^\mu \, , \tag{7.59}$$

then since (7.56) is gauge invariant we must require

$$\partial_\mu \omega^\mu = \partial_\mu \omega'^\mu \, . \tag{7.60}$$

[25]Shiing-Shen Chern, 1911-2004, Chinese, American after 1960. James Harris Simons, 1938-, American.

This necessary condition may be verified for an infinitesimal gauge transformation by setting $\delta A^a{}_\nu = -(D_\nu\lambda)^a$ in (7.58) which then gives, using the Bianchi identity (7.53) again,

$$\begin{aligned}\delta\omega^\mu = & -\varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big((D_\sigma\lambda)^a A^b{}_\rho\big) - \varepsilon^{\mu\nu\sigma\rho} g_{ab}\,(D_\nu\lambda)^a F^b{}_{\sigma\rho} \\ = & \,\varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a (D_\sigma A)^b{}_\rho - \lambda^a F^b{}_{\sigma\rho}\big) \\ = & -\varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a \partial_\sigma A^b{}_\rho\big)\,. \end{aligned} \tag{7.61}$$

Hence it is evident that the result in (7.61) satisfies

$$\partial_\mu \delta\omega^\mu = 0\,. \tag{7.62}$$

In three dimensions the identities for ω^μ may be applied to

$$\mathcal{L}_{CS} = \varepsilon^{\nu\sigma\rho} g_{ab}\big(A^a{}_\nu \partial_\sigma A^b{}_\rho + \tfrac{1}{3} f^b{}_{cd} A^a{}_\nu A^c{}_\sigma A^d{}_\rho\big)\,, \tag{7.63}$$

which defines the Chern-Simons lagrangian for gauge fields. For an infinitesimal gauge transformation, by virtue of (7.61), $\mathcal{L}_{CS}$ becomes a total derivative since

$$\delta A^a{}_\nu = -(D_\nu\lambda)^a \quad \Rightarrow \quad \delta\mathcal{L}_{CS} = -\varepsilon^{\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a\partial_\sigma A^b{}_\rho\big)\,, \tag{7.64}$$

so that the corresponding action is invariant. Under a general variation

$$\delta \int \mathrm{d}^3x\ \mathcal{L}_{CS} = \int \mathrm{d}^3x\ \varepsilon^{\nu\sigma\rho} g_{ab}\,\delta A^a{}_\nu F^b{}_{\sigma\rho}\,, \tag{7.65}$$

so that the dynamical equations are

$$F^a{}_{\mu\nu} = 0\,, \tag{7.66}$$

so the connection A_μ is 'flat' since the associated curvature is zero (Cherns-Simons theory is thus similar to three dimensional pure gravity where the Einstein equations require that the Riemann curvature tensor vanishes). In a Chern-Simons theory there are no perturbative degrees of freedom, as in the case of Yang-Mills theory, but topological considerations play a crucial role.

Topology also becomes relevant as the Chern-Simons action is not necessarily invariant under all gauge transformations if they belong to topological classes which cannot be continuously connected to the identity. To discuss this further it is much more natural again to use the language of forms, expressing all results in terms of $A(x) = A^a{}_\mu(x) t_a\,\mathrm{d}x^\mu$ a Lie algebra matrix valued connection one-form, $[t_a, t_b] = f^c{}_{ab} t_c$ as in (5.60), and replacing the group invariant scalar product by the matrix trace. For any set of such Lie algebra matrices $\{X_1, \dots, X_n\}$ the trace $\mathrm{tr}(X_1 \dots X_n)$ is invariant under the action of adjoint group transformations $X_r \to g X_r g^{-1}$ for all r. Since the wedge product is associative and the trace is invariant under cyclic permutations we have

$$\begin{aligned}\mathrm{tr}\big(\underbrace{A\wedge\cdots\wedge A}_{n}\big) &= \mathrm{tr}\big((\underbrace{A\wedge\cdots\wedge A}_{n-1})\wedge A\big) = (-)^{n-1}\mathrm{tr}\big(A\wedge(\underbrace{A\wedge\cdots\wedge A}_{n-1})\big) \\ &= 0 \qquad \text{for } \ n \ \text{ even}\,. \end{aligned} \tag{7.67}$$

The Chern-Simons theory is then defined in terms of the three-form

$$\omega = \mathrm{tr}\big(A \wedge \mathrm{d}A + \tfrac{2}{3}\, A \wedge A \wedge A\big) = \mathrm{tr}\big(A \wedge F - \tfrac{1}{3}\, A \wedge A \wedge A\big)\,, \tag{7.68}$$

with the two-form curvature F as in (7.33). It is easy to see that

$$\mathrm{d}\omega = \mathrm{tr}\big(\mathrm{d}A \wedge \mathrm{d}A + 2\,\mathrm{d}A \wedge A \wedge A\big) = \mathrm{tr}\big(F \wedge F\big)\,, \tag{7.69}$$

which is equivalent to (7.56) and (7.57). For a finite gauge transformation

$$A' = gAg^{-1} + g\mathrm{d}g^{-1}\,, \qquad F' = gFg^{-1}\,, \tag{7.70}$$

so that, from (7.68),

$$\begin{aligned} \omega' = \omega + \mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge (F - A \wedge A)\big) - \mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g \wedge A\big) \\ - \tfrac{1}{3}\,\mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g\big)\,. \end{aligned} \tag{7.71}$$

Using

$$\mathrm{d}g^{-1}g = -g^{-1}\mathrm{d}g\,, \qquad \mathrm{d}\big(g^{-1}\mathrm{d}g\big) = -g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\,, \tag{7.72}$$

we get

$$\omega' = \omega + \mathrm{d}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge A\big) + \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big)\,. \tag{7.73}$$

In this discussion $g^{-1}\mathrm{d}g$ is unchanged under $g \to g_0 g$, for any fixed g_0, and so defines a left invariant one-form. If b^r are coordinates on the associated group manifold $\mathcal{M}_G$ then $g^{-1}(b)\mathrm{d}g(b) = \omega^a(b)t_a$ where $\omega^a(b)$ are the one forms defined in the general analysis of Lie groups in (5.48).

Since, using (7.72),

$$\mathrm{d}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big) = -\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big) = 0\,, \tag{7.74}$$

by virtue (7.67), we have

$$\mathrm{d}\omega' = \mathrm{d}\omega\,, \tag{7.75}$$

which is equivalent to (7.60). However although $\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big)$ is therefore a closed three-form it need not be exact so that its integration over a three manifold $\mathcal{M}_3$ may not vanish, in which case we would have

$$\int_{\mathcal{M}_3} \omega' \neq \int_{\mathcal{M}_3} \omega\,, \tag{7.76}$$

for some $g(x)$. The Cherns-Simons action is not then gauge invariant for such gauge transformations g.

To discuss

$$I = \int_{\mathcal{M}_3} \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big)\,, \tag{7.77}$$

we note that for a variation of g, since

$$\delta(g^{-1}\mathrm{d}g) = g^{-1}\mathrm{d}(\delta g\, g^{-1})\, g\,, \tag{7.78}$$

then

$$\delta\,\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = \mathrm{tr}\big(\mathrm{d}(\delta g\, g^{-1})\wedge \mathrm{d}g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\big)$$
$$= \mathrm{d}\,\mathrm{tr}\big(\delta g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\big)\,, \tag{7.79}$$

since $\mathrm{d}(\mathrm{d}g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}) = -\mathrm{d}^2(\mathrm{d}g\, g^{-1}) = 0$. Hence, for arbitrary smooth variations δg,

$$\delta I = 0\,, \tag{7.80}$$

so that I is a topological invariant, only when $g(x)$ can be continuously transformed to the identity must $I = 0$.

If we consider $g(\theta) \in SU(2)$ with coordinates θ^r, $r = 1, 2, 3$ then

$$\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = \rho(\theta)\,\mathrm{d}^3\theta\,, \tag{7.81}$$

The integration measure in (7.81) is defined in terms of the left invariant Lie algebra one forms so that for $g(\theta') = g_0 g(\theta)$ we have

$$\rho(\theta')\,\mathrm{d}^3\theta' = \rho(\theta)\,\mathrm{d}^3\theta\,. \tag{7.82}$$

Up to a sign, depending just on the sign of $\det[\partial'^r/\partial\theta^s]$, this is identical with the requirements for an invariant integration measure described in section 5.7. To check the normalisation we assume that near the origin, $\boldsymbol{\theta} \approx 0$, then $g(\theta) \approx I + i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}$ and hence

$$\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) \approx \tfrac{1}{3}i^3\,\mathrm{tr}\big(\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\wedge\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\wedge\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\big)$$
$$= \tfrac{2}{3}\,\varepsilon_{ijk}\,\mathrm{d}\theta^i\wedge\mathrm{d}\theta^j\wedge\mathrm{d}\theta^k = 4\,\mathrm{d}^3\theta\,, \tag{7.83}$$

assuming (5.21) and standard formulae for the Pauli matrices in (2.12) with (2.14). Thus $\rho(0) = 4$ and the results for the group integration volume for $SU(2)$ in (5.153) then imply, integrating over $\mathcal{M}_{SU(2)} \simeq S^3$,

$$\int_{\mathcal{M}_{SU(2)}} \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = 8\pi^2\,. \tag{7.84}$$

In general the topological invariant defined by (7.77), for a compact 3-manifold $\mathcal{M}_3$, corresponds to the index of the map defined by $g(x)$ from $\mathcal{M}_3$ to a subgroup $SU(2) \subset G$, *i.e.* the number of times the map covers the $SU(2)$ subgroup for $x \in \mathcal{M}_3$. The result (7.84) then requires that in general

$$I = 8\pi^2 n \qquad \text{for} \qquad n \in \mathbb{Z}\,. \tag{7.85}$$

In the functional integral approach to quantum field theories the action only appears in the form e^{iS}. In consequence S need only be defined up to integer multiples of 2π. Hence despite the fact that the action is not invariant under all gauge transformations a well defined quantum gauge Chern-Simons theory is obtained, on a compact 3-manifold $\mathcal{M}_3$, by employing as the action

$$S_{\mathrm{CS}} = \frac{k}{4\pi}\int_{\mathcal{M}_3} \mathrm{tr}\big(A\wedge\mathrm{d}A + \tfrac{2}{3}\,A\wedge A\wedge A\big)\,, \qquad k\in\mathbb{Z}\,, \tag{7.86}$$

so that, unlike Yang-Mills theory, the coupling is quantised. There is no requirement for k to be positive, the cubic terms become effectively small, and the theory is weakly coupled, when k is large.

7.3 Gauge Invariants and Wilson Loops

Only gauge invariant quantities have any significance in gauge field theories. Although it is necessary in non abelian gauge theories to solve the dynamical equations for gauge dependent fields, or in a quantum theory, to integrate over the gauge fields, only for gauge invariants is a well defined calculational result obtained. For abelian gauge theories this is a much less significant issue. The classical dynamical equations only involve $F_{\mu\nu}$ which is itself gauge invariant, (7.14). However even in this case the associated quantum field theory, QED, requires a much more careful treatment of gauge issues.

For a non abelian gauge theory $F_{\mu\nu} = F^a{}_{\mu\nu} t_a$ is a matrix belonging to a Lie algebra representation for the gauge group which transforms homogeneously under gauge transformations as in (7.34). The same transformation properties further apply to products of F's, at the same space-time point, and also to the gauge covariant derivatives $D_{\alpha_1} \dots D_{\alpha_r} F_{\mu\nu}$. Since $[D_\alpha, D_\beta]F_{\mu\nu} = [F_{\alpha\beta}, F_{\mu\nu}]$ the indices $\alpha_1, \dots, \alpha_n$ may be symmetrised to avoid linear dependencies. A natural set of gauge invariants, for pure gauge theories, is then provided by the matrix traces of products of F's, with arbitrarily many symmetrised covariant derivatives, at the same point,

$$\operatorname{tr}\big(D_{\alpha_{11}} \dots D_{\alpha_{1r_1}} F_{\mu_1\nu_1}\, D_{\alpha_{21}} \dots D_{\alpha_{2r_2}} F_{\mu_2\nu_2} \,\dots\, D_{\alpha_{s1}} \dots D_{\alpha_{sr_s}} F_{\mu_s\nu_s}\big)\,. \tag{7.87}$$

Such matrix traces may also be further restricted to a trace over a symmetrised product of the Lie algebra matrices, since any commutator may be simplified by applying the Lie algebra commutation relations, and also to just one of the s invariants, in the above example, related by cyclic permutation as the traces satisfy $\operatorname{tr}(X_1 \dots X_s) = \operatorname{tr}(X_s X_1 \dots X_{s-1})$. If the gauge group G has no $U(1)$ factors then $\operatorname{tr}(t_a) = 0$. The simplest example of such an invariant then involves just two F's, which include the energy momentum tensor as shown in (7.51). In general there are also derivative relations since

$$\partial_\mu \operatorname{tr}(X_1 \dots X_s) = \sum_{i=1}^{s} \operatorname{tr}(X_1 \dots\, D_\mu X_i \,\dots X_s)\,. \tag{7.88}$$

However, depending on the gauge group, the traces in (7.87) are not independent for arbitrary products of F's, even when no derivatives are involved. To show this we may consider the identity

$$\det(I - X) = e^{\operatorname{tr}\ln(I-X)}\,, \tag{7.89}$$

which is easy to demonstrate, for arbitrary diagonaliseable matrices X, since both sides depend only on the eigenvalues of X and the exponential converts the sum over eigenvalues provided by the trace into a product which gives the determinant. Expanding the right hand side gives

$$\begin{aligned}\det(I - X) &= e^{-\sum_{r\ge1}\operatorname{tr}(X^r)/r} \\ &= 1 - \operatorname{tr}(X) + \tfrac{1}{2}\big(\operatorname{tr}(X)^2 - \operatorname{tr}(X^2)\big) - \tfrac{1}{6}\big(\operatorname{tr}(X)^3 - 3\operatorname{tr}(X)\operatorname{tr}(X^2) + 2\operatorname{tr}(X^3)\big) + \dots\,.\end{aligned} \tag{7.90}$$

If X is a $N \times N$ matrix then $\det(I - X)$ is at most $\mathrm{O}(X^N)$ so that terms which are of higher order on the right hand side must vanish identically[26]. If $N=2$ this gives the relation

$$\operatorname{tr}(X^3) = \tfrac{3}{2}\operatorname{tr}(X)\operatorname{tr}(X^2) - \tfrac{1}{2}\operatorname{tr}(X)^3\,, \tag{7.91}$$

and if $N=3$, and also we require $\operatorname{tr}(X)=0$, the relevant identity becomes

$$\operatorname{tr}(X^4) = \tfrac{1}{2}\operatorname{tr}(X^2)^2\,. \tag{7.92}$$

In general $\operatorname{tr}(X^n)$ when $n>N$ is expressible in terms of products of $\operatorname{tr}(X^s)$ for $s \le N$.

For $G=SU(N)$ and taking t_a to belong to the fundamental representation these results are directly applicable to simplifying symmetrised traces appearing in (7.87) since the results for $\operatorname{tr}(X^n)$ are equivalent to relations for $\operatorname{tr}(t_{(a_1}\dots t_{a_N)})$.

7.3.1 Wilson Loops

The gauge field A_μ is a connection introduced to ensure that derivatives of gauge dependent fields transform covariantly under gauge transformations. It may be used, as with connections in differential geometry, to define 'parallel transport' of gauge dependent fields along a path in space-time between two points, infinitesimally for $x \to x+\mathrm{d}x$ this gives $\mathrm{d}x^\mu D_\mu\phi(x)$, where ϕ is a field belonging to a representation space for the gauge group G and D_μ is the gauge covariant derivative for this representation. Any continuous path $\Gamma_{x,y}$ linking the point y to x may be parameterised by $x^\mu(t)$ where $x^\mu(0)=y^\mu, x^\mu(1)=x^\mu$. For all such paths there is an associated element of the gauge group $\mathcal{G}$, as in (7.2), which is obtained by integrating along the path $\Gamma_{x,y}$. For the particular matrix representation $\mathcal{R}$ of G determined by ϕ this group element corresponds to $P(\Gamma_{x,y}) \in \mathcal{R}$ where $P(\Gamma_{x,y})\phi(y)$ transforms under local gauge transformations $g(x)\in\mathcal{R}$ belonging to G_x while $\phi(y)$ transforms as in (7.5) for $g(y)$ belonging to G_y.

For simplicity we consider an abelian gauge theory first. In this case $P(\Gamma_{x,y}) \in U(1)$ and under gauge transformations transforms as a local field at x and its conjugate at y. For a representation specified by a charge q as in (7.5), this is defined in terms of the differential equation

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t} - iq\,\dot{x}^\mu(t)A_\mu\big(x(t)\big)\Big)P(t,t') = 0\,, \qquad P(t,t)=1\,, \quad \dot{x}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\,, \tag{7.93}$$

which has a solution,

$$P(t,t') = e^{iq\int_{t'}^{t}\mathrm{d}\tau\,\dot{x}^\mu(\tau)A_\mu(x(\tau))}\,. \tag{7.94}$$

We then require

$$P(\Gamma_{x,y}) = P(1,0) = e^{iq\int_{\Gamma_{x,y}}\mathrm{d}x^\mu A_\mu(x)} \in U(1)\,, \tag{7.95}$$

[26]Equivalently if $F(z)=\det(I-zX) = 1+\sum_{r=1}^{N} a_r(X)\,z^r$ then

$$-\frac{F'(z)}{F(z)} = \operatorname{tr}\big(X(I-zX)^{-1}\big) = \sum_{r=0}^{\infty} z^r \operatorname{tr}(X^{r+1})\,,$$

and expanding the left hand side determines $\operatorname{tr}(X^n)$ for all n solely in terms of a_r, $r=1,\dots N$ which are also expressible in terms of $\operatorname{tr}(X^n)$ for $n\le N$.

which is independent of the particular parameterisation of the path $\Gamma_{x,y}$. Under the abelian gauge transformation in (7.11)

$$P(\Gamma_{x,y}) \underset{e^{i\alpha}}{\longrightarrow} P(\Gamma_{x,y})\, e^{iq\int_{\Gamma_{x,y}} \mathrm{d}x^\mu \partial_\mu \alpha(x)} = e^{iq\,\alpha(x)}\, P(\Gamma_{x,y})\, e^{-iq\,\alpha(y)}\,, \tag{7.96}$$

demonstrating that, for ϕ transforming under gauge transformations as in (7.5),

$$P(\Gamma_{x,y})\phi(y) \underset{e^{i\alpha}}{\longrightarrow} e^{iq\alpha(x)} P(\Gamma_{x,y})\phi(y)\,. \tag{7.97}$$

If Γ is a closed path, with a parameterisation $x^\mu(t)$ such that $x^\mu(1) = x^\mu(0) = x^\mu \in \Gamma$, then $\Gamma = \Gamma_{x,x}$ for any x on Γ. It is evident from (7.96) that $P(\Gamma)$ is gauge invariant. In this abelian case $P(\Gamma)$ may be expressed just in terms of the gauge invariant curvature in (7.14) using Stokes' theorem

$$P(\Gamma) = e^{iq\oint_\Gamma \mathrm{d}x^\mu A_\mu(x)} = e^{\frac{1}{2}iq\int_S \mathrm{d}S^{\mu\nu} F_{\mu\nu}(x)}\,, \tag{7.98}$$

for S any surface with boundary Γ and $\mathrm{d}S^{\mu\nu} = -\mathrm{d}S^{\nu\mu}$ the orientated surface area element (in three dimensions the identity is $\oint_\Gamma \mathrm{d}\mathbf{x}\cdot\mathbf{A} = \int_S \mathrm{d}\mathbf{S}\cdot\mathbf{B}$, $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ with $\mathrm{d}S_i = \frac{1}{2}\varepsilon_{ijk}\mathrm{d}S^{jk}$).

For the non abelian case (7.93) generalises to a matrix equation

$$\Big(I\,\frac{\mathrm{d}}{\mathrm{d}t} + A(t)\Big)P(t,t') = 0\,, \qquad A(t) = \dot{x}^\mu A_\mu(x(t))\,, \quad P(t,t) = I\,, \tag{7.99}$$

where $A(t)$ is a matrix belonging to the Lie algebra for a representation $\mathcal{R}$ of G. (7.99) may also be expressed in an equivalent integral form

$$P(t,t') = I - \int_{t'}^{t} \mathrm{d}\tau\; A(\tau)P(\tau,t')\,. \tag{7.100}$$

Solving this iteratively gives

$$\begin{aligned} P(t,t') &= I + \sum_{n\geq 1}(-1)^n \int_{t'}^{t}\mathrm{d}t_1 \int_{t'}^{t_1}\mathrm{d}t_2 \ldots \int_{t'}^{t_{n-1}}\mathrm{d}t_n\; A(t_1)A(t_2)\ldots A(t_n) \\ &= I + \sum_{n\geq 1}(-1)^n \frac{1}{n!}\prod_{r=1}^{n}\int_{t'}^{t}\mathrm{d}t_r\; \mathcal{T}\big\{A(t_1)A(t_2)\ldots A(t_n)\big\}\,. \end{aligned} \tag{7.101}$$

where $\mathcal{T}$ denotes that the non commuting, for differing t, $A(t)$ are t-ordered so that

$$\mathcal{T}\big\{A(t)A(t')\big\} = \begin{cases} A(t)A(t')\,, & t \geq t'\,, \\ A(t')A(t)\,, & t < t'\,. \end{cases} \tag{7.102}$$

The final expression can be simply written as a $\mathcal{T}$-ordered exponential

$$P(t,t') = \mathcal{T}\Big\{e^{-\int_{t'}^{t}\mathrm{d}\tau\; A(\tau)}\Big\}\,. \tag{7.103}$$

The corresponding non abelian generalisation of (7.95) is then

$$P(\Gamma_{x,y}) = P(1,0) = \mathcal{P}\Big\{e^{-\int_{\Gamma_{x,y}}\mathrm{d}x^\mu\, A_\mu(x)}\Big\} \in \mathcal{R}\,, \tag{7.104}$$

with $\mathcal{P}$ denoting path-ordering along the path Γ (this is equivalent to t-ordering with the particular parameterisation $x^\mu(t)$). These satisfy the group properties

$$P(\Gamma_{x,y})\, P(\Gamma_{y,z}) = P(\Gamma_{x,y} \circ \Gamma_{y,z})\,, \tag{7.105}$$

where $\Gamma_{x,y} \circ \Gamma_{y,z}$ denotes path composition, and, if $\mathcal{R}$ is a unitary representation

$$P(\Gamma_{x,y})^{-1} = P(\Gamma^{-1}_{y,x}) = P(\Gamma_{x,y})^\dagger\,, \tag{7.106}$$

with $\Gamma^{-1}_{y,x}$ the inverse path to $\Gamma_{x,y}$.

For a gauge transformation as in (7.26), $g(x) \in \mathcal{R}$, then in (7.99)

$$A(t) \underset{g}{\to} g(t)A(t)g(t)^{-1} - \dot{g}(t)g(t)^{-1}\,,\ g(t) = g(x(t)) \quad \Rightarrow \quad P(t) \underset{g}{\to} g(t)P(t,t')g(t')^{-1}\,, \tag{7.107}$$

and hence

$$P(\Gamma_{x,y}) \underset{g}{\longrightarrow} g(x)P(\Gamma_{x,y})\, g(y)^{-1}\,. \tag{7.108}$$

For $\Gamma = \Gamma_{x,x}$ a closed path then we may obtain a gauge invariant by taking the trace

$$W(\Gamma) = \mathrm{tr}\big(P(\Gamma_{x,x})\big)\,. \tag{7.109}$$

$W(\Gamma)$ is a *Wilson*[27] *loop.* It depends on the path Γ and also on the particular representation $\mathcal{R}$ of the gauge group. Wilson loops form a natural, but over complete, set of non local gauge invariants for any non abelian gauge theory. They satisfy rather non trivial identities reflecting the particular representation and gauge group. Subject to these the gauge field can be reconstructed from Wilson loops for arbitrary closed paths up to a gauge transformation. The associated gauge groups elements for paths connecting two points, as given in (7.104), may also be used to construct gauge invariants involving local gauge dependent fields at different points. For the field ϕ, transforming as in (7.5), $\phi(x)^\dagger P(\Gamma_{x,y})\phi(y)$ is such a gauge invariant, assuming the gauge transformation g is unitary so that (7.5) also implies $\phi(x)^\dagger \to \phi(x)^\dagger g(x)^{-1}$.

If a closed loop Γ is shrunk to a point then the Wilson loop $W(\Gamma)$ can be expanded in terms of local gauge invariants, of the form shown in (7.87), at this point. As an illustration we consider a rectangular closed path with the associated Wilson loop

$$W(\Box) = \mathrm{tr}\big(P(\Gamma_{x,x+be_j})\, P(\Gamma_{x+be_j,x+ae_i+be_j})\, P(\Gamma_{x+ae_i+be_j,x+ae_i})\, P(\Gamma_{x+ae_i,x})\big)\,, \tag{7.110}$$

where here Γ are all straight line paths and e_i, e_j are two orthogonal unit vectors. To evaluate $W(\Box)$ as $a, b \to 0$ it is convenient to use operators $\hat{\mathrm{x}}^\nu, \hat{\partial}_\mu$ with the commutation relations

$$\big[\hat{\mathrm{x}}^\mu, \hat{\mathrm{x}}^\nu\big] = 0\,, \qquad \big[\hat{\partial}_\mu, \hat{\partial}_\nu\big] = 0\,, \qquad \big[\hat{\partial}_\mu, \hat{\mathrm{x}}^\nu\big] = \delta_\mu{}^\nu\,, \tag{7.111}$$

which have a representation, acting on vectors $|x\rangle$, $x \in \mathbb{R}^4$, where

$$\hat{\mathrm{x}}^\mu|x\rangle\,, \qquad \hat{\partial}_\mu|x\rangle = -\partial_\mu|x\rangle\,. \tag{7.112}$$

[27]Kenneth Geddes Wilson, 1936-, American. Nobel prize 1982.

In terms of these operators, since $\hat{\mathrm{x}}^\nu \, e^{-te^\mu \hat{D}_\mu} = e^{-te^\mu \hat{D}_\mu}(\hat{\mathrm{x}}^\nu + te^\nu)$,

$$e^{-te^\mu \hat{D}_\mu}|x\rangle = |x(t)\rangle \, P(\Gamma_{x(t),x}) \, , \qquad \hat{D}_\mu = \hat{\partial}_\mu + A_\mu(\hat{\mathrm{x}}) \, , \quad x^\nu(t) = x^\nu + te^\nu \, , \tag{7.113}$$

which defines $P(\Gamma_{x(t),x})$ for the straight line path $\Gamma_{x(t),x}$ from x to $x(t)$, with $P(\Gamma_{x,x}) = I$. To verify that $P(\Gamma_{x(t),x})$ agrees with (7.103) we note that

$$\frac{\partial}{\partial t} e^{-te^\mu \hat{D}_\mu}|x\rangle = -e^\mu \hat{D}_\mu \, e^{-te^\mu \hat{D}_\mu}|x\rangle = \Big(\frac{\partial}{\partial t}\, |x(t)\rangle - |x(t)\rangle \, e^\mu A_\mu(x(t))\Big) P(\Gamma_{x(t),x}) \, , \tag{7.114}$$

using (7.112) as well as (7.113). It is then evident that (7.114) reduces to

$$\frac{\partial}{\partial t} \, P(\Gamma_{x(t),x}) = -e^\mu A_\mu(x(t)) \, P(\Gamma_{x(t),x}) \, , \tag{7.115}$$

which is identical to (7.93). For the rectangular closed path in (7.110)

$$\begin{aligned}
&|x\rangle \, P(\Gamma_{x,x+be_j}) \, P(\Gamma_{x+be_j,x+ae_i+be_j}) \, P(\Gamma_{x+ae_i+be_j,x+ae_i}) \, P(\Gamma_{x+ae_i,x}) \\
&= e^{b\hat{D}_j} \, e^{\alpha \hat{D}_i} \, e^{-b\hat{D}_j} \, e^{-\alpha \hat{D}_i}|x\rangle \\
&= e^{ab\,[\hat{D}_j,\hat{D}_i] - \frac{1}{2}a^2 b\,[[\hat{D}_j,\hat{D}_i],\hat{D}_i]| + \frac{1}{2}ab^2\,[\hat{D}_j,[\hat{D}_j,\hat{D}_i]] + \dots} \, |x\rangle \\
&= |x\rangle \, e^{-ab\, F_{ij}(x) - \frac{1}{2}a^2 b\, D_i F_{ij}(x) - \frac{1}{2}ab^2 D_j \, F_{ij}(x) + \dots} \, ,
\end{aligned} \tag{7.116}$$

using the Baker Cambell Hausdorff formula described in 5.4.2 and $[\hat{D}_i, \hat{D}_j] = F_{ij}(\hat{\mathrm{x}})$. Hence, for a N-dimensional representation with $\mathrm{tr}(t_a) = 0$, the leading approximation to (7.110) is just

$$\begin{aligned}
W(\Box) = N &+ \tfrac{1}{2}a^2 b^2 \big(1 + \tfrac{1}{2}a\partial_i + \tfrac{1}{2}b\partial_j + \tfrac{1}{6}a^2\partial_i^2 + \tfrac{1}{6}b^2\partial_j^2 + \tfrac{1}{4}ab\partial_i\partial_j\big) \, \mathrm{tr}\big(F_{ij}F_{ij}\big) \\
&- \tfrac{1}{24}a^4 b^2 \, \mathrm{tr}\big(D_i F_{ij} D_i F_{ij}\big) - \tfrac{1}{24}a^2 b^4 \, \mathrm{tr}\big(D_j F_{ij} D_j F_{ij}\big) \\
&- \tfrac{1}{6}a^3 b^3 \, \mathrm{tr}\big(F_{ij}F_{ij}F_{ij}\big) + \dots \, , \quad \text{no sums on } \; i,j \, .
\end{aligned} \tag{7.117}$$

For completeness we also consider how $P(\Gamma_{x,y})$ changes under variations in the path $\Gamma_{x,y}$. For this purpose the path Γ is now specified by $x^\mu(t,s)$, depending continuously on the additional variable s, which includes possible variations in the end points at $t = 0, 1$. If we define t, s covariant derivatives on these paths by

$$D_t = I \, \frac{\partial}{\partial t} + A_t(t) \, , \;\; D_s = I \, \frac{\partial}{\partial s} + A_s(t) \, , \quad A_t(t) = \frac{\partial x^\mu}{\partial t} A_\mu(x) \, , \;\; A_s(t) = \frac{\partial x^\mu}{\partial s} A_\mu(x) \, , \tag{7.118}$$

leaving the dependence on s implicit, then

$$[D_t, D_s] = F(t) = \frac{\partial x^\mu}{\partial t} \frac{\partial x^\nu}{\partial s} \, F_{\mu\nu}(x) \, . \tag{7.119}$$

With the definitions in (7.118), (7.99) becomes $D_t P(t) = 0$. Acting with D_s gives

$$D_t D_s P(t,t') = F(t) P(t,t') \, , \qquad D_s P(t,t) = A_s(t) \, , \tag{7.120}$$

which has a straightforward solution giving

$$\frac{\mathrm{d}}{\mathrm{d}s} P(1,0) + A_s(1) P(1,0) - P(1,0) A_s(0) = \int_0^1 \mathrm{d}t \; P(1,t) F(t) P(t,0) \, . \tag{7.121}$$

The result (7.121) may be recast as

$$\delta_\Gamma P(\Gamma_{x,y}) + \delta x^\nu A_\nu(x)\, P(\Gamma_{x,y}) - P(\Gamma_{x,y})\, \delta y^\nu A_\nu(y) \\ = \int_{\Gamma_{x,y}} \mathrm{d}z^\mu \; P(\Gamma_{x,z})\, F_{\mu\nu}(z)\delta x^\nu(z)\, P(\Gamma_{z,y})\,, \tag{7.122}$$

where

$$\Gamma_{x,y} = \Gamma_{x,z} \circ \Gamma_{z,y} \qquad \text{for} \quad z \in \Gamma_{x,y}\,. \tag{7.123}$$

For a Wilson loop

$$\delta_\Gamma W(\Gamma) = \oint_\Gamma \mathrm{d}x^\mu \; \mathrm{tr}\big(F_{\mu\nu}(x)\delta x^\nu(x)\, P(\Gamma_{x,x})\big)\,. \tag{7.124}$$

For a pure Chern-Simons theory then, as a consequence of the dynamical equation (7.66), there are no local gauge invariants and also Wilson loops are invariant under smooth changes of the loop path. The Wilson loop $W(\Gamma) \neq N$ only if it is not contractable to a point.

8 Integrations over Spaces Modulo Group Transformations

In a functional integration approach to quantum gauge field theories it is necessary to integrate over the non trivial space of gauge fields modulo gauge transformations, as in (7.4) with the definitions (7.1) and (7.2). This often becomes rather involved with somewhat formal manipulations of functional integrals but the essential ideas can be illustrated in terms of well defined finite dimensional integrals.

To this end we consider n-dimensional integrals of the form

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \; f(x)\,, \tag{8.1}$$

for classes of functions f which are invariant under group transformations belonging to a group G,

$$f(x) = f(x^g)\,, \qquad \text{for} \quad x \underset{g}{\rightarrow} x^g \;\; \text{for all} \;\; g \in G\,. \tag{8.2}$$

Necessarily we require

$$(x^{g_1})^{g_2} = x^{g_1 g_2}\,, \qquad (x^g)^{g^{-1}} = x\,, \tag{8.3}$$

and also we assume, under the change of variable $x \to x^g$,

$$\mathrm{d}^n x = \mathrm{d}^n x^g\,. \tag{8.4}$$

The condition (8.4) is an essential condition on the integration measure in (8.1), which is here assumed for simplicity to be the standard translation invariant measure on $\mathbb{R}^n$. If the group transformation g acts linearly on x then it is necessary that $G \subset Sl(n,\mathbb{R}) \ltimes T_n$, which contains the n-dimensional translation group T_n.

For any x the action of the group G generates the orbit $\mathrm{Orb}(x)$ and those group elements which leave x invariant define the stability group H_x,

$$\mathrm{Orb}(x) = \{x^g\}\,, \qquad H_x = \{h : x^h = x\}\,. \tag{8.5}$$

Clearly two points on the same orbit have isomorphic stability groups since

$$H_{x^g} = g^{-1} H_x g \simeq H_x \subset G \,. \tag{8.6}$$

We further require that for arbitrary x, except perhaps for a lower dimension subspace, the stability groups are isomorphic so that $H_x \simeq H$. Defining the manifold $\mathcal{M}$ to be formed by the equivalence classes $[x] = \{x/\sim\}$, where $x^g \sim x$, or equivalently by the orbits $\mathrm{Orb}(x)$, then $\mathcal{M} \simeq \mathbb{R}^n/(G/H)$. We here assume that G, and also in general H, are Lie groups, and further that H is compact. In this case $\mathcal{M}$ has a dimension which is less than n. Although $\mathbb{R}^n$ is topologically trivial, $\mathcal{M}$ may well have a non trivial topology.

In the integral (8.1), with a G-invariant function f, the integration may then be reduced to a lower dimensional integration over $\mathcal{M}$, by factoring off the invariant integration over G. To achieve this we introduce 'gauge-fixing functions' $P(x)$ on $\mathbb{R}^n$ such that,

$$\begin{aligned} &\text{for all } x \in \mathbb{R}^n \text{ then } P(x^g) = 0 \quad \text{for some } g \in G \,, \\ &\text{if } P(x_0) = 0 \text{ then } P(x_0{}^g) = 0 \quad \Rightarrow \quad g = h \in H \,, \ x_0{}^h = x_0 \,. \end{aligned} \tag{8.7}$$

In consequence the independent functions $P(x) \in \mathbb{R}^{\hat{n}}$ where $\hat{n} = \dim G - \dim H$. The solutions of the gauge fixing condition may be parameterised in terms of coordinates θ^r, $r = 1, \dots, n - \hat{n}$, so that

$$P\big(x_0(\theta)\big) = 0 \quad \Rightarrow \quad \theta^r \ \text{ coordinates on } \ \mathcal{M} \,, \quad \dim \mathcal{M} = n - \hat{n} \,. \tag{8.8}$$

For any $P(x)$ an associated function $\Delta(x)$ is defined by integrating over the G-invariant measure, as discussed in 5.7, according to

$$\int_G \mathrm{d}\rho(g) \; \delta^{\hat{n}}\big(P(x^g)\big) \; \Delta(x) = 1 \,. \tag{8.9}$$

Since by construction $\mathrm{d}\rho(g) = \mathrm{d}\rho(g'g)$ then it is easy to see that

$$\Delta(x^g) = \Delta(x) \quad \text{for all} \quad g \in G \,. \tag{8.10}$$

Using (8.9) in (8.1), and interchanging orders of integration, gives

$$\begin{aligned} \int_{\mathbb{R}^n} \mathrm{d}^n x \; f(x) &= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x \; \delta^{\hat{n}}\big(P(x^g)\big) \, \Delta(x) \; f(x) \\ &= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x^g \; \delta^{\hat{n}}\big(P(x^g)\big) \, \Delta(x^g) \; f(x^g) \\ &= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x \; \delta^{\hat{n}}\big(P(x)\big) \, \Delta(x) \; f(x) \,. \end{aligned} \tag{8.11}$$

using the invariance conditions (8.2), (8.4) and (8.10), and in the last line just changing the integration variable from x^g to x. For integration over $\mathcal{M}$ we then have a measure, which is expressible in terms of the coordinates θ^r, given by

$$\mathrm{d}\mu(\theta) = \mathrm{d}^n x \; \delta^{\hat{n}}\big(P(x)\big) \, \Delta(x) \,. \tag{8.12}$$

To determine $\Delta(x)$ in (8.9) then, assuming (8.7), if

$$g(\alpha, h) = \exp(\alpha)\, h\,, \qquad \alpha \in \mathfrak{g}/\mathfrak{h}\,, \tag{8.13}$$

we define a linear operator D, which may depend on x_0, such that

$$x_0{}^{g(\alpha,h)} = x_0 + D(x_0)\alpha\,, \quad \text{for} \ \ \alpha \approx 0\,, \qquad D(x_0) : \mathfrak{g}/\mathfrak{h} \to \mathbb{R}^n\,. \tag{8.14}$$

If $\{T_{\hat a}\}$ is a basis for $\mathfrak{g}/\mathfrak{h}$ (if $\mathfrak{g}$ has a non degenerate Killing form κ then $\kappa(\mathfrak{h}, T_{\hat a}) = 0$ for all $\hat a$ and we may write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}/\mathfrak{h}$) then

$$\alpha = \alpha^{\hat a} T_{\hat a}\,, \tag{8.15}$$

and, with the decomposition in (8.13),

$$\mathrm{d}\rho(g) \approx \prod_{\hat a=1}^{\hat n} \mathrm{d}\alpha^{\hat a}\, \mathrm{d}\rho_H(h) \qquad \text{for} \qquad \alpha^{\hat a} \approx 0\,, \tag{8.16}$$

for $\mathrm{d}\rho_H(h)$ the invariant integration measure on H. For x near x_0 we define the linear operator P' by

$$P(x_0 + y) = P'(x_0)y \quad \text{for} \quad y \approx 0\,, \qquad P'(x_0) : \mathbb{R}^n \to \mathbb{R}^{\hat n}\,. \tag{8.17}$$

Then in (8.9), with (8.16),

$$\begin{aligned} \int_G \mathrm{d}\rho(g)\, \delta^{\hat n}\big(P(x^g)\big) = \int_G \mathrm{d}\rho(g)\, \delta^{\hat n}\big(P(x_0{}^g)\big) &= V_H \int \mathrm{d}^{\hat n}\alpha\, \delta^{\hat n}\big(P'(x_0)D(x_0)\alpha\big) \\ &= V_H\, \frac{1}{\big|\det P'(x_0)D(x_0)\big|}\,, \qquad V_H = \int_H \mathrm{d}\rho_H(h)\,. \end{aligned} \tag{8.18}$$

Hence in (8.9)

$$\Delta(x) = \frac{1}{V_H}\, \big|\det P'(x_0)D(x_0)\big| \qquad \text{for} \qquad x = x_0{}^g\,. \tag{8.19}$$

In a quantum gauge field theory context $\det P'(x_0)D(x_0)$ is the *Faddeev-Popov*[28] *determinant*. The determinant is non vanishing except at points x_0 such that $P(x_0{}^g) = 0$ has solutions for $g \approx e$ and $g \notin H$ and the gauge fixing condition $P(x) = 0$ does not sufficiently restrict g. The resulting measure, since

$$P(x) = 0 \qquad \Rightarrow \qquad x = x(\theta, \alpha) = x_0(\theta)^{g(\alpha,h)}\,, \tag{8.20}$$

from (8.12) becomes, with a change of variables $x \to \theta, \alpha$,

$$\mathrm{d}\mu(\theta) = \frac{1}{V_H}\, \mathrm{d}^n x\, \delta^{\hat n}\big(P(x)\big)\, \big|\det M(\theta)\big|\,, \qquad M(\theta) = P'(x_0(\theta))D'(x_0(\theta))\,. \tag{8.21}$$

Note that

$$\delta^{\hat n}\big(P(x(\theta,\alpha))\big)\, \big|\det M(\theta)\big| = \delta^{\hat n}(\alpha)\,, \tag{8.22}$$

[28]Ludvig Dmitrievich Faddeev, 1934-, Russian. Viktor Nikolaevich Popov, Russian.

and therefore the measure over $\mathcal{M}$ may also be expressed in terms of the Jacobian from θ, α to x since

$$\mathrm{d}\mu(\theta) = \mathrm{d}^{n-\hat{n}}\theta \, \left| \det \left[\frac{\partial x}{\partial \theta}, \frac{\partial x}{\partial \alpha} \right] \right|_{\alpha=0} . \tag{8.23}$$

With these results, for G compact, (8.11) gives

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \, f(x) = V_G \int_{\mathbb{R}^n} \mathrm{d}^n x \, \delta^{\hat{n}}\big(P(x)\big) \, \Delta(x) \, f(x) = V_G \int_{\mathcal{M}} \mathrm{d}\mu(\theta) \, f\big(x_0(\theta)\big) . \tag{8.24}$$

As an extension we consider the situation when there is a discrete group W, formed by transformations $\theta \to \theta^{g_i}$, such that

$$W = \big\{ g_i : x_0(\theta^{g_i}) = x_0(\theta)^{g(g_i)}, \, g(g_i) \in G \big\} . \tag{8.25}$$

It follows that $M(\theta^{g_i}) = M(\theta)$ and $\mathrm{d}\mu(\theta^{g_i}) = \mathrm{d}\mu(\theta)$. Since the stability group H leaves x_0 invariant $g(g_i)$ is not unique, hence in general it is sufficient that $g(g_i)g(g_j) = g(g_i g_j)h$ for $h \in H$. In many cases it is possible to restrict the coordinates $\{\theta^r\}$ so that W becomes trivial but it is also often natural not to impose such constraints on the θ^r's and to divide (8.21) by $|W|$ to remove multiple counting so that

$$\mathrm{d}\mu(\theta) = \frac{1}{|W| \, V_H} \, \mathrm{d}^n x \, \delta^{\hat{n}}\big(P(x)\big) \, \big|\det M(\theta)\big| , \tag{8.26}$$

8.1 Integrals over Spheres

As a first illustration of these methods we consider examples where the group G is one of the compact matrix groups $SO(n)$, $U(n)$ or $Sp(n)$ and the orbits under the action of group transformations are spheres.

For the basic integral over $x \in \mathbb{R}^n$ in (8.1), where $x = (x^1, \dots, x^n)$, we then consider

$$f(x) = F(x^2) , \tag{8.27}$$

where $x^2 = x^i x^i$ is the usual flat Euclidean metric. In this case we take $G = SO(n)$ which acts as usual $x \underset{R}{\to} x' = Rx$, regarding x here as an n-component column vector, for any $R \in SO(n)$. Since $\det R = 1$ of course $\mathrm{d}^n x' = \mathrm{d}^n x$. The orbits under the action of $SO(n)$ are all x with $x^2 = r^2$ fixed and so are spheres S^{n-1} for radii r. A representative point on any such sphere may be chosen by restricting to the intersection with the positive 1-axis or

$$x_0 = r(1, 0, \dots, 0, 0) , \qquad r > 0 . \tag{8.28}$$

In this case the stability group, for all $r > 0$, $H \simeq SO(n-1)$ since matrices leaving x_0 in (8.28) invariant have the form

$$R(\hat{R}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{R} \end{pmatrix} , \qquad \hat{R} \in SO(n-1) . \tag{8.29}$$

Note that $\dim SO(n) = \frac{1}{2}n(n-1)$ so that in this example $\hat{n} = \dim SO(n) - \dim SO(n-1) = n-1$, and therefore $n - \hat{n} = 1$ corresponding to the single parameter r.

Corresponding to the choice (8.28) the corresponding gauge fixing condition, corresponding to $\delta^{\hat{n}}(P(x))$, is

$$\mathcal{F}(x) = \theta(x^1) \prod_{i=2}^{n} \delta(x^i) \, . \tag{8.30}$$

The condition $x^1 > 0$ may be omitted but then there is a residual group $W \simeq \mathbb{Z}_2$ corresponding to reflections $x^1 \to -x^1$. For the generators of $SO(n)$ given by (5.218) we have

$$S_{s1} x_0 = r(0, \dots, \underbrace{1}_{s'\text{th place}}, \dots, 0) \, , \qquad s = 2, \dots, n \, , \tag{8.31}$$

so that in (8.13) we may take

$$\alpha = \sum_{s=2}^{n} \alpha_s S_{s1} \, , \tag{8.32}$$

so that

$$\exp(\alpha) x_0 = r(1, \alpha_2, \dots, \alpha_n) \qquad \text{for} \quad \alpha \approx 0 \, . \tag{8.33}$$

For the measure we assume a normalisation such that

$$\mathrm{d}\rho_{SO(n)}(R) \approx \mathrm{d}^{n-1}\alpha \, \mathrm{d}\rho_{SO(n-1)}(\hat{R}) \quad \text{for} \quad R = \exp(\alpha) R(\hat{R}) \, , \quad \alpha \approx 0 \, , \tag{8.34}$$

where $R(\hat{R})$ is given in (8.29). With the gauge fixing function in (8.30)

$$\int_{SO(n)} \mathrm{d}\rho_{SO(n)}(R) \, \mathcal{F}(Rx) = V_{SO(n-1)} \int \mathrm{d}^{n-1}\alpha \prod_{s=2}^{n} \delta\big(\alpha_s |x^1|\big) = V_{SO(n-1)} \, \frac{1}{|x^1|^{n-1}} \, . \tag{8.35}$$

Hence

$$\Delta(x) = \frac{1}{V_{SO(n-1)}} \, r^{n-1} \, , \qquad x^2 = r^2 \, , \quad r > 0 \, . \tag{8.36}$$

With this (8.24) becomes

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \; F(x^2) = V_{SO(n)} \int_{\mathbb{R}^n} \mathrm{d}^n x \; \mathcal{F}(x) \, \Delta(x) \, F(x^2) = \frac{V_{SO(n)}}{V_{SO(n-1)}} \int_0^\infty \mathrm{d}r \; r^{n-1} F(r^2) \, . \tag{8.37}$$

Of course this is just the same result as obtained by the usual separation of angular variables for functions depending on the radial coordinate r.

For a special case

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \; e^{-x^2} = \pi^{\frac{1}{2}n} = \frac{V_{SO(n)}}{V_{SO(n-1)}} \int_0^\infty \mathrm{d}r \; r^{n-1} e^{-r^2} = \frac{V_{SO(n)}}{V_{SO(n-1)}} \, \tfrac{1}{2} \, \Gamma(\tfrac{1}{2}n) \, , \tag{8.38}$$

giving

$$\frac{V_{SO(n)}}{V_{SO(n-1)}} = S_n = \frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)} \, , \tag{8.39}$$

where S_n is the volume of S^{n-1}. Since $V_{SO(2)} = 2\pi$, or $V_{SO(1)} = 1$, in general

$$V_{SO(n)} = 2^{n-1} \frac{\pi^{\frac{1}{4}n(n+1)}}{\prod_{i=1}^{n} \Gamma\big(\frac{1}{2}i\big)} \, . \tag{8.40}$$

For the corresponding extension to the complex case we consider integrals over $\mathbb{C}^n \simeq \mathbb{R}^{2n}$, of real dimension $2n$, with coordinates $Z = (z_1, \dots, z_n)$, $z_i \in \mathbb{C}$. The analogous integrals are then

$$\int_{\mathbb{C}^n} \mathrm{d}^{2n} Z \, F(\bar{Z}Z) \, , \qquad \bar{Z}Z = \sum_{i=1}^{n} |z_i|^2 \, , \tag{8.41}$$

and where

$$\mathrm{d}^{2n} Z = \prod_{i=1}^{n} \mathrm{d}^2 z_i \, , \qquad \mathrm{d}^2 z = \mathrm{d}x \mathrm{d}y \quad \text{for} \quad z = x + iy \, . \tag{8.42}$$

In this case we may take $G = U(n) \subset O(2n)$ where the transformations act $Z \underset{U}{\to} UZ$ for $U \in U(n)$ so that $\bar{Z}Z$ is invariant, as is also $\mathrm{d}^{2n}Z$. As in the discussion for $SO(n)$ we may take on each orbit

$$Z_0 = r(1, 0, \dots, 0, 0) \, , \qquad r > 0 \, . \tag{8.43}$$

The stability group $H \simeq U(n-1)$ corresponding to matrices

$$U(\hat{U}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{U} \end{pmatrix} , \qquad \hat{U} \in U(n-1) \, . \tag{8.44}$$

In this case $\dim U(n) = n^2$ so that $\hat{n} = \dim U(n) - \dim U(n-1) = 2n - 1$. The orbits are just specified again by the single variable r.

Corresponding to (8.43) the gauge fixing condition becomes

$$\mathcal{F}(Z) = \theta(\mathrm{Re}\, z_1) \delta(\mathrm{Im}\, z_1) \prod_{i=2}^{n} \delta^2(z_i) \, , \qquad \delta^2(z) = \delta(x)\delta(y) \, , \quad z = x + iy \, . \tag{8.45}$$

In terms of the generators defined in (5.212) we let

$$\alpha = i\alpha_1 \, R^1{}_1 + \sum_{s=2}^{n} \big(\alpha_s \, R^s{}_1 - \alpha_s{}^* R^1{}_s\big) \, , \qquad \alpha_1 \in \mathbb{R} \, , \quad \alpha_s \in \mathbb{C} \, , \; s \geq 2 \, . \tag{8.46}$$

Hence

$$\exp(\alpha) Z_0 = r(1 + i\alpha_1, \alpha_2, \dots, \alpha_n) \, , \qquad \alpha \approx 0 \, , \tag{8.47}$$

and we take

$$\mathrm{d}\rho_{U(n)}(U) \approx \mathrm{d}\alpha_1 \textstyle\prod_{s=2}^{n} \mathrm{d}^2 \alpha_s \; \mathrm{d}\rho_{U(n-1)}(\hat{U}) \quad \text{for} \quad U = \exp(\alpha) U(\hat{U}) \, , \quad \alpha \approx 0 \, . \tag{8.48}$$

With these results

$$\int \mathrm{d}\rho_{U(n)}(U) \; \mathcal{F}(UZ) = V_{U(n-1)} \, \frac{1}{|z_1|^{2n-1}} \, , \tag{8.49}$$

which implies

$$\Delta(Z) = \frac{1}{V_{U(n-1)}}\, r^{2n-1}\,, \qquad \bar{Z}Z = r^2\,, \quad r > 0\,. \tag{8.50}$$

Finally

$$\int_{\mathbb{C}^n} \mathrm{d}^{2n}Z\; F(\bar{Z}Z) = V_{U(n)} \int_{\mathbb{C}^n} \mathrm{d}^{2n}Z\; \mathcal{F}(Z)\, \Delta(Z)\, F(\bar{Z}Z) = \frac{V_{U(n)}}{V_{U(n-1)}} \int_0^\infty \mathrm{d}r\; r^{2n-1} F(r^2)\,. \tag{8.51}$$

Corresponding to (8.39), (8.51) requires

$$\frac{V_{U(n)}}{V_{U(n-1)}} = S_{2n}\,. \tag{8.52}$$

Taking $V_{U(1)} = 2\pi$ we have, with our normalisation,

$$V_{U(n)} = 2^n\, \frac{\pi^{\frac{1}{2}n(n+1)}}{\prod_{i=1}^n \Gamma(i)}\,. \tag{8.53}$$

Since $U(n) \simeq SU(n) \otimes U(1)/\mathbb{Z}_n$

$$V_{U(n)} = \frac{2\pi}{n}\, V_{SU(n)}\,. \tag{8.54}$$

A very similar discussion applies in terms of quaternionic numbers which are relevant for $Sp(n)$. For $Q = (q_1, \dots, q_n) \in \mathbb{H}^n$ the relevant integrals are

$$\int_{\mathbb{H}^n} \mathrm{d}^{4n}Q\; F(\bar{Q}Q)\,, \qquad \bar{Q}Q = \sum_{i=1}^n |q_i|^2\,, \tag{8.55}$$

and where

$$\mathrm{d}^{4n}Q = \prod_{i=1}^n \mathrm{d}^4 q_i\,, \qquad \mathrm{d}^4 q = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v \quad \text{for} \quad z = x + iy + ju + kv\,. \tag{8.56}$$

$\bar{Q}Q$ is invariant under $Q \underset{M}{\to} MQ$ for $M \in Sp(n) \subset SO(4n)$, regarded as $n \times n$ quaternionic unitary matrices M satisfying (1.62). As before we choose

$$Q_0 = r(1, 0, \dots, 0, 0)\,, \qquad r > 0\,. \tag{8.57}$$

The stability group $H \simeq Sp(n-1)$ corresponding to quaternionic matrices where M is expressible in terms of $\hat{M} \in Sp(n-1)$ in an identical fashion to (8.44). We now have $\dim Sp(n) = n(2n+1)$ so that $\hat{n} = \dim Sp(n) - \dim Sp(n-1) = 4n - 1$.

The associated gauge fixing condition becomes

$$\mathcal{F}(Q) = \theta(\mathrm{Re}\, q_1)\delta^3(\mathrm{Im}\, q_1) \prod_{i=2}^n \delta^4(q_i)\,, \qquad \delta^4(q) = \delta(x)\delta(y)\delta(u)\delta(v)\,, \quad q = x + iy + iu + iv\,. \tag{8.58}$$

In terms of the generators defined in (5.212) we let

$$\alpha = \alpha_1 \, R^1{}_1 + \sum_{s=2}^{n} \big(\alpha_s \, R^s{}_1 - \overline{\alpha_s} \, R^1{}_s \big) \, , \qquad \alpha_s \in \mathbb{H} \, , \quad \mathrm{Re}\, \alpha_1 = 0 \, . \tag{8.59}$$

and

$$\mathrm{d}\rho_{Sp(n)}(M) \approx \mathrm{d}^3\alpha_1 \textstyle\prod_{s=2}^{n} \mathrm{d}^4\alpha_s \; \mathrm{d}\rho_{Sp(n-1)}(\hat{M}) \qquad \alpha \approx 0 \, . \tag{8.60}$$

Hence we find

$$\Delta(Q) = \frac{1}{V_{Sp(n-1)}} \, r^{4n-1} \, , \qquad \bar{Q}Q = r^2 1 \, , \quad r > 0 \, . \tag{8.61}$$

The integral in (8.55) becomes

$$\int_{\mathbb{H}^n} \mathrm{d}^{4n}Q \; F(\bar{Q}Q) = V_{Sp(n)} \int_{\mathbb{H}^n} \mathrm{d}^{4n}Q \; \mathcal{F}(Q) \, \Delta(Q) \, F(\bar{Q}Q) = \frac{V_{Sp(n)}}{V_{Sp(n-1)}} \int_0^\infty \mathrm{d}r \; r^{4n-1} F(r^2) \, , \tag{8.62}$$

and corresponding to (8.39), (8.62) requires

$$\frac{V_{Sp(n)}}{V_{Sp(n-1)}} = S_{4n} \, . \tag{8.63}$$

Since $Sp(1) = \{q : |q|^2 = 1\}$, with the group property depending on $|q_1 q_2| = |q_1| \, q_2|$, the group manifold is just S^3 and

$$V_{Sp(1)} = \int \mathrm{d}^4 q \; \delta(|q| - 1) = S_4 = 2\pi^2 \, , \tag{8.64}$$

just as in (5.153). Hence

$$V_{Sp(n)} = 2^n \, \frac{\pi^{n(n+1)}}{\prod_{i=1}^{n} \Gamma(2i)} \, . \tag{8.65}$$

The results for the group volumes in (8.40), (8.53) and (8.65) depend on the conventions adopted in the normalisation of the group invariant integration measure which are here determined by (8.34), (8.48) and (8.60) in conjunction with (8.32), (8.46) and (8.59) respectively.

8.2 Integrals over Symmetric and Hermitian Matrices

A class of finite dimensional group invariant integrals which are rather more similar to gauge theories are those which involve integrals over real symmetric or complex hermitian matrices.

For the real case for $n \times n$ symmetric matrices X the relevant integrals are of the form

$$\int \mathrm{d}^{\frac{1}{2}n(n+1)} X \; f(X) \, , \quad X = X^T \, , \quad d^{\frac{1}{2}n(n+1)} X = \prod_{i=1}^{n} \mathrm{d}X_{ii} \prod_{1 \le i < j \le n} \mathrm{d}X_{ij} \, , \tag{8.66}$$

and we assume the invariance

$$f(X) = f(RXR^{-1})\,, \qquad R \in SO(n)\,. \tag{8.67}$$

The measure $\mathrm{d}^{\frac{1}{2}n(n+1)}X$ is invariant under $X \to RXR^{-1}$. A standard result in the discussion of matrices is that any symmetric matrix such as X may be diagonalised so that

$$RXR^{-1} = \Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}\,, \tag{8.68}$$

where λ_i are the eigenvalues of X. If $\{\lambda_i\}$ are all different there is no continuous Lie subgroup of $SO(n)$ such that $R\Lambda R^{-1} = \Lambda$ since

$$\dim\{X : X = X^T\} - \dim SO(n) = \tfrac{1}{2}n(n+1) - \tfrac{1}{2}n(n-1) = n\,, \tag{8.69}$$

corresponding to the number of independent λ_i. The orbits of X under the action of $SO(n)$ are then determined by the eigenvalues $\{\lambda_i\}$. For any $SO(n)$ invariant function as in (8.67) we may write

$$f(X) = \hat{f}(\lambda)\,, \qquad \lambda = (\lambda_1, \dots \lambda_n)\,. \tag{8.70}$$

However there is a discrete stability group for Λ. The diagonal matrices corresponding to reflections in the i-direction

$$\qquad\qquad i \tag{8.71}$$

$$R_i = i\begin{pmatrix} \begin{smallmatrix}1&0\\0&1\end{smallmatrix} & \dots\dots\dots\dots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & \begin{smallmatrix}1&&\\&-1&\\&&1\end{smallmatrix} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots\dots & \begin{smallmatrix}1&0\\0&1\end{smallmatrix} \end{pmatrix} \in O(n)\,, \quad i = 1, \dots, n\,, \qquad R_i^2 = I_n\,. \tag{8.72}$$

generate the discrete group

$$\{R_{a_1\dots a_n} = R_1{}^{a_1} \dots R_{n-1}{}^{a_n} : a_i = 0, 1\} \simeq \mathbb{Z}_2{}^{\otimes n}\,, \tag{8.73}$$

such that for any element

$$R_{a_1\dots a_n}\Lambda R_{a_1\dots a_n}{}^{-1} = \Lambda\,. \tag{8.74}$$

Furthermore for any permutation $\sigma \in \mathcal{S}_n$ there are corresponding matrices $R_\sigma \in O(n)$, such that $(R_\sigma)_{ij}x_j = x_{\sigma(i)}$. The matrices $\{R_\sigma\}$ form a faithful representation of $\mathcal{S}_n$ and

$$R_\sigma \Lambda R_\sigma{}^{-1} = \Lambda_\sigma = \begin{pmatrix} \lambda_{\sigma(1)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_{\sigma(n)} \end{pmatrix}\,. \tag{8.75}$$

For the permutation $(i\,i{+}1)$ the associated matrix is

$$i \tag{8.76}$$

$$R_{(i\,i+1)} = i \begin{pmatrix} \begin{smallmatrix}1&0\\0&1\end{smallmatrix} & \cdots\cdots\cdots\cdots & & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \begin{smallmatrix}0&1\\1&0\end{smallmatrix} & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots\cdots\cdots\cdots & & \begin{smallmatrix}1&0\\0&1\end{smallmatrix} \end{pmatrix}, \qquad i = 1,\dots,n-1\,, \tag{8.77}$$

which may be used to determine R_σ for any σ by group multiplication. Since $R_\sigma R_i R_\sigma{}^{-1} = R_{\sigma(i)}$ the groups generated by permutations and reflections may be identified as the semi-direct product $\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\otimes n}$. However $\det R_\sigma = P_\sigma = \pm 1$, with P_σ defined in (6.88), and also $\det R_{a_1\dots a_n} = (-1)^{\sum_i a_i}$. Restricting to the subgroup belonging to $SO(n)$, having determinant one, we then take for the group W, as defined in (8.25),

$$W = \big(\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\otimes n}\big)/\mathbb{Z}_2\,, \qquad |W| = 2^{n-1}n!\,. \tag{8.78}$$

It is possible to restrict W to $\mathbb{Z}_2{}^{\otimes n-1}$, formed by $R_{a_1\dots a_n}$ with $\sum_i a_i$ even, by requiring that the eigenvalues in (8.68) are ordered so that $\lambda_i \le \lambda_{i+1}$. However the choice of W in (8.78) is generally more convenient.

Taking $X_0 = \Lambda$, as in (8.68), the corresponding gauge fixing condition is

$$\mathcal{F}(X) = \prod_{1\le i<j\le n} \delta(X_{ij})\,. \tag{8.79}$$

For a rotation

$$R(\alpha) = \exp(\alpha)\,, \qquad \alpha = -\alpha^T\,, \tag{8.80}$$

and the group invariant integration is then assumed to be normalised such that, for R as in (8.80),

$$\mathrm{d}\rho_{SO(n)}\big(R(\alpha)\big) \approx \prod_{1\le i<j\le n} \mathrm{d}\alpha_{ij}\,, \qquad \alpha \approx 0\,. \tag{8.81}$$

With these assumptions, applying (8.9),

$$\frac{1}{\Delta(X)} = \int_{SO(n)} \mathrm{d}\rho_{SO(n)}(R)\; \mathcal{F}\big(RXR^{-1}\big) = \prod_{1\le i<j\le n} \int \mathrm{d}\alpha_{ij}\, \delta\big(\alpha_{ij}(\lambda_j - \lambda_i)\big)\,, \tag{8.82}$$

so that

$$\Delta(X) = |\hat\Delta(\lambda)| \qquad \text{for} \qquad \hat\Delta(\lambda) = \prod_{1\le i<j\le n} (\lambda_i - \lambda_j)\,. \tag{8.83}$$

The resulting $SO(n)$ invariant integration over symmetric matrices becomes

$$\begin{aligned} \int \mathrm{d}^{\frac12 n(n+1)}X\; f(X) &= \frac{V_{SO(n)}}{|W|} \int \mathrm{d}^{\frac12 n(n+1)}X\; \mathcal{F}(X)\Delta(X)\, f(X) \\ &= \frac{V_{SO(n)}}{2^{n-1}n!} \int \mathrm{d}^n\lambda\; |\hat\Delta(\lambda)|\, \hat f(\lambda)\,. \end{aligned} \tag{8.84}$$

Since the normalisations chosen in (8.80) and (8.81) are compatible with those assumed previously we may use (8.40) for $V_{SO(n)}$.

For the particular example

$$f(X) = e^{-\frac{1}{2}\kappa\,\mathrm{tr}(X^2)}\,, \qquad \mathrm{tr}(X^2) = \sum_{i=1}^{n} X_{ii}{}^2 + 2\sum_{1\le i<j\le n} X_{ij}{}^2 = \sum_{i=1}^{n}\lambda_i^2\,, \tag{8.85}$$

then

$$\int \mathrm{d}^{\frac{1}{2}n(n+1)}X\; e^{-\frac{1}{2}\kappa\,\mathrm{tr}(X^2)} = 2^{\frac{1}{2}n}\left(\frac{\pi}{\kappa}\right)^{\frac{1}{4}n(n+1)}. \tag{8.86}$$

Using (8.40) this defines a normalised probability measure for the eigenvalues for a Gaussian ensemble of symmetric real matrices

$$\mathrm{d}\mu(\lambda)_{\text{symmetric matrices}} = \frac{\kappa^{\frac{1}{4}n(n+1)}}{2^{\frac{3}{2}n}\prod_{i=1}^{n}\Gamma\big(1+\frac{1}{2}i\big)}\prod_{i=1}^{n}\mathrm{d}\lambda_i\;|\hat{\Delta}(\lambda)|\,e^{-\frac{1}{2}\kappa\sum_i\lambda_i^2}\,. \tag{8.87}$$

There is a corresponding discussion for complex hermitian $n\times n$ matrices when the integrals are of the form

$$\int \mathrm{d}^{n^2}X\; f(X)\,, \quad X = X^\dagger\,, \quad d^{n^2}X = \prod_{i=1}^{n}\mathrm{d}X_{ii}\prod_{1\le i<j\le n}\mathrm{d}^2X_{ij}\,, \tag{8.88}$$

where f satisfies

$$f(X) = f(UXU^{-1})\,, \qquad U\in U(n)\,. \tag{8.89}$$

Just as before hermitian matrices may be diagonalised

$$UXU^{-1} = \Lambda\,, \tag{8.90}$$

where the diagonal elements of Λ are the eigenvalues of X as in (8.68). In this case there is a non trivial continuous subgroup of $U(n)$ leaving Λ invariant formed by the diagonal matrices

$$U_0(\beta) = \begin{pmatrix} e^{i\beta_1} & 0 & \dots & 0 \\ 0 & e^{i\beta_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & e^{i\beta_n}\end{pmatrix}, \tag{8.91}$$

and hence we may identify

$$H \simeq U(1)^{\otimes n}\,. \tag{8.92}$$

In addition we may identify $W = \mathcal{S}_n$ formed by $\{R_\sigma\}\subset U(n)$ which permute the eigenvalues in Λ.

The gauge fixing condition restricting X to diagonal form is now

$$\mathcal{F}(X) = \prod_{1\le i<j\le n}\delta^2(X_{ij})\,. \tag{8.93}$$

In this case we may write for arbitrary $U \in U(n)$,

$$U(\alpha,\beta) = \exp(\alpha)U_0(\beta)\,, \qquad \alpha = -\alpha^\dagger\,, \quad \alpha_{ii} = 0 \text{ all } i\,, \tag{8.94}$$

and the group invariant integration is then assumed to be normalised such that, for U as in (8.94),

$$\mathrm{d}\rho_{U(n)}\big(U(\alpha,\beta)\big) \approx \prod_{1\le i<j\le n} \mathrm{d}^2\alpha_{ij} \prod_{i=1}^{n} \mathrm{d}\beta_i\,, \qquad \alpha \approx 0\,,\ 0 \le \beta_i < 2\pi\,. \tag{8.95}$$

With these assumptions

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(U)\; \mathcal{F}\big(UXU^{-1}\big) = (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^2\alpha_{ij}\, \delta^2\big(\alpha_{ij}(\lambda_j - \lambda_i)\big)\,. \tag{8.96}$$

Since

$$\delta^2(\lambda z) = \frac{1}{|\lambda|^2}\,\delta^2(z)\,, \tag{8.97}$$

this gives

$$\Delta(X) = \frac{1}{(2\pi)^n} \prod_{1\le i<j\le n} (\lambda_i - \lambda_j)^2 = \frac{1}{(2\pi)^n}\,\hat{\Delta}(\lambda)^2\,. \tag{8.98}$$

The result for $U(n)$ invariant integration over hermitian matrices becomes

$$\begin{aligned}\int \mathrm{d}^{n^2}X\; f(X) &= \frac{V_{U(n)}}{n!} \int \mathrm{d}^{n^2}X\; \mathcal{F}(X)\Delta(X)\, f(X) \\ &= \frac{V_{U(n)}}{n!\,(2\pi)^n} \int \mathrm{d}^n\lambda\; \hat{\Delta}(\lambda)^2\, \hat{f}(\lambda)\,,\end{aligned} \tag{8.99}$$

where we may use (8.53) for $V_{U(n)}$.

For a Gaussian function

$$\int \mathrm{d}^{n^2}X\; e^{-\frac{1}{2}\kappa\,\mathrm{tr}(X^2)} = 2^{\frac{1}{2}n}\Big(\frac{\pi}{\kappa}\Big)^{\frac{1}{2}n^2}\,, \qquad \mathrm{tr}(X^2) = \sum_{i=1}^{n} X_{ii}{}^2 + 2\sum_{1\le i<j\le n} |X_{ij}|^2 = \sum_{i=1}^{n} \lambda_i^2\,. \tag{8.100}$$

Using (8.53) this defines a normalised probability measure for the eigenvalues for a Gaussian ensemble of hermitian matrices

$$\mathrm{d}\mu(\lambda)_{\text{hermitian matrices}} = \frac{\kappa^{\frac{1}{2}n^2}}{(2\pi)^{\frac{1}{2}n}\,\prod_{i=1}^{n} i!} \prod_{i=1}^{n} \mathrm{d}\lambda_i\; \hat{\Delta}(\lambda)^2\, e^{-\frac{1}{2}\kappa\sum_i \lambda_i{}^2}\,. \tag{8.101}$$

Extending this to quaternionic hermitian $n\times n$ matrices the relevant integrals are

$$\int \mathrm{d}^{n(2n-1)}X\; f(X)\,, \quad X = \bar{X}\,, \quad d^{n(2n-1)}X = \prod_{i=1}^{n} \mathrm{d}X_{ii} \prod_{1\le i<j\le n} \mathrm{d}^4X_{ij}\,, \tag{8.102}$$

where $\bar{X}$ is defined by (1.61) and integration over quaternions is given by (8.56). f is now assumed to satisfy

$$f(X) = f(MXM^{-1}) \,, \qquad M \in U(n,\mathbb{H}) \,. \tag{8.103}$$

Such quaternionic matrices may be diagonalised so that, for a suitable $M \in U(n,\mathbb{H})$,

$$MXM^{-1} = \Lambda \,, \qquad \lambda_i = \bar{\lambda}_i \,. \tag{8.104}$$

Using the correspondence with $2n \times 2n$ complex matrices provided by (1.63) and (1.64), when $M \to \mathcal{M} \in USp(2n,\mathbb{C})$ and $X \to \mathcal{X}$ where

$$\mathcal{X} = \mathcal{X}^\dagger \,, \qquad \mathcal{X} = -J\mathcal{X}^T J \,. \tag{8.105}$$

The eigenvalues of $\mathcal{X}$ must then be $\pm\lambda_i$, $i = 1, \dots n$, and (8.104) is equivalent to the matrix theorem that the $2n \times 2n$ antisymmetric matrix $\mathcal{X}J$ may be reduced to a canonical form in terms of $\{\lambda_i\}$,

$$\mathcal{M}\,\mathcal{X}J\,\mathcal{M}^T = \begin{pmatrix} \begin{smallmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{smallmatrix} & & 0 & & \\ 0 & & \begin{smallmatrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{smallmatrix} & & \\ & & & \ddots & \\ & & & & \begin{smallmatrix} 0 & \lambda_n \\ -\lambda_n & 0 \end{smallmatrix} \end{pmatrix} \quad \text{for} \ \ \mathcal{M} \in U(2n) \,. \tag{8.106}$$

In (8.104) the subgroup of $U(n,\mathbb{H})$ leaving Λ invariant is formed by quaternionic matrices

$$M_0(q) = \begin{pmatrix} q_1 & 0 & \dots & 0 \\ 0 & q_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & q_n \end{pmatrix} , \qquad |q_i| = 1 \,, \tag{8.107}$$

giving

$$H \simeq U(1,\mathbb{H})^{\otimes n} \,. \tag{8.108}$$

As before $W = \mathcal{S}_n$ formed by $\{R_\sigma\} \subset U(n,\mathbb{H})$ which permute the diagonal elements in Λ.

The gauge fixing condition restricting X to diagonal form is now

$$\mathcal{F}(X) = \prod_{1\le i<j\le n} \delta^4(X_{ij}) \,. \tag{8.109}$$

In this case we may write for arbitrary $M \in U(n,\mathbb{H})$,

$$M(\alpha,q) = \exp(\alpha) M_0(q) \,, \qquad \alpha = -\bar{\alpha} \,, \quad \alpha_{ii} = 0 \text{ all } i \,, \tag{8.110}$$

and the group invariant integration is then assumed to be normalised such that, for M as in (8.110),

$$\mathrm{d}\rho_{U(n,\mathbb{H})}\big(M(\alpha,q)\big) \approx \prod_{1\le i<j\le n} \mathrm{d}^4\alpha_{ij} \prod_{i=1}^n \mathrm{d}^4 q_i \, \delta(|q_i| - 1) \,, \qquad \alpha \approx 0 \,. \tag{8.111}$$

With these assumptions and using (8.64)

$$\int_{U(n,\mathbb{H})} \mathrm{d}\rho_{U(n,\mathbb{H})}(M)\, \mathcal{F}\big(MXM^{-1}\big) = (2\pi^2)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4\alpha_{ij}\, \delta^4\big(\alpha_{ij}(\lambda_j-\lambda_i)\big)\,. \tag{8.112}$$

In this case

$$\Delta(X) = \frac{1}{(2\pi^2)^n} \prod_{1\le i<j\le n} (\lambda_i-\lambda_j)^4 = \frac{1}{(2\pi^2)^n}\, \hat\Delta(\lambda)^4\,. \tag{8.113}$$

The result for $U(n,\mathbb{H})$ invariant integration over quaternion hermitian matrices becomes

$$\begin{aligned}\int \mathrm{d}^{n(2n-1)}X\; f(X) &= \frac{V_{Sp(n)}}{n!} \int \mathrm{d}^{n^2}X\; \mathcal{F}(X)\Delta(X)\, f(X)\\ &= \frac{V_{Sp(n)}}{n!\,(2\pi^2)^n} \int \mathrm{d}^n\lambda\; \hat\Delta(\lambda)^4\, \hat f(\lambda)\,,\end{aligned} \tag{8.114}$$

where we may use (8.65) for $V_{Sp(n)}$.

For the Gaussian integral

$$\int \mathrm{d}^{n(2n-1)}X\; e^{-\frac12\kappa\,\mathrm{tr}(X^2)} = 2^{\frac12 n}\Big(\frac{\pi}{\kappa}\Big)^{\frac12 n(2n-1)}\,, \quad \mathrm{tr}(X^2) = \sum_{i=1}^n X_{ii}{}^2 + 2\sum_{1\le i<j\le n} |X_{ij}|^2 = \sum_{i=1}^n \lambda_i^2\,. \tag{8.115}$$

Using (8.65) we therefore obtain a normalised probability measure for the eigenvalues for a Gaussian ensemble of hermitian quaternionic matrices

$$\mathrm{d}\mu(\lambda)_{\text{hermitian quaternionic matrices}} = \Big(\frac{2}{\pi}\Big)^{\frac12 n} \frac{\kappa^{\frac12 n(2n-1)}}{\prod_{i=1}^n (2i)!} \prod_{i=1}^n \mathrm{d}\lambda_i\; \hat\Delta(\lambda)^4\, e^{-\frac12\kappa\sum_i \lambda_i^2}\,. \tag{8.116}$$

8.2.1 Large n Limits

The results for the eigenvalue measure $\mathrm{d}\mu(\lambda)$, given by (8.87), (8.101) and (8.116) for a Gaussian distribution of real symmetric and hermitian complex and quaternion matrices, can be simplified significantly in a limit when n is large. In each case the distribution has the form

$$\mathrm{d}\mu(\lambda) = N_n\, \mathrm{d}^n\lambda\; e^{-W(\lambda)}\,, \qquad W(\lambda) = \frac12\kappa\sum_i \lambda_i^2 - \tfrac12\beta \sum_{i,j,i\ne j} \ln|\lambda_i-\lambda_j|\,, \tag{8.117}$$

where $\beta = 1,2,4$ and we may order the the eigenvalues so that

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n\,. \tag{8.118}$$

For a minimum $W(\lambda)$ is stationary when

$$\kappa\,\lambda_i = \beta \sum_{j\ne i} \frac{1}{\lambda_i-\lambda_j}\,. \tag{8.119}$$

In the large n limit we may approximate λ_i by a smooth function,

$$\lambda_i \to \lambda(x) \, , \quad x = \frac{i}{n} \, , \qquad \sum_{i=1}^{n} = n \int_0^1 \mathrm{d}x = n \int \mathrm{d}\lambda \, \rho(\lambda) \, , \quad \rho(\lambda) = \frac{\mathrm{d}x}{\mathrm{d}\lambda} > 0 \, , \tag{8.120}$$

where $\rho(\lambda)$ determines the eigenvalue distribution and is normalised since

$$\int \mathrm{d}\lambda \, \rho(\lambda) = \int_0^1 \mathrm{d}x = 1 \, . \tag{8.121}$$

As $n \to \infty$ the distribution is dominated by $\lambda(x)$ such that $W(\lambda)$ is close to its minimum. The minimum is determined by (8.119) or, taking the large n limit,

$$\frac{\kappa}{n\beta} \, \lambda = P \int \mathrm{d}\mu \, \rho(\mu) \, \frac{1}{\lambda - \mu} \, , \tag{8.122}$$

where P denotes that the principal part prescription is used for the singularity in the integral at $\mu = \lambda$.

(8.122) is an integral equation for ρ. To solve this we define the function

$$F(z) = \int_{-R}^{R} \mathrm{d}\mu \, \rho(\mu) \, \frac{1}{z - \mu} \sim \frac{1}{z} \quad \text{as} \quad z \to \infty \, , \tag{8.123}$$

using (8.121) and assuming

$$\rho(\mu) > 0 \, , \quad |\mu| < R \, , \qquad \rho(\mu) = 0 \, , \quad |\mu| > R \, . \tag{8.124}$$

$F(z)$ is analytic in z save for a cut along the real axis from $-R$ to R. The integral equation requires

$$F(\mu \pm i\epsilon) = \frac{\kappa}{n\beta} \, \mu \mp i\pi \, \rho(\mu) \, , \qquad |\mu| < R \, . \tag{8.125}$$

Requiring $F(z) = \mathrm{O}(z^{-1})$ for large z this has the unique solution

$$F(z) = \frac{\kappa}{n\beta} \big(z - \sqrt{z^2 - R^2} \big) \, . \tag{8.126}$$

The large z condition in (8.123) requires

$$R^2 = \frac{2n\beta}{\kappa} \, . \tag{8.127}$$

This then gives

$$\rho(\lambda) = \frac{2}{\pi R^2} \, \sqrt{R^2 - \lambda^2} \, . \tag{8.128}$$

This is Wigner's semi-circle distribution and is relevant for nuclear energy levels.

8.3 Integrals over Compact Matrix Groups

Related to the discussion of integrals over group invariant functions of symmetric or hermitian matrices there is a corresponding treatment for integrals over functions of matrices belonging to the fundamental representation for $SO(n)$, $U(n)$ or $Sp(n)$. For simplicity we consider the unitary case first.

For matrices $U \in U(n)$ the essential integral to be considered is then defined in terms of the n^2-dimensional group invariant measure by

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(U)\, f(U)\,, \tag{8.129}$$

where

$$f(U) = f(VUV^{-1}) \qquad \text{for all} \qquad V \in U(n)\,. \tag{8.130}$$

Just as for hermitian matrices U can be diagonalised so that

$$VUV^{-1} = U_0(\theta)\,, \qquad \theta = (\theta_1, \dots, \theta_n)\,, \tag{8.131}$$

where U_0 is defined in (8.91). For θ_i all different V is arbitrary up to $V \sim VU_0(\beta)$, for any $\beta = (\beta_1, \dots, \beta_n)$ so that the associated stability group $H = U(1)^{\otimes n}$. The remaining discrete symmetry group in this case is then

$$W_{U(n)} \simeq \mathcal{S}_n\,, \tag{8.132}$$

since, for any permutation $\sigma \in \mathcal{S}_n$, there is a $R_\sigma \in O(n)$ such that

$$R_\sigma U_0(\theta) R_\sigma{}^{-1} = U_0(\theta_\sigma)\,, \qquad \theta_\sigma = (\theta_{\sigma(1)}, \dots, \theta_{\sigma(n)})\,. \tag{8.133}$$

Thus we use the gauge fixing condition

$$\mathcal{F}(U) = \prod_{1 \le i < j \le n} \delta^2(U_{ij})\,. \tag{8.134}$$

Using the same results as given in (8.94) and (8.95) we then get

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(V)\, \mathcal{F}\big(VUV^{-1}\big) = (2\pi)^n \prod_{1 \le i < j \le n} \int \mathrm{d}^2\alpha_{ij}\, \delta^2\big(\alpha_{ij}(e^{i\theta_j} - e^{i\theta_i})\big)\,, \tag{8.135}$$

so that, using (8.97),

$$\begin{aligned} \Delta(U) = \frac{1}{(2\pi)^n} \prod_{1 \le i < j \le n} |e^{i\theta_j} - e^{i\theta_i}|^2 &= \frac{1}{(2\pi)^n} \prod_{1 \le i < j \le n} \big(2 \sin \tfrac{1}{2}(\theta_i - \theta_j)\big)^2 \\ &= \frac{1}{(2\pi)^n}\, \hat{\Delta}(e^{i\theta}) \hat{\Delta}(e^{-i\theta})\,, \end{aligned} \tag{8.136}$$

with the definition (8.83). The basic formula (8.26) then gives an integration measure over the θ_i's

$$\mathrm{d}\mu_{U(n)}(\theta) = \frac{1}{n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \prod_{1 \le i < j \le n} \big(2 \sin \tfrac{1}{2}(\theta_i - \theta_j)\big)^2\,, \qquad 0 \le \theta_i \le 2\pi\,. \tag{8.137}$$

By restricting $f(U)=1$ in (8.129) it is clear that this integration measure is normalised, $\int \mathrm{d}\mu_{U(n)}(\theta)=1$, since $V_{U(n)}$ may be factored from both sides.

To reduce to $SU(n)$ we let $\theta_i=\theta+\hat\theta_i$, $i=1,\dots,n-1$, $\theta_n=\theta-\sum_{i=1}^{n-1}\hat\theta_i$, where now $0\le\hat\theta_i\le 2\pi$ and $0\le\theta\le 2\pi/n$ and also $\prod_{i=1}^n \mathrm{d}\theta_i = n\mathrm{d}\theta\prod_{i=1}^{n-1}\mathrm{d}\hat\theta_i$. The θ integral may then be factored off, corresponding to the decomposition $U(n)\simeq SU(n)\otimes U(1)/\mathbb{Z}_n$, or equivalently θ_n is no longer an independent variable but determined by $\sum_i\theta_i=0$. For any R_σ if $\det R_\sigma=-1$ we may define $\hat R_\sigma = e^{\pi i/n}R_\sigma$ and otherwise $\hat R_\sigma=R_\sigma$ so that $\{\hat R_\sigma\}\subset SU(n)$ and also $\hat R_\sigma U_0(\theta){R_\sigma}^{-1}=U_0(\theta_\sigma)$. Hence, as in (8.132), we still have

$$W_{SU(n)}\simeq \mathcal{S}_n\,. \tag{8.138}$$

Restricting (8.137) to $SU(n)$ we then obtain

$$\mathrm{d}\mu_{SU(n)}(\theta)=\frac{1}{n!\,(2\pi)^{n-1}}\prod_{i=1}^{n-1}\mathrm{d}\theta_i\prod_{1\le i<j\le n}\big(2\sin\tfrac12(\theta_i-\theta_j)\big)^2\,,\qquad \theta_n=-\sum_{i=1}^{n-1}\theta_i\,. \tag{8.139}$$

For real orthogonal matrices in a similar fashion

$$\int_{SO(n)}\mathrm{d}\rho_{SO(n)}(R)\,f(R)\,,\qquad f(R)=f(SRS^{-1})\quad\text{for all}\quad S\in SO(n)\,. \tag{8.140}$$

In this case it is necessary to distinguish between even and odd n. For any $R\in SO(2n)$ it can be transformed to

$$SRS^{-1}=R_0(\theta)=\begin{pmatrix} r(\theta_1) & 0 & \dots & 0\\ 0 & r(\theta_2) & & \vdots\\ \vdots & & \ddots & \vdots\\ 0 & \dots\dots\dots & & r(\theta_n)\end{pmatrix}\,,\qquad S\in SO(2n)\,, \tag{8.141}$$

where $R_0(\theta)$ is written as a $n\times n$ matrix of 2×2 blocks with

$$r(\theta)=\left(\begin{smallmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{smallmatrix}\right). \tag{8.142}$$

In (8.141) $S\sim SR_0(\beta)$, for arbitrary $\beta=(\beta_1,\dots,\beta_n)$, so that the stability group for $R_0(\theta)$ is then $SO(2)^{\otimes n}$. The discrete group defined by $2n\times 2n$ matrices $\{S\}\in O(2n)$ such that $SR_0(\theta)S^{-1}=R_0(\theta')$ is $\mathcal{S}_n\ltimes{\mathbb{Z}_2}^{\otimes n}$, with the permutation group $\mathcal{S}_n$ formed by $\{R_\sigma\otimes I_2\}$ and ${\mathbb{Z}_2}^{\otimes n}$ generated by

$$R_i = {}_i\!\begin{pmatrix} I_2 & \dots & & & & \dots & 0\\ \vdots & \ddots & & & & & \vdots\\ \vdots & & I_2 & & & & \vdots\\ \vdots & & & \sigma_3 & & & \vdots\\ \vdots & & & & I_2 & & \vdots\\ \vdots & & & & & \ddots & \vdots\\ 0 & \dots & & & & \dots & I_2\end{pmatrix}^{\!\!\;i}\in O(2n)\,,\quad i=1,\dots,n\,,\qquad R_i^2=I_{2n}\,, \tag{8.143}$$

since $\sigma_3\, r(\theta)\, \sigma_3 = r(2\pi - \theta)$, for $\sigma_3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Restricting to the subgroup formed my matrices with determinant one

$$W_{SO(2n)} \simeq \big(\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\otimes n}\big)/\mathbb{Z}_2\,. \tag{8.144}$$

Writing $R \in SO(2n)$ in terms of $2{\times}2$ blocks R_{ij}, $i,j = 1,\dots n$, the gauge fixing condition is then taken as

$$\mathcal{F}(R) = \prod_{1\le i<j\le n} \delta^4(R_{ij})\,, \tag{8.145}$$

with the definitions

$$\delta^4(A) = \delta(a)\delta(b)\delta(c)\delta(d)\,, \quad \mathrm{d}^4 A = \mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c\,\mathrm{d}d \qquad \text{for} \qquad A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right). \tag{8.146}$$

For a general rotation $S \in SO(2n)$ we may write

$$S = e^A S_0(\beta)\,, \qquad A^T = -A\,, \qquad A_{ii} = 0 \ \text{ all } i\,, \tag{8.147}$$

and then

$$\mathrm{d}\rho_{SO(2n)}(S) \approx \prod_{1\le i<j\le n} \mathrm{d}^4 A_{ij} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad A \approx 0\,. \tag{8.148}$$

Using (8.148) is then sufficient to obtain

$$\int_{SO(2n)} \mathrm{d}\rho_{SO(2n)}(S)\ \mathcal{F}\big(SR_0(\theta)S^{-1}\big) = (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4 A_{ij}\ \delta^4\big(A_{ij} r(\theta_j) - r(\theta_i) A_{ij}\big)\,. \tag{8.149}$$

With

$$\delta^4\big(Ar(\theta) - r(\theta')A\big) = \frac{1}{4(\cos\theta - \cos\theta')^2}\,\delta^4(A)\,, \tag{8.150}$$

we then get for $SO(2n)$

$$\Delta(R) = \frac{1}{(2\pi)^n} \prod_{1\le i<j\le n} \big(2(\cos\theta_i - \cos\theta_j)\big)^2 = \frac{1}{(2\pi)^n}\,\big(\hat{\Delta}(2\cos\theta)\big)^2\,, \tag{8.151}$$

where $\hat{\Delta}$ is defined by (8.83).

Combining the ingredients the measure for integration reduces in the $SO(2n)$ case to an integral over the n θ_i's given by

$$\mathrm{d}\mu_{S0(2n)}(\theta) = \frac{1}{2^{n-1} n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i\ \big(\hat{\Delta}(2\cos\theta)\big)^2\,. \tag{8.152}$$

For $SO(2n{+}1)$ (8.141) may be modified, by introducing one additional row and column, to

$$SRS^{-1} = R_0(\theta) = \begin{pmatrix} r(\theta_1) & 0 & \dots & 0 & 0 \\ 0 & r(\theta_2) & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & & & r(\theta_n) & 0 \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}, \qquad S \in SO(2n+1)\,, \tag{8.153}$$

with $r(\theta)$ just as in (8.142). Instead of (8.143) we may now take

$$
R_i = i \begin{pmatrix} I_2 & \cdots & & \overset{i}{} & \cdots & 0 & 0 \\ \vdots & \ddots & & & & & \vdots \\ \vdots & & I_2 & & & & \vdots \\ \vdots & & & \sigma_3 & & & \vdots \\ \vdots & & & & I_2 & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & & & & & I_2 & 0 \\ 0 & \cdots & & & \cdots & & -1 \end{pmatrix} \in SO(2n+1)\,, \quad i = 1, \dots, n\,, \tag{8.154}
$$

and in a similar fashion, for any permutation $\sigma \in \mathcal{S}_n$, there is a $R_\sigma \in SO(2n+1)$, with the matrix R_σ having $1, -1$ in the bottom right hand corner according to whether σ is even,odd, such that $R_\sigma R_0(\theta) R_\sigma{}^{-1} = R_0(\theta_\sigma)$. Hence

$$
W_{SO(2n+1)} \simeq \mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\otimes n}\,. \tag{8.155}
$$

In this case $R \in SO(2n+1)$ is expressible in terms of 2×2 blocks R_{ij}, $i, j = 1, \dots n$, 2×1 blocks $R_{i\, n+1}$ and also 1×2 blocks $R_{n+1\, i}$ for $i = 1, \dots n$. The gauge fixing condition is now

$$
\mathcal{F}(R) = \prod_{1 \le i < j \le n} \delta^4(R_{ij}) \prod_{i=1}^{n} \delta^2(R_{i\, n+1})\,, \tag{8.156}
$$

with $\delta^2\big(\begin{smallmatrix} a \\ b \end{smallmatrix}\big) = \delta(a)\, \delta(b)$, similarly to (8.146). Expressing $S \in SO(2n+1)$ in the same form as (8.147) we now have

$$
\mathrm{d}\rho_{SO(2n+1)}(S) \approx \prod_{1 \le i < j \le n} \mathrm{d}^4 A_{ij} \prod_{i=1}^{n} \mathrm{d}^2 A_{i\, n+1} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad A \approx 0\,, \tag{8.157}
$$

so that

$$
\begin{aligned}
&\int_{SO(2n+1)} \mathrm{d}\rho_{SO(2n+1)}(S)\; \mathcal{F}\big(S R_0(\theta) S^{-1}\big) \\
&= (2\pi)^n \prod_{1 \le i < j \le n} \int \mathrm{d}^4 A_{ij}\; \delta^4\big(A_{ij} r(\theta_j) - r(\theta_i) A_{ij}\big) \prod_{i=1}^{n} \int \mathrm{d}^2 A_{i\, n+1}\; \delta^2\big((I_2 - r(\theta_i)) A_{i\, n+1}\big)\,.
\end{aligned} \tag{8.158}
$$

In the $SO(2n+1)$ case this implies

$$
\Delta(R) = \frac{1}{(2\pi)^n} \big(\hat{\Delta}(2\cos\theta)\big)^2 \prod_{i=1}^{n} \big(2 \sin \tfrac{1}{2}\theta_i\big)^2\,, \tag{8.159}
$$

and in consequence

$$
\mathrm{d}\mu_{S0(2n+1)}(\theta) = \frac{1}{2^n n!\, (2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \big(2 \sin \tfrac{1}{2}\theta_i\big)^2 \, \big(\hat{\Delta}(2\cos\theta)\big)^2\,. \tag{8.160}
$$

The remaining case to consider is for integrals over $M \in Sp(n) \simeq U(n,\mathbb{H})$ of the form

$$\int_{Sp(n)} \mathrm{d}\rho_{Sp(n)}(M)\, f(M)\,, \qquad f(M) = f(NMN^{-1}) \quad \text{for all} \quad N \in Sp(n)\,. \tag{8.161}$$

By a suitable transformation the quaternion matrix M can be reduced to the diagonal form

$$NMN^{-1} = M_0(\theta) = \begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & e^{i\theta_n} \end{pmatrix}, \qquad N \in Sp(n)\,, \tag{8.162}$$

As before $N \sim NM_0(\beta)$ so the stability group is $U(1)^{\otimes n}$. The remaining discrete group generated by $R_\sigma 1 \in Sp(n)$, for $\sigma \in \mathcal{S}_n$ and 1 the unit quaternion, and also by

$$N_i = \begin{matrix} \\ \\ \\ i \\ \\ \\ \end{matrix}\overset{\displaystyle i}{\begin{pmatrix} 1 & \dots\dots\dots\dots & & & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & & j & \vdots \\ \vdots & & & 1 & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots\dots\dots\dots & & & 1 \end{pmatrix}} \in Sp(n)\,, \quad i = 1,\dots,n\,. \tag{8.163}$$

In this case $N_i^2 = N_0(\beta)$, with $e^{i\beta_i} = -1$. $e^{i\beta_j} = 1$, $j \neq i$, so that N_i corresponds to a $\mathbb{Z}_2$ symmetry. Hence for $Sp(n)$ we have

$$W_{Sp(n)} \simeq \mathcal{S}_n \ltimes \mathbb{Z}_2^{\otimes n}\,. \tag{8.164}$$

For the $Sp(n)$ case we take

$$\mathcal{F}(M) = \prod_{1\le i<j\le n} \delta^4(M_{ij}) \prod_{i=1}^{n} \delta^2(M_{ii})\,, \tag{8.165}$$

where, for any quaternion q, $\delta^4(q)$ is defined as in (8.58) and also here

$$\delta^2(q) = \delta(u)\delta(v) \qquad \text{for} \qquad q = x + iy + ju + kv\,. \tag{8.166}$$

Writing then, for any $N \in Sp(n)$,

$$N = e^\alpha M_0(\beta)\,, \qquad \alpha_{ij} = -\bar\alpha_{ji} \in \mathbb{H}\,,\ i \neq j\,, \quad \alpha_{ii} = ju_i + kv_i\,, \tag{8.167}$$

we have

$$\mathrm{d}\rho_{Sp(n)}(N) \approx \prod_{1\le i<j\le n} \mathrm{d}^4\alpha_{ij} \prod_{i=1}^{n} \mathrm{d}^2\alpha_{ii} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad \alpha \approx 0\,, \tag{8.168}$$

so that

$$\begin{aligned} &\int_{Sp(n)} \mathrm{d}\rho_{Sp(n)}(N)\, \mathcal{F}\big(NM_0(\theta)N^{-1}\big) \\ &= (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4\alpha_{ij}\, \delta^4\big(\alpha_{ij}\, e^{i\theta_j} - e^{i\theta_i}\, \alpha_{ij}\big) \prod_{i=1}^{n} \int \mathrm{d}^2\alpha_{ii}\, \delta^2\big(\alpha_{ii}\, e^{i\theta_i} - e^{i\theta_i}\, \alpha_{ii}\big)\,. \end{aligned} \tag{8.169}$$

For this case we may use

$$
\begin{aligned}
\delta^4\big(\alpha\, e^{i\theta} - e^{i\theta'}\,\alpha\big) &= \frac{1}{4(\cos\theta-\cos\theta')^2}\,\delta^4\big(\alpha\big)\,,\\
\delta^2\big(\alpha\, e^{i\theta} - e^{i\theta}\,\alpha\big) &= \frac{1}{4\sin^2\theta}\,\delta^2\big(\alpha\big) \quad \text{for} \quad \alpha = ju+kv\,,
\end{aligned}
\tag{8.170}
$$

to obtain

$$
\Delta(M) = \frac{1}{(2\pi)^n}\,\big(\hat\Delta(2\cos\theta)\big)^2 \prod_{i=1}^n \big(2\sin\theta_i\big)^2\,. \tag{8.171}
$$

Hence

$$
\mathrm{d}\mu_{Sp(n)}(\theta) = \frac{1}{2^n n!\,(2\pi)^n} \prod_{i=1}^n \mathrm{d}\theta_i\,\big(2\sin\theta_i\big)^2\,\big(\hat\Delta(2\cos\theta)\big)^2\,. \tag{8.172}
$$

As special cases we have $\mathrm{d}\mu_{Sp(1)}(\theta) = \mathrm{d}\mu_{SU(2)}(\theta)$, $\mathrm{d}\mu_{SO(3)}(\theta) = 2\,\mathrm{d}\mu_{Sp(1)}(\frac{1}{2}\theta)$ and also, from $SO(4) \simeq (Sp(1)\otimes Sp(1))/\mathbb{Z}_2$, $\mathrm{d}\mu_{SO(4)}(\theta_1-\theta_2,\theta_1+\theta_2) = 2\,\mathrm{d}\mu_{Sp(1)}(\theta_1)\,\mathrm{d}\mu_{Sp(1)}(\theta_2)$ with, from $SO(5) \simeq Sp(2)/\mathbb{Z}_2$, $\mathrm{d}\mu_{SO(5)}(\theta_1-\theta_2,\theta_1+\theta_2) = 2\,\mathrm{d}\mu_{Sp(2)}(\theta_1,\theta_2)$, and, from $SO(6)\simeq SU(4)/\mathbb{Z}_2$, $\mathrm{d}\mu_{SO(6)}(\theta_2+\theta_3,\theta_3+\theta_1,\theta_1+\theta_2) = 2\,\mathrm{d}\mu_{SU(4)}(\theta_1,\theta_2,\theta_3)$.

8.4 Integration over a Gauge Field and Gauge Fixing

An example where the reduction of a functional integral over a gauge field $A\in\mathcal{A}$ can be reduced to $\mathcal{A}/\mathcal{G}$, where $\mathcal{G}$ is a the gauge group, in an explicit fashion arises in just one dimension. We then consider a gauge field $A(t)$ with the gauge transformation, following (7.26),

$$
A(t) \underset{g}{\longrightarrow} A(t)^{g(t)} = g(t)A(t)g(t)^{-1} - \partial_t g(t)\, g(t)^{-1}\,, \tag{8.173}
$$

where here we take

$$
A(t) = -A(t)^\dagger \in \mathfrak{u}(n)\,, \qquad g(t)\in U(n)\,. \tag{8.174}
$$

The essential functional integral has the form

$$
\int \mathrm{d}[A]\; f(A)\,, \qquad f(A^g) = f(A)\,, \tag{8.175}
$$

where we restrict to $t\in S^1$ by requiring the fields to satisfy the periodicity conditions

$$
A(t) = A(t+\beta)\,, \qquad g(t) = g(t+\beta)\,. \tag{8.176}
$$

In one dimension there are no local gauge invariants. However if we define

$$
U = \mathcal{T}\big\{e^{-\int_0^\beta \mathrm{d}t\, A(t)}\big\} \in U(n)\,, \tag{8.177}
$$

where $\mathcal{T}$ denotes t-ordering, then, as a consequence of the discussion in 7.3 and the periodicity requirement (8.176), the gauge invariant function f in (8.175) should have the form

$$
f(A) = \hat f(U) \qquad \text{where} \qquad \hat f(U) = \hat f(g\,Ug^{-1}) \quad \text{for all} \quad g\in U(n)\,. \tag{8.178}
$$

In particular

$$P_\beta(U) = \mathrm{tr}(U)\,, \tag{8.179}$$

is gauge invariant, being just the Wilson loop for the circle S^1 arising from imposing periodicity in t. $P_\beta(U)$ is a *Polyakov*[29] *loop.*

The general discussion for finite group invariant integrals can be directly applied to the functional integral (8.175). It is necessary to choose a convenient gauge fixing condition. For any $A(t)$ there is a gauge transformation $g(t)$ such that

$$A(t)^{g(t)} = iX\,, \qquad X^\dagger = X\,. \tag{8.180}$$

In consequence we may choose a gauge condition $\partial_t A(t) = 0$ or equivalently take

$$\mathcal{F}[A] = \delta'[A]\,, \tag{8.181}$$

where $\delta'[A]$ is a functional δ-function, δ' denoting the exclusion of constant modes. For a general Fourier expansion on S^1

$$A(t) = iX + \sum_{n\neq 0} A_n\, e^{2\pi int/\beta}\,, \qquad X^\dagger = X\,, \quad A_n{}^\dagger = -A_{-n}\,, \tag{8.182}$$

where X is a hermitian and A_n are complex $n\times n$ matrices, then

$$\delta'[A] = \prod_{n>0} \mathcal{N}_n\, \delta^{2n^2}\big(A_n\big)\,. \tag{8.183}$$

$\mathcal{N}_n$ is a normalisation factor which is chosen later. With the expansion (8.182) the functional integral can also be defined by taking

$$\mathrm{d}[A] = \mathrm{d}^{n^2}X \prod_{n>0} \frac{1}{\mathcal{N}_n}\, \mathrm{d}^{2n^2}A_n\,. \tag{8.184}$$

The integral (8.9) defining the Faddeev Popov determinant then becomes

$$\int_{\mathcal{G}} \mathrm{d}\mu(g)\ \delta'[A^g] \quad \text{where} \quad A(t) = (iX)^{g(t)} \quad \text{for some} \quad g(t)\,, \tag{8.185}$$

and where $\mathrm{d}\mu(g)$ is the invariant measure for the gauge group $\mathcal{G}$. From (8.173) for an infinitesimal gauge transformation

$$(iX)^{g(t)} = iX + i[\lambda(t), X] - \partial_t\lambda(t) \qquad \text{for} \quad g(t) \approx I + \lambda(t)\,, \quad \lambda(t)^\dagger = -\lambda(t)\,. \tag{8.186}$$

If

$$g(t) = g_0\big(I + \lambda(t)\big) \quad \text{for} \quad \lambda(t)\approx 0\,, \quad \lambda(t) = \sum_{n\neq 0}\lambda_n\, e^{2\pi int/\beta}\,, \quad \lambda_n{}^\dagger = -\lambda_{-n}\,, \tag{8.187}$$

then we may take

$$\mathrm{d}\mu(g) \approx \mathrm{d}\rho_{U(n)}(g_0)\ \mathrm{d}[\lambda]\,, \qquad \mathrm{d}[\lambda] = \prod_{n>0} \mathrm{d}^{2n^2}\lambda_n\,. \tag{8.188}$$

[29] Alexandre M. Polyakov, 1945-, Russian.

Hence from (8.185) we define

$$\int_{\mathcal{G}} \mathrm{d}\mu(g)\; \delta'\big[(iX)^g\big] = \frac{V_{U(n)}}{\Delta(X)}\,, \tag{8.189}$$

where

$$\begin{aligned}\frac{1}{\Delta(X)} &= \int \mathrm{d}[\lambda]\; \delta'\big[i[\lambda, X] - \partial_t \lambda\big] = \prod_{n>0} \mathcal{N}_n \int \mathrm{d}^{2n^2}\lambda_n\; \delta^{2n^2}\Big(\frac{2\pi n}{i\beta}\lambda_n - i[X,\lambda_n]\Big)\\ &= \prod_{n>0} \int \mathrm{d}^{2n^2}\lambda_n\; \delta^{2n^2}\Big(\lambda_n + \frac{\beta}{2\pi n}[X,\lambda_n]\Big) \quad \text{for} \quad \mathcal{N}_n = \Big(\frac{\beta}{2\pi n}\Big)^{2n^2}\,,\end{aligned} \tag{8.190}$$

which gives

$$\Delta(X) = \prod_{n>0} \bigg(\det\Big(I_{n^2} + \frac{\beta}{2\pi n}\, X^{\mathrm{ad}}\Big)\bigg)^2\,. \tag{8.191}$$

The essential functional integral in (8.175) then reduces to just an integral over hermitian matrices X,

$$\int \mathrm{d}[A]\; f(A) = \frac{1}{V_{U(n)}} \int \mathrm{d}^{n^2}X\; \Delta(X)\, f(iX)\,. \tag{8.192}$$

There is a remaining invariance under $X \to gXg^{-1}$ for constant $g \in U(n)$. This may be used to diagonalise X so that $gXg^{-1} = \Lambda$ where Λ is the diagonal matrix in terms of the eigenvalues $\lambda_1, \dots, \lambda_n$, as in (8.68). In terms of these

$$\text{eigenvalues}\{X^{\mathrm{ad}}\} = \lambda_i - \lambda_j\,, \quad i,j = 1,\dots,n\,. \tag{8.193}$$

Hence

$$\det\Big(I_{n^2} + \frac{\beta}{2\pi n}\, X^{\mathrm{ad}}\Big) = \prod_{1\le i<j\le n} \Big(1 - \frac{(\lambda_i - \lambda_j)^2\beta^2}{4\pi^2 n^2}\Big)\,. \tag{8.194}$$

Using

$$\prod_{n>0}\Big(1 - \frac{\theta^2}{\pi^2 n^2}\Big) = \frac{\sin\theta}{\theta}\,, \tag{8.195}$$

we get

$$\Delta(X) = \prod_{1\le i<j\le n} \bigg(\frac{\sin\frac{1}{2}(\lambda_i - \lambda_j)\beta}{\frac{1}{2}(\lambda_i - \lambda_j)\beta}\bigg)^2\,. \tag{8.196}$$

As a consequence of (8.99) we further express (8.192) in terms of an integral over the eigenvalues $\{\lambda_i\}$ using

$$\frac{1}{V_{U(n)}} \int \mathrm{d}^{n^2}X \to \frac{1}{n!\,(2\pi)^n} \int \mathrm{d}^n\lambda \prod_{1\le i<j\le n} (\lambda_i - \lambda_j)^2\,. \tag{8.197}$$

Using this in conjunction (8.196) in (8.192) gives finally

$$\int \mathrm{d}[A]\; f(A) = \frac{1}{\beta^{n^2}} \int \mathrm{d}\mu_{U(n)}(\beta\lambda)\; f(i\Lambda)\,, \tag{8.198}$$

with the measure for integration over $U(n)$ determined by (8.137).

Although the freedom of constant gauge transformations has been used in transforming $X \to \Lambda$ there is also a residual gauge freedom given by

$$g(t) = e^{2\pi i r t/\beta} I \,, \quad r = 0, \pm 1, \pm 2, \dots \qquad \Rightarrow \qquad \Lambda^{g(t)} = \Lambda - \frac{2\pi r}{\beta} I \,. \tag{8.199}$$

For this to be a symmetry for $f(iX) = f(i\Lambda)$ we must have

$$f(iX) = \hat{f}(e^{-i\beta X}) \,, \tag{8.200}$$

where $\hat{f}$ is defined in terms of the line integral over t in (8.178). The final result (8.198) shows that the functional integral over $A(t)$ reduces after gauge fixing just to invariant integration over the unitary matrix U.