Group Theory Lecture Notes

Hugh Osborn

latest update: November 23, 2023

Based on part III lectures Symmetries and Groups, Michaelmas Term 2008, revised and extended at various times subsequently

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.
P. Ramond, *Group Theory, A Physicists Survey,* CUP (2010).
A relatively gentle physics motivated treatment, and includes discussion of finite groups.
A. Zee, *Group Theory in a Nutshell for Physicists.* Princeton University Press (2016).
Quite lengthy, comprehensive with many physics applications, some nice anecdotal remarks.
P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Lie Groups*, Princeton University Press (2009), http://birdtracks.eu
Idiosyncratic, but full of material not found elswhere. Great for doing calculations.

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 references 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. It introduces anti-unitary representations. 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 Isopspin 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, 1929-2019, 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 $1, \ldots m$, 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. $\mathbb{1}$ is the unit matrix, on occasion $\mathbb{1}_n$ denotes the $n \times n$ unit matrix.

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

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

If some indices are to be omitted from symmetrisation or antisymmetrisation they are surrounded by $|\ldots|$, thus $T_{[i|k|j]} = \frac{1}{2}(T_{ikj} - T_{jki})$.

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 $\{v_r\}$, $r = 1, \ldots, \dim \mathcal{V}$ 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 and we require the equivalence relations $(v_1 + v'_2, v_2) \sim (v_1, v_2) + (v'_1, v_2)$, $(v_1, v_2 + v'_2) \sim (v_1, v_2) + (v_1, v'_2)$, $c(v_1, v_2) \sim (c\, v_1, v_2) \sim (v_1, c\, v_2)$ so as to extend the vector space properties from $\mathcal{V}_1, \mathcal{V}_2$ to $\mathcal{V}_1 \otimes \mathcal{V}_2$. An arbitrary vector in $\mathcal{V}_1 \otimes \mathcal{V}_2$ is then a linear combination $v = \sum_{r,s} a_{rs}\, (v_{1r}, v_{2s})$ so that $\dim(\mathcal{V}_1 \otimes \mathcal{V}_2) = \dim \mathcal{V}_1 \ \dim \mathcal{V}_2$. The tensor product $\mathcal{V} \otimes \mathcal{V}$ can be decomposed into symmetric snd antisymmetric subspaces $\vee^2 \mathcal{V}$ and $\wedge^2 \mathcal{V}$ with bases $(v_r, v_s) + (v_s, v_r)$ and $(v_r, v_s) - (v_s, v_r)$ respectively, corresponding to $a_{rs} = a_{(rs)}$ and $a_{rs} = a_{[rs]}$. $\dim \vee^2 \mathcal{V} = \frac{1}{2} \dim \mathcal{V}(\dim \mathcal{V} + 1)$ and $\dim \wedge^2 \mathcal{V} = \frac{1}{2} \dim \mathcal{V}(\dim \mathcal{V} - 1)$.

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 $v_i \in \mathcal{V}_i$, and where $(v_1 + v'_1) + (v_2 + v'_2) = (v_1 + v_2) + (v'_1 + v'_2)$ and $c(v_1 + v_2) = c\, v_1 + c\, v_2$. Equivalently it has a basis $\{v_{1r}, v_{2s}\}$ and an arbitrary vector in $\mathcal{V}_1 \otimes \mathcal{V}_2$ has the form $v = \sum_r a_r\, v_{1r} + \sum_s b_s\, v_{2s}$ so that $\dim(\mathcal{V}_1 \oplus \mathcal{V}_2) = \dim \mathcal{V}_1 + \dim \mathcal{V}_2$.

1 Introduction, Definitions and Examples

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

For any $g \in G$ the smallest integer m such that $g^m = e$ is the *order* of g.

Two groups $G = \{g_i\}$ and $G' = \{g'_j\}$ are *isomorphic*, $G \simeq G'$, if there is a one to one correspondence $\theta : g_i \leftrightarrow g'_j$ between the elements consistent with the group multiplication rules. Even if $G \simeq G'$ there is not necessarily a unique choice for θ but of course we must have $\theta : e \leftrightarrow e'$.

A crucial consequence of the basic group axioms is

$$\{g_i g\} = \{g_i\} \quad \text{for any} \quad g \quad \text{since } g_j g = g_i g \quad \Rightarrow \quad g_j = g_i \,, \tag{1.6}$$

which implies for a finite group

$$\textstyle\sum_i f(g_i) = \sum_i f(g_i g) \,. \tag{1.7}$$

1.1.1 Subgroups and Cosets

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. A *proper* subgroup H is when $H \neq G$ and is denoted $H < G$. For any subgroup H there is an equivalence relation between $g_i, g'_i \in G$,

$$g_i \sim g'_i \quad \Leftrightarrow \quad g'_i = g_i h \quad \text{for} \quad h \in H \,. \tag{1.8}$$

Each equivalence class $\{g_i\}$ defines a *left coset* and has $|H|$ elements. Of course $\{e\} = H$. There can also be right cosets where in (1.8) we take $g_i = h\, g'_i$ instead. The cosets form the *coset space* G/H,

$$G/H \simeq G/\sim \; = \{gH : g \in G, g \sim g' \;\; \text{if } g' = gh, h \in H\} \,. \tag{1.9}$$

Since each coset is distinct

$$\dim G/H = |G/H| = |G|/|H| \,. \tag{1.10}$$

The fact that, for any subgroup $H \subset G$, $|H|$ divides $|G|$ is *Lagrange's theorem.*[1] The *index* of the subgroup H in G is the number of cosets in G/H, denoted $G : H = |G|/|H|$. The index is also a divisor of $|G|$. In general left and right cosets are different.

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

1.1.2 Normal Subgroups, Quotient Group, Simple Groups and Composition Series

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

$$gNg^{-1} = N \quad \text{for all} \quad g \in G \,. \tag{1.11}$$

This may be denoted by $N \lhd G$ (or $G \rhd N$). In this case G/N becomes a group since for $g'_i = g_i h_i, g'_j = g_j h_j$, with $h_i, h_j \in N$, then $g'_i g'_j = g_i g_j h$ for some $h \in N$. $Q = G/N$ is then the *quotient group*, or sometimes the *factor group*. The group G is an *extension* of Q by N. The quotient group is expressible in terms of cosets by

$$G/N \simeq \{gN/\sim \, : g \in G, g \sim gh, h \in N\} \,, \tag{1.12}$$

where elements of the quotient group satisfy

$$(gN)\,(g'N) = (gg'N)\,, \quad (gN)^{-1} = (g^{-1}N)\,, \quad e = (N)\,, \tag{1.13}$$

with the group multiplication rule and inverse following from $Ng = gN$, for N a normal subgroup, and $N^2 = N$. In general the quotient group Q is not a subgroup of G. For an abelian group all subgroups are necessarily normal subgroups. A normal subgroup N is maximal if there is no $N' \neq N, G$ such that $N \lhd N' \lhd G$. As shown later there can be more than one maximal normal subgroup.

For a normal subgroup the left and right cosets are identical since $gh = h'g$ for any $g \in G$ and for any $h \in N$ there is a corresponding $h' \in N$.

If H is a subgroup of G and $|G|/|H| = 2$ then H has to be a normal subgroup since the right coset other than H has to be equal to the left coset. In this case the quotient group $G/H \simeq \mathbb{Z}_2$ and for $g, g' \in G$, $g, g' \notin H$ then $gg' \in H$. In this case H is sometimes called the *halving subgroup.*

A group G is *simple* if the only normal subgroups are G and the trivial subgroup formed by the identity $\{e\}$ by itself. Simple groups are the building blocks for finite groups. If N is

[1] Joseph-Louis Lagrange, born Giuseppe Luigi Lagrangia, 1736-1813, French, after Italian.

a maximal normal subgroup of G then the quotient G/N is a simple group. To verify this if N is not maximal there is a normal subgroup $N' \lhd G$ such that N is a subgroup of N'.. Since N is a normal subgroup of G it must also be a normal subgroup of N', $N \lhd N'$. The corresponding quotient groups can be expressed as

$$G/N \simeq \{gN/\sim : g \in G, g \sim gh, h \in N\}\,, \quad N'/N \simeq \{gN/\sim : h' \in N', h' \sim h'h, h \in N\}\,. \quad (1.14)$$

Directly

$$(gN)\, h'N\, (gN)^{-1} = (gh'g^{-1})N \in N'/N\,, \quad (1.15)$$

since N' is a normal subgroup. Thus $N'/N \lhd G/N$. Conversely if G/N has a non trivial normal subgroup then there is a normal subgroup $N' \lhd G$ with N a non trivial normal subgroup of N'. Hence if N is maximal G/N is simple.

For any G there is a *composition series* of successive maximal normal subgroups $N_i \lhd N_{i-1}$, $N_0 = G$ where

$$N_n \lhd \cdots \lhd N_1 \lhd G\,, \qquad N_n = \{e\}\,, \quad (1.16)$$

and all quotients N_{i-1}/N_i are simple groups. The composition series is not necessarily unique but all composition series for G have the same length n. Of course simple groups themselves have length one.

1.1.3 Direct Product of Groups

For two groups G_1, G_2 we may define a *direct product* group $G_1 \times G_2$ formed by pairs of elements $\{(g_1, g_2)\}$, belonging to (G_1, G_2), which is defined by the rules

$$(g_1, g_2)(g'_1, g'_2) = (g_1 g'_1, g_2 g'_2)\,, \quad (g_1, g_2)^{-1} = (g_1^{-1}, g_2^{-1})\,, \quad e = (e_1, e_2)\,. \quad (1.17)$$

So long as it is clear which elements belong to G_1 and which to G_2 we may write the elements of $G_1 \times G_2$ as just $g_1 g_2 = g_2 g_1$ and $e = e_1 e_2$. For finite groups $|G_1 \times G_2| = |G_1|\,|G_2|$. In any direct product $G_1 \times G_2$ then $G_1 \simeq \{(g_1, e_2)\}$, $G_2 \simeq \{(e_1, g_2)\}$ are both normal subgroups. For the direct product $G \times G$ then G is of course a subgroup but there is also the *diagonal* subgroup G formed from elements $\{(g, g)\}$ which is not a normal subgroup of $G \times G$.

1.2 Cyclic, Dihedral and Permutation Groups

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

1.2.1 Cyclic Group

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. The cyclic group C_n is also defined by the complex numbers $e^{2\pi i r/n}$, $r = 0, \dots, n-1$, of modulus one, under multiplication. Clearly it is abelian and $C_n \simeq \mathbb{Z}_n$. Abstractly $\mathbb{Z}_n$, C_n both can be defined by

$$\mathbb{Z}_n \simeq C_n = \left\{a^r : r = 0, 1, \dots, n,\ a^0 = a^n = e\right\}. \quad (1.18)$$

Obviously $\mathbb{Z}_1$ is the trivial one element group. For p prime $\mathbb{Z}_p$ has no subgroups, since p has no divisors, and hence $\mathbb{Z}_p$ is simple. If $n = pq$ then $\mathbb{Z}_p$ and $\mathbb{Z}_q$ are normal subgroups and $\mathbb{Z}_{pq}/\mathbb{Z}_p \simeq \mathbb{Z}_q$. If p, q are coprime (no common factors) $\mathbb{Z}_{pq} \simeq \mathbb{Z}_p \times \mathbb{Z}_q$ and both $\mathbb{Z}_p$, $\mathbb{Z}_q$ are maximal normal subgroups.

For further illustration we consider $\mathbb{Z}_2 \times \mathbb{Z}_4$ where $\mathbb{Z}_2 = \{e, b\}$ with $b^2 = e$ and $\mathbb{Z}_4 = \{e, a, a^2, a^3\}$ with $a^4 = e$. This has proper subgroups, which by Lagrange's theorem can only have order 2 or 4,

$$\mathbb{Z}_2 = \{e, b\},\ \{e, a^2\},\ \{e, ba^2\},\ \ \mathbb{Z}_4 = \{e, a, a^2, a^3\},\ \{e, ba, a^2, ba^3\},\ \ K_4 = \{e, b, a^2, ba^2\}\,. \tag{1.19}$$

The group $K_4 = \mathbb{Z}_2 \times \mathbb{Z}_2$ is not cyclic, it is the *Klein*[2], group, it has 4 elements all but the identity of order 2. The list in (1.19) is not just a sum of direct products of subgroups of $\mathbb{Z}_2$ and $\mathbb{Z}_4$, as demonstrated by the subgroups $\{e, ba^2\}$ and $\{e, ba, a^2, ba^3\}$.

For $\mathbb{Z}_{12}$ there are three possible composition series $\mathbb{Z}_1 \triangleleft \mathbb{Z}_2 \triangleleft \mathbb{Z}_6 \triangleleft \mathbb{Z}_{12}$, $\mathbb{Z}_1 \triangleleft \mathbb{Z}_2 \triangleleft \mathbb{Z}_4 \triangleleft \mathbb{Z}_{12}$ or $\mathbb{Z}_1 \triangleleft \mathbb{Z}_3 \triangleleft \mathbb{Z}_6 \triangleleft \mathbb{Z}_{12}$, all of length 3. For $\mathbb{Z}_n$ the sizes of the quotient groups in the composition series correspond to the prime factors of n and the length of the composition series is the number of prime factors, allowing for multiplicity.

1.2.2 Dihedral Group

The dihedral group D_n, of order $2n$, is the symmetry group for a regular n-sided polygon and is formed by rotations a through angles $2\pi r/n$ together with reflections b. In general

$$D_n = \left\{a^r, a^r b : r = 0, 1, \ldots, n-1,\ a^0 = a^n = e,\ b^2 = e,\ ab = ba^{n-1}\right\}. \tag{1.20}$$

For any r $(a^r b)^2 = e$. For $n > 2$ the group is non abelian since $ba \neq ab$, note that $D_2 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$. In general $\{e, b\} \simeq \mathbb{Z}_2$ is a subgroup of D_n.

The centre of D_n depends on whether n is even or odd, $\mathcal{Z}(D_{2n+1}) = \{e\}$ whereas $\mathcal{Z}(D_{2n}) = \{e, a^n\} \simeq \mathbb{Z}_2$.

The normal subgroups of D_n also depend on n. If k divides n, $k|n$, then the abelian group $\mathbb{Z}_k = \left\{a^{n/k\, r} : r = 0, 1, \ldots, k-1\right\}$ is a normal subgroup. This includes $\mathbb{Z}_n$. Also D_{2n} has normal subgroups D_n given by $\left\{a^{2r}, a^{2r}b : r = 0, 1, \ldots, n-1\right\}$ and $\left\{a^{2r}, a^{2r+1}b : r = 0, 1, \ldots, n-1\right\}$.

For H an abelian group then a generalised dihedral group can be defined by all elements $\{(h, e),\ (h, b)\}$ for $h \in H$ and $\{e, b\}$ forming the group $\mathbb{Z}_2$ so that $b^2 = e$. Group multiplication is defined so that $(h_1, e)(h_2, e) = (h_1 h_2, e),\ (h_1, e)(h_2, b) = (h_1 h_2, b),\ (h_1, b)(h_2, e) = (h_1 {h_2}^{-1}, b),\ (h_1, b)(h_2, b) = (h_1 {h_2}^{-1}, e)$. Note that $(h, e)^{-1} = (h^{-1}, e),\ (h, b)^{-1} = (h, b)$.

1.2.3 Symmetric Group

A frequently occurring group is the permutation or *symmetric group* $\mathcal{S}_n$ on n objects. It is easy to see that the order of $\mathcal{S}_n$ is $n!$. For $n = 3$ $\mathcal{S}_3 \simeq D_3$, as this is the symmetry group of an equilateral triangle under permutations of the vertices.

[2] Felix Klein, 1829-1925, German.

The elements of the permutation group can be decomposed into *cycles*. Acting on $\{1,2,\dots,n\}$ the 2-cycle (ij), for $i \neq j$, takes $i \leftrightarrow j$, the 3-cycle (ijk), with i,j,k all different, takes $i \to j \to k \to i$ and so on for the arbitrary p-cycle $(i_1,\dots,i_p)$, such that $p \leq n$ and $i_r \neq i_s$, $1 \leq i_r \leq n$, which generates cyclic permutations of $\{i_1,\dots,i_p\}$. Trivially the 1-cycle (i) leaves i invariant. Clearly $(i_1,\dots,i_p)^p = e$ and for any one of the $\binom{n}{p}$ choices for the p $\{i_r\}$ there are $(p-1)!$ choices for the p-cycle involving $\{i_r\}$ since any p-cycle is invariant under cyclic permutations. For i,j,k,l all distinct $(ij)(kl) = (kl)(ij)$ and also $(ij)(jk) = (ijk)$. In general a p-cycle can be written as a product of $p-1$ 2-cycles

$$(i_1,\dots,i_p) = (i_1 i_2)(i_2 i_3)\dots(i_{p-1}i_p)\,. \tag{1.21}$$

To verify the decomposition of the action of some $g \in \mathcal{S}_n$ into cycles we may consider for an arbitrary $i \in \{1,2,\dots,n\}$ all $g^r i$, $r = 1,2,\dots$. For some minimal p we must have $g^p i = i$. The action of g then generates a p-cycle $(i_1,\dots,i_p)$ where $i_1 = i$. Then for $j \in \{1,2,\dots,n\}$, $j \notin \{i_1,\dots,i_p\}$ acting repeatedly with g generates a new q-cycle $(j_1,\dots,j_q)$ for some q and $j_1 = j$. Continuing in this fashion any element of $\{1,2,\dots,n\}$ belongs to some cycle, if $gk = k$ then the decomposition involves the 1-cycle (k). Thus we may write $g = g_{(i_1\dots i_p)(j_1\dots j_q)\dots}$. The identity e corresponds to the n 1-cycles $(1)(2)\dots(n)$. Clearly g and $g^{-1} = g_{(i_p\dots i_1)(j_q\dots j_1)\dots}$ have the same cycle decomposition. If h corresponds to a permutation σ where $\sigma\{1,2,\dots,n\} = \{\sigma(1),\sigma(2),\dots,\sigma(n)\}$ then $h\,g_{(i_1\dots i_p)(j_1\dots j_q)\dots}h^{-1} = g_{(\sigma(i_1)\dots\sigma(i_p))(\sigma(j_1)\dots\sigma(j_q))\dots}$.

Elements in $\mathcal{S}_n$ which are given by products of non overlapping p_i-cycles, $i = 1,\dots,r$, with $1 \leq p_r \leq \dots \leq p_2 \leq p_1 \leq n$ and $\sum_{i=1}^r p_i = n$ form a subset denoted by $[p_1,p_2,\dots p_r]$. The identity is obtained for $r = n$ and $p_i = 1$, $i = 1,\dots,n$. Subsets in which one or more p_i are different are distinct. To count the number of permutations belonging to each cycle type we first assume the p_i are all different so that $p_i > p_{i+1}$. Then there are $n!/\prod_i p_i!$ ways of assigning $\{1,2,\dots,n\}$ to the different cycles but each cycle is invariant under p_i cyclic permutations so there are $(p_i-1)!$ possibilities for each p_i-cycle. This gives $\prod_{i=1}^r (p_i-1)!\,n!/p_i! = n!/\prod_{i=1}^r p_i$ different permutations. Suppose more generally there are j_i p_i-cycles so that $\sum_{i=1}^r j_i\,p_i = n$. Then the previous argument gives a factor ${p_i}^{j_i}$ but there must also be a factor $j_i!$ in the denominator since permutations between the different p_i-cycles give the same permutation. The number of elements in $\mathcal{S}_n$ which belong to the subset corresponding to cycles $[p_1(j_1),p_2(j_2),\dots,p_r(j_r)]$, where $p_i(j_i)$ means p_i is repeated j_i times, is then

$$N_{[p_1(j_1),\dots,p_r(j_r)]} = \frac{n!}{\prod_{i=1}^r {p_i}^{j_i} j_i!}\,, \quad p_i > p_{i+1}\,,\ j_i \geq 1\,, 1 \leq r \leq n\,, \quad \sum_{i=1}^r j_i p_i = n\,. \tag{1.22}$$

Since the total number of permutations is $n!$ we must have

$$n! = \sum_{r=1}^n \sum_{j_1,j_2,\dots j_r \geq 1} \sum_{p_1>p_2>\dots p_r \geq 1} \delta_{n,\sum_{i=1}^r j_i p_i}\, N_{[p_1(j_1),\dots,p_r(j_r)]}\,. \tag{1.23}$$

Note that $N_{[1(n)]} = 1$. Each choice of integer $p_i > p_{i+1} > \dots p_r > 0$, with p_i repeated j_i times,

$$\underline{\lambda} = [p_1(j_1),p_2(j_2),\dots,p_r(j_r)]\,, \qquad \sum_{i=1}^r j_i\,p_i = n\,, \quad r = 1,2,\dots\,, \tag{1.24}$$

corresponds to a *partition* of n. The number of possible partitions increases very rapidly with n.

All $n!$ elements of $\mathcal{S}_n$ can be obtained from products of the $n-1$ 2-cycles $\sigma_i = (i\,i{+}1)$ where the σ_i may be defined abstractly by requiring

$$\big\{\sigma_i,\ i=1,\dots,n-1 : {\sigma_i}^2 = e,\ (\sigma_i\sigma_{i+1})^3 = e,\ i=1,\dots,n-2,\ \sigma_i\sigma_j = \sigma_j\sigma_i\,,\ |i-j|>1\big\}\,. \tag{1.25}$$

The condition $(\sigma_i\sigma_{i+1})^3 = e$ can be alternatively expressed as

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}\,. \tag{1.26}$$

The conditions in (1.25) are sufficient to define the permutation group $\mathcal{S}_n$ in terms of all possible distinct products of σ_i and can be used to determine the group multiplication. For each possible product imposing the relations (1.25) ensure the number of σ_i in the product is unchanged mod 2. Thus for

$$\sigma = \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_p} \qquad \operatorname{sgn}(\sigma) = \varepsilon_\sigma = (-1)^p\,, \tag{1.27}$$

ε_σ, the *sign* or *signature* of σ, does not depend on the particular decomposition of σ. Even/odd elements of $\mathcal{S}_n$ correspond to $\varepsilon_\sigma = \pm 1$. Crucially

$$\varepsilon_\sigma\,\varepsilon_{\sigma'} = \varepsilon_{\sigma\sigma'}\,, \qquad \varepsilon_{\sigma^{-1}} = \varepsilon_\sigma\,. \tag{1.28}$$

The product $\sigma_1\sigma_2\dots\sigma_{r-1}$ corresponds to the r-cycle $(1\,2\,\dots\,r)$ and must in consequence satisfy

$$(\sigma_1\sigma_2\dots\sigma_{r-1})^r = e\,, \quad r = 2,3,\dots,n\,. \tag{1.29}$$

This may be verified as a consequence of just the relations (1.25) by induction from

$$(\sigma_1\dots\sigma_r)^s = (\sigma_1\dots\sigma_{r-1})^s\,\sigma_r\sigma_{r-1}\dots\sigma_{r-s+1}\,, \quad s = 1,\dots,r\,, \tag{1.30}$$

which, subject to (1.29), implies $(\sigma_1\dots\sigma_r)^r = (\sigma_1\dots\sigma_r)^{-1}$ and hence (1.29) for $r \to r+1$. In turn (1.30) follows inductively starting from $s=1$ by multiplying (1.30) on the right successively by $\sigma_1\dots\sigma_r$ and using

$$\sigma_r\sigma_{r-1}\dots\sigma_{r-s+1}\,(\sigma_1\dots\sigma_r) = (\sigma_1\dots\sigma_{r-1})\,\sigma_r\dots\sigma_{r-s}\,, \quad s = 1,\dots,r\,, \tag{1.31}$$

or, commuting σ_i, σ_j for $|i-j|>1$,

$$\begin{aligned}\sigma_r\sigma_{r-1}\dots\sigma_{r-s+1}\,\sigma_{r-s}\sigma_{r-s+1}\dots\sigma_r &= \sigma_{r-s}\big(\sigma_r\sigma_{r-1}\dots\sigma_{r-s+2}\,\sigma_{r-s+1}\sigma_{r-s+2}\dots\sigma_r\big)\sigma_{r-s}\\ &= \dots = \sigma_{r-s}\sigma_{r-s+1}\dots\sigma_{r-1}\,\sigma_r\sigma_{r-1}\dots\sigma_{r-s}\,,\end{aligned} \tag{1.32}$$

repeatedly using (1.26). This argument easily extends to showing

$$(\sigma_i\sigma_{i+1}\dots\sigma_{i+p-2})^p = e\,, \tag{1.33}$$

so that $\sigma_i\sigma_{i+1}\dots\sigma_{i+p-2}$ is a p-cycle.

As an illustration of how more general permutations can be generated in terms of products of $\{\sigma_i\}$ then

$$\tilde{\sigma}_j = \begin{cases} \sigma_i \sigma_{i+1} \dots \sigma_j \sigma_{j-1} \dots \sigma_i \,, & j > i \,, \\ \sigma_i \,, & j = i \,, \\ \sigma_j \sigma_{j+1} \dots \sigma_{i-1} \sigma_{i-2} \dots \sigma_j \,, & j < i-1 \,, \end{cases} \tag{1.34}$$

satisfies $\tilde{\sigma}_j{}^2 = e$ and corresponds to the 2-cycle $(i\, j{+}1)$.

An alternative expression for arbitrary elements belonging to $\mathcal{S}_n$ can be obtained in terms of products of cycles where

$$a_r = \sigma_1 \dots \sigma_{r-1} \,, \quad r = 2, \dots, n \,, \quad a_1 = e \,, \quad a_r{}^r = e \,, \quad \varepsilon_{a_r} = (-1)^{r-1} \,, \tag{1.35}$$

and then an arbitrary $\sigma \in \mathcal{S}_n$ can be written as

$$\sigma = a_n{}^{r_n}\, a_{n-1}{}^{r_{n-1}} \dots a_2{}^{r_2} \,, \qquad r_i = 0, 1, \dots, i-1 \,. \tag{1.36}$$

It is easy to see that there are $n!$ possibilities and arbitrary products can be brought to the form (1.36) by repeatedly using

$$a_s\, a_t = a_t^2\, a_{t-1}{}^{t-2}\, a_{s-1} \,, \quad s = 2, 3, \dots, t-1 \,, \quad t = 3, \dots n \,, \tag{1.37}$$

as well as $a_r{}^r = e$. Of course this rule preserves the signature. For just $n = 3$ then taking $a_2 = b$, a_3 reproduces the group multiplication rules defining $D_3 \simeq \mathcal{S}_3$ in (1.20).

1.2.4 Alternating Group

The alternating group A_n is the normal subgroup of $\mathcal{S}_n$ formed by even permutations. It has $n!/2$ elements and for $n \geq 5$ is simple since, as discussed further later, there are then no normal subgroups apart from the identity. In general $\mathcal{S}_n / A_n \simeq \mathbb{Z}_2$.

As a consequence of (1.21) the alternating group A_n can only contain single p-cycles with p odd or products of distinct p-cycles, $p_1, p_2, \dots$, with $\sum_i (p_i - 1)$ even. Every 3-cycle in $\mathcal{S}_n$ is contained in A_n since for i, j, k distinct

$$(ijk) = (ik)(ij) \,. \tag{1.38}$$

Furthermore any product of 2-cycles is expressible as a product of 3-cycles. If the 2-cycles are not distinct this has just been shown, otherwise with i, j, k, l all different.

$$(ij)(kl) = (ij)(ik)(ki)(kl) = (jik)(ikl) \,. \tag{1.39}$$

As a consequence A_n may be generated in terms of 3-cycles just as $\mathcal{S}_n$ is generated by 2-cycles.

To verify this for A_n it is sufficient to consider the 3-cycles $s_i = (12\, i{+}2)$, $i = 1, 2, \dots, n{-}2$, where for $i \neq j$, $s_i s_j = (1i)(2j)$. The group elements s_i, $= 1, \dots, n-2$ obey the abstract relations

$$s_i^3 = e \,, \qquad i \neq j, \quad (s_i s_j)^2 = e \quad \Rightarrow \quad s_i s_j = s_j{}^{-1} s_i^{-1} = s_j{}^2 s_i^2 \,. \tag{1.40}$$

Starting from $A_3 = \{e, s_1, {s_1}^2\}$ the groups A_n may be defined inductively by

$$A_{n+1} = \left\{A_n, A_n\, s_{n-1}, A_n\, {s_{n-1}}^2, A_n\, s_{n-1} s_i^2 : i = 1, \dots, n-2\right\}. \tag{1.41}$$

It is necessary to check $A_{n+1}\, s_{n-1} = A_{n+1}$, which follows from (1.40) since $s_{n-1}\, s_i^2\, s_{n-1} = s_i^2 s_{n-1}\, s_i^2$ and $A_n s_i = A_n$, and also for $j = 1, \dots, n-2$, $A_{n+1} s_j = A_{n+1}$, which is a consequence of $s_{n-1}\, s_j = s_j^2 {s_{n-1}}^2$, ${s_{n-1}}^2 s_j = s_j s_{n-1} s_j^2$ and for $j \neq i$, $s_{n-1} s_i^2\, s_j = s_i^2 s_j s_{n-1} s_i^2$.

All 3-cycles are conjugate in A_n so long as $n \geq 5$. It must be true that there is a $\sigma \in \mathcal{S}_n$ such that

$$\sigma(ijk)\sigma^{-1} = (i'j'k'), \tag{1.42}$$

for arbitrary distinct i, j, k and i', j', k'. If σ is even then $\sigma \in A_n$, otherwise if $n \geq 5$ we can take $\sigma \to \sigma' = \sigma(lm)$ where l, m are different from i, j, k to achieve the same result and $\sigma' \in A_n$. In a similar fashion all n-cycles, for n odd, are conjugate in A_{n+2}. However 5-cycles in A_5 are not necessarily conjugate. There are 6 $\mathbb{Z}_5$ subgroups $\{e, \sigma_i, \sigma_i^2, \sigma_i^3, \sigma_i^4\}$ generated by a 5-cycle σ_i. Of these two sets of 3 are each conjugate.

As special cases $A_3 \simeq \mathbb{Z}_3$ and A_4 is the symmetry group, without reflections, of a regular tetrahedron. The group is formed by $2\pi/3, 4\pi/3$ rotations about axes from each of 4 vertices to the centre of the opposite face and also rotations of π about the three lines joining the mid points of opposite edges. Thus, apart from the identity, A_4 is composed of 8 different 3-cycles and 3 products of two distinct 2-cycles which, with the identity, form a normal subgroup.

For arbitrary n any normal subgroup $N \lhd A_n$ which contains a 3-cycle must, as a consequence of (1.42), also contain all 3-cycles and so $N = A_n$. For $n \geq 5$ it is possible to show that N must contain a 3-cycle and therefore $N = A_n$ and A_n is simple so long as $n \geq 5$. To show this let $\sigma \in N$ be a non trivial group element containing p 1-cycles. Then it is possible to choose $\tau \in A_n$, $\tau^{-1}\sigma\tau \in N$ since $\sigma \in N$, such that $\sigma' = \sigma^{-1}\tau^{-1}\sigma\tau \in N$ contains p' 1-cycles with $p' > p$. This process may be continued until $p' = n-3$ and then σ' is just a 3-cycle. As an illustration if $\sigma = (12)(34)$ then taking $\tau = (12)(35)$ we have $\sigma' = (345)$ or if $\sigma = (12345)$ and $\tau = (123)$ then $\sigma' = (125)$.

1.3 Orbit Stabiliser Theorem

An important result which has many applications arises when a group G acts on a space $X = \{x\}$ so that for any $g \in G$ there is an action $x \to gx$. For any particular $x \in X$ the *stabiliser group*, or *little group*, G_x is defined by those elements of G which leave x invariant, $G_x = \{h : h \in G,\, hx = x\}$. It is easy to see that G_x is a subgroup of G. The *orbit* of x is the set of points in X obtained by the action of G, $O_x = \{x' : x' = gx\}$. O_x can be identified with the coset G/G_x. This is the *orbit stabiliser theorem* and we have for a finite group G the dimension of the orbit

$$\dim O_x = |G|/|G_x|, \tag{1.43}$$

which is an integer by Lagrange's theorem. Clearly for $x' \in O_x$ then $G_{x'} \simeq G_x$ since $hx = x$, $x' = gx$ implies $h'x' = x'$ for $h' = ghg^{-1}$. In general the space X may be decomposed into orbits under the action of G.

1.4 Further Definitions

Here we give some supplementary definitions connected with groups which play a crucial role in the theory of groups and introduces important notation.

1.4.1 Automorphisms and Semi-Direct Product

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

$$\varphi(g_i)\varphi(g_j) = \varphi(g_i g_j) \quad \text{for all} \quad g_i, g_j \in G\,, \tag{1.44}$$

so that $G_\varphi = \{\varphi(g_i)\} \simeq G$. Clearly we must have $\varphi(e) = e$ and $\varphi(g^{-1}) = \varphi(g)^{-1}$. In general for any fixed $g \in G$ we may define an *inner automorphism* by $\varphi_g(g_i) = g g_i g^{-1}$, otherwise the automorphism is *outer*. It is straightforward to see that the set of all automorphisms of G itself forms a group $\mathrm{Aut}\, G$ which must include the group of inner autmorphisms $\mathrm{Inn}\, G = G/\mathcal{Z}(G)$ as a normal subgroup, the quotient $\mathrm{Out}\, G = \mathrm{Aut}\, G/\mathrm{Inn}\, G$ defines the outer automorphism group. For an abelian group there are no non trivial inner automorphisms but there can be non trivial outer automorphisms, e.g. for $\mathbb{Z}_3$ take $\{e, a, a^2\} \to \{e, a^2, a\}$. In this case $\mathrm{Aut}\, \mathbb{Z}_3 = \mathbb{Z}_2$ and $\mathbb{Z}_3/\mathcal{Z}(\mathbb{Z}_3) = \{e\}$ the trivial one element group. In a similar fashion $\mathrm{Aut}\, \mathbb{Z}_n = \mathbb{Z}_{n-1}$ whenever n is prime.

There is also an *antiautomorphism* group $\mathrm{Anti}\, G$ which is defined by all maps $g_i \to \varphi(g_i)$ such that

$$\varphi(g_i)\varphi(g_j) = \varphi(g_j g_i) \quad \text{for all} \quad g_i, g_j \in G\,. \tag{1.45}$$

Clearly $\mathrm{Anti}\, G = \mathrm{Aut}\, G$ if G is abelian.

If $H \subset \mathrm{Aut}\, G$, so that for any $h \in H$ and any $g \in G$ we have $g \underset{h}{\to} \varphi_h(g)$ with

$$\varphi_h(g_1)\varphi_h(g_2) = \varphi_h(g_1 g_2)\,, \tag{1.46}$$

and

$$\varphi_{h_1}(\varphi_{h_2}(g)) = \varphi_{h_1 h_2}(g)\,, \quad \varphi_h(e) = e\,, \quad \varphi_e(g) = g\,, \quad \varphi_{h^{-1}}(g) = \varphi_h{}^{-1}(g)\,, \tag{1.47}$$

we may define a new group called the *semi-direct product* of H with G, denoted $H \ltimes G$, or $G \rtimes H$. 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

$$(h, g)(h', g') = (hh', g\,\varphi_h(g'))\,, \quad (h, g)^{-1} = (h^{-1}, \varphi_{h^{-1}}(g^{-1}))\,. \tag{1.48}$$

From this it follows that

$$(h, g)(e, g')(h, g)^{-1} = (e, g\varphi_h(g')g^{-1}) \quad \text{for any} \quad g, g' \in G\,,\ h \in H\,. \tag{1.49}$$

Consequently the subgroup $\{(e, g)\} \simeq G$ is a normal subgroup of $H \ltimes G$ and hence $H \simeq (H \ltimes G)/G$.

It is often convenient to write the elements of $H \ltimes G$ as simple products so that we may abbreviate $(h,g) \to hg = \varphi_h(g)h$.

As a simple illustration we have $D_n \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_n$ where $\mathbb{Z}_2 = \{e, b : b^2 = e\}$ and also $\mathbb{Z}_n = \{a^r : r = 0, \dots n-1,\ a^n = e\}$. The semi-direct product is then defined by taking, for any $g = a^r \in \mathbb{Z}_n$, $\varphi_b(g) = g^{-1} = b\, g\, b^{-1}$. $\mathbb{Z}_n$ is a normal subgroup of of D_n.

The case of a maximal semi-direct product of G with $\operatorname{Aut} G$

$$G \rtimes \operatorname{Aut} G = \operatorname{Hol} G\,, \tag{1.50}$$

has a special name, the *holomorph* of G. From the above example $D_3 \simeq \operatorname{Hol} \mathbb{Z}_3$.

1.4.2 Wreath Products and Central Products

Another semi-direct product is obtained if we take $G \to G_{n\times} = G \times G \times \cdots \times G$, the n-fold direct product, with $H = \mathcal{S}_n$ permuting the elements of each of the factors so that, for any $g = (g_1, g_2, \dots, g_n) \in G_{n\times}$,

$$g_\sigma \equiv \varphi_\sigma(g_1, g_2, \dots, g_n) = \big(g_{\sigma(1)}, g_{\sigma(2)}, \dots, g_{\sigma(n)}\big)\,, \quad \sigma \in \mathcal{S}_n\,. \tag{1.51}$$

This then defines $\mathcal{S}_n \ltimes G_{n\times}$. This is an example of a *wreath product*, denoted $G \wr \mathcal{S}_n$, and has order $n!|\,G|^n$. A particular example is $B_n = \mathbb{Z}_2 \wr \mathcal{S}_n$, the hyperoctahedral group, which is the symmetry group of the n-dimensional hypercube. As special cases $B_2 = D_4$, the symmetry group of the square, and $B_3 = \mathcal{S}_4 \times \mathbb{Z}_2$, the symmetry group of the cube.

If the permutations acting on $G_{n\times}$ are restricted to a subgroup of $\mathcal{S}_n$ there is a corresponding wreath product, for just cyclic permutations of $(g_1, \dots, g_n)$ this is $G \wr \mathbb{Z}_n$.

Another prescription for combining groups is obtained essentially by taking the direct product and dividing by some common elements. If G_1, G_2 are two groups then for $H_1 \subset \mathcal{Z}(G_1)$, $H_2 \subset \mathcal{Z}(G_2)$ then for $H_1 \simeq H_2$ with the isomorphism θ such that $h_1 \leftrightarrow h_2$ then $\{(h_1, h_2)\} = H$, where $H \simeq H_1 \simeq H_2$ is a normal subgroup of $H_1 \times H_2$. The quotient $(G_1 \times G_2)/H$ defines the *central product*.

A rather trivial example arises if $G_1 = \mathbb{Z}_4$, $G_2 = \mathbb{Z}_6$. In this case we can take $H = \mathbb{Z}_2$ and $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\mathbb{Z}_2 \simeq \mathbb{Z}_{12}$.

1.4.3 Conjugacy Classes

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*

$$\mathcal{C}_s = \{g_i : g_i \sim g_{i'} = g g_i g^{-1},\ g \in G\}\,, \qquad s = 1, \dots, N_{\text{char}}\,. \tag{1.52}$$

Different conjugacy classes are distinct $\mathcal{C}_s \cap \mathcal{C}_{s'} = \varnothing$, $s \neq s'$, and $G = \cup_s \mathcal{C}_s$, $|G| = \sum_{s=1}^{N_{\text{char}}} d_s$ for $d_s = \dim \mathcal{C}_s$. For any $g \in G$, $g\, \mathcal{C}_s\, g^{-1} = \mathcal{C}_s$. Clearly the identity is in a conjugacy class $\mathcal{C}_1$ by itself. In general

$$g_i, g_j \in \mathcal{C}_s \quad \Rightarrow \quad g_i^{-1}, g_j^{-1} \in \mathcal{C}_{\bar{s}}\,, \qquad |\mathcal{C}_{\bar{s}}| = |\mathcal{C}_s|\,, \tag{1.53}$$

where $\mathcal{C}_{\bar{s}} = \mathcal{C}_s$ if $g, g^{-1} \in \mathcal{C}_s$. It is sometimes convenient in this case to write $\mathcal{C}_{\bar{s}} = \mathcal{C}_s{}^{-1}$.

For any $h \in G$ the associated conjugacy class containing h is given by

$$\mathcal{C}(h) = \{ghg^{-1} : g \in G\} . \tag{1.54}$$

G also contains a subgroup defined by

$$C_G(h) = \{g : ghg^{-1} = h,\ g \in G\} , \tag{1.55}$$

and then, by the orbit stabiliser theorem,

$$\dim \mathcal{C}(h) = |G|/|C_G(h)| . \tag{1.56}$$

By Lagrange's theorem $|C_G(h)|$ divides $|G|$ so that the dimensions of any conjugacy class must also divide $|G|$.

Under group multiplication the product of two conjugacy classes must be expressible in terms of a union of conjugacy classes so that there is a multiplication rule

$$\mathcal{C}_s \, \mathcal{C}_t = \bigcup_u c_{st}{}^u \, \mathcal{C}_u , \tag{1.57}$$

where $c_{st}{}^u$ takes the values $0, 1, 2, \dots$. Since $\mathcal{C}_s \, g = g \, \mathcal{C}_s$ for any $g \in G$ then $\mathcal{C}_s \, \mathcal{C}_t = \mathcal{C}_t \, \mathcal{C}_s$ and thus $c_{st}{}^u = c_{ts}{}^u$. Furthermore $c_{1t}{}^u = \delta_t{}^u$, $c_{s\bar{s}}{}^1 = d_s$ and $\sum_u c_{st}{}^u d_u = d_s \, d_t$.

For an abelian group all elements have their own conjugacy class necessarily of dimension 1. The elements in a conjugacy class have similar properties such as $g_i{}^n = e$ for the same n for all $g_i \in \mathcal{C}_s$. Any normal subgroup is composed of conjugacy classes which must include $\mathcal{C}_1$. For $\mathcal{S}_3$ which has elements $\{e, a, a^2, b, ab, a^2b\}$, where $b = (12)$, $a = (123)$, there are three conjugacy classes $\{e\}$, $\{a, a^2\}$, $\{b, ab, a^2b\}$. $\{e, a, a^2\}$ forms a normal subgroup which is isomorphic to $\mathbb{Z}_3$.

For the dihedral group D_n, as defined in (1.20), the conjugacy classes are different according to whether n is even or odd. Labelling them by their size these are

$$\begin{aligned}
&\mathcal{C}_1 = \{e\} ,\ \mathcal{C}_{2,r} = \left\{a^r,\, a^{n-r}\right\} ,\ r = 1, \dots \tfrac{1}{2}(n-1) ,\ \mathcal{C}_n = \left\{a^r b : r = 0, 1, \dots, n-1\right\} , \quad n \text{ odd} ,\\
&\mathcal{C}_{1,1} = \{e\} ,\ \mathcal{C}_{1,2} = \left\{a^{\frac{1}{2}n}\right\} ,\ \mathcal{C}_{2,r} = \left\{a^r,\, a^{n-r}\right\} ,\ r = 1, \dots, \tfrac{1}{2}(n-2) ,\\
&\quad \mathcal{C}_{\frac{1}{2}n,1} = \left\{a^{2r} b : r = 0, 1, \dots, \tfrac{1}{2}n - 1\right\} ,\ \mathcal{C}_{\frac{1}{2}n,2} = \left\{a^{2r+1} b : r = 0, 1, \dots, \tfrac{1}{2}n - 1\right\} , \quad n \text{ even} .
\end{aligned} \tag{1.58}$$

There are then $\frac{1}{2}(n+3)$ conjugacy classes for n odd, $\frac{1}{2}(n+6)$ for n even. The conjugacy classes are all self inverse in that each conjugacy class $\mathcal{C}$ contains the inverse for each group element in $\mathcal{C}$.

Under multiplication for n odd

$$\begin{aligned}
&\mathcal{C}_{2,r} \, \mathcal{C}_{2,s} = \begin{cases} \mathcal{C}_{2,|r-s|} \cup \mathcal{C}_{2,r+s} \ \text{ or } \ \mathcal{C}_{2,|r-s|} \cup \mathcal{C}_{2,n-r-s} & r \neq s \\ 2\,\mathcal{C}_1 \cup \mathcal{C}_{2,2r} \ \text{ or } \ 2\,\mathcal{C}_1 \cup \mathcal{C}_{2,n-2r} & r = s \end{cases} ,\\
&\mathcal{C}_{2,r} \, \mathcal{C}_n = \mathcal{C}_n \, \mathcal{C}_{2,r} = 2\,\mathcal{C}_n , \qquad \mathcal{C}_n \, \mathcal{C}_n = n\,\mathcal{C}_1 \cup_r \mathcal{C}_{2,r} .
\end{aligned} \tag{1.59}$$

Results for n even are similar.

For n even D_n has an external automorphism which interchanges the two conjugacy classes involving b. The automorphism group is generated by $c : a \to a^{n-1}$, $b \to ab$. Clearly c^2 is the identity and $\mathrm{Out}\, D_n = \{e, c : c^2 = e\} = \mathbb{Z}_2$. Correspondingly for n even $\mathbb{Z}_2 \ltimes D_n \simeq D_{2n}$ since for $\mathbb{Z}_2 = \{e, c\}$ and D_n as in (1.20) then $D_{2n} = \{a^r,\, a^r c\,, a^r b\,,\, a^r b\, c : r = 0, \dots, n-1\}$ for $c\,a\,c = a^{n-1}$, $c\,b\,c = ab$. Defining $a' = abc,\ b' = ac$ then $a'^{\,2n} = b'^{\,2} = e$ and a', b' satisfy the conditions to generate D_{2n}. Conversely for n odd $\mathbb{Z}_2 \times D_n \simeq D_{2n}$ where the extra element c commutes with a, b. This follows from $\mathbb{Z}_2 \times \mathbb{Z}_n \simeq \mathbb{Z}_{2n}$ for n odd.

The decomposition of any $\sigma \in \mathcal{S}_n$ into non overlapping, or disjoint, cycles is unchanged under conjugation. For $\sigma \in \mathcal{S}_n$ given by a product of non overlapping cycles all other $\sigma' \in \mathcal{S}_n$ expressible as a product of the same cycles can be obtained from σ by some permutation in $\mathcal{S}_n$ and in consequence σ' can be obtained from σ by conjugation. If σ' has a different expression in terms of cycles then it cannot. Hence for an identical decomposition into cycles, up to ordering, σ, σ' belong to the same conjugacy class. The different conjugacy classes of $\mathcal{S}_n$ are then labelled $\mathcal{C}_{[p_1(j_1), \dots, p_r(j_r)]}$ with $p_1, p_2, \dots, p_r$ together with $j_1, \dots, j_r$ corresponding to a partition of n such that $1 \le p_r < \dots < p_2 < p_1 \le n$ with $\sum_i j_i p_i = n$. The dimensions of the conjugacy classes for $\mathcal{S}_n$ are given by the general formula (1.22).

As a consequence of (1.28) $\varepsilon_{\sigma\sigma'\sigma^{-1}} = \varepsilon_{\sigma'}$ so that the sign of all group elements in $\mathcal{S}_n$ belonging to a particular conjugacy class is the same and is given by

$$\sigma \in \mathcal{C}_{[p_1(j_1), \dots, p_r(j_r)]}\,, \qquad \varepsilon_\sigma = (-1)^{\sum_i j_i (p_i - 1)}\,. \tag{1.60}$$

Since there equal numbers of σ with $\varepsilon_\sigma = \pm 1$ we must have

$$\sum_{r=1}^{n} \sum_{j_1, j_2, \dots j_r \ge 1} \sum_{p_1 > p_2 > \dots p_r \ge 1} \delta_{n, \sum_{i=1}^r j_i p_i}\ N_{[p_1(j_1), \dots, p_r(j_r)]}\, (-1)^{\sum_i j_i (p_i - 1)} = 0\,. \tag{1.61}$$

The elements of the alternating group A_n correspond to those $\sigma \in \mathcal{S}_n$ with a cycle decomposition in which the number of even n-cycles is also even and $\varepsilon_\sigma = 1$. However the conjugacy classes in A_n are not just determined by the cycle decomposition as for $\mathcal{S}_n$ since only even permutations are generated by conjugation in A_n. Conjugacy classes in $\mathcal{S}_n$ may split into two on reduction to A_n.

The conjugacy classes for $\mathcal{S}_n$ can be identified with *Young*[3] *diagrams* which are formed by n boxes in rows of decreasing size according to the partition of n. Those for $\mathcal{S}_5$ are

$$\begin{gathered}
[5]\ \ydiagram{5}\,, \quad [4,1]\ \ydiagram{4,1}\,, \quad [3,2]\ \ydiagram{3,2}\,, \quad [2(2),1]\ \ydiagram{2,2,1}\,, \\
[3,1(2)]\ \ydiagram{3,1,1}\,, \quad [2,1(3)]\ \ydiagram{2,1,1,1}\,, \quad [1(5)]\ \ydiagram{1,1,1,1,1}\,.
\end{gathered} \tag{1.62}$$

Such diagrams play an important role in discussions of the permutation group.

[3]Alfred Young, 1873-1940, British, 10th wrangler

As a illustration we consider the cases $n = 3, 4, 5$. For $\mathcal{S}_3$ there are just three conjugacy classes with dimensions

$$N_{[3]} = 2\,, \qquad N_{[2,1]} = \binom{3}{2} = 3\,, \qquad N_{[1(3)]} = 1\,. \tag{1.63}$$

The different conjugacy classes are $\mathcal{C}_{[1(3)]} = \{e\}$, $\mathcal{C}_{[3]} = \{a, a^2\}$, $\mathcal{C}_{[2,1]} = \{b, ba, ba^2\}$ where $a^3 = b^2 = e$, $ba = a^2 b$. $\mathcal{C}_{[2,1]}$ contains all odd permutations so $A_3 = \{e, a, a^2\}$. However this is abelian so each element has its own conjugacy class. Similarly for $\mathcal{S}_4$

$$N_{[4]} = 3! = 6\,, \quad N_{[3,1]} = 2\binom{4}{3} = 8\,, \quad N_{[2(2)]} = \tfrac{1}{2}\binom{4}{2} = 3\,, \quad N_{[2,1(2)]} = \binom{4}{2} = 6\,, \quad N_{[1(4)]} = 1\,. \tag{1.64}$$

The elements of the alternating group A_4 belong to the conjugacy classes $\mathcal{C}_{[1(4)]}$, $\mathcal{C}_{[3,1]}$ and $\mathcal{C}_{[2(2)]}$ in $\mathcal{S}_4$. The elements of $\mathcal{C}_{[2(2)]}$, $a = (12)(34)$, $b = (13)(24)$, $ab = ba = (14)(23)$, together e form a normal subgroup of both $\mathcal{S}_4$ and A_4, $a^2 = b^2 = e$, which is isomorphic to D_2. However the cycles $(123)(4)$ and $(124)(3)$ belonging to A_4 are not conjugate in A_4 since the permutation linking them involves one 2-cycle. Hence $\mathcal{C}_{[3,1]}$ decomposes into two equal conjugacy classes of size 4.

For $\mathcal{S}_5$ corresponding to (1.62)

$$\begin{aligned} &N_{[5]} = 4! = 24\,, \quad N_{[4,1]} = 3\binom{5}{4} = 30\,, \quad N_{[3,2]} = 2\binom{5}{2} = 20\,, \quad N_{[3,1(2)]} = 2\binom{5}{3} = 20\,, \\ &N_{[2(2),1]} = \tfrac{1}{2}\binom{5}{2}\binom{3}{2} = 15\,, \quad N_{[2,1(3)]} = \binom{5}{2} = 10\,, \quad N_{[1(5)]} = 1\,. \end{aligned} \tag{1.65}$$

A_5 is given by elements belonging to the conjugacy classes $\mathcal{C}_{[3,1(2)]}$, $\mathcal{C}_{[5]}$, $\mathcal{C}_{[2(2),1]}$ as well as $\mathcal{C}_{[1(5)]}$ in $\mathcal{S}_5$. However $\mathcal{C}_{[5]}$ splits into two conjugacy classes for A_5 each of size 12 and the dimensions of the five conjugacy classes are then $1 + 20 + 12 + 12 + 15 = 60$. In this case none of the dimensions including 1 for $\mathcal{C}_{[1(5)]}$ add up to a divisor of $60 = |A_5|$. Hence A_5 has no non trivial normal subgroups and is simple. It is the smallest non abelian group with this property.

An alternative labelling of conjugacy classes for $\mathcal{S}_n$ instead of $[p_1(j_1), \dots, p_r(j_r)]$ for $r = 1, \dots, n$ is obtained by taking

$$k_m = \begin{cases} j_i, & p_i = m \text{ for some } i\,, \\ 0, & p_i \neq m \text{ for any } i\,, \end{cases} \qquad m = 1, \dots, n\,, \quad \textstyle\sum_{m=1}^{n} m k_m = n\,, \tag{1.66}$$

where $N_{[p_1(j_1),\dots,p_r(j_r)]} = n!/\prod_{m=1}^{n} m^{k_m} m!$ and $\varepsilon_\sigma = (-1)^{(m-1)k_m}$. For a function on conjugacy classes of $\mathcal{S}_n$, $f(k_1, \dots, k_n)$, the corresponding sum over all classes is simply given by

$$\sum_{k_1,k_2,\dots,k_n \geq 0} \delta_{n,\sum_m m k_m} \, \frac{n!}{\prod_{m=1}^{n} m^{k_m} m!} \, f(k_1, \dots k_n)\,. \tag{1.67}$$

As an example for $\underline{q} = (q_1, q_2, \dots)$ there is a generating function obtained by summing over n

$$F(u, \underline{q}) = \sum_{n=0}^{\infty} u^n \prod_{m=1}^{n} \bigg(\sum_{k_m \geq 0} \frac{{q_m}^{k_m}}{m^{k_m} m!} \bigg) \delta_{n,\sum_m m k_m} = \exp\bigg(\sum_{m=1}^{\infty} \frac{u^m q_m}{m} \bigg)\,. \tag{1.68}$$

If $q_m = f(x^m)$ for some function f then

$$F(u,\underline{q}) = \mathrm{PE}(u,x;f) = \exp\bigg(\sum_{m=1}^{\infty}\frac{u^m f(x^m)}{m}\bigg) = 1 + u\,f(x) + \sum_{m\geq 2} u^m \vee^m f(x)\,, \tag{1.69}$$

is the plethystic exponential formed from the function f and

$$\vee^2 f(x) = \tfrac{1}{2}\big(f(x)^2 + f(x^2)\big)\,, \quad \vee^3 f(x) = \tfrac{1}{6}\big(f(x)^3 + 3\,f(x^2)f(x) + 2f(x^3)\big)\,. \tag{1.70}$$

For a constant function $f(x) = c$

$$\exp\bigg(\sum_{m=1}^{\infty}\frac{u^m c}{m}\bigg) = \frac{1}{(1-u)^c} = \sum_{r=0}^{\infty}\tfrac{1}{r!}\,c(c+1)\dots(c+r-1)\,u^r\,. \tag{1.71}$$

Including the sign factor for odd permutations

$$\mathrm{PE}_A(u,x;f) = \exp\bigg(\sum_{m=1}^{\infty}(-1)^{m-1}\frac{u^m f(x^m)}{m}\bigg) = 1 + u\,f(x) + \sum_{m\geq 2} u^m \wedge^m f(x)\,, \tag{1.72}$$

with

$$\wedge^2 f(x) = \tfrac{1}{2}\big(f(x)^2 - f(x^2)\big)\,, \quad \wedge^3 f(x) = \tfrac{1}{6}\big(f(x)^3 - 3\,f(x^2)f(x) + 2f(x^3)\big)\,. \tag{1.73}$$

1.4.4 Centre, Normaliser, Centraliser and Commutator 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 $\mathcal{Z}(G) \simeq G$. The centre is a normal subgroup and the quotient group $G/\mathcal{Z}(G)$ is referred to as the *inner automorphism group.*

For a subset of a group $H \subset G$, not necessarily a subgroup, 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 called the *normaliser* of H in G, written $N_G(H)$. If H is a subgroup then clearly $H \lhd N_G(H)$ and if H is a normal subgroup, $N_G(H) = G$.

The subgroup of G formed by elements $\{g\}$ such that $ghg^{-1} = h$ for all $h \in H$ forms the *centraliser* $C_G(H)$. Necessarily $C_G(H) \subset N_G(H)$.

For any two elements of G, g,h, then

$$[g,h] = g^{-1}h^{-1}g\,h\,, \tag{1.74}$$

is the *commutator* of g,h. If G is abelian then $[g,h] = e$ for all g,h. More generally if $[g,h] = e$ then g and h commute. In general $[g,h]^{-1} = [h,g]$ and for any $g' \in G$ then $g'[g,h]g'^{-1} = [g'gg'^{-1}, ghg'^{-1}]$. The product of two commutators need not be a commutator (this can only arise if the order of G is at least 96) but there is a subgroup of G, the *commutator* or *derived subgroup* $G' = [G,G]$ formed from arbitrary products of commutators. From the above $g[G,G]g^{-1} = [G,G]$ for any $g \in G$ so $[G,G]$ is a normal subgroup. For any $g_1, g_2 \in G$

then $g_1 g_2 = g_2 g_1 [g_1, g_2]$ so that the quotient group $G/[G,G]$ is abelian. For any normal subgroup $N \lhd G$ then the quotient G/N is abelian only if $[G,G] < N$. Of course if G is abelian $[G,G] = \{e\}$, the trivial group.

A group is *perfect* if $G = [G,G]$. A non abelian simple group must be perfect since $[G,G]$ is a normal subgroup and $[G,G] \neq \{e\}$ if G is non abelian. The converse is not necessary.

As an example we may consider the dihedral group D_n as presented in (1.20). In this case $[D_n, D_n] = \{e, a, a^2, \dots, a^{n-1}\} = \mathbb{Z}_n$. The quotient is just $\mathbb{Z}_2$.

The notion of the commutator subgroup may be extended to define the *derived series* where $G^{(n)} = [G^{(n-1)}, G^{(n-1)}]$ where $G^{(0)} = G$ and $n = 1, 2, \dots$. Evidently $\dots \lhd G^{(n)} \lhd G^{(n-1)} \lhd \dots G^{(1)} \lhd G$. For a finite group this series must terminate in a perfect group or $G^{(n)} = \{e\}$ for some n. In this case G is *solvable*.

1.4.5 Double Coset

An extension of the notion of a coset is a *double coset*. If H, K are subgroups $H, K \subset G$ then the equivalence relation (1.8) can be extended to

$$g_i \sim g'_i \quad \Leftrightarrow \quad g_i = k g'_i h \;\; \text{for} \;\; h \in H \,, \;\; k \in K \,, \tag{1.75}$$

where under this equivalence relation $\{g_i\}$ defines a double coset and these double cosets form $K \backslash G / H$. In general there is a one to one correspondence of $K \backslash G / H$ with $H \backslash G / K$ since $g_i^{-1} = h^{-1} g'^{-1}_i k^{-1}$. The potential extension of Lagrange's theorem is not valid since in general the cosets belonging to $K \backslash G / H$ have dimensions which need not divide $|G|$. If $G = D_3$ and $H = \{e, b\}$, $K = \{e, ba\}$, which are both $\mathbb{Z}_2$ subgroups, then $K \backslash G / H$ is comprised of $\{e, b, ba, a^2\}$ and $\{a, ba^2\}$ whereas $H \backslash G / K$ is formed from $\{e, b, ba, a\}$ and $\{a^2, ba^2\}$.

1.4.6 Goursat's Lemma

A nice application of the basic definitions of group theory due to *Goursat*[4] shows how subgroups of direct products are obtained. For $G = G_1 \times G_2$ and a subgroup $H < G$ then we may define two subgroups of G_1 by

$$H_1 = \{h_1\} \,, \quad (h_1, h_2) \in H \text{ for any } h_2 \,, \quad N_1 = \{n_1\} \,, \quad (n_1, e_2) \in H \,. \tag{1.76}$$

Since $(h_1, h_2)(n_1, e_2)(h_1, h_2)^{-1} = (h_1 n_1 {h_1}^{-1}, e_2)$ $h_1 n_1 {h_1}^{-1} \in N_1$ for any $h_1 \in H_1$ so that N_1 is a normal subgroup of H_1, $N_1 \lhd H_1$. There is then a quotient group $Q_1 = H_1/N_1$. In a similar fashion we can define H_2, N_2 and Q_2. The aim of the lemma is to show $Q_1 \simeq Q_2$.

To show this assume (h_1, h_2), $(h_1, {h_2}') \in H$. Then $(h_1, h_2)^{-1}(h_1, {h_2}') = (e_1, {h_2}^{-1} {h_2}')$ so that ${h_2}^{-1} {h_2}' \in N_2$ or h_2, ${h_2}'$ both correspond to the same $q_2 \in Q_2$. Extending this to h_1, h_1' ensures

$$(h_1, h_2), \, ({h_1}', {h_2}') \in H \quad \Leftrightarrow \quad ({h_1}', {h_2}') = (h_1 n_1, h_2 n_2) \;\; \text{for some} \;\; n_1 \in N_1 \,, n_2 \in N_2 \,. \tag{1.77}$$

[4] Édouard Goursat, 1858-1936, French.

This implies that any $(h_1,h_2)\in H$ can be expressed as (q_1n_1,q_2n_2) where q_2 determines q_1 and vice versa. In consequence there is a unique mapping $\varphi: q_1\to q_2$ with inverse $\varphi^{-1}: q_2\to q_1$ and this preserves the group properties so that $Q_1\simeq Q_2$.

The subgroups $H < G_1\times G_2$ are therefore determined by subgroups H_1,H_2 of G_1,G_2 which have normal subgroups N_1, N_2 such that $H_1/N_1\simeq H_2/N_2\simeq Q$. If $H_1=N_1$ and $H_2=N_2$, so the quotient groups are trivial, $H=H_1\times H_2$. From this result the order of H, $|H|=|N_1|\,|N_2|\,|Q|=|H_1|\,|H_2|/|Q|$.

As an illustration if we consider the subgroups of $\mathbb{Z}_2\times\mathbb{Z}_4$ in (1.19) which are not simple direct products we have

$$\begin{aligned}&H=\{e,ba^2\}\,,\quad H_1=\{e,b\}\,,\quad N_1=\{e\}\,,\quad H_2=\{e,a^2\}\,,\quad N_2=\{e\}\,,\\&H=\{e,ba,a^2,ba^3\}\,,\quad H_1=\{e,b\}\,,\ N_1=\{e\}\,,\quad H_2=\{e,a,a^2,a^3\}\,,\ N_2=\{e,a^2\}\,,\end{aligned}\tag{1.78}$$

In both examples $H_1/N_1\simeq H_2/N_2\simeq\mathbb{Z}_2$. For $G\times G$ the diagonal subgroup corresponds to taking $H_1=H_2=G$ and $N_1=N_2=\{e\}$ with the map φ obtained by taking $g_1=g_2$.

1.5 Quaternion Groups

Quaternions are defined in terms $\{i,j,k\}$ which have the properties

$$i^2=j^2=k^2=-1\,,\qquad ij=-ji=k\,,\quad jk=-kj=i\,,\quad ki=-ik=j\,,\tag{1.79}$$

were famously discovered, or invented, by Hamilton[5] on 16th October 1843 and are an extension of the usual complex numbers $\mathbb{C}$. They can be defined by requiring i, j to satisfy just the relations $i^2=j^2=(ij)^2=-1$. Then

$$\mathbb{H}=\big\{q: q=x_0\,1+x_1\,i+x_2\,j+x_3\,k,\ x_i\in\mathbb{R}\big\}\,,\tag{1.80}$$

where q has a conjugate $\bar q=x_0\,1-x_1\,i-x_2\,j-x_3\,k$ and $q\bar q=|q|^21$, $|q|^2=\sum_r {x_r}^2$. If $\bar q=-q$, so that $x_0=0$, then q is *imaginary*. A unit quaternion has $|q|=1$. A crucial property of quaternions is that they satisfy $|q\,q'|=|q|\,|q'|$ and so form a *division algebra*.

Any quaternion can be expressed as a product of two imaginary quaternions. To show this suppose $q=q_0\,1+\mathrm{Im}\,q$ and choose an imaginary quaternion r such that

$$\mathrm{Im}\,q\,r+r\ \mathrm{Im}\,q=0\,,\quad \bar r=-r\,,\quad r^2=-1\quad\Rightarrow\quad r\,q\,r=-\bar q\,.\tag{1.81}$$

For $r=r_1\,i+r_2\,j+r_3\,k$ this just requires (r_1,r_2,r_3) is a unit 3-vector orthogonal to (x_1,x_2,x_3). Then another imaginary quaternion s is given by

$$s=-r\,q\quad\Rightarrow\quad \bar s=-s\,,\quad s^2=-|q|^2\,1\,,\quad q=r\,s\,.\tag{1.82}$$

Furthermore any unit quaternion q can be written as a commutator as in (1.74). For $q=q_0\,1+q_1\,u$, $\bar u=-u$, $u^2=-1$, ${q_0}^2+{q_1}^2=1$ a unit quaternion square root can be defined

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

by $\sqrt{q} = a\,1 + b\,u$ for $a^2 = \frac{1}{2}(q_0 + |q|)$, $2\,a\,b = q_1$ and then taking $\sqrt{q} = r\,s$ for imaginary unit quaternions r, s implies

$$q = r\,s\,r\,s = r^{-1}s^{-1}r\,s\,. \tag{1.83}$$

Quaternion multiplication is associative and any non zero quaternion q has a unique inverse $q^{-1} = \bar{q}/|q|^2$ so that $\mathbb{H}^* = \{q : q \in \mathbb{H}, q \neq 0\}$ forms a group. The subgroups of $\mathbb{H}^*$ define several important groups. Since the product of two unit quaternions is also a unit quaternion there is an infinite continuous non abelian group

$$\mathrm{Q} = \{q : q \in \mathbb{H},\, |q| = 1\}\,, \qquad q^{-1} = \bar{q}\,. \tag{1.84}$$

The centre $\mathcal{Z}(\mathrm{Q}) \simeq \{1, -1\}$ and the quotient $\mathrm{Q}/\{1,-1\} = \{q : |q| = 1, q \sim -q\}$. From (1.83) $\mathrm{Q} \simeq [\mathbb{H}^*, \mathbb{H}^*]$ and the associated quotient group $\mathbb{H}^*/[\mathbb{H}^*, \mathbb{H}^*] = \{|q| : q \in \mathbb{H}^*\}$ is just the group formed by positive real numbers under multiplication.

Q contains abelian subgroups $\{e^{\theta\,\hat{q}} = \cos\theta\,1 + \sin\theta\,\hat{q} : \hat{q} = \alpha\,i + \beta\,j + \gamma\,k,\ |\hat{q}| = 1,\ 0 \leq \theta < 2\pi\}$.

There are also finite subgroups of Q. It is easy to see that

$$\begin{aligned} C_n &= \left\{e^{2\pi r/n\,i} : r = 0, 1, \dots, n-1\right\}, \qquad n = 1, 2, \dots\,, \\ \mathrm{Q}_{4n} &= \left\{e^{\pi r/n\,i},\, e^{\pi r/n\,i}j : r = 0, 1, \dots, 2n-1\right\}, \quad n = 1, 2, \dots\,, \end{aligned} \tag{1.85}$$

form groups of order n, $4n$ respectively. Clearly $C_n \simeq \mathbb{Z}_n$ and form the *cyclic* groups. The groups Q_{4n} are referred to as *dicyclic* or *binary dihedral* or *generalised quaternion* groups. They may also be denoted, for later convenience, by $2D_n$. For $n = 1$ the group is not dicyclic but we may identify $2D_1 \simeq \mathbb{Z}_4$. For $n = 2$ this group becomes the non abelian quaternion group

$$\mathrm{Q}_8 = \left\{\pm 1,\, \pm i,\, \pm j,\, \pm k\right\}. \tag{1.86}$$

Q_{4n} contains C_{2n} as a normal subgroup. The conjugacy classes for Q_{4n} are very similar to those for D_{2n} in (1.58)

$$\begin{aligned} &\mathcal{C}_{1,1} = \{1\}\,, \quad \mathcal{C}_{1,2} = \{-1\}\,, \quad \mathcal{C}_{2,r} = \left\{e^{\pi r/n\,i},\, e^{\pi(n-r)r/n\,i}\right\},\ r = 1, \dots, n-1\,, \\ &\mathcal{C}_{n,1} = \left\{e^{\pi 2r/n\,i}j : r = 0, 1, \dots, n-1\right\}, \quad \mathcal{C}_{n,2} = \left\{e^{\pi(2r+1)/n\,i}j : r = 0, 1, \dots, n-1\right\}. \end{aligned} \tag{1.87}$$

Q_8 can be decomposed into five conjugacy classes which each have one or two elements, $\mathcal{C}_{1,1} = \{1\}, \mathcal{C}_{1,2} = \{-1\}, \mathcal{C}_{2,1} = \{\pm i\}, \mathcal{C}_{2,2} = \{\pm j\}, \mathcal{C}_{2,3} = \{\pm k\}$. Q_{4n} contains a normal subgroup $\{1, -1\}$ and the quotient $\mathrm{Q}_{4n}/\{1,-1\} \simeq D_n$ which is not a subgroup of Q_{4n}.

The resemblance of Q_{4n} to the dihedral group may be shown by defining it by the conditions

$$a^{2n} = e\,, \qquad a^n = b^2\,, \qquad b\,a\,b^{-1} = a^{-1}\,, \tag{1.88}$$

where in terms of quaternions $a = e^{\pi/n\,i}$, $b = j$ and of course $e = 1$ while $a^n = b^2 = -1$.

There are further interesting finite groups which can be generated using two imaginary unit quaternions u, v, $u^2 = v^2 = -1$. If we consider arbitrary products of u, v then we must have for a finite group

$$(uv)^n = 1 \quad \text{for some minimal } n\,. \tag{1.89}$$

The group $\{\pm 1, \pm v, \pm(uv)^r, \pm(uv)^r v : r = 1, \dots, n-1\}$ subject to (1.89), so that $(vu)^r = (uv)^{n-r}$, is isomorphic to Q_{4n} in (1.85). More general groups are obtained by considering products of r and s where initially we take

$$r = u\,, \quad s = e^{\pi/3\,v} = \tfrac{1}{2}\big(1 + \sqrt{3}\,v\big)\,, \qquad r^2 = s^3 = -1\,, \tag{1.90}$$

and then requiring for finiteness that, for some $n = 1, 2, \dots,$

$$t = rs = \pm\cos\pi/n\,1 + \sin\pi/n\,z\,, \quad z = -\bar{z},\ z^2 = -1 \quad \Rightarrow \quad t^n = -(\pm 1)^n\,, \tag{1.91}$$

where $\cos\pi/n \geq 0$. Assuming $uv = \cos\phi\,1 + \sin\phi\,y$, with y an imaginary unit quaternion,

$$\tfrac{1}{2}\sqrt{3}\,\cos\phi = \pm\cos\pi/n \quad \Rightarrow \quad -\tfrac{1}{2}\sqrt{3} \leq \cos\pi/n \leq \tfrac{1}{2}\sqrt{3}\,. \tag{1.92}$$

This is only possible if $n = 2, 3, 4, 5, 6$ but for $n = 2$, $\cos\phi = 0$ and the corresponding group is isomorphic to Q_{12} and for $n = 6$ then we may take $\cos\phi = \pm 1$ and hence $v = \mp u$ and the group isomorphic to C_{12}. The interesting cases are then[6]

$$n = 3\,, \ \ \cos\tfrac{1}{3}\pi = \tfrac{1}{2}\,, \qquad n = 4\,, \ \ \cos\tfrac{1}{4}\pi = \tfrac{1}{\sqrt{2}}\,, \qquad n = 5\,, \ \ \cos\tfrac{1}{5}\pi = \tfrac{1}{4}\big(\sqrt{5} + 1\big)\,, \tag{1.93}$$

and we may take

$$\begin{gathered} s = \tfrac{1}{2}(1 + i + j + k)\,, \quad n = 3\,, \ r = i\,, \ t^3 = 1\,, \quad n = 4\,, \ r = \tfrac{1}{\sqrt{2}}(i + j)\,, \ t^4 = -1\,, \\ n = 5\,, \ r = \tfrac{1}{2}(i + \sigma\,j + \tau\,k)\,, \ \sigma = \tfrac{1}{2}\big(\sqrt{5} - 1\big)\,, \ \tau = \tfrac{1}{2}\big(\sqrt{5} + 1\big)\,, \ t^5 = 1\,. \end{gathered} \tag{1.94}$$

1.5.1 Tetrahedral, Octahedral and Icosahedral Groups

The group corresponding to $n = 3$ can be obtained by noting that Q_8 has an external automorphism generated by $(i, j, k) \to (j, k, i) = \frac{1}{4}(1+i+j+k)(i, j, k)(1-i-j-k)$ so that

$$\mathrm{T} = \mathbb{Z}_3 \ltimes \mathrm{Q}_8 = \mathrm{Q}_8 \cup \big\{\tfrac{1}{2}(\pm 1 \pm i \pm j \pm k)\big\}\,, \tag{1.95}$$

where there are 16 possible choices of $\pm$ (each $\pm$ is independent) so that T has order 24 and is referred to as the *binary tetrahedral group*, also denoted as $2T$. For the group elements not in Q_8

$$\big(\tfrac{1}{2}(-1 + q)\big)^3 = 1\,, \quad \big(\tfrac{1}{2}(1 + q)\big)^6 = 1\,, \quad q \in \{\,\pm i \pm j \pm k\}\,. \tag{1.96}$$

There are seven conjugacy classes comprising T given by

$$\begin{gathered} \mathcal{C}_1 = \{1\}\,, \qquad \mathcal{C}_2 = \{-1\}\,, \qquad \mathcal{C}_4 = \{\pm i, \pm j, \pm k\}\,, \\ \mathcal{C}_3 = \big\{\tfrac{1}{2}(-1 + q)\big\}\,, \quad \mathcal{C}'_3 = \big\{\tfrac{1}{2}(-1 - q)\big\}\,, \quad \mathcal{C}_6 = -\mathcal{C}'_3\,, \quad \mathcal{C}'_6 = -\mathcal{C}_3\,, \\ q \in \{i + j + k, -i - j + k, -i + j - k, i - j - k\}\,, \end{gathered} \tag{1.97}$$

where here we label the conjugacy classes by the order of the group elements comprising them. These have dimensions 1+1+6+4+4+4+4 = 24. Without any specific choice of r, s other

[6]The real part of $(e^{i\pi/5})^5 + 1 = 0$ leads to a quintic equation for $x = \cos\pi/5$ which can be factorised as $(x+1)(4x^2 - 2x - 1)^2 = 0$. The other roots determine $\cos\pi = -1$ and $\cos 3\pi/5 = \frac{1}{4}(-\sqrt{5} + 1)$.

than satisfying $r^2 = s^3 = -1$, $(rs)^3 = 1$ the elements of T are determined by the conjugacy classes $\mathcal{C}'_3 = \{-s, rs, rsr, srs\}$, $\mathcal{C}_3 = \{s^2, sr, rs^2, s^2r\}$, $\mathcal{C}_4 = \{\pm r, \pm srs^2, \pm s^2rs\}$ which along with $\mathcal{C}_1, \mathcal{C}_2$ and $-\mathcal{C}'_3$, $-\mathcal{C}_3$ give the whole group. In this case $T = \mathrm{T}/\{1,-1\} \simeq A_4$ where for $q \in \mathrm{T}$, T is obtained by assuming $q \sim -q$. Hence T has four conjugacy classes inherited from those for T which correspond to $\mathcal{C}_1 = \{1\}, \mathcal{C}_2 = \{i,j,k\}, \mathcal{C}_3 = \{\frac{1}{2}(-1+q)\}, \mathcal{C}'_3 = \{\frac{1}{2}(-1-q)\}$, with four possible q as in (1.97), and where in this case $\mathcal{C}_n$ is composed of quaternions with $q^n = \pm 1$. The dimensions of each class are then $1+3+4+4 = 12$.

The binary tetrahedral group T has a $\mathbb{Z}_2$ automorphism which is generated by taking $(i,\,j,\,k) \to (-i,\,k,\,j) = -\frac{1}{2}(j+k)(i,\,j,\,k)(j+k)$ linking the conjugacy classes $\mathcal{C}_3, \mathcal{C}'_3$ and also $\mathcal{C}_6, \mathcal{C}'_6$ in (1.97). Hence for $n = 4$ the *binary octahedral group*, O or $2O$, is given by

$$\mathrm{O} = \mathbb{Z}_2 \ltimes \mathrm{T} = \mathrm{T} \cup \Big\{\tfrac{1}{\sqrt{2}}(\pm q \pm q') : q, q' = (1,i),\,(1,j),\,(1,k),\,(i,j),\,(j,k),\,(k,j)\Big\}\,, \qquad (1.98)$$

which is of order 48. There are also seven conjugacy classes for O given, in a similar notation, by

$$\begin{aligned}
&\mathcal{C}_1 = \{1\}\,, \qquad \mathcal{C}_2 = \{-1\}\,, \qquad \mathcal{C}_4 = \Big\{\pm i, \pm j, \pm k\Big\}\,,\\
&\mathcal{C}_3 = \Big\{\tfrac{1}{2}(-1 \pm i \pm j \pm k)\Big\}\,, \qquad \mathcal{C}_6 = -\mathcal{C}_3\,,\\
&\mathcal{C}'_4 = \Big\{\tfrac{1}{\sqrt{2}}(\pm i \pm j),\, \tfrac{1}{\sqrt{2}}(\pm j \pm k),\, \tfrac{1}{\sqrt{2}}(\pm i \pm k)\Big\}\,,\\
&\mathcal{C}_8 = \Big\{\tfrac{1}{\sqrt{2}}(\pm 1 \pm i),\, \tfrac{1}{\sqrt{2}}(\pm 1 \pm j),\, \tfrac{1}{\sqrt{2}}(\pm 1 \pm k)\Big\}\,,
\end{aligned} \qquad (1.99)$$

with dimensions $1+1+6+8+8+12+12 = 48$. The group $O = \mathrm{O}/\{1,-1\} \simeq \mathcal{S}_4$ and is composed of five conjugacy classes $\mathcal{C}_1 = \{1\}, \mathcal{C}_2 = \{i,j,k\}, \mathcal{C}'_2 = \Big\{\frac{1}{\sqrt{2}}(q \pm q') : q, q' = (i,j),\,(j,k),\,(k,i)\Big\}, \mathcal{C}_3 = \Big\{\frac{1}{2}(-1+q) : q \in \{\pm i \pm j \pm k\}\Big\}, \mathcal{C}_4 = \Big\{\frac{1}{\sqrt{2}}(1 \pm q) : q = i,\,j,\,k\Big\}$ where $1+3+6+8+6 = 24$.

For $n = 5$ there is a group I, the *binary icosahedral group* also denoted as $2I$, which can be defined through the conjugacy classes

$$\begin{aligned}
&\mathcal{C}_1 = \{1\}\,, \qquad \mathcal{C}_2 = \{-1\}\,,\\
&\mathcal{C}_4 = \Big\{\pm i,\, \pm j,\, \pm k,\, \tfrac{1}{2}(\pm i \pm \sigma j \pm \tau k),\, \tfrac{1}{2}(\pm \tau i \pm j \pm \sigma k),\, \tfrac{1}{2}(\pm \sigma i \pm \tau j \pm k)\Big\}\,,\\
&\mathcal{C}_6 = \Big\{\tfrac{1}{2}(1 \pm i \pm j \pm k),\, \tfrac{1}{2}(1 \pm \tau i \pm \sigma j),\, \tfrac{1}{2}(1 \pm \sigma i \pm \tau k),\, \tfrac{1}{2}(1 \pm \tau j \pm \sigma k)\Big\}\,,\ \mathcal{C}_3 = -\mathcal{C}_6\,,\\
&\mathcal{C}_5 = \Big\{\tfrac{1}{2}(\sigma \pm i \pm \tau j),\, \tfrac{1}{2}(\sigma \pm j \pm \tau k),\, \tfrac{1}{2}(\sigma \pm k \pm \tau i)\Big\}\,, \qquad \mathcal{C}_{10} = -\mathcal{C}_5\,,\\
&\mathcal{C}'_5 = \Big\{\tfrac{1}{2}(-\tau \pm i \pm \sigma k),\, \tfrac{1}{2}(-\tau \pm j \pm \sigma i),\, \tfrac{1}{2}(\tau \pm k \pm \sigma j)\Big\}\,, \qquad \mathcal{C}'_{10} = -\mathcal{C}'_5\,,
\end{aligned} \qquad (1.100)$$

with dimensions $1+1+30+20+20+12+12+12+12 = 120$.[7] In this case the quotient $I = \mathrm{I}/\{1,-1\} \simeq A_5$ with five conjugacy classes, $\mathcal{C}_1, \mathcal{C}_3, \mathcal{C}_5, \mathcal{C}'_5$, as in (1.100), and also $\mathcal{C}_2 = \Big\{i, j, k, \frac{1}{2}(i \pm \sigma j \pm \tau k), \frac{1}{2}(\pm \tau i + j \pm \sigma k), \frac{1}{2}(\pm \sigma i \pm \tau j + k)\Big\}$. Thus $1+20+12+12+15 = 60$. The closure of (1.100) under multiplication depends on $\sigma\tau = 1$, $\tau^2 - \tau - 1 = 0$ implying $\tau - \sigma = 1$, $\tau^2 + \sigma^2 = 3$. There are two solutions, one is given in (1.94), the other is obtained by taking $\tau \leftrightarrow -\sigma$. The binary icosahedral group contains the subgroup $\mathrm{T} = 2T$, which may be generated by taking $r = \frac{1}{2}(i + \sigma j + \tau k)$, $s = \frac{1}{2}(1 + \tau i + \sigma j)$, $rs = \frac{1}{2}(-1 + \tau j + \sigma k)$, where

[7] Although I has order 120 it is not isomorphic to $\mathcal{S}_5$ or $\mathbb{Z}_2 \times A_5$.

$r^2 = s^3 = -1$, $(rs)^3 = 1$. The cosets t^rT, $r = 1,2,3,4$, for $t = \frac{1}{2}(\sigma + i + \tau j)$, $t^5 = 1$, give, with T, the full group I. This verifies $|\mathrm{I}| = 5|\mathrm{T}| = 120$.

More generally, if instead of (1.90), we were to take $s = \cos\pi/m\, 1 + \sin\pi/m\, v$, so that $s^m = -1$, then requiring (1.91) leads to $-\sin\pi/m \le \cos\pi/n \le \sin\pi/m$. If $n = 2$ this gives Q_{4m}. Otherwise for $m > 3$ the only possibilities are $m = 4$, $n = 3,4$ and $m = 5,6$, $n = 3$. This just leads to groups already discussed above other than if $m = n = 4$ when $\cos\phi = \pm 1$ and the group becomes C_8. In summary the finite quaternion groups which contain the element -1 are just

$2G$	$2I$	$2O$	$2T$	$2D_n$	$2C_n$
order	120	48	24	$4n$	$2n$
normal subgroups	$\mathbb{Z}_2$	$2T, 2D_2, \mathbb{Z}_2$	$2D_2, \mathbb{Z}_2$	$2D_{\frac{1}{2}n}\ 2\vert n$	$2C_k\ k\vert n$
				$2C_k\ k\vert n, C_k\ k\vert n\ k\,\text{odd}$	$C_k\ k\vert n\ k\,\text{odd}$
$G = 2G/\{1,-1\}$	$I \simeq A_5$	$O \simeq \mathcal{S}_4$	$T \simeq A_4$	D_n	C_n

(1.101)

where for completeness $2C_n = \left\{e^{\pi r/n\, i} : r = 0,1,\dots,2n-1\right\} \simeq C_{2n}$ and $k|n$ denotes k divides n. The only remaining finite group is C_n for n odd, as defined in (1.85). The groups T, O, I are related to the symmetries of the tetrahedron, cube or equivalently octahedron, dodecahedron or equivalently icosahedron (in crystallographic literature I is denoted by Y). Abstractly they can be generated by elements r, s, $t = rs$ satisfying $r^l = s^m = t^n = e$. With labels (n,m,l) then D_n, T, O, I correspond to $(n,2,2)$, $(3,3,2)$, $(4,3,2)$, $(5,3,2)$ where for any such group P

$$\frac{1}{l} + \frac{1}{m} + \frac{1}{n} = 1 + \frac{2}{|P|}\,. \qquad (1.102)$$

The labels for D_n, T, O, I are the unique integers n,m,l, up to a reordering, such that the left hand side in (1.102) is > 1.

Under conjugation the real part of any quaternion group element is invariant so the imaginary parts of the unit quaternions in each conjugacy class (1.97), (1.99) and (1.100) form closed sets under conjugation by any element of T, O and I respectively. Under conjugation the normal $\mathbb{Z}_2$ subgroups $\{1,-1\}$ leave any quaternion invariant so the imaginary parts define points in three dimensional space which are closed under the action of T, O and I. For T from (1.97) this gives the four points $(1,1,1)$, $(-1,-1,1)$, $(-1,1,-1)$, $(1,-1,-1)$ which form the vertices of a tetrahedron and also the six points $(\pm1,0,0)$, $(0,\pm1,0)$, $(0,0,\pm1)$ correspond to the midpoints of the tetrahedron edges and form the vertices of an octahedron. For O using (1.99) the same octahedron reappears and also $(\pm1,\pm1,\pm1)$ giving the eight vertices of a cube. For I the results in (1.100) lead to two sets of 12 points $(\pm1,\pm\tau,0)$, $(0,\pm1,\pm\tau)$, $(\pm\tau,0,\pm1)$ and $(\pm1,0,\pm\sigma)$, $(\pm\sigma,\pm1,0)$, $(0,\pm\sigma,\pm1)$ which, since $\sigma = 1/\tau$, are both the vertices of icosahedron (where 5 triangles meet at each vertex). The 20 points $\frac{1}{2}(\pm1,\pm1,\pm1)$, $\frac{1}{2}(\pm\tau,\pm\sigma,0)$, $(0,\pm\tau,\pm\sigma)$, $\frac{1}{2}(\pm\sigma,0,\pm\tau)$ arising from the conjugacy class $\mathcal{C}_6$ form the vertices of a dodecahedron (where 3 pentagons meet at each vertex). The remaining cases are, for O from $\mathcal{C}'_4$, $\frac{1}{\sqrt{2}}(\pm1,\pm1,0)$, $\frac{1}{\sqrt{2}}(0,\pm1,\pm1)$, $\frac{1}{\sqrt{2}}(\pm1,0,\pm1)$ and, for I from the conjugacy class $\mathcal{C}_4$, $(\pm1,\pm1,\pm1)$, $\frac{1}{2}(\pm1,\pm\sigma,\pm\tau)$, $\frac{1}{2}(\pm\tau,\pm1,\pm\sigma)$, $\frac{1}{2}(\pm\sigma,\pm\tau,\pm1)$ which correspond to the midpoints of the 12 and 30 edges respectively. These are the vertices of the cuboctahedron (this has 8 triangle and 6 square faces, two of each meeting at each vertex) and icosidodecahedron (this has 20 triangle and 12 pentagon faces, with two each meeting at each vertex).

1.6 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. The commutator subgroup for $Gl(n,\mathbb{R})$, $[Gl(n,\mathbb{R}),Gl(n,\mathbb{R})]\simeq Sl(n,\mathbb{R})$. It is easy to see that $A^{-1}B^{-1}AB$ has unit determinant for any $A,B\in Gl(n,\mathbb{R})$ and any element in $Sl(n,\mathbb{R})$ can be obtained as a commutator. The same of course applies for $Gl(n,\mathbb{C})$ and $Sl(n,\mathbb{C})$. For $M\in Gl(n,\mathbb{R})$ there are n^2 real parameters while for $M\in Sl(n,\mathbb{R})$, with one condition, there are n^2-1. The same applies for $Gl(n,\mathbb{C})$ and $Sl(n,\mathbb{C})$ although the parameters are then complex.

The trivial one dimensional case for $Gl(1,\mathbb{R})\simeq\mathbb{R}$ where for $x,y\in\mathbb{R}$ the group operation is just addition, the identity is 0 and the inverse of x is $-x$.

Various matrix groups which are subgroups of $Gl(n,\mathbb{R})$ or $Gl(n,\mathbb{C})$ are obtained by requiring a bilinear or sesquilinear quadratic form $\langle x,y\rangle$ for $x,y\in\mathbb{R}^n$ or $x,y\in\mathbb{C}^n$ is invariant under the group action on x,y. For a sesquilinear form $\langle x,y\rangle$ then $\langle x,y+y'\rangle=\langle x,y\rangle+\langle x,y'\rangle$, $\langle x+x',y\rangle=\langle x,y\rangle+\langle x',y\rangle$ and $\langle\alpha\,x,\beta\,y\rangle=\alpha^*\beta\langle x,y\rangle$ for $\alpha,\beta\in\mathbb{C}$. The sesquilinear form is hemitian if $\langle x,y\rangle=\langle y,x\rangle^*$.

Continuous such matrix groups of frequent interest are

1.6.1 Orthogonal

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

$$M^TM=\mathbb{1}\,. \tag{1.103}$$

This set of matrices is closed under multiplication since $(M_1M_2)^T={M_2}^T{M_1}^T$. For $SO(n)$ $\det M=1$. The invariant quadratic form is just $\langle x,x\rangle=x^Tx$. With $x\in\mathbb{R}^n$ this is positive definite. 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 (1.103) provides $\frac{1}{2}n(n+1)$ conditions. Hence $O(n)$, and also $SO(n)$, have $\frac{1}{2}n(n-1)$ parameters. For n even $\pm\mathbb{1}\in SO(n)$ and these form the centre of the group so long as $n\geq 2$. Thus $\mathcal{Z}(SO(2n))\simeq\mathbb{Z}_2$, $n=2,3,\ldots$, while $\mathcal{Z}(SO(2n+1))=\{\mathbb{1}\}$, $n=1,2,\ldots$ is trivial although $\mathcal{Z}(O(2n+1))=\{\pm\mathbb{1}\}\simeq\mathbb{Z}_2$, $n=1,2,\ldots$ and $O(2n+1)\simeq\mathbb{Z}_2\times SO(2n+1)$. For the trivial one dimensional case $O(1)\simeq\mathbb{Z}_2$.

1.6.2 Unitary

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

$$M^\dagger M = \mathbb{1}\,. \tag{1.104}$$

Closure follows from $(M_1 M_2)^\dagger = M_2{}^\dagger M_1{}^\dagger$. For $SU(n)$ $\det M = 1$. The invariant quadratic form $\langle x, x\rangle = x^\dagger x$ for $x \in \mathbb{C}^n$ is hermitian. 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. (1.104) requires $|\det M| = 1$ so imposing $\det M = 1$ now provides one additional condition so that $SU(n)$ has $n^2 - 1$ parameters. The centre of $U(n)$ or $SU(n)$ consists of all elements proportional to the identity (this follows from Schur's lemma shown later) so that $\mathcal{Z}(SU(n)) = \{e^{2r\pi i/n}\mathbb{1} : r = 0, \dots n-1\} \simeq \mathbb{Z}_n$, while $\mathcal{Z}(U(n)) = \{e^{i\alpha}\mathbb{1} : 0 \le \alpha < 2\pi\} \simeq U(1)$.

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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad 0 \le \theta < 2\pi\,, \tag{1.105}$$

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

$$e^{i\theta}\,, \qquad 0 \le \theta < 2\pi\,. \tag{1.106}$$

Topologically $U(1) \simeq S^1$, the circle.

1.6.3 Symplectic

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

$$M^T J_{2n} M = J_{2n}\,, \tag{1.107}$$

where J_{2n} is a $2n \times 2n$ antisymmetric matrix with the standard form

$$J_{2n} = \begin{pmatrix} \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & 0 & & \\ 0 & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & & \\ & & \ddots & \\ & & & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \end{pmatrix}. \tag{1.108}$$

In this case $M^T J_{2n} M$ is antisymmetric so that (1.107) 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 $\langle v', v\rangle = -\langle v, v'\rangle = v'^T J_{2n} v$ so that $\langle x, x\rangle = 0$. For an orthonormal basis $\{e_i\}$ we may define $J_{ij} = \langle e_i, e_j\rangle$.

The condition (1.107) requires $\det M = 1$ so there are no further restrictions as for $O(n)$ and $U(n)$. To show this we define the *Pfaffian*[8] for $2n \times 2n$ antisymmetric matrices A by

$$\mathrm{Pf}(A) = \frac{1}{2^n n!}\,\varepsilon_{i_1 \dots i_{2n}} A_{i_1 i_2} \dots A_{i_{2n-1} i_{2n}}\,, \tag{1.109}$$

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

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

$$\det A = \mathrm{Pf}(A)^2 \,, \tag{1.110}$$

and it is easy to see that

$$\mathrm{Pf}(J_{2n}) = 1 \,. \tag{1.111}$$

The critical property here is that

$$\mathrm{Pf}(M^T A M) = \det M \, \mathrm{Pf}(A) \quad \text{since} \quad \varepsilon_{i_1 \dots i_{2n}} M_{i_1 j_1} \dots M_{i_{2n} j_{2n}} = \det M \, \varepsilon_{j_1 \dots j_{2n}} \,. \tag{1.112}$$

Applying (1.112) with $A = J_{2n}$ to the definition of symplectic matrices in (1.107) shows that we must have $\det M = 1$.

Since both $\pm\mathbb{1}$ belong to $Sp(2n, \mathbb{R})$ then the centre $\mathcal{Z}(Sp(2n, \mathbb{R})) \simeq \mathbb{Z}_2$.

For $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in Sp(2, \mathbb{R})$ the condition (1.107) requires just $ad - bc = 1$ or $\det M = 1$. Hence $Sp(2, \mathbb{R}) \simeq Sl(2, \mathbb{R})$.

1.6.4 Quaternion Matrix Groups

Matrix groups can also be extended to quaternions where a $n \times n$ quaternionic matrix M has the form

$$M = a\, 1 + b\, i + c\, j + d\, k \,, \quad a, b, c, d \ \text{ real } n \times n \ \text{ matrices} \,, \tag{1.113}$$

and the adjoint is

$$\bar{M} = a^T 1 - b^T i - c^T j - d^T k \,. \tag{1.114}$$

Matrix multiplication of non singular, or invertible, $n \times n$ quaternionic matrices defines $Gl(n, \mathbb{H})$ since quaternions obey the crucial associativity property.[9] The notion of a determinant with the usual properties is problematic for quaternionic matrices but $Sl(n, \mathbb{H})$ can be defined as the commutator group $[Gl(n, \mathbb{H}), Gl(n, \mathbb{H})]$. A definition analogous to a determinant due to Dieudonné[10] is based on the quotient group $Gl(n, \mathbb{H})/[Gl(n, \mathbb{H}), Gl(n, \mathbb{H})]$ which is a one dimensional abelian group. This can be expressed, for $M \in Gl(n, \mathbb{H})$, in terms of a real $|M| \geq 0$, $|M| = |MK| = |KM|$ for $K \in [Gl(n, \mathbb{H}), Gl(n, \mathbb{H})]$, satisfying the group properties $|M_1 M_2| = |M_1||M_2|$, where $|\bar{M}| = |M|$ and $|rM| = |r|^n |M|$ for $r \in \mathbb{R}$. In general $|M| = 0$ if and only if M is singular, so that there is a quaternionic column vector v such that $Mv = 0$, and there is no inverse.[11] For $M \in Sl(n, \mathbb{H})$ then $|M| = 1$ which is preserved

[9] The inverse is a little more complicated than the usual matrix inverse since quaternions do not commute. A single non zero quaternion has an inverse. For $M \in Gl(2, \mathbb{H})$ then writing $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ ca^{-1} & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} a & 0 \\ 0 & d' \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & a^{-1}b \\ 0 & 1 \end{smallmatrix}\right)$ with $a \neq 0$ and $d' = d - ca^{-1}b \neq 0$ then $M^{-1} = \left(\begin{smallmatrix} 1 & -a^{-1}b \\ 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & d'^{-1} \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 \\ -ca^{-1} & 1 \end{smallmatrix}\right)$ is both a right and left inverse for M. Alternatively for $d \neq 0$, $M = \left(\begin{smallmatrix} 1 & bd^{-1} \\ 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} a' & 0 \\ 0 & d \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 \\ d^{-1}c & 1 \end{smallmatrix}\right)$, $a' = a - bd^{-1}c$ leads to an equivalent expression for M^{-1}. These results can be generalised to larger square quaternionic matrices.

[10] Jean Dieudonné, 1906-92, French. Founder member of Bourbaki.

[11] For $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in Gl(2, \mathbb{H})$ then $|M| = |a|\,|d - ca^{-1}b| = |d|\,|a - bd^{-1}c|$. Note that if $a \to a + q\,c$, $b \to b + q\,d$ or $b \to b + a\,q$, $d \to d + c\,q$, $q \in \mathbb{H}$, which correspond to $M \to QM$ or $M \to MQ$ for $Q = \left(\begin{smallmatrix} 1 & q \\ 0 & 1 \end{smallmatrix}\right)$, $|M|$ is invariant. For $|M| = 0$ and $|a| \neq 0$, when it is necessary that $d - ca^{-1}b = 0$, we can take for the zero eigenvalue vector $v = \left(\begin{smallmatrix} a^{-1}b \\ -1 \end{smallmatrix}\right)$.

under matrix multiplication. While $M \in Gl(n,\mathbb{H})$ has $4n^2$ real parameters, $M \in Sl(n,\mathbb{H})$ has $4n^2 - 1$.

For a single quaternion the group Q defined in (1.84) is isomorphic to SU(2). To show this we relate the quaternions to 2×2 matrices according to

$$1 \to \mathbb{1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \quad i \to I = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right), \quad j \to J = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \quad k \to K = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right). \tag{1.115}$$

Then

$$q = x_0\,1 + x_1\,i + x_2\,j + x_3\,k, \quad \leftrightarrow \quad \mathcal{Q} = \begin{pmatrix} x_0 + x_1\,i & x_3 + x_4\,i \\ -x_3 + x_4\,i & x_0 - x_1\,i \end{pmatrix}, \quad x_0, x_1, x_2, x_3 \in \mathbb{R}, \tag{1.116}$$

ensures

$$q_1 q_2 \;\leftrightarrow\; \mathcal{Q}_1 \mathcal{Q}_2\,, \qquad \bar{q} \leftrightarrow \mathcal{Q}^\dagger\,, \qquad |q|^2 \leftrightarrow \det \mathcal{Q}\,. \tag{1.117}$$

Furthermore

$$J\mathcal{Q} = \mathcal{Q}^* J\,. \tag{1.118}$$

Note that from (1.116) $\det \mathcal{Q} = {x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2 = |q|^2$ so that $SU(2)$ can be identified with points on S^3.

For any quaternion matrix M there is an associated $2n\times 2n$ complex matrix $\mathcal{M}$ obtained by replacing quaternions by 2×2 matrices as in (1.115)

$$M \to \mathcal{M}\,, \quad \bar{M} \to \mathcal{M}^\dagger\,, \quad \mathbb{1}_n 1 \to \mathbb{1}_{2n}\,, \quad \mathbb{1}_n j \to \tilde{J}_{2n} \quad \Rightarrow \quad \mathcal{M}^* = -\tilde{J}_{2n}\,\mathcal{M}\tilde{J}_{2n}\,, \tag{1.119}$$

where $M_1 M_2 \to \mathcal{M}_1 \mathcal{M}_2$, $M^{-1} \to \mathcal{M}^{-1}$. With M as in (1.113)

$$\mathcal{M} = \begin{pmatrix} e & f \\ -f^* & e^* \end{pmatrix}, \quad e = a + b\,i\,, \; f = c + d\,i\,, \qquad \tilde{J}_{2n} = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}. \tag{1.120}$$

With the usual notion of a determinant for complex matrices $\det \mathcal{M} = |M|^2$.

The unitary quaternion matrix group $U(n,\mathbb{H})$ is defined by $n\times n$ matrices of quaternions where for any $M \in U(n,\mathbb{H})$ with an associated adjoint $\bar{M}$, defined as in (1.113) and (1.114),

$$\bar{M}M = \mathbb{1}_n\,1\,, \tag{1.121}$$

for $\mathbb{1}_n$ the unit $n\times n$ matrix. From (1.116) $U(1,\mathbb{H}) \simeq SU(2)$. 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 (1.121) provides $n(2n-1)$ conditions so that $U(n,\mathbb{H})$ has $n(2n+1)$ parameters. The condition (1.121) requires $|M| = 1$. From (1.116) $U(1,\mathbb{H}) \simeq SU(2)$. In this case under the map (1.119) $\mathcal{M}$ satisfies, since $\mathcal{M}^* = (\mathcal{M}^T)^{-1}$, satisfies (1.107) as well as $\mathcal{M}^\dagger\mathcal{M} = \mathbb{1}_{2n}$ and so $\mathcal{M} \in Sp(2n,\mathbb{C}) \cap U(2n)$. These properties are preserved under multiplication and define the group $USp(2n)$ which is therefore isomorphic to $U(n,\mathbb{H})$. This group may also be denoted by $Sp(n)$.

1.6.5 Heisenberg Group

The Heisenberg[12] group H is defined in terms of upper triangular 3×3 matrices

$$A(a,b,c) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a,b,c \in \mathbb{R}, \tag{1.122}$$

where under matrix multiplication $A(a,b,c)A(a,b',c') = A(a+a', b+b', c+c'+ab')$ and the inverse $A(a,b,c)^{-1} = A(-a,-b,-c+ab)$. This group is connected with the position momentum commutation relations in quantum mechanics. The Heisenberg group is non abelian and since $A(a,b,c)A(a,b',c')A(a,b,c)^{-1} = A(0,0,c'+ab'-ba')$ then $\{A(0,0,c)\} \simeq \mathbb{R}$ forms its centre $\mathcal{Z}(H)$. Clearly $\{A(a,0,0)\} \simeq \{A(0,b,0)\} \simeq \mathbb{R}$ are abelian subgroups. There is a discrete infinite subgroup by restricting $a,b,c \in \mathbb{Z}$.

1.6.6 Compact and Non Compact

The matrix groups $SO(n)$, $SU(n)$ and $U(n,\mathbb{H})$ are *compact*, which will be defined precisely later but for the moment can be taken to mean that the natural paramete rs vary over a finite range. On the other hand $Gl(n,\mathbb{R})$, $Sl(n,\mathbb{R})$, as well as their complex counterparts, are non compact. Any one dimensional continuous subgroup of a compact continuous group must be isomorphic to $U(1)$ while a non compact continuous group will have at least one dimensional subgroups isomorphic to $\mathbb{R}$. $Sp(2n,\mathbb{R})$ is non compact, which is evident since matrices of the form

$$\begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}, \qquad -\infty < \theta < \infty\,. \tag{1.123}$$

belong to $Sp(2,\mathbb{R})$ and form a one dimensional subgroup isomorphic to $\mathbb{R}$. The Heisenberg group is clearly non compact.

The group $USp(2n) \simeq U(n,\mathbb{H})$ is compact.

There are also various extensions of the orthogonal and unitary groups to non compact groups which arise frequently in physics. Suppose g is the diagonal $(n+m)\times(n+m)$ matrix defined by

$$g = \begin{pmatrix} \mathbb{1}_n & 0 \\ 0 & -\mathbb{1}_m \end{pmatrix}, \tag{1.124}$$

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

$$M^T g\, M = g\,. \tag{1.125}$$

The invariant form in this case is $\langle v',v\rangle = v'^T gv$ is no longer positive. 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 (1.123).

[12] Werner Karl Heisenberg, 1901-76, German. Nobel prize 1932.

2 Representations

In physical applications of groups representations play a crucial role. 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

$$D(g_1)D(g_2) = D(g_1 g_2)\,, \tag{2.1}$$
$$D(e) = \mathbb{1}\,, \tag{2.2}$$
$$D(g^{-1}) = D(g)^{-1}\,, \tag{2.3}$$

where $\mathbb{1}$ denotes the unit matrix. If the matrices $D(g)$ are $n \times n$ the representation has *dimension n*. For each matrix group the definition of course provides a representation which is termed the *fundamental representation*.

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 there exist group elements h, not equal to e, where $D(h) = \mathbb{1}$. For all such h then $\{h\} = H$ and 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 $\bar{D}(g) = D(g)^*$ since complex conjugation preserves matrix multiplication. The matrices $(D(g)^{-1})^T$ also define a representation.

Since

$$\det\big(D(g_1)D(g_2)\big) = \det D(g_1)\,\det D(g_2)\,, \quad \det \mathbb{1} = 1\,, \quad \det D(g)^{-1} = \big(\det D(g)\big)^{-1}\,, \tag{2.4}$$

$\{\det D(g)\}$ 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

$$D'(g) = S D(g) S^{-1} \quad \text{for all} \quad g \in G\,, \tag{2.5}$$

where $D(g) \to D'(g)$ is a *similarity transformation.*

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

$$g g_i = \sum_j g_j D_{\mathrm{reg},ji}(g)\,, \tag{2.6}$$

where $[D_{\mathrm{reg},ji}(g)]$ are representation matrices with a 1 in each column and row and with all other elements zero. In general

$$D_{\mathrm{reg},ji}(g) = \begin{cases} \delta_{ji}\,, & g = e\,, \\ 0\,, j = i\,, & g \neq e\,. \end{cases} \tag{2.7}$$

As an example for $D_3 = \{e, a, a^2, b, ba, ba^2\}$, where $a^3 = b^2 = e, ab = ba^2$, then

$$D_{\mathrm{reg}}(a) = \begin{pmatrix} 0&0&1&0&0&0 \\ 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&1 \\ 0&0&0&1&0&0 \end{pmatrix}, \qquad D_{\mathrm{reg}}(b) = \begin{pmatrix} 0&0&0&1&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&1 \\ 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&1&0&0&0 \end{pmatrix}. \tag{2.8}$$

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

$$v \underset{g}{\to} v^g = D(g)v \,. \tag{2.9}$$

Transformations as in (2.5) 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

$$D(g)u \in \mathcal{U} \quad \text{for all} \quad u \in \mathcal{U} \,, \tag{2.10}$$

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 (2.10),

$$D(g) = \begin{pmatrix} \hat{D}(g) & B(g) \\ 0 & C(g) \end{pmatrix} \quad \text{for} \quad u = \begin{pmatrix} \hat{u} \\ 0 \end{pmatrix}, \tag{2.11}$$

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 (2.11) 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

$$D(R, a) = \begin{pmatrix} R & a \\ 0 & 1 \end{pmatrix}, \tag{2.12}$$

with the group multiplication rule

$$D(R_1, a_1)D(R_2, a_2) = D(R_1 R_2, R_1 a_2 + a_1) \,, \tag{2.13}$$

which has the abelian subgroup T_n for $R = \mathbb{1}$. The group defined by (2.13) 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

$$SD(g)S^{-1} = \begin{pmatrix} D_1(g) & 0 & & \\ 0 & D_2(g) & & \\ & & \ddots & \\ & & & D_k(g) \end{pmatrix}, \tag{2.14}$$

and where $D_r(g)$ form irreducible representations for each r. This can be written as

$$D(g) \simeq \bigoplus_{r=1}^{k} D_r(g)\,, \tag{2.15}$$

Thus for $\mathcal{R}$ the representation given by the matrices $D(g)$ and $\mathcal{R}_r$ corresponding to the irreducible representation matrices $D_r(g)$ then $\mathcal{R}$ is decomposed as

$$\mathcal{R} = \mathcal{R}_1 + \cdots + \mathcal{R}_k\,. \tag{2.16}$$

If there are N_G inequivalent irreducible representations they may be labelled $\mathcal{R}_r$, $r = 1, \dots n_G$, and then in general a particular irreducible representation $\mathcal{R}_r$ may appear more than once, with multiplicity m_r, in the decomposition (2.16), and (2.15) can be reduced to just

$$D(g) \simeq \bigoplus_{r=1}^{N_G} m_r\, D_r(g)\,, \tag{2.17}$$

Representations of a finite group are always completely reducible since, as shown later, representations can be shown to be equivalent to unitary representations formed from unitary matrices. This is known as *Maschke's threorem.*[13]

There is always a trivial one dimensional representation $\mathcal{R}_1$ which is irreducible and is given by

$$D_1(g) = 1 \quad \text{for all} \quad g \in G\,. \tag{2.18}$$

As an example from (2.8)

$$SD_{\text{reg}}(a)S^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & s & 0 & 0 & 0 & 0 \\ 0 & 0 & \bar{s} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & s & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{s} \end{pmatrix}, \qquad S^{-1}D_{\text{reg}}(b)S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix},$$
$$S^{-1} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & \bar{s} & s & -1 & -\bar{s} & -s \\ 1 & s & \bar{s} & -1 & -s & -\bar{s} \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & s & \bar{s} & 1 & s & \bar{s} \\ 1 & \bar{s} & s & 1 & \bar{s} & s \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad s = e^{2\pi i/3} = -\tfrac{1}{2}(1 - \sqrt{3}\,i)\,. \tag{2.19}$$

In consequence this representation decomposes into two one-dimensional representations and two equivalent two-dimensional ones.

[13] Heinrich Maschke, 1853-1908, German.

2.1 Schur's Lemmas

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

Schur's Lemmas.[14] If $D_1(g), D_2(g)$ form two irreducible representations then *(i)*

$$SD_1(g) = D_2(g)S\,, \tag{2.20}$$

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

$$SD(g) = D(g)S\,, \tag{2.21}$$

for all g for an irreducible representation $D(g)$ then $S \propto \mathbb{1}$.

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 $\mathcal{V}_1 \underset{S}{\rightarrow} \mathcal{V}_2$. 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 (2.20). 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 (2.20) 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 (2.21) 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 $S = \lambda I$.

To obtain *(ii)* it is necessary in general that $\mathcal{V}_\lambda$ is a complex vector space so that the reduction to irreducible representations must allow for complex representations. For an abelian group all matrices $D(g)$ commute and can be simultaneously diagonalised so they are reducible to one dimensional complex representations.

2.2 Induced Representations

A representation of a group G also gives a representation when restricted to a subgroup H. If the representation for G is irreducible the restricted representation for H need not be.

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

$$v \underset{h}{\rightarrow} D(h)v\,, \qquad h \in H\,, \quad v \in \mathcal{V}\,, \tag{2.22}$$

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

[14] Issai Schur, 1875-1941, Russian, worked in Germany, forced to leave in 1939 for Palestine.

$h \in H$. 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

$$g g_i = g_j h \qquad \text{for some} \quad h \in H\,, \quad i,j = 1,\dots,N\,, \quad N = |G|/|H|\,. \tag{2.23}$$

Assuming (2.23) determines h the induced representation is defined so that that under the action of a group transformation $g \in G$,

$$v_i \underset{g}{\to} D(h) v_j\,, \qquad v_i = (g_i, v)\,, \quad D(h) v_j = (g_j, D(h) v)\,. \tag{2.24}$$

In (2.24) 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 $h \in H$,

$$D_{ji}(g) = \begin{cases} D(h)\,, & g_j^{-1} g\, g_i = h \in H\,, \\ 0\,, & g_j^{-1} g\, g_i \notin H\,. \end{cases} \tag{2.25}$$

To show that (2.24) is in accord with the group multiplication rule we consider a subsequent transformation g' so that

$$v_i \underset{g}{\to} D(h) v_j \underset{g'}{\to} D(h') D(h) v_k = D(h'h) v_k \quad \text{for} \quad g' g_j = g_k h' \quad \Rightarrow \quad (g'g) g_i = g_k h' h\,. \tag{2.26}$$

The dimension of the induced representation of G is then $|G|/|H| \dim \mathcal{R}_H$ with $\mathcal{R}_H$ the representation defined by $\{D(h)\}$.

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. This shows that the induced representation is not in general irreducible.

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

$$v \underset{a}{\to} e^{\frac{2\pi k i}{n}} v\,. \tag{2.27}$$

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 (2.27) then with $v_2 = (b, v)$ (2.24) requires, using $ab = ba^{-1}$,

$$(v_1, v_2) \underset{a}{\to} \big(e^{\frac{2\pi k i}{n}} v_1, e^{-\frac{2\pi k i}{n}} v_2\big) = (v_1, v_2) A_k\,, \qquad (v_1, v_2) \underset{b}{\to} (v_2, v_1) = (v_1, v_2) B\,, \tag{2.28}$$

for 2×2 complex matrices A_k, B,

$$A_k = \begin{pmatrix} e^{\frac{2\pi k i}{n}} & 0 \\ 0 & e^{-\frac{2\pi k i}{n}} \end{pmatrix}\,, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \tag{2.29}$$

which satisfy $A_k{}^n = \mathbb{1}$, $B^2 = \mathbb{1}$, $A_k B = B A_k{}^{n-1}$ and so give a two dimensional representation of D_n for each k. By considering $A_k \to B A_k B$ it is clear that the representation for $k \to n-k$

is equivalent to that in (2.29). For n even we may take $k = 1,\dots,(n-2)/2$, for n odd $k = 1,\dots,(n-1)/2$ to give inequivalent two dimensional irreducible representations. If $k = 0$ in (2.29) $A = \mathbb{1}$ then the matrix B is reducible so that there are two one dimensional representations corresponding to taking $b \to \pm 1$. For n even then taking $k = n/2$ and $A = -\mathbb{1}$ there are similarly two more one dimensional representations. The representations are then given by, for n odd,

$$\begin{aligned}
\mathcal{R}_{1,1}: \quad & (a^r, a^r b) && \to && (1,1) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{1,2}: \quad & (a^r, a^r b) && \to && (1,-1) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{2,k}: \quad & (a^r, a^r b) && \to && (A_k{}^r, A_k{}^r B) \ , \quad r = 0,\dots,n-1\,,\ k = 1,\dots,\tfrac{1}{2}(n-1)\,,
\end{aligned} \tag{2.30}$$

and for n even

$$\begin{aligned}
\mathcal{R}_{1,1}: \quad & (a^r, a^r b) && \to && (1,1) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{1,2}: \quad & (a^r, a^r b) && \to && (1,-1) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{1,3}: \quad & (a^r, a^r b) && \to && ((-1)^r, (-1)^r) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{1,4}: \quad & (a^r, a^r b) && \to && ((-1)^r, -(-1)^r) \ , \quad r = 0,\dots,n-1\,, \\
\mathcal{R}_{2,k}: \quad & (a^r, a^r b) && \to && (A_k{}^r, A_k{}^r B) \ , \quad r = 0,\dots,n-1\,,\ k = 1,\dots,\tfrac{1}{2}(n-1)\,,
\end{aligned} \tag{2.31}$$

The number of representations match the number of conjugacy classes in (1.58). Corresponding to (2.29) there is an equivalent basis

$$R\,A_k\,R^{-1} = \begin{pmatrix} \cos\frac{2\pi k}{n} & -\sin\frac{2\pi k}{n} \\ \sin\frac{2\pi k}{n} & \cos\frac{2\pi k}{n} \end{pmatrix}, \qquad R\,B\,R^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.32}$$

where $R = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ -i & i \end{smallmatrix}\right)$, $R^{-1} = R^\dagger$.

A very similar construction works for the dicyclic group Q_{4n} as defined in (1.85). This is similar to D_{2n} where we take $a^{2n} = e$, $ab = ba^{-1}$ but now $a^n = b^2$. Thus

$$(v_1, v_2) \underset{a}{\to} \big(e^{\frac{\pi k i}{n}} v_1, e^{-\frac{\pi k i}{n}} v_2\big) = (v_1, v_2) A_k\,, \quad A_k = \begin{pmatrix} e^{\frac{\pi k i}{n}} & 0 \\ 0 & e^{-\frac{\pi k i}{n}} \end{pmatrix}, \quad k = 0,1,\dots,n\,. \tag{2.33}$$

Hence a^n gives $(v_1, v_2) \to (-1)^k (v_1, v_2)$ so that instead of (2.28) for the action of b we may require

$$(v_1, v_2) \underset{b}{\to} \big((-1)^k v_2, v_1\big) = (v_1, v_2) B_k\,, \qquad B_k = \begin{pmatrix} 0 & 1 \\ (-1)^k & 0 \end{pmatrix}, \quad k = 0,\dots,n\,. \tag{2.34}$$

For $k = 0$, $A_0 = \mathbb{1}$ and $B_0 \sim \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, for $k = n$, $A_n = -\mathbb{1}$ and $B_n \sim \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, n even, $B_n \sim \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, n odd. Hence there are four one dimensional irreducible representations and $n-1$ two dimensional ones. For $k = 1$ the representation (2.33) and (2.34) is equivalent to using the quaternion representation in (1.115) in (1.85). The representations follow a similar pattern to the dihedral case, for n odd

$$
\begin{array}{llcll}
\mathcal{R}_{1,1}: & (a^r, a^r b) & \to & (1,1) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,2}: & (a^r, a^r b) & \to & (1,-1) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,3}: & (a^r, a^r b) & \to & ((-1)^r, (-1)^r i) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,4}: & (a^r, a^r b) & \to & ((-1)^r, -(-1)^r i) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{2,k}: & (a^r, a^r b) & \to & ({A_k}^r, {A_k}^r B_k) \ , & r = 0,\dots,2n-1\,,\ k = 1,\dots,n-1\,,
\end{array} \tag{2.35}
$$

and for n even

$$
\begin{array}{llcll}
\mathcal{R}_{1,1}: & (a^r, a^r b) & \to & (1,1) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,2}: & (a^r, a^r b) & \to & (1,-1) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,3}: & (a^r, a^r b) & \to & ((-1)^r, (-1)^r) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{1,4}: & (a^r, a^r b) & \to & ((-1)^r, -(-1)^r) \ , & r = 0,\dots,2n-1\,, \\
\mathcal{R}_{2,k}: & (a^r, a^r b) & \to & ({A_k}^r, {A_k}^r B_k) \ , & r = 0,\dots,2n-1\,,\ k = 1,\dots,n-1\,.
\end{array} \tag{2.36}
$$

For k even the representations can be brought to a real form by a similarity transformation as in (2.32).

2.3 Unitary Representations

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

$$
D(g)^\dagger = D(g^{-1}) = D(g)^{-1}\,. \tag{2.37}
$$

For such representation then the usual scalar product on $\mathcal{V}$ is invariant, for transformations as in (2.9) ${v_1}^{g\dagger}{v_2}^g = {v_2}^\dagger v_1$. 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

$$
S = \sum_i D(g_i)^\dagger D(g_i)\,, \tag{2.38}
$$

where the sum is over all elements of the group $G = \{g_i\}$. As a consequence of (1.7)

$$
\begin{aligned}
SD(g)^{-1} = SD(g^{-1}) &= \sum_i D(g_i)^\dagger D(g_i g^{-1}) \\
&= \sum_i D(g_i g)^\dagger D(g_i) \\
&= D(g)^\dagger \sum_i D(g_i)^\dagger D(g_i) = D(g)^\dagger S\,,
\end{aligned} \tag{2.39}
$$

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 $D'(g)^\dagger D'(g) = \mathbb{1}$).

2.3.1 Infinite Dimensional Unitary Representations

Representations can be infinite dimensional when they are typically expressed in the terms of the group action on spaces of functions. Such representations arise when considering unitary representations of non compact groups. As an example we consider unitary representations for the Heisenberg group as define by products of matrices of the form (1.122). Acting on complex square integrable functions $\{f\}$ on $\mathbb{R}$ belonging to L^2 we define an action $f \to T_{(a,b,c)} f$, for a, b, c real, by

$$T_{(a,b,c)} f(x) = e^{i(bx+c)} f(x+a)\,. \tag{2.40}$$

It is an exercise to check this satisfies the group multiplication rules and that this forms a unitary representation.

2.3.2 Real and Pseudo-Real Representations

For any unitary complex representation the conjugate representation may be equivalent so that

$$D(g)^* = C D(g) C^{-1} \quad \text{for all} \quad g \in G\,. \tag{2.41}$$

Assuming the representation is unitary so that $D(g)^* = D(g)^{-1T}$ then using (2.41) and its transpose for $D(g)^{-1}$

$$C^{-1T} C\, D(g)\, C^{-1} C^T = D(g) \quad \Rightarrow \quad \big[D(g),\, C^{-1} C^T\big] = 0\,. \tag{2.42}$$

For an irreducible representation Schur's lemma requires

$$C^{-1} C^T = \alpha \mathbb{1} \quad \Rightarrow \quad C^T = \alpha\, C \quad \Rightarrow \quad C = \alpha\, C^T \quad \Rightarrow \quad \alpha^2 = 1\,. \tag{2.43}$$

Hence $\alpha = \pm 1$. For $\alpha = 1$ there is an S such that $S^T C\, S = \mathbb{1}$ and then $S^{-1} D(g) S = D'(g)$ is a real representation. Otherwise $C^T = -C$ and the representation is *pseudo-real*. For C to be invertible it must be even dimensional and it is then reducible to the form (1.108).

A prescription for C is obtained by taking

$$C = \sum_i D(g_i)^T U\, D(g_i)\,, \tag{2.44}$$

for arbitrary U since then, using (1.7) again,

$$C\, D(g) = \sum_i D(g_i g^{-1})^T U\, D(g_i) = D(g)^{-1T}\, C\,. \tag{2.45}$$

Thus for real, pseudo-real representations $C^T = \pm C$ so that

$$\sum_i D(g_i)^T U D(g_i) = \pm \sum_i D(g_i)^T U^T D(g_i) \quad \Rightarrow \quad \sum_i D_{rs}(g_i) D_{uv}(g_i) = \pm \sum_i D_{us}(g_i) D_{rv}(g_i)\,. \tag{2.46}$$

For a complex representation there can be no matrix C as defined in (2.44) for any U. This then requires in this case

$$\sum_i D_{rs}(g_i) D_{uv}(g_i) = 0\,. \tag{2.47}$$

For the two dimensional representation of the quaternion groups considered in 1.5 defined by (1.115) then $q \to Q$, $\bar{q} \to Q^\dagger = Q^{-1}$ so that the representation is unitary and also irreducible since the only 2×2 matrix commuting with I, J, K is proportional to $\mathbb{1}$. Also as $J(I,J,K)J = (I^T, J^T, K^T)$ then Q satisfies (2.41) with $C \to J = -J^T$. Hence these representations are pseudo-real.

In general any pseudo-real representation can be written in terms of matrices of quaternions, with the quaternions having real coefficients.

2.4 Orthogonality Relations

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

$$S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = \frac{1}{|G|} \sum_i D^{(\mathcal{R}')}_{rv}(g_i{}^{-1}) D^{(\mathcal{R})}_{us}(g_i) \, , \tag{2.48}$$

where $D^{(\mathcal{R})}(g), D^{(\mathcal{R}')}(g)$ are the matrices corresponding to the irreducible representation $\mathcal{R}, \mathcal{R}'$. Then

$$\begin{aligned} S^{(\mathcal{R}',\mathcal{R})}_{rt,uv} D^{(\mathcal{R})}_{ts}(g) &= \frac{1}{|G|} \sum_i D^{(\mathcal{R}')}_{rv}(g g_i{}^{-1}) D^{(\mathcal{R})}_{us}(g_i) = D^{(\mathcal{R}')}_{rt}(g) \, S^{(\mathcal{R}',\mathcal{R})}_{ts,uv} \, , \\ S^{(\mathcal{R}',\mathcal{R})}_{rs,uw} D^{(\mathcal{R}')}_{wv}(g) &= \frac{1}{|G|} \sum_i D^{(\mathcal{R}')}_{rv}(g_i{}^{-1}) D^{(\mathcal{R})}_{us}(g g_i) = D^{(\mathcal{R})}_{uw}(g) \, S^{(\mathcal{R}',\mathcal{R})}_{rs,wv} \, , \end{aligned} \tag{2.49}$$

for any $g \in G$. The proof of (2.49) follows essentially since $\{g_i\} = \{g_i g\}$. Schur's lemmas then requires that $S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = 0$ unless $\mathcal{R}' = \mathcal{R}$ when $S^{(\mathcal{R}',\mathcal{R})}_{rs,uv}$ must be proportional to δ_{rs} and also δ_{uv}. Hence we must have

$$S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = \frac{1}{n_{\mathcal{R}}} \, \delta_{\mathcal{R}'\mathcal{R}} \, \delta_{rs} \, \delta_{uv} \, , \tag{2.50}$$

where $n_{\mathcal{R}} = \dim \mathcal{R}$ is the dimension of the representation $\mathcal{R}$. The constant in (2.50) is determined by considering $S^{(\mathcal{R},\mathcal{R})}_{ru,us} = \sum_i D^{(\mathcal{R})}_{rs}(e) = \delta_{rs}$.

2.5 Characters

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

$$\chi_{\mathcal{R}}(g) = \mathrm{tr}\big(D^{(\mathcal{R})}(g)\big) \, . \tag{2.51}$$

Since traces are unchanged under cyclic permutations $\chi_{\mathcal{R}}(g' g g'^{-1}) = \chi_{\mathcal{R}}(g)$ so that the character depends only on the conjugacy classes of each element. Hence we may write $\chi(g_i) \equiv \chi(\mathcal{C}_s)$ for any $g_i \in \mathcal{C}_s$ where for N_{char} different conjugacy classes in G, $s = 1, \dots, N_{\mathrm{char}}$. With previous conventions in (1.53) for $g_i \in \mathcal{C}_s$, $\chi(g_i{}^{-1}) = \chi(\mathcal{C}_{\bar{s}})$. Similarly the character is unchanged when calculated for any representations related by an equivalence transformation as in (2.5). Since for a finite group any representation is equivalent to a unitary one we must also have

$$\chi_{\mathcal{R}}(g^{-1}) = \chi_{\mathcal{R}}(g)^* \quad \text{or} \quad \chi_{\mathcal{R}}(\mathcal{C}_{\bar{s}}) = \chi_{\mathcal{R}}(\mathcal{C}_s)^* \,. \tag{2.52}$$

For real or pseudo-real representations due to (2.41) characters are real. If the character is complex then $\chi_{\mathcal{R}}{}^* = \chi_{\mathcal{R}'}$ for $\mathcal{R}' \neq \mathcal{R}$ the conjugate representation and necessarily there are group elements such that $\mathcal{C}(g) \neq \mathcal{C}(g^{-1})$. As a special case

$$\chi_{\mathcal{R}}(e) = \mathrm{tr}_{\mathcal{R}}(\mathbb{1}) = \dim \mathcal{R} = n_{\mathcal{R}} \,. \tag{2.53}$$

For $D^{(\mathcal{R})}(g)$ equivalent to a unitary representation, as is the case for any finite group, then, since $D^{(\mathcal{R})}(g)$ has n_R eigenvalues of modulus one, $|\chi_{\mathcal{R}}(g)| \le n_{\mathcal{R}}$ and $|\chi_{\mathcal{R}}(g)| = n_{\mathcal{R}} > 1$ only for $D^{(\mathcal{R})}(g) = \pm \mathbb{1}_{n_{\mathcal{R}}}$. For any direct product group $G_1 \times G_2$ then $\chi_{G_1\times G_2}((g_1, g_2)) = \chi_{G_1}(g_1)\,\chi_{G_2}(g_2)$.

For the regular representation defined by (2.6) then by (2.7)

$$\chi_{\mathrm{reg}}(e) = |G| \,, \qquad \chi_{\mathrm{reg}}(g) = 0 \,, \quad g \neq e \,. \tag{2.54}$$

For two representations $\mathcal{R}$, $\mathcal{R}'$ there is a scalar product for the associated characters defined by

$$\langle \chi_{\mathcal{R}'}, \chi_{\mathcal{R}} \rangle = \frac{1}{|G|} \sum_{g\in G} \chi_{\mathcal{R}'}(g)^* \, \chi_{\mathcal{R}}(g) = \frac{1}{|G|} \sum_{s=1}^{N_{\mathrm{char}}} \chi_{\mathcal{R}'}(\mathcal{C}_s)^* \, d_s \, \chi_{\mathcal{R}}(\mathcal{C}_s) \,, \quad d_s = \dim \mathcal{C}_s \,. \tag{2.55}$$

The characters for the irreducible representations play a crucial role. For irreducible representations $\{\mathcal{R}_r : r = 1, \dots, N_G\}$ the corresponding characters $\chi_r(g) = \chi_{\mathcal{R}_r}(g)$. For the singlet representation (2.18) $\chi_1(g) = 1$. As a consequence of the orthogonality relation for irreducible representations (2.50) then using (2.52) for two irreducible representations $\mathcal{R} = \mathcal{R}_r$, $\mathcal{R}' = \mathcal{R}_{r'}$ this reduces to

$$\langle \chi_{r'}, \chi_r \rangle = \frac{1}{|G|} \sum_{g\in G} \chi_{r'}(g)^* \chi_r(g) = \frac{1}{|G|} \sum_{s=1}^{N_{\mathrm{char}}} \chi_{r'}(\mathcal{C}_s)^* \, d_s \, \chi_r(\mathcal{C}_s) = \delta_{r'r} \,. \tag{2.56}$$

Equivalently

$$X_{rs} = \chi_r(\mathcal{C}_s) \,, \quad D_{ss'} = d_s \, \delta_{ss'} \quad \Rightarrow \quad X \, D \, X^\dagger = |G| \, \mathbb{1}_{n_G} \,. \tag{2.57}$$

where X is a $N_G \times N_{\mathrm{char}}$ matrix and D is a diagonal $N_{\mathrm{char}} \times N_{\mathrm{char}}$ matrix. (2.57) requires the number of irreducible representations $N_G \le N_{\mathrm{char}}$. From (2.18) and (2.53)

$$\chi_1(\mathcal{C}_s) = 1 \quad \text{all } s \,, \qquad \chi_r(e) = \dim \mathcal{R}_r = n_r \,. \tag{2.58}$$

For any $f(g)$ satisfying $f(g) = f(g' g g'^{-1})$ for all $g' \in G$, so that $f(g) = f(\mathcal{C}_s)$ for all $g \in \mathcal{C}_s$, we may then define $\langle f, \chi_{\mathcal{R}} \rangle = \sum_{s=1}^{N_{\mathrm{char}}} f(\mathcal{C}_s)^* d_s \, \chi_{\mathcal{R}}(\mathcal{C}_s)/|G|$, as in (2.56). An important result, demonstrated later, is

$$\langle f, \chi_r \rangle = 0 \;\; \text{for all } r \quad \Rightarrow \quad f(g) = 0 \,. \tag{2.59}$$

If $N_G > N_{\rm char}$ there is a non zero vector such that $\sum_r v_r X_{rs} = 0$. Since this contradicts (2.59) we must have

$$N_G = N_{\rm char} \, . \tag{2.60}$$

Hence X is a non singular square matrix and using $|G|\, X^{-1} = D\, X^\dagger$ equivalently

$$X^\dagger X = |G|\, D^{-1} \, , \qquad \sum_{r=1}^{N_G} \chi_r(\mathcal{C}_s)\chi_r(\mathcal{C}_{s'})^* = |G|/d_s\, \delta_{ss'} \, . \tag{2.61}$$

For $s = s' = 1$ then $\mathcal{C}_1 = \{e\}$ is the conjugacy class containing just the identity by itself, $d_1 = 1$, so that

$$\sum_{r=1}^{N_G} {n_r}^2 = |G| \, . \tag{2.62}$$

This plays an important role in constraining irreducible representations for finite groups. As an illustration for the dihedral group D_n then $\sum_{k=1}^{(n-2)/2} 2^2 + 1 + 1 + 1 + 1 = 2n$ for n even and $\sum_{k=1}^{(n-1)/2} 2^2 + 1 + 1 = 2n$ for n odd. For $\mathrm{Q}_{4n} = 2D_n$ then $(n-1)2^2 + 4 \times 1 = 4n$.

Characters distinguish the different possible representations. From (2.46) and (2.47), setting $s = u$, $r = v$ and summing, for an irreducible such representation (note that if $g, g' \in \mathcal{C}_s$ then $g^2, g'^2 \in \mathcal{C}_{s'}$ for some $\mathcal{C}_{s'}$)

$$\mathcal{F}_{FS} = \frac{1}{|G|} \sum_{g \in G} \chi_r(g^2) = \begin{cases} \langle \chi_r, \chi_r \rangle \\ -\langle \chi_r, \chi_r \rangle \\ \quad 0 \end{cases} = \begin{cases} 1 & \text{real} \\ -1 & \text{pseudo-real} \\ 0 & \text{complex} \end{cases} . \tag{2.63}$$

This formula characterising the different possibilities for representations was first obtained jointly by Frobenius[15] and Schur in 1906 and is sometimes referred to as the *Frobenius-Schur indicator*.

Using the orthogonality results then for an arbitrary representation $\mathcal{R}$ then decomposing into irreducible representations as in (2.16)

$$\chi_{\mathcal{R}}(g) = \sum_{r=1}^{N_G} m_r\, \chi_r(g) \, , \tag{2.64}$$

and the multiplicity of the irreducible representation $\mathcal{R}_r$ is then given by

$$m_r = \langle \chi_r, \chi_{\mathcal{R}} \rangle \, . \tag{2.65}$$

Applying this to the regular representation which is completely reducible

$$D_{\rm reg}(g) \simeq \bigoplus_{r=1}^{N_G} m_{{\rm reg},r}\, D_{\mathcal{R}_r}(g) \, , \quad \chi_{\rm reg}(g) = \sum_{r=1}^{N_G} m_{{\rm reg},r}\, \chi_r(g) \, , \tag{2.66}$$

and using (2.52) and (2.54)

$$m_{{\rm reg},r} = \langle \chi_r, \chi_{\rm reg} \rangle = \frac{1}{|G|}\, \chi_r(e)\, \chi_{\rm reg}(e) = n_r \, . \tag{2.67}$$

[15] Ferdinand Georg Frobenius, 1849-1917, German. A pioneer of representation theory.

We also then have

$$|G| = \chi_{\rm reg}(e) = \sum_{r=1}^{N_G} m_{{\rm reg},r}\,\chi_r(e) = \sum_{r=1}^{N_G} {n_r}^2\,, \tag{2.68}$$

which reproduces (2.62).

To prove (2.59) we define for any representation $\mathcal{R}$

$$T_{\mathcal{R}} = \frac{1}{|G|}\sum_i f(g_i)^*\, D_{\mathcal{R}}(g_i)\,. \tag{2.69}$$

Using the group property and $f(gg_ig^{-1}) = f(g_i)$ it is easy to see that

$$D_{\mathcal{R}}(g)^{-1}\, T_{\mathcal{R}}\, D_{\mathcal{R}}(g) = T_{\mathcal{R}} \quad \text{or} \quad T_{\mathcal{R}}\, D_{\mathcal{R}}(g) = D_{\mathcal{R}}(g)\, T_{\mathcal{R}} \quad \text{for all} \quad g\,. \tag{2.70}$$

Applying this for the irreducible representation $\mathcal{R}_r$ and using Schur's lemma we must have

$$T_{\mathcal{R}_r} = c\,\mathbb{1}_{n_r}\,, \tag{2.71}$$

with $\mathbb{1}_{n_r}$ the identity matrix in this representation. However

$$\langle f, \chi_{\mathcal{R}}\rangle = \frac{1}{|G|}\sum_i f(g_i)^*\,\chi_{\mathcal{R}}(g_i) = \mathrm{tr}(T_{\mathcal{R}}) = 0\,, \tag{2.72}$$

by virtue of the assumption (2.59). Hence in (2.71) $c = 0$ so that $T_{\mathcal{R}_r} = 0$ for all irreducible representations. Since the regular representation can be decomposed into irreducible representations as in (2.66)

$$T_{\mathcal{R}_{\rm reg}} = 0 \quad \Rightarrow \quad \sum_i f(g_i)^*\, g_i = 0\,, \tag{2.73}$$

from the definition (2.6) since $g_i = \sum_j g_j D_{{\rm reg},jk}(g_i) {g_k}^{-1}$. Since all g_i are independent this is only possible if $f(g_i) = 0$ for all i.

For an induced representation as in (2.25) if for the subgroup representation

$$\chi(h) = \mathrm{tr}\big(D(h)\big)\,, \tag{2.74}$$

then

$$\chi_{\text{induced rep.}}(g) = \sum_i \chi({g_i}^{-1} g\, g_i)\big|_{{g_i}^{-1} g\, g_i \in H}\,. \tag{2.75}$$

If this is applied to the case when $H = \{e\}$, giving the regular representation, we get (2.54).

2.5.1 Further Constraints on Dimensions of Irreducible Representations

The dimensions of irreducible representations, n_r, $r = 1, \dots, N_G$, for finite groups are constrained by (2.62). A further condition is that n_r must divide $|G|$. This is somewhat non trivial to prove.

For any representation $\mathcal{R}$ and conjugacy class $\mathcal{C}$ we may define

$$T_{\mathcal{C}}{}^{\mathcal{R}} = \sum_{g\in\mathcal{C}} D^{(\mathcal{R})}(g)\,. \tag{2.76}$$

Since $g\,\mathcal{C}g^{-1} = \mathcal{C}$ then

$$D^{(\mathcal{R})}(g)\,T_{\mathcal{C}}{}^{\mathcal{R}}D^{(\mathcal{R})}(g)^{-1} = T_{\mathcal{C}}{}^{\mathcal{R}} \quad \Rightarrow \quad \big[T_{\mathcal{C}}{}^{\mathcal{R}},\, D^{(\mathcal{R})}(g)\big] = 0 \quad \text{for all} \quad g \in G\,. \tag{2.77}$$

By Schur's lemma if $\mathcal{R}$ is irreducible $T_{\mathcal{C}}{}^{\mathcal{R}}$ is proportional to the identity so that

$$T_{\mathcal{C}}{}^{\mathcal{R}} = \frac{d_{\mathcal{C}}}{n_{\mathcal{R}}}\,\chi_{\mathcal{R}}(\mathcal{C})\,\mathbb{1}_{n_{\mathcal{R}}}\,, \tag{2.78}$$

where the coefficient is determine by taking the trace. Furthermore summing over the various conjugacy classes $\mathcal{C}_s$ and using the orthogonality of characters

$$R_{\mathcal{R}} = \sum_s \frac{|G|}{d_s}\,T_{\mathcal{C}_s}{}^{\mathcal{R}}T_{\mathcal{C}_{\bar{s}}}{}^{\mathcal{R}} = \frac{|G|^2}{n_{\mathcal{R}}{}^2}\,\mathbb{1}_{n_{\mathcal{R}}}\,. \tag{2.79}$$

As a consequence of (1.57)

$$T_{\mathcal{C}_s}{}^{\mathcal{R}_r}\,T_{\mathcal{C}_t}{}^{\mathcal{R}_r} = \sum_u c_{st}{}^u\,T_{\mathcal{C}_u}{}^{\mathcal{R}_r} \quad \Rightarrow \quad \chi_r(\mathcal{C}_s)\,\chi_r(\mathcal{C}_t) = n_r \sum_u \frac{d_u}{d_s\,d_t}\,c_{st}{}^u\,\chi_u(\mathcal{C}_u)\,, \tag{2.80}$$

so that

$$\sum_r \frac{1}{n_r}\,\chi_r(\mathcal{C}_s)\chi_r(\mathcal{C}_t)\chi_r(\mathcal{C}_{\bar{u}}) = \frac{|G|}{d_s\,d_t}\,c_{st}{}^u\,. \tag{2.81}$$

Showing that n_r divides $|G|$ depends on applying these results to the regular representation. From (2.66) and (2.67) this is expressed as a direct sum over irreducible representations $D_{\text{reg}}(g) \simeq \bigoplus_{r=1}^{N_G} n_r\,D_{\mathcal{R}_r}(g)$. Then

$$T_{\text{reg},\mathcal{C}} = \sum_{g\in\mathcal{C}} D_{\text{reg}}(g) \simeq \bigoplus_{r=1}^{N_G} n_r \left(\frac{d_{\mathcal{C}}}{n_r}\,\chi_r(\mathcal{C})\,\mathbb{1}_{n_{\mathcal{R}}}\right), \tag{2.82}$$

and

$$R_{\text{reg}} = \sum_s \frac{|G|}{d_s}\,T_{\text{reg},\mathcal{C}_s}T_{\text{reg},\mathcal{C}_{\bar{s}}} \simeq \bigoplus_{r=1}^{N_G} n_r \left(\frac{|G|^2}{n_r{}^2}\,\mathbb{1}_{n_r}\right), \tag{2.83}$$

using orthogonality of characters again. The eigenvalues of R_{reg} are then $|G|^2/n_r{}^2$, which has multiplicity n_r and are necessarily rational. The eigenvalues are determined by the roots of $\det(\lambda\,\mathbb{1}_{|G|} - R_{\text{reg}}) = 0$. Since $|G|/d_s$ are integers the elements of R_{reg} are necessarily integers as $D_{\text{reg}}(g)$ has elements which are only 1 or 0. Hence $\det(\lambda\,\mathbb{1}_{|G|} - R_{\text{reg}})$ can be expanded as a polynomial in λ with integer coefficients and leading term $\lambda^{|G|}$. The rational root theorem states that in this situation any root which is rational must an integer which is a factor of $\det(R_{\text{reg}})$.[16] Hence $(|G|/n_r)^2$ and therefore, since integer square roots are either integers or irrational, $|G|/n_r$ are integers for all r.

[16]For a polynomial $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$ with a_i integers then for a rational root $P(p/q) = 0$ where p, q are integers with no common factors. The equation determining this root can be rewritten as $q(a_{n-1}p^{n-1} + \cdots + a_0 q^{n-1}) = -p^n$. Since q then divides p^n and does not contain p as a factor this forces $q = 1$.

2.6 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* $\mathcal{R}_1 \times \mathcal{R}_2$ in terms of the direct product of the representation 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 $\mathcal{R}_1 \times \mathcal{R}_2$ 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

$$\mathcal{R}_p \times \mathcal{R}_q = \mathcal{R}_q \times \mathcal{R}_p = \sum_r n_{pq,r} \, \mathcal{R}_r \, , \tag{2.84}$$

where $n_{pq,r}$ are integers, which may be zero, and $n_{pq,r} > 1$ if the representation $\mathcal{R}_r$ occurs more than once. For non finite groups there are infinitely many irreducible representations but the sum in (2.84) 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 $\mathrm{tr}_{\mathcal{V}_p \otimes \mathcal{V}_q}\big(D^{(\mathcal{R}_p)}(g) \otimes D^{(\mathcal{R}_q)}(g)\big) = \mathrm{tr}_{\mathcal{V}_p}\big(D^{(\mathcal{R}_p)}(g)\big) \, \mathrm{tr}_{\mathcal{V}_q}\big(D^{(\mathcal{R}_q)}(g)\big)$, so that, in terms of the characters $\chi_r(g) = \mathrm{tr}_{\mathcal{V}_r}\big(D^{(\mathcal{R}_r)}(g)\big)$, (2.84) is equivalent to

$$\chi_p(g)\chi_q(g) = \sum_r n_{pq,r} \, \chi_r(g) \, . \tag{2.85}$$

Using (2.56) the coefficients $n_{pq,r}$ can be determined by

$$n_{pq,r} = \frac{1}{|G|} \sum_i \chi_r(g_i)^* \chi_p(g_i)\chi_q(g_i) = \frac{1}{|G|} \sum_s d_s \, \chi_r(\mathcal{C}_{\bar{s}})\chi_p(\mathcal{C}_s)\chi_q(\mathcal{C}_s) \, . \tag{2.86}$$

The result (2.84) is exactly equivalent to the decomposition of the associated representation spaces, with the same expansion for $\mathcal{V}_p \otimes \mathcal{V}_q$ into a direct sum of irreducible spaces $\mathcal{V}_r$. If $\mathcal{R}_p \otimes \mathcal{R}_q$ contains the trivial or singlet representation then it is possible to construct a scalar product $\langle v, v' \rangle$ between vectors $v \in \mathcal{V}_p, v' \in \mathcal{V}_q$ which is invariant under group transformations, $\langle D^{(\mathcal{R}_p)}(g)v, D^{(\mathcal{R}_q)}(g)v' \rangle = \langle v, v' \rangle$.

2.6.1 Symmetric and Antisymmetric Products

The tensor product of vector spaces $\mathcal{V}_r$, $r = 1, \dots, n$ where $\mathcal{V}_i, \mathcal{V}_j \simeq \mathcal{V}$ for all i, j can be decomposed into representation spaces for the permutation group $\mathcal{S}_n$. The simplest cases are the one dimensional totaily symmetric or antisymmetric representations of $\mathcal{S}_n$ which can be denoted by

$$\begin{aligned} \vee^n V &= \big(\underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{n}\big)_{\mathrm{sym}} \, , \qquad && \dim \vee^n V = \tfrac{1}{n!} \dim \mathcal{V}(\dim \mathcal{V} + 1) \dots (\dim \mathcal{V} + n - 1) \, , \\ \wedge^n \mathcal{V} &= \big(\underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{n}\big)_{\mathrm{antisym}} \, , \qquad && \dim \wedge^n V = \tfrac{1}{n!} \dim \mathcal{V}(\dim \mathcal{V} - 1) \dots (\dim \mathcal{V} - n + 1) \, . \end{aligned} \tag{2.87}$$

$\vee^n V$ is also commonly denoted as $\mathrm{Sym}^n V$. The action of permutations commutes with the group action on $\mathcal{V}$ so that $\vee^n V$ and $\wedge^n \mathcal{V}$ both form representation spaces, in general reducible, for the group G with a matrix representation $D(g)$ of dimension $\dim \mathcal{V}$. The representation matrices have the form, for the symmetric, antisymmetric cases respectively,

$$D_{(r_1|s_1}(g) D_{r_2|s_2}(g) \dots D_{r_n)s_n}(g) \,, \quad D_{[r_1|s_1}(g) D_{r_2|s_2}(g) \dots D_{r_n]s_n}(g) \,, \tag{2.88}$$

involving a sum over $n!$ permutations $\sigma \in \mathcal{S}_n$.

The corresponding characters

$$\begin{aligned} \chi_{\vee^n}(g) &= D_{(r_1|r_1}(g) D_{r_2|r_2}(g) \dots D_{r_n)r_n}(g) \,, \\ \chi_{\wedge^n}(g) &= D_{[r_1|r_1}(g) D_{r_2|r_2}(g) \dots D_{r_n]r_n}(g) \,, \end{aligned} \tag{2.89}$$

can be reduced to the characters $\chi(g) = \mathrm{tr}\, D(g)$ using the decomposition of any permutation σ into cycles. Thus for a cycle decomposition $\sigma = [p_1, p_2, \dots, p_r]$, $\sum_{i=1}^r p_i = n$, then one such permutation of indices $r_1, r_2, \dots, r_n$ generates a contribution of the form

$$\begin{aligned} & D_{\sigma(r_1)r_1}(g) D_{\sigma(r_2)r_2}(g) \dots D_{\sigma(r_n)r_n}(g)\big|_{\sigma(r_1,\dots,r_n)=(r_1\dots r_{p_1})(r_{p_1+1}\dots r_{p_1+p_2})\dots(r_{n-p_r+1}\dots r_n)} \\ & \quad = \chi(g^{p_1})\, \chi(g^{p_2}) \dots \chi(g^{p_r}) \,. \end{aligned} \tag{2.90}$$

Any permutation with the same cycle decomposition generates an identical expression so that summing over all possible cycles (2.89) becomes

$$\begin{aligned} \chi_{\vee^n}(g) &= \tfrac{1}{n!} \sum_{r=1}^n \sum_{j_1, j_2, \dots j_r \ge 1} \sum_{p_1 > p_2 > \dots p_r \ge 1} \delta_{n, \sum_{i=1}^r j_i p_i}\, N_{[p_1(j_1), \dots, p_r(j_r)]} \prod_{i=1}^n \chi(g^{p_i})^{j_i} \,, \\ \chi_{\wedge^n}(g) &= \tfrac{1}{n!} \sum_{r=1}^n \sum_{j_1, j_2, \dots j_r \ge 1} \sum_{p_1 > p_2 > \dots p_r \ge 1} \delta_{n, \sum_{i=1}^r j_i p_i}\, N_{[p_1(j_1), \dots, p_r(j_r)]} (-1)^{\sum_i j_i (p_i - 1)} \prod_{i=1}^n \chi(g^{p_i})^{j_i} \,, \end{aligned} \tag{2.91}$$

with the numbers for j_i p_i-cycles in each cycle decomposition given by (1.22). Results for $\chi_{\vee^2}(g)$, $\chi_{\vee^3}(g)$ are just as in (1.70) and (1.71) is equivalent to the result for $\dim \vee^n V$ in (2.87). Similarly $\chi_{\wedge^2}(g)$, $\chi_{\wedge^3}(g)$ are essentially of the form in (1.73) so that

$$\begin{aligned} \left.\begin{matrix} \chi_{\vee^2 \mathcal{V}^n}(g) \\ \chi_{\wedge^2 \mathcal{V}^n}(g) \end{matrix}\right\} &= \tfrac{1}{2}\big(\chi_{\mathcal{V}^n}(g)^2 \pm \chi_{\mathcal{V}^n}(g^2) \big) \,, \\ \left.\begin{matrix} \chi_{\vee^3 \mathcal{V}^n}(g) \\ \chi_{\wedge^2 \mathcal{V}^n}(g) \end{matrix}\right\} &= \tfrac{1}{6}\big(\chi_{\mathcal{V}^n}(g)^3 \pm 3\, \chi_{\mathcal{V}^n}(g) \chi_{\mathcal{V}^n}(g^2) + 2\, \chi_{\mathcal{V}^n}(g^3) \big) \,. \end{aligned} \tag{2.92}$$

2.7 Character Tables

Knowing the irreducible representations the corresponding characters can readily be computed and can be represented as tables composed of the elements of the square matrix X.

For abelian groups the irreducible representations are one dimensional as are also the characters which each correspond to a single group element. For $\mathbb{Z}_n$, or equivalently C_n, defined by $\{a^r : a^n = e\}$, we can take

$$X_{kr} = \chi_k(a^r) = e^{2kr\pi i/n} \,, \quad k, r = 0, 1, \dots, n-1 \,, \tag{2.93}$$

where k labels the representation and r the conjugacy class. It is easy to verify that $XX^\dagger = N\,\mathbb{1}_N$. The character table is just

$\mathbb{Z}_n$	$\mathcal{C}_{1,r}$	$\mathcal{F}_{FS}$
$\mathcal{R}_{1,k}$	$e^{2kr\pi i/n}$	$1,\ k=0,\tfrac{1}{2}n(n \text{ even}),\ 0$ otherwise

$$, \quad k,r = 0,\dots,n-1\,. \tag{2.94}$$

For the dihedral group D_n the conjugacy classes are listed in (1.58) and the irreducible representations in (2.30) and (2.31) with (2.29). The character tables are then, for n odd,

D_n n odd	$\mathcal{C}_1$	$\mathcal{C}_{2,r}$	$\mathcal{C}_n$	$\mathcal{F}_{FS}$
$\mathcal{R}_{1,1}$	1	1	1	1
$\mathcal{R}_{1,2}$	1	1	-1	1
$\mathcal{R}_{2,k}$	2	$2\cos 2kr\pi/n$	0	1

$$, \quad k,r = 1,\dots,\tfrac{1}{2}(n-1)\,, \tag{2.95}$$

where $\chi_{2,0}(\mathcal{C}) = \chi_{1,1}(\mathcal{C}) + \chi_{1,2}(\mathcal{C})$. For n even,

D_n n even	$\mathcal{C}_{1,1}$	$\mathcal{C}_{1,2}$	$\mathcal{C}_{2,r}$	$\mathcal{C}_{\frac{1}{2}n,1}$	$\mathcal{C}_{\frac{1}{2}n,2}$	$\mathcal{F}_{FS}$
$\mathcal{R}_{1,1}$	1	1	1	1	1	1
$\mathcal{R}_{1,2}$	1	1	1	-1	-1	1
$\mathcal{R}_{1,3}$	1	$(-1)^{n/2}$	$(-1)^r$	1	-1	1
$\mathcal{R}_{1,4}$	1	$(-1)^{n/2}$	$(-1)^r$	-1	1	1
$\mathcal{R}_{2,k}$	2	$2(-1)^k$	$2\cos 2kr\pi/n$	0	0	1

$$, \quad k,r = 1,\dots,\tfrac{1}{2}(n-2)\,. \tag{2.96}$$

In this case $\chi_{2,0}(\mathcal{C}) = \chi_{1,1}(\mathcal{C}) + \chi_{1,2}(\mathcal{C})$ and $\chi_{2,n/2}(\mathcal{C}) = \chi_{1,3}(\mathcal{C}) + \chi_{1,4}(\mathcal{C})$. The representations are all real.

For the dicyclic groups the conjugacy classes are given in (1.87) and irreducible representations in (2.35) and (2.36) with (2.33) and (2.34). In this case

Q_{4n} n odd	$\mathcal{C}_{1,1}$	$\mathcal{C}_{1,2}$	$\mathcal{C}_{2,r}$	$\mathcal{C}_{n,1}$	$\mathcal{C}_{n,2}$	$\mathcal{F}_{FS}$
$\mathcal{R}_{1,1}$	1	1	1	1	1	1
$\mathcal{R}_{1,2}$	1	1	1	-1	-1	1
$\mathcal{R}_{1,3}$	1	-1	$(-1)^r$	i	$-i$	0
$\mathcal{R}_{1,4}$	1	-1	$(-1)^r$	$-i$	i	0
$\mathcal{R}_{2,k}$	2	$2(-1)^k$	$2\cos kr\pi/n$	0	0	$(-1)^k$

$$, \quad k,r = 1,\dots,n-1\,, \tag{2.97}$$

and

Q_{4n} n even	$\mathcal{C}_{1,1}$	$\mathcal{C}_{1,2}$	$\mathcal{C}_{2,r}$	$\mathcal{C}_{n,1}$	$\mathcal{C}_{n,2}$	$\mathcal{F}_{FS}$
$\mathcal{R}_{1,1}$	1	1	1	1	1	1
$\mathcal{R}_{1,2}$	1	1	1	-1	-1	1
$\mathcal{R}_{1,3}$	1	1	$(-1)^r$	1	-1	1
$\mathcal{R}_{1,4}$	1	1	$(-1)^r$	-1	1	1
$\mathcal{R}_{2,k}$	2	$2(-1)^k$	$2\cos kr\pi/n$	0	0	$(-1)^k$

$$, \quad k,r = 1,\dots,n-1\,. \tag{2.98}$$

For both n even and odd $\chi_{2,0}(\mathcal{C}) = \chi_{1,1}(\mathcal{C}) + \chi_{1,2}(\mathcal{C})$, $\chi_{2,n}(\mathcal{C}) = \chi_{1,3}(\mathcal{C}) + \chi_{1,4}(\mathcal{C})$. The representations $\mathcal{R}_{2,k}$ for $n \geq 2$ and k odd are pseudo real. For $n = 1$ dropping the $\mathcal{C}_{2,r}$ column and $\mathcal{R}_{2,k}$ row the character table is identical to that for $\mathbb{Z}_4$.

These character tables satisfy the required orthogonality conditions.

2.8 Molien Series

A nice application of the results for characters is a formula due to Molien.[17] Suppose a group G acts on a representation space $\mathcal{V}^n$, of dimension n, so that for $x \in \mathcal{V}^n$ and $g \in G$ there is a linear action $x \to g\,x$. In terms of coordinates x_s, $s = 1, \dots, n$ this becomes a $n \times n$ dimensional representation of G given by

$$(g\,x)_s = \textstyle\sum_t x_t\, D_{ts}(g)\,. \tag{2.99}$$

An important question is then to determine possible G invariant homogeneous polynomials of degree p, $P(\lambda x) = \lambda^p P(x)$, $P(g\,x) = P(x)$. Let m_p be the number of such invariant polynomials. Then there is a generating function

$$M_G(\mathcal{V}^n, t) = \sum_{p\geq 0} m_p\, t^p = \frac{1}{|G|} \sum_{g\in G} 1\Big/ \det\big(\mathbb{1} - t\, D(g)\big)\,, \tag{2.100}$$

which is the Molien series.

The determinant in (2.100) can be expanded

$$\begin{aligned} 1\Big/\det\big(\mathbb{1} - t\,D(g)\big) &= \exp\big(-\operatorname{tr}\ln(\mathbb{1} - t\,D(g)\big) = \exp\bigg(\sum_{m\geq 1} \frac{t^m}{m} \operatorname{tr} D(g)^m\bigg) \\ &= \exp\bigg(\sum_{m\geq 1} \frac{t^m}{m}\, \chi_{\mathcal{V}^n}(g^m)\bigg) = \mathrm{PE}\big(t, g; \chi_{\mathcal{V}^n}\big) = 1 + \chi_{\mathcal{V}^n}(g) + \sum_{m\geq 2} u^n\, \chi_{\vee^m \mathcal{V}^n}(g)\,, \end{aligned} \tag{2.101}$$

where the plethystic exponential is defined in (1.69) and $\chi_{\vee^m\mathcal{V}^n}(g) = \vee^m \chi_{\mathcal{V}^n}(g)$ is the character for the representation obtained from $\{D(g)\}$ acting on the m-fold symmetric tensor product space $\vee^m \mathcal{V}^n$ as given in (2.91). As a result (2.100) can be expressed as

$$M_G(\mathcal{V}^n, t) = \frac{1}{|G|} \sum_{g\in G} \mathrm{PE}\big(t, g; \chi_{\mathcal{V}^n}\big)\,. \tag{2.102}$$

To verify this result (2.100) we consider all possible homogeneous polynomials of degree p. There are

$$N_p = \frac{1}{p!}(n)_p\,, \qquad (n)_p = \frac{\Gamma(n+p)}{\Gamma(n)} = \begin{cases} 1\,, & p = 0 \\ n(n+1)\dots(n+p-1)\,, & p \geq 1 \end{cases}, \tag{2.103}$$

such polynomials (here $(n)_p$ is the Pochhammer symbol). A basis of degree p polynomials $\{\mathcal{P}_\alpha(x) : \alpha = 1, \dots, N_p\}$ defines a N_p-dimensional representation of G

$$\mathcal{P}_\alpha(g\,x) = \textstyle\sum_{\beta=1}^{N_p} \mathcal{P}_\beta(x)\, \mathcal{D}_{\beta\alpha}(g)\,. \tag{2.104}$$

[17] Theodor Georg Andreas Molien, 1861-1941, Baltic German, Russian nationality.

The representation $\{\mathcal{D}(g)\}$ is formed from a p-fold symmetrised tensor product of the representation $\{D(g)\}$ and we must have

$$\begin{aligned} &\text{Eigenvalues}\, D(g) = \{\lambda_a(g) : a = 1, \dots, n\} \\ \Rightarrow \quad &\text{Eigenvalues}\, \mathcal{D}(g) = \big\{ \textstyle\prod_{a=1}^n \lambda_a(g)^{d_a} : d_a \geq 0, \, \sum_{a=1}^n d_a = p \big\} . \end{aligned} \tag{2.105}$$

There are N_p possible choices for $\{d_a\}$ so that the number of eigenvalues is equal to the dimension of the representation. For a particular $\mathcal{D}(g)$ then the corresponding character is obtained by summing over all the eigenvalues

$$\chi_{\mathcal{D}}(g) = \mathrm{tr}\big(\mathcal{D}(g)\big) = \textstyle\prod_{a=1}^n \sum_{d_a \geq 0} \lambda_a(g)^{d_a} \, \delta_{\sum_a d_a , p} \, . \tag{2.106}$$

The representation $\{\mathcal{D}(g)\}$ can be decomposed by using characters

$$\chi_{\mathcal{D}}(g) = \sum_{i=1}^{n_G} a_{s,p} \, \chi_s(g) \, , \qquad a_{s,p} = \frac{1}{|G|} \sum_{g \in G} \chi_s(g)^* \chi_{\mathcal{D}}(g) \, . \tag{2.107}$$

Hence now

$$\begin{aligned} \sum_{p \geq 0} a_{s,p} \, t^p &= \frac{1}{|G|} \sum_{g \in G} \chi_s(g)^* \prod_{a=1}^n \sum_{p \geq 0} \sum_{d_a \geq 0} \big(t \, \lambda_a(g)\big)^{d_a} \, \delta_{\sum_a d_a , p} \\ &= \frac{1}{|G|} \sum_{g \in G} \chi_s(g)^* \prod_{a=1}^n \sum_{d_a \geq 0} \big(t \, \lambda_a(g)\big)^{d_a} = \frac{1}{|G|} \sum_{g \in G} \chi_s(g)^* \prod_{a=1}^n \big(\mathbb{1} - t \, \lambda_a(g)\big)^{-1} \\ &= \frac{1}{|G|} \sum_{g \in G} \chi_s(g)^* \, 1 \Big/ \det\big(\mathbb{1} - t \, D(g)\big) . \end{aligned} \tag{2.108}$$

The number of invariant polynomials is equal to the number of singlet representations contained in the decomposition of $\{\mathcal{D}(g)\}$ so that $m_p = a_{1,p}$. For the trivial singlet representation $\chi_1(g) = 1$. In this case (2.108) reduces to (2.100).

2.8.1 Anticommuting Molien Series

From a physics perspective it is also interesting to consider invariants under the group action of G on a n-dimensional Grassmanian manifold $\mathbb{M}^n$ which is defined in terms of anticommuting coordinates $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ with $\theta_s \theta_t = -\theta_t \theta_s$. The group action is identical to (2.99). A basis of degree p polynomials in θ, $\tilde{\mathcal{P}}_\alpha(\theta)$ satisfying (2.104) with $x \to \theta$ may also be constructed but now $N_p = \binom{n}{p}$ and we most have $p \leq n$, higher degree polynomials vanish. If $v_a(\theta)$, linear in θ, is an eigenvector of $D(g)$ with eigenvalues λ_a, $a = 1, \dots, n$ then the eigenvectors of $\mathcal{D}(g)$ are just $\prod_{a=1}^n v_a(\theta)^{d_a}$ with $d_a = 0, 1$, since $v_a(\theta)^2 = 0$, and $\sum_a d_a = p$. Hence

$$\text{Eigenvalues}\, \mathcal{D}(g) = \big\{ \textstyle\prod_{a=1}^n \lambda_a(g)^{d_a} : d_a = 0, 1, \, \sum_{a=1}^n d_a = p \big\} , \tag{2.109}$$

and instead of (2.106)

$$\chi_{\mathcal{D}}(g) = \mathrm{tr}\big(\mathcal{D}(g)\big) = \textstyle\prod_{a=1}^n \sum_{d_a = 0}^1 \lambda_a(g)^{d_a} \, \delta_{\sum_a d_a , p} \, . \tag{2.110}$$

The decomposition of $\chi_{\mathcal{D}}(g)$ into irreducible representations remains as in (2.107) and

$$\begin{aligned}\sum_{p\geq 0} \tilde{a}_{s,p}\, t^p &= \frac{1}{|G|}\sum_{g\in G} \chi_s(g)^* \prod_{a=1}^{n} \sum_{p\geq 0}\sum_{d_a=0}^{1} \big(t\,\lambda_a(g)\big)^{d_a}\, \delta_{\sum_a d_a,\,p} \\ &= \frac{1}{|G|}\sum_{g\in G} \chi_s(g)^* \prod_{a=1}^{n} \sum_{d_a=0}^{1} \big(t\,\lambda_a(g)\big)^{d_a} = \frac{1}{|G|}\sum_{g\in G} \chi_s(g)^* \prod_{a=1}^{n} \big(\mathbb{1} + t\,\lambda_a(g)\big) \\ &= \frac{1}{|G|}\sum_{g\in G} \chi_s(g)^* \det\big(\mathbb{1} + t\,D(g)\big)\,. \end{aligned} \tag{2.111}$$

The number of invariant polynomials of degree p is $\tilde{m}_p = a_{1,p}$ and the generating function becomes now

$$\tilde{M}_G(\mathbb{M}^n, t) = \sum_{p\geq 0} \tilde{m}_p\, t^p = \frac{1}{|G|}\sum_{g\in G} \det\big(\mathbb{1} + t\,D(g)\big)\,, \tag{2.112}$$

The first order contribution in and expansion in t, determining the number of linear invariants, is the same as in (2.100) and is given by $\sum_g \chi(g)$ which will be zero for an irreducible representation. In general (2.112) is a polynomial of degree n.

In an analogous fashion to (2.101)

$$\begin{aligned}\det\big(\mathbb{1} + t\,D(g)\big) &= \exp\bigg(\sum_{m\geq 1} (-1)^{m-1}\frac{t^m}{m}\,\chi_{\mathcal{V}^n}(g^m)\bigg) \\ &= \mathrm{PE}_f\big(t, g; \chi_{\mathcal{V}^n}\big) = 1 + \chi_{\mathcal{V}^n}(g) + \sum_{m\geq 2} u^n\, \chi_{\wedge^m \mathcal{V}^n}(g)\,, \end{aligned} \tag{2.113}$$

with $\chi_{\wedge^m \mathcal{V}^n}(g) = \wedge^m \chi_{\mathcal{V}^n}(g)$ the character for the representation acting on the m-fold antisymmetric tensor product space $(\mathcal{V}^n \otimes \cdots \otimes \mathcal{V}^n)_{\text{antisym.}}$.

2.8.2 Examples of Molien Series

As a first example we consider $\mathbb{Z}_n$ which is generated by $a^n = e$. Acting on $\mathbb{R}^n$ there is a corresponding action given by cyclic permutations of the coordinates $(x_1, x_2, \ldots, x_{n-1}, x_n) \to (x_2, x_3, \ldots, x_n, x_1)$. This is just an n-cycle in the group of all permutations $\mathcal{S}_n$. The associated $n \times n$ representation matrix A_n, where ${A_n}^n = \mathbb{1}_n$, is then of the form

$$A_n = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \tag{2.114}$$

In general

$$\det\big(\mathbb{1} - t\,{A_n}^r\big) = (1-t^m)^{n/m}, \quad \text{for } m,\ m|n, \text{ the smallest integer such that } r\,m = 0 \mod n\,. \tag{2.115}$$

For $r = 1$, $m = n$ this can be worked out directly. Other cases can be found by generalising results such as ${A_4}^2 \simeq A_2 \oplus A_2$ ${A_6}^2 \simeq A_3 \oplus A_3$, ${A_6}^3 \simeq A_2 \oplus A_2 \oplus A_2$ together with $\det(A \oplus B) = \det A \det B$.

A special case of the Molien formula for this case is

$$M_{\mathbb{Z}_n}(\mathbb{R}^n,t) = \frac{1}{n}\bigg(\frac{1}{(1-t)^n} + \frac{n-1}{1-t^n}\bigg), \qquad n \text{ prime}. \tag{2.116}$$

This representation is reducible. Restricting to just the one dimensional irreducible representation

$$M_{\mathbb{Z}_n}(\mathbb{C},t) = \frac{1}{n}\sum\nolimits_{r=0}^{n-1}\frac{1}{1-\exp(\frac{2r\pi i}{n})\,t} = \frac{1}{1-t^n}. \tag{2.117}$$

The summation can be calculated by expanding in t and using $\sum_{r=0}^{n-1}\exp(\frac{2qr\pi i}{n}) = n\,\delta_{q,pn}$ for $p = 0,1,2,\dots$. For $z \in \mathbb{C}$ the invariants are just z^{rn}, $n = 1,2,\dots$ which are just products of a single fundamental invariant z^n. For $n=2$ we may restrict $\mathbb{C}$ to $\mathbb{R}$.

Other general results are obtained by considering the two dimensional representation of the cyclic group C_n generated by the real 2×2 rotation matrix $A = \begin{pmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n}\end{pmatrix}$ satisfying $A^n = \mathbb{1}_2$. Each element has its own conjugacy class and the Molien formula gives[18]

$$M_{C_n}(\mathbb{R}^2,t) = \frac{1}{n}\sum\nolimits_{r=0}^{n-1}\frac{1}{1-2\cos\frac{2r\pi}{n}\,t+t^2} = \frac{1+t^n}{(1-t^2)(1-t^n)}. \tag{2.118}$$

For $(x,y)\in\mathbb{R}^2$ expanding (2.118) in powers of t then gives the number of invariants under the action of C_n on x,y. These can all be expressed in terms of sums and products of a finite number of fundamental invariants. There are two primary invariants $p_1 = x^2+y^2$, $p_2 = Re(x+i\,y)^n$ which correspond to the factors $1-t^2$ and $1-t^n$ in the denominator. There is also a secondary invariant $q = Im(x+i\,y)^n$. For $n=2$, $p_2 = x^2-y^2$ and $q = 2xy$. Both p_2 and q are invariant under $x+i\,y \to e^{2\pi/n\,i}(x+i\,y)$ but $q^2 = {p_1}^n - {p_2}^2$. Of course the role of p_2 and q can be interchanged. Any invariant is then obtained by sums and products of p_1, p_2 together also with terms linear in q.

For the dihedral group, with the conjugacy classes in (1.58) and the representation given by (2.29) or (2.32) for $k=1$, the Molien formula becomes

$$M_{D_n}(\mathbb{R}^2,t) = \begin{cases} \frac{1}{2n}\Big(\frac{1}{(1-t)^2} + \frac{1}{(1+t)^2} + \frac{n}{1-t^2} + 2\sum_{r=1}^{\frac12 n-1}\frac{1}{1-2\cos\frac{2r\pi}{n}\,t+t^2}\Big), & n \text{ even}, \\ \frac{1}{2n}\Big(\frac{1}{(1-t)^2} + \frac{n}{1-t^2} + 2\sum_{r=1}^{\frac12(n-1)}\frac{1}{1-2\cos\frac{2r\pi}{n}\,t+t^2}\Big), & n \text{ odd}. \end{cases} \tag{2.119}$$

In either case[19]

$$M_{D_n}(\mathbb{R}^2,t) = \frac{1}{(1-t^2)(1-t^n)}. \tag{2.120}$$

[18]Expanding in t the sum becomes $\frac{1}{n}\sum_{r=0}^{n-1}\sum_{p,q\ge0}t^{p+q}\exp(\frac{2(p-q)r\pi i}{n}) = \sum_{s=-\infty}^{\infty}\sum_{q\ge0,-ns}t^{2q+ns}$ which is readily evaluated.

[19]For $n=2m$ the basic sum can be reduced to

$$\begin{aligned}\frac{1}{2m}\sum\nolimits_{r=0}^{m}\frac{1}{1-2\cos\frac{r\pi}{m}\,t+t^2} &= \frac{1}{4m}\sum_{p,q\ge0}t^{p+q}\sum\nolimits_{r=0}^{m}\Big(\exp(\tfrac{(p-q)r\pi i}{m}) + \exp(-\tfrac{(p-q)r\pi i}{m})\Big) \\ &= \frac{1}{4m}\sum_{p,q\ge0}t^{p+q}\big(\delta_{p,q} + (-1)^{p-q}\big) + \frac12\sum_{s=-\infty}^{\infty}\sum_{q\ge0,-2ms}t^{2q+2ms},\end{aligned}$$

which is then straightforward.

In the real basis (2.32) and, for $(x,y)\in\mathbb{R}^2$, p_1, p_2 as in the $\mathbb{Z}_2$ case are still invariants but q is no longer since it changes sign under $y\to -y$.

A further illustrative example is given by the dicyclic, or binary dihedral, group Q_{4n}. The conjugacy classes are given in (1.87) and a two dimensional irreducible complex representation is given in (2.33) and (2.34) setting $k=1$. For this representation $\det(\mathbb{1}-t\,A^r) = 1-2\cos\frac{r\pi}{n}\,t+t^2$, $\det(\mathbb{1}-t\,A^rB)=1+t^2$, and

$$\begin{aligned} M_{\mathrm{Q}_{4n}}(\mathbb{C}^2,t) &= \frac{1}{4n}\bigg(\frac{1}{(1-t)^2}+\frac{1}{(1+t)^2}+\frac{2\,n}{1+t^2}+2\textstyle\sum_{r=1}^{n-1}\displaystyle\frac{1}{1-2\cos\frac{r\pi}{n}\,t+t^2}\bigg) \\ &= \frac{1+t^{2n+2}}{(1-t^4)(1-t^{2n})}\,. \end{aligned} \tag{2.121}$$

There are two primary invariants, for $(x,y)\in\mathbb{C}^2$ these are $p_1=x^2y^2$, $p_2=x^{2n}+y^{2n}$, and also $q=xy(x^{2n}-y^{2n})$ is a secondary invariant since $q^2=p_1\,p_2{}^2-4\,p_2{}^{n+1}$. All invariants are formed from sums of products of p_1,p_2 together with similar expressions linear in q.

There are more invariants when considering the symmetric group $\mathcal{S}_n$ acting on the $\mathbb{R}^n$ by permuting the coordinates $(x_1,x_2,\dots,x_n)\to(x_{\sigma(1)},x_{\sigma(2)},\dots,x_{\sigma(n)})$, $\sigma\in\mathcal{S}_n$. The sum in (2.100) can be reduced to a sum over conjugacy classes which are given by products of p_i-cycles and with the results in subsection 1.4.3 for the numbers of group elements in each conjugacy class together with using, from (2.115) for $n=p_i$, $r=1$, that in any conjugacy class containing $\sigma\in\mathcal{S}_n$ each p_i-cycle contributes a factor $1-t^{p_i}$ to $\det[\mathbb{1}_n-t\,D(g)]$ so that

$$\det[\mathbb{1}_n-t\,D(g)] = \textstyle\prod_{i=1}^r(1-t^{p_i})\,, \qquad g\in\mathcal{C}_{[p_1,\dots,p_r]}\,, \quad \textstyle\sum_{i=1}^r p_i=n\,. \tag{2.122}$$

The general formula (2.100) then becomes, with a sum over conjugacy classes as in (1.23),

$$M_{\mathcal{S}_n}(\mathbb{R}^n,t) = \frac{1}{n!}\sum_{r=1}^n\ \sum_{j_1,j_2,\dots j_r\ge 1}\ \sum_{p_1>p_2>\dots p_r\ge 1}\delta_{n,\sum_{i=1}^r j_ip_i}\ N_{[p_1(j_1),\dots,p_r(j_r)]}\frac{1}{\prod_{i=1}^r(1-t^{p_i})^{j_i}}\,. \tag{2.123}$$

Applying this for $n=3,4,5$ with the results given in (1.63), (1.64), (1.65)

$$\begin{aligned} M_{\mathcal{S}_3}(\mathbb{R}^3,t) &= \tfrac16\big(\tfrac{2}{1-t^3}+\tfrac{3}{(1-t^2)(1-t)}+\tfrac{1}{(1-t)^3}\big)\,, \\ M_{\mathcal{S}_4}(\mathbb{R}^4,t) &= \tfrac{1}{24}\big(\tfrac{6}{1-t^4}+\tfrac{8}{(1-t^3)(1-t)}+\tfrac{3}{(1-t^2)^2}+\tfrac{6}{(1-t^2)(1-t)^2}+\tfrac{1}{(1-t)^4}\big)\,, \\ M_{\mathcal{S}_5}(\mathbb{R}^5,t) &= \tfrac{1}{120}\big(\tfrac{24}{1-t^5}+\tfrac{30}{(1-t^4)(1-t)}+\tfrac{20}{(1-t^3)(1-t^2)}+\tfrac{20}{(1-t^3)(1-t)^2} \\ &\qquad +\tfrac{15}{(1-t^2)^2(1-t)}+\tfrac{10}{(1-t^2)(1-t)^3}+\tfrac{1}{(1-t)^5}\big)\,. \end{aligned} \tag{2.124}$$

The generating function obtained from (2.123) using (1.69) is expressible as a plethystic exponential

$$\sum_{n=0}^\infty u^n M_{\mathcal{S}_n}(\mathbb{R}^n,t) = \mathrm{PE}(u,t;f)\,, \qquad f(t)=\frac{1}{1-t}\,. \tag{2.125}$$

By expanding in t for this $f(t)$, $\sum_{m=1}^\infty\frac{u^m}{m}f(t^m) = -\sum_{r=0}^\infty\ln(1-u\,t^r)$, and hence

$$\sum_{n=0}^\infty u^n M_{\mathcal{S}_n}(\mathbb{R}^n,t) = \prod_{r=0}^\infty\frac{1}{1-u\,t^r} = \sum_{n=0}^\infty u^n\prod_{r=1}^\infty\sum_{s_r\ge 0}t^{\sum_{r\ge1}s_r r}\Big|_{\sum_{r\ge1}s_r\le n}\,. \tag{2.126}$$

In each case the results are in accord with the general formula

$$M_{\mathcal{S}_n}(\mathbb{R}^n,t) = \prod_{r=1}^{n} \frac{1}{1-t^r}\,. \tag{2.127}$$

This expression can be interpreted as showing that for every $p \le n$ there is a new homogeneous symmetric polynomial of degree p, other such polynomials can be represented as sums, with rational coefficients, of products of symmetric polynomials of lower degree. For instance we may consider $p_p(x) = \sum_{a=1}^{n} x_a{}^p$, $p = 1,\dots,n$ as a set of primary polynomial invariants. For $n = 2$ (2.123) coincides with (2.116).

For the anticommuting case

$$\tilde{M}_{\mathcal{S}_n}(\mathbb{M}^n,t) = \sum_{r=0}^{n} t^r\,. \tag{2.128}$$

so there one invariant for any $p \le n$.

The n-dimensional representation of $\mathcal{S}_n$ is reducible since setting all x_a equal defines an invariant subspace. An irreducible $(n-1)$-dimensional representation is obtained by imposing the linear condition $\sum_a x_a = 0$, which of course is invariant under $\mathcal{S}_n$. In this case

$$M_{\mathcal{S}_n}(\mathbb{R}^{n-1},t) = \prod_{r=2}^{n} \frac{1}{1-t^r}\,. \tag{2.129}$$

The lack of the linear $1-t$ factor reflects the absence of the $\sum_a x_a$ invariant. (2.129) is identical with (2.120) for $n = 3$.

For the alternating group A_n conjugacy classes involving odd numbers of 2-cycles are removed. Following the results in 1.4.3 for $n = 3,4,5$ the general expression is

$$M_{A_n}(\mathbb{R}^n,t) = \big(1 + t^{\frac{1}{2}n(n-1)}\big)\prod_{r=1}^{n} \frac{1}{1-t^r}\,. \tag{2.130}$$

With less symmetry there is an extra secondary invariant $q(x) = \prod_{1\le b<a\le n}(x_a - x_b)$, of degree $\frac{1}{2}n(n-1)$, as well as the n primary invariants which are present for $\mathcal{S}_n$. $q_n(x)$ changes sign under odd permutations in $\mathcal{S}_n$. However $q_n(x)^2$ is invariant under all $\mathcal{S}_n$ permutations and can be expressed in terms of $\{p_p(x)\}$. This explains the presence of the single factor in the numerator in (2.130). There is a similar reduction as in going from (2.123) to (2.129) on restricting to an irreducible representation on $\mathbb{R}^{n-1}$.

2.8.3 Molien Series and Wreath Products

The wreath product, whose action is defined in (1.51), plays an important role when groups act on spaces with symmetry conditions imposed. The simplest illustration is $\mathbb{Z}_2 \wr \mathbb{Z}_2$ where $\mathbb{Z}_2 \times \mathbb{Z}_2$ acts on $(x,y) \in \mathbb{R}^2$ by reflecting x and y and there is also a $\mathbb{Z}_2$ action obtained by interchanging x and y. On this two dimensional space the representation of $\mathbb{Z}_2 \wr \mathbb{Z}_2$ is obtained by taking

$$\begin{gathered} d_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \quad d_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \quad d_3 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \quad d_4 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \\ d_5 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \quad d_6 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), \quad d_7 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \quad d_8 = \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right), \end{gathered} \tag{2.131}$$

with d_1, d_2, d_3, d_4 representing $\mathbb{Z}_2 \times \mathbb{Z}_2$. Applying (2.100)

$$\begin{aligned} M_{\mathbb{Z}_2 \wr \mathbb{Z}_2}(\mathbb{R}^2, t) &= \frac{1}{8} \sum_{i=1}^{8} \frac{1}{\det(\mathbb{1} - t\, d_i)} = \frac{1}{8}\bigg(\frac{1}{(1-t)^2} + \frac{1}{(1+t)^2} + \frac{4}{1-t^2} + \frac{2}{1+t^2} \bigg) \\ &= \frac{1}{(1-t^2)(1-t^4)} \, . \end{aligned} \tag{2.132}$$

The fundamental invariants are just $x^2 + y^2$ and $x^4 + y^4$.

More generally for $G \wr \mathcal{S}_n$ the Molien formula allows for significant simplifications. For a representation $d(g)$ of G acting on $\mathbb{R}^k$ then (2.100) gives

$$M_{G \wr \mathcal{S}_n}\big(\mathbb{R}^{nk}, t\big) = \frac{1}{|G|^n n!} \sum_{\sigma \in \mathcal{S}_n} \sum_{g \in G_{n\times}} \frac{1}{\det\big(\mathbb{1} - t\, D_\sigma(g_\sigma)\big)} \, , \tag{2.133}$$

where D_σ is the representation of $G \wr \mathcal{S}_n$ acting on $\mathbb{R}^{nk}$ formed from the n-fold tensor product of $d(g)$. This representation depends on the decomposition of σ into non overlapping cycles,

$$\sigma = \sigma_{p_1} \dots \sigma_{p_r} \in \mathcal{C}_{[p_1, \dots, p_r]} \quad \Rightarrow \quad g_\sigma = g_{(p_1)} \cup \dots g_{(p_r)} \, , \quad g_{(p)} = \big(g_{(p)1}, \dots, g_{(p)p}\big) \tag{2.134}$$

with σ_p, ${\sigma_p}^p = e$, generating cyclic permutations of the group elements in $g_{(p)}$. In this case $D_\sigma(g_\sigma) \simeq D_{p_1}\big(g_{(p_1)}\big) \oplus \dots \oplus D_{p_r}\big(g_{(p_r)}\big)$ with

$$D_p\big(g_{(p)}\big) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & d_1 \\ d_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & d_3 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d_p & 0 \end{pmatrix} , \quad p \geq 2 \, , \quad D_1\big(g_{(1)}\big) = d_1 \, , \quad d_i = d\big(g_{(p)i}\big) \, . \tag{2.135}$$

Since

$$\begin{pmatrix} \mathbb{1}_k & 0 & 0 & \cdots & 0 & -t\, d_1 \\ -t\, d_2 & \mathbb{1}_k & 0 & \cdots & 0 & 0 \\ 0 & -t\, d_3 & \mathbb{1}_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -t\, d_p & \mathbb{1}_k \end{pmatrix} \begin{pmatrix} \mathbb{1}_k & 0 & 0 & \cdots & 0 & 0 \\ t\, d_2 & \mathbb{1}_k & 0 & \cdots & 0 & 0 \\ t^2\, d_3 d_2 & 0 & \mathbb{1}_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ t^{p-1} d_p \dots d_2 & 0 & 0 & \cdots & 0 & \mathbb{1}_k \end{pmatrix} = \begin{pmatrix} \mathbb{1}_k - t^p d_1 d_p \dots d_2 & 0 & 0 & \cdots & 0 & -t\, d_1 \\ 0 & \mathbb{1}_k & 0 & \cdots & 0 & 0 \\ 0 & -t\, d_3 & \mathbb{1}_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -t\, d_p & \mathbb{1}_k \end{pmatrix} , \tag{2.136}$$

evaluating the determinant becomes straightforward giving

$$\det\big(\mathbb{1}_{pk} - t\, D_p(g_{(p)})\big) = \det\big(\mathbb{1}_k - t^p d_p \dots d_1\big) = \det\big(\mathbb{1}_k - t^p d(g_{(p)p} \dots g_{(p)1})\big) \, . \tag{2.137}$$

Using

$$\textstyle\sum_{(g_1, \dots, g_p) \in G_{p\times}} f(g_p \dots g_1) = |G|^{p-1} \sum_{g \in G} f(g) \, , \tag{2.138}$$

the Molien sum in (2.133) can be reduced to a sum over conjugacy classes of $\mathcal{S}_n$

$$\begin{aligned} M_{G \wr \mathcal{S}_n}(\mathbb{R}^{nk}, t) &= \frac{1}{n!} \sum_{r=1}^{n} \sum_{j_1, j_2, \dots j_r \geq 1} \sum_{p_1 > p_2 > \dots p_r \geq 1} \delta_{n, \sum_{i=1}^r j_i p_i} \prod_{i=1}^{r} \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det\big(\mathbb{1}_k - t^{p_i} d(g)\big)} \\ &= \frac{1}{n!} \sum_{r=1}^{n} \sum_{j_1, j_2, \dots j_r \geq 1} \sum_{p_1 > p_2 > \dots p_r \geq 1} \delta_{n, \sum_{i=1}^r j_i p_i} \prod_{i=1}^{r} M_G(\mathbb{R}^k, t^{p_i}) \, . \end{aligned} \tag{2.139}$$

From (1.69)

$$\sum_{n \geq 0} u^n M_{G \wr \mathcal{S}_n}(\mathbb{R}^{nk}, t) = \mathrm{PE}\big(u, t; M_G(\mathbb{R}^k)\big) \, . \tag{2.140}$$

Various special cases are easily obtained. From (2.116) and (2.123), (2.123)

$$M_{C_m \wr \mathcal{S}_n}(\mathbb{C}^n, t) = M_{\mathcal{S}_n}(\mathbb{C}^n, t^m) = \prod_{r=1}^{n} \frac{1}{1 - t^{mr}} \,. \tag{2.141}$$

For $m = 2$, taking $\mathbb{C}^n \to \mathbb{R}^n$, this is just the Molien series for the symmetry group of the n-dimensional hypercube. For $n = 2$ this reduces to (2.132). Using (2.120) with $N_{[2]} = N_{[1,1]} = 1$,

$$M_{D_m \wr \mathbb{Z}_2}(\mathbb{R}^4, t) = \frac{1}{2}\Big(M_{D_m}\big(\mathbb{R}^2, t\big)^2 + M_{D_m}\big(\mathbb{R}^2, t^2\big)\Big) = \frac{1 + t^{m+2}}{(1-t^2)(1-t^4)(1-t^m)(1-t^{2m})} \,. \tag{2.142}$$

It is an exercise to determine the primary and secondary invariants in this case.

2.9 Symmetries in Quantum Mechanics, Projective and Anti-Unitary Representations

A symmetry of a physical system is defined as a set of transformation 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

$$|\langle\phi|\psi\rangle|^2 = |\langle\phi'|\psi'\rangle|^2 \quad \text{for all} \quad |\psi\rangle, |\phi\rangle \in \mathcal{H} \,. \tag{2.143}$$

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[20] proved that there is an operator U such that

$$U|\psi\rangle = |\psi'\rangle \,, \tag{2.144}$$

and either

$$\langle\phi'|\psi'\rangle = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle \,, \quad U\big(a_1|\psi_1\rangle + a_2|\psi_2\rangle\big) = a_1 U|\psi_1\rangle + a_2 U|\psi_2\rangle \,, \tag{2.145}$$

so that U is *unitary linear*, or

$$\langle\phi'|\psi'\rangle = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle^* \,, \quad U\big(a_1|\psi_1\rangle + a_2|\psi_2\rangle\big) = {a_1}^* U|\psi_1\rangle + {a_2}^* U|\psi_2\rangle \,, \tag{2.146}$$

and U is *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). This requirement also apparent since if x, p are the position,

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

momentum operators then the action of time reversal requires $TxT^{-1} = x$, $TpT^{-1} = -p$ and for this to be compatible with the fundamental commutation relation $[x, p] = i\,1$ it is necessary that T is anti-linear.

In the simplest case if G is a symmetry for a physical system with a Hamiltonian H we must require

$$U[g]HU[g]^{-1} = H \quad \text{for all} \quad g \in G\,. \tag{2.147}$$

If H has energy levels with degeneracy so that

$$H|\psi_r\rangle = E|\psi_r\rangle\,, \quad r = 1, \dots, n\,, \tag{2.148}$$

then it is easy to see that

$$H\,U[g]|\psi_r\rangle = E\,U[g]|\psi_r\rangle\,. \tag{2.149}$$

For a symmetry group $G = \{g\}$ there are then unitary operators $U[g]$ where we require $U[e] = 1$, $U[g^{-1}] = U[g]^{-1}$. If the unitary operators satisfy the usual group multiplication rules $U(g_i)U(g_j) = U(g_ig_j)$ then

$$U[g]|\psi_r\rangle = \sum_{s=1}^{n} |\psi_s\rangle D_{sr}(g)\,, \tag{2.150}$$

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.

2.9.1 Projective Representations

However in quantum mechanics, because of the freedom of complex phases, we may relax the product rule and require only

$$U[g_i]U[g_j] = e^{i\gamma(g_i,g_j)}U[g_ig_j]\,. \tag{2.151}$$

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,

$$\gamma(g_i, g_jg_k) + \gamma(g_j, g_k) = \gamma(g_ig_j, g_k) + \gamma(g_i, g_j)\,. \tag{2.152}$$

There are always solutions to (2.152) of the form

$$\gamma(g_i, g_j) = \alpha(g_ig_j) - \alpha(g_i) - \alpha(g_j)\,, \tag{2.153}$$

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 (2.151). For most groups there are no non trivial solutions for $\gamma(g_i, g_j)$ so the extra freedom allowed by (2.151) may be neglected so there is no need to consider projective representations, although there are some cases when it is essential.

As an extension of the above we may consider wave functions $\Psi(X)$ depending on variables $X \in M$ on which there is a group action $X \to gX$ for $g \in G$. There is then an induced action on Ψ given by

$$\Psi(X) \to \Psi_g(X) = e^{i\phi_g(X)}\Psi(g^{-1}X)\,, \tag{2.154}$$

where we allow for a phase ϕ_g depending on g. Group multiplication requires

$$e^{i\phi_{g_1 g_2}(X)} = e^{i\phi_{g_1}(X)}\, e^{i\phi_{g_2}(g_1{}^{-1}X)}\,. \tag{2.155}$$

A trivial solution is obtained if

$$e^{i\phi_g(X)} = e^{i(\chi(X)-\chi(g^{-1}X))}\,, \tag{2.156}$$

since then we may redefine $\Psi(X) \to e^{i\chi(X)}\Psi(X)$ and eliminate the phase from the transformation (2.154).

Phases satisfying (2.152) or (2.155) modulo the trivial solutions (2.153) or (2.156) correspond to cohomology classes $H^2(G, U(1))$ or $H^1(G, U(1))$ of the group G.[21]

The complex phase in (2.151) may be restricted to a representation of a discrete subgroup of $U(1)$. Thus we may have $U[g_i]U[g_j] = A(g_i, g_j)\,U[g_i g_j]$ with $A(g_i, g_j) \in C_n$ and $A(g_i, g_j g_k)\,A(g_j, g_k) = A(g_i g_j, g_k)\,A(g_i, g_j)$ with the trivial arbitrariness, corresponding to redefinitions of $U(g_i)$, $A(g_i, g_j) \sim A(g_i, g_j)B(g_i g_j)B(g_i)^{-1}B(g_j)^{-1}$ for $B(g) \in C_n$. The existence of non trivial $A(g_i, g_j)$ corresponds to $H^2(G, \mathbb{Z}_n)$. Assuming $U(e) = 1$ then $A(e, g_i) = A(g_i, e) = 1$ for all g_i.

As an illustration we may consider the group $D_2 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ which has four elements $\{e, a, b, ab\}$ with $a^2 = b^2 = e$, $ab = ba$. There is a two dimensional projective representation, corresponding to non trivial $H^2(D_2, \mathbb{Z}_2)$, formed by taking $D(a) = I$, $D(b) = J$, $D(ab) = K$ in terms of the quaternion representation matrices in (1.115). For this example $A(g_i, g_j)$ is given by the table

$$\begin{array}{c|cccc} A & e & a & b & ab \\ \hline e & 1 & 1 & 1 & 1 \\ a & 1 & -1 & 1 & -1 \\ b & 1 & -1 & -1 & 1 \\ ab & 1 & 1 & -1 & -1 \end{array}\,. \tag{2.157}$$

2.9.2 Anti-Unitary Representations

Anti-unitary representations are possible for a group $G = G_0 \cup G_1$ where G_0 is a normal subgroup and for any $h \in G_0$, $a, a' \in G_1$ then ha, $ah \in G_1$, $aa' \in G_0$. This requires $\dim G_0 =$

[21] For real functions of n group elements $g_i \in G$, $\varphi_n(g_1, \dots, g_n) \in C^n$, we may define $\mathrm{d} : C^n \to C^{n+1}$ by

$$(\mathrm{d}\varphi_n)(g_1, \dots, g_{n+1}) = \varphi_n(g_2, \dots, g_{n+1}) - \varphi_n(g_1 g_2, g_3, \dots, g_{n+1}) + \varphi_n(g_1, g_2 g_3, \dots, g_{n+1}) \\ - \dots + (-1)^n \varphi_n(g_1, \dots, g_n g_{n+1}) + (-1)^{n+1}\varphi_n(g_1, \dots, g_n).$$

Then $\mathrm{d}^2 = 0$. Define $Z^n = \ker \mathrm{d} \cap C^n$, or $\{\varphi_n : \mathrm{d}\varphi_n = 0\}$, and the cohomology class $H^n(G, \mathbb{R})$ is defined by $H^n = Z^n/\mathrm{d}C^{n-1}$. The elements of $H^n(G, U(1))$ are given by $e^{i\,\varphi_n(g_1,\dots,g_n)}$ with $\varphi_n(g_1, \dots, g_n) \sim \varphi_n(g_1, \dots, g_n) + 2\pi$.

$\dim G_1$. Acting on states $\{|\psi_r\rangle\}$ then $U(a)$ for $a \in G_1$ can act anti-linearly so that

$$U[h] \sum_{r=1}^{n} |\psi_r\rangle a_r = \sum_{r,s=1}^{n} |\psi_s\rangle D_{sr}(h) a_r \,, \quad U[a] \sum_{r=1}^{n} |\psi_r\rangle a_r = \sum_{r,s=1}^{n} |\psi_s\rangle D_{sr}(a) a_r{}^* \,, \tag{2.158}$$

with both $D(h)$, $D(a)$ unitary matrices. Hence

$$D(hg) = D(h)D(g) \,, \quad D(ag) = D(a)D(g)^* \quad \text{for all} \quad h \in G_0,\, a \in G_1,\, g \in G \,, \tag{2.159}$$

so that $D(a^2) = D(a)D(a)^*$, and

$$D(a^{-1}) = D(a)^{-1*} = D(a)^T \,. \tag{2.160}$$

As was first described by Wigner this defines a *co-representation* extending the standard representation of G_0. With these results

$$D'(h) = D(a^{-1}ha) = \big(D(a)^{-1}D(h)D(a)\big)^* \,, \tag{2.161}$$

also forms a representation of G_0. For a co-representation the matrix similarity transform (2.5) becomes

$$D(h) \to SD(h)S^{-1} \,, \qquad D(a) \to SD(a)S^{-1*} \,. \tag{2.162}$$

The matrices remain unitary if S is unitary. For $S = \alpha\,\mathbb{1}$, $|\alpha| = 1$, then $D(a) \sim \alpha^2 D(a)$. The co-representation is reducible if under the transformation (2.162) the matrices $D(h)$, $D(a)$ can be made block diagonal for all h, a.

The different possibilities for anti-unitary representations can be obtained by assuming a decomposition

$$D(h) = \begin{pmatrix} M(h) & 0 \\ 0 & \bar{M}(h) \end{pmatrix} , \qquad D(a) = \begin{pmatrix} 0 & N(a) \\ \bar{N}(a) & 0 \end{pmatrix} , \tag{2.163}$$

where $M(h)$, $\bar{M}(h)$, $N(a)$, $\bar{N}(a)$ are all unitary matrices with

$$\bar{N}(a^{-1}) = N(a)^T \,, \quad N(a)\bar{N}(a)^* = M(a^2) \,, \quad \bar{N}(a)N(a)^* = \bar{M}(a^2) \tag{2.164}$$

Necessarily $\{M(h)\}$ and $\{\bar{M}(h)\}$ define representations $\mathcal{R}_0$ and $\bar{\mathcal{R}}_0$ of G_0 which have equal dimension d_0, and the co-representation $\mathcal{R}$ of G defined by (2.163) then has dimension $2d_0$. From (2.161)

$$M(a^{-1}ha) = \big(\bar{N}(a)^{-1}\bar{M}(h)\bar{N}(a)\big)^* \,, \qquad \bar{M}(a^{-1}ha) = \big(N(a)^{-1}M(h)N(a)\big)^* \,. \tag{2.165}$$

Representations with $M \leftrightarrow \bar{M}$ and $N \leftrightarrow \bar{N}$ are equivalent. To obtain an irreducible representation for G we require that the representations $\mathcal{R}_0$, $\bar{\mathcal{R}}_0$ are irreducible.

For the decomposition (2.163) a special case arises when $\mathcal{R}_0$ and $\bar{\mathcal{R}}_0$ are equivalent representations, $\mathcal{R}_0 \simeq \bar{\mathcal{R}}_0$, and, by a choice of basis, we may take $M(h) = \bar{M}(h)$ for all h. Then (2.165) requires

$$\bar{N}(a)^{-1}M(h)\bar{N}(a) = N(a)^{-1}M(h)N(a) \quad \Rightarrow \quad \big[M(h),\, N(a)\bar{N}(a)^{-1}\big] = 0 \text{ for all } h \,. \tag{2.166}$$

Assuming $\mathcal{R}_0$ is an irreducible representation then by Schur's lemma $\bar{N}(a)N(a)^{-1} \propto \mathbb{1}$ and hence

$$\bar{N}(a) = \alpha\, N(a) \quad \Rightarrow \quad \bar{N}(a)^{-1} = \alpha^*\, N(a)^{-1} \quad \Rightarrow \quad \alpha^*\alpha = 1\,, \tag{2.167}$$

since $N(a)$, $\bar{N}(a)$ are unitary matrices. Substituting in (2.164) with $M = \bar{M}$

$$\alpha\, \bar{N}(a)N(a)^* = M(a^2)\,, \qquad \alpha = \alpha^*\,. \tag{2.168}$$

Hence $\alpha^2 = 1$ and there are two possibilities.

$$\begin{aligned} \mathrm{I}:\quad & \bar{N}(a) = N(a)\,, \quad D(h) = M(h)\mathbb{1}_2\,, \quad D(a) = N(a)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\,, \\ \mathrm{II}:\quad & \bar{N}(a) = -N(a)\,, \quad D(h) = M(h)\mathbb{1}_2\,, \quad D(a) = N(a)\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\,. \end{aligned} \tag{2.169}$$

The remaining case arises when

$$\mathrm{III.} \quad \mathcal{R}_0 \not\simeq \bar{\mathcal{R}}_0\,, \quad \text{(i)} \ \ \bar{\mathcal{R}}_0 \simeq {\mathcal{R}_0}^*\,, \quad \text{(ii)} \ \ \bar{\mathcal{R}}_0 \not\simeq {\mathcal{R}_0}^*\,. \tag{2.170}$$

The first case is reducible since we may easily construct a real S such that $S\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)S^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ so that

$$\big\{D(h), D(a)\big\} \to \big\{M(h), N(a)\big\} \oplus \big\{M(h), -N(a)\big\}\,. \tag{2.171}$$

The two anti-unitary representations of dimension d_0 in this decomposition for case I are equivalent. Otherwise in cases II, III there is only a single irreducible representation of dimension $2d_0$. If in (2.165) $a \to a'$ then $\bar{M}(h) \to \bar{M}'(h) = S\bar{M}(h)S^{-1}$ corresponds, up to an equivalence, to the same representation $\bar{\mathcal{R}}_0$. If the group G is abelian then only IIIi is possible in (2.170). The anti-unitary irreducible representations of G are then determined by the unitary irreducible representations of G_0.

In the trivial case where $G = \{e, a\} \simeq \mathbb{Z}_2$ and $M(h), N(a) \to 1$ then we may identify $U(a) = T$, the anti-unitary time reversal operator and cases I or II arise according to whether $T^2 = 1$ or -1 and we take $\bar{N}(a) \to 1$ or -1 respectively. As described later $T^2 = (-1)^F$ where F is the fermion number. For time reversal invariant systems with an odd number of spin-$\frac{1}{2}$ particles there is then a twofold degeneracy. This applies in atomic physics with just an external electric field, since electric fields are invariant under time reversal, and is termed Kramers[22] degeneracy.

A restriction of general anti-unitary representations arises if the action of $a \in G_1$ generates an inner automorphism on G_0 so that

$$a^{-1}ha = fhf^{-1}\,, \quad f \in G_0 \ \text{ for all } \ h \in G_0\,, \quad \Rightarrow \quad [af, h] = 0\,, \quad (af)^2 \in \mathcal{Z}(G_0)\,,. \tag{2.172}$$

If the centre of G_0, $\mathcal{Z}(G_0)$, is trivial then $(af)^2 = e$. In this case $G \simeq G_0 \times \mathbb{Z}_2$ and we may identify $T = U[a_0 f]$ as the time reversal operator for some particular $a_0 \in G_1$. For such an anti-unitary representation following (2.159)

$$D_T D(h)^* = D(h) D_T\,, \qquad D_T {D_T}^* = \pm\mathbb{1}\,, \qquad D_T = D(a_0 f)\,, \tag{2.173}$$

allowing $(a_0 f)^2 = e$ to be projectively represented so that $T^2 = \pm 1$. For any $a \in G_1$

$$D(a) = D(a{a_0}^{-1}) D_T\,. \tag{2.174}$$

[22]Hans Kramer, 1894-1952, Dutch.

D_T is assumed to be unitary, ${D_T}^* = {D_T}^{-1T}$, so that this implies

$$D(h)^* = CD(h)C^{-1}\,, \quad C = \pm C^T\,, \qquad C = {D_T}^{-1}\,. \tag{2.175}$$

In case I of the Wigner classification $\{D(h)\}$ forms an irreducible representation which is then real or pseudo-real according to whether $T^2 = 1$ or -1. For case II $D(h)$ decomposes into two identical irreducible representations $M(h)$ and we can take

$$D_T = \begin{pmatrix} 0 & K \\ -K & 0 \end{pmatrix}, \qquad KM(h)^* = M(h)K\,, \qquad KK^* = \mp \mathbb{1}\,, \tag{2.176}$$

In this case we can therefore take $C = K^{-1}$ with $C = \mp C^T$ so that the representation defined by $\{M(h)\}$ is pseudo-real or real depending on whether $T^2 = -1$ or 1. Finally for case III $D(h)$ decomposes into two different irreducible representations $M(h)$, $\bar{M}(h)$ and we can take

$$D_T = \begin{pmatrix} 0 & K \\ \bar{K} & 0 \end{pmatrix}, \quad K\bar{M}(h)^* = M(h)K\,,\ \bar{K}M(h)^* = \bar{M}(h)\bar{K}\,,\ \ K\bar{K}^* = \bar{K}K^* = \pm\mathbb{1}\,. \tag{2.177}$$

Since M, $\bar{M}$ are inequivalent the representations are necessarily complex with $\bar{M}$ equivalent to the conjugate representation formed from M. The type of representation of G_0, real, pseudo-real or complex, then determines the associated anti-unitary representation of G to correspond to cases I, II or III.

Extending results from the standard discussions of representations to the anti-unitary case is more involved. The complex character

$$\chi(g) = \mathrm{tr}\big(D(g)\big)\,, \tag{2.178}$$

is well defined if $g \in G_1$, as a consequence of (2.162), under changes of basis only if S is restricted to be real. Restricting to $g \in G_0$ the character is independent of the choice of basis and since, from (2.159) and (2.160),

$$D(hgh^{-1}) = D(h)D(g)D(h)^{-1}\,, \ \ D(ag^{-1}a^{-1}) = D(a)D(g)^TD(a)^{-1}\,, \quad h \in G_0\,, a \in G_1\,, \tag{2.179}$$

the character $\chi(g)|_{g\in G_0}$ is invariant for any g belonging to an extended conjugacy class defined by $\bar{\mathcal{C}}_G(g) = \{hgh^{-1}, ag^{-1}a^{-1} : h \in G_0, a \in G_1\}$. In general $\bar{\mathcal{C}}_G(g) \supset \mathcal{C}_{G_0}(g)$ but may also include additional conjugacy classes $\mathcal{C}_{G_0}(g_1)$ if $ag^{-1}a^{-1} = g_1 \notin \mathcal{C}_{G_0}(g)$ for some a. If $a'g^{-1}a'^{-1} = g_2$ for some $a' \neq a$ then necessarily $g_2 \in \mathcal{C}_{G_0}(g_1)$ since $aa'^{-1}g_2a'a^{-1} = g_1$ and $aa'^{-1} \in G_0$. If there is any $a \in G_1$ such that $ag^{-1}a^{-1} = g$ for some $a \in G_1$ then $\bar{\mathcal{C}}_G(g) = \mathcal{C}_{G_0}(g)$, $g \in G_0$, if there is no such a, so that $ag^{-1}a^{-1} = g_1 \notin \mathcal{C}_{G_0}(g)$ for any a, then $\bar{\mathcal{C}}_G(g) = \mathcal{C}_{G_0}(g) \cup \mathcal{C}_{G_0}(g_1)$.

For the case $G = G_0 \times \mathbb{Z}_2$ then for $g \in G_0$ $\bar{\mathcal{C}}_G(g) = \mathcal{C}_{G_0}(g)$ if $g^{-1} \in \mathcal{C}_{G_0}(g)$, otherwise $\bar{\mathcal{C}}_G(g) = \mathcal{C}_{G_0}(g) \cup \mathcal{C}_{G_0}(g^{-1})$.

From (2.163) and (2.165)

$$\chi_{\mathcal{R}}(h) = \begin{cases} 2\,\chi_{\mathcal{R}_0}(h)\,, & h = a^{-1}h^{-1}a \ \text{ for some } a \in G_1\,, \\ \chi_{\mathcal{R}_0}(h) + \chi_{\bar{\mathcal{R}}_0}(h)\,, & \chi_{\bar{\mathcal{R}}_0}(h) = \chi_{\mathcal{R}_0}(a^{-1}h^{-1}a)\,,\ h \neq a^{-1}h^{-1}a \ \text{ for any } a \in G_1\,. \end{cases} \tag{2.180}$$

If the irreducible representations for G_0 are labelled $\mathcal{R}_{0,r}$, with $\bar{\mathcal{R}}_{0,r} = \mathcal{R}_{0,\bar{r}}$, with χ_r the corresponding characters and $\mathcal{C}_s$ are the conjugation classes for G_0, with $r, s = 1, \dots, n_0$, then for G the irreducible anti-unitary representations are then $\mathcal{R}_{G,j}$ with associated characters $\chi_{G,j}$ and the extended conjugation classes $\bar{\mathcal{C}}_{G,k}$ are then determined in terms of these

$$
\bar{\mathcal{C}}_{G,k} = \begin{cases} \mathcal{C}_s \,, & h \in \mathcal{C}_s, \ a^{-1}h^{-1}a \in \mathcal{C}_{\bar{s}} \,, \ \mathcal{C}_s = \mathcal{C}_{\bar{s}} \,, \\ \mathcal{C}_s \cup \mathcal{C}_{\bar{s}} \,, & h \in \mathcal{C}_s, \ a^{-1}h^{-1}a \in \mathcal{C}_{\bar{s}} \,, \ \mathcal{C}_s \neq \mathcal{C}_{\bar{s}} \,, \end{cases}
$$

$$
\chi_{G,j}(h)\big|_{h \in \bar{\mathcal{C}}_{G,k}} \equiv \chi_{G,j}(\bar{\mathcal{C}}_{G,k}) = \begin{cases} \chi_r(\mathcal{C}_s) \,, \ 2\,\chi_r(\mathcal{C}_s) \,, & \text{I, II} \,, \\ \chi_r(\mathcal{C}_s) + \chi_{\bar{r}}(\mathcal{C}_s) \,, \ \chi_{\bar{r}}(\mathcal{C}_s) = \chi_r(\mathcal{C}_{\bar{s}}) \,, & \text{III} \,. \end{cases} \tag{2.181}
$$

The number of irreducible representations for G is equal to the number of extended conjugacy classes so that $j, k = 1, \dots, \bar{n}$. In general $n_0 - \bar{n} = \bar{n}_{\text{III}}$. For IIIi the anti-unitary characters are real. From the orthogonality relations for characters of G_0 (2.56) we may directly obtain

$$
\frac{1}{|G_0|} \sum_{h \in G_0} \chi_{G,j'}(h)^* \chi_{G,j}(h) = \frac{1}{|G_0|} \sum_{k=1}^{\bar{n}} \bar{d}_k \, \chi_{G,j'}(\bar{\mathcal{C}}_{G,k})^* \chi_{G,j}(\bar{\mathcal{C}}_{G,k}) = t_j \, \delta_{j'j} \,, \quad \bar{d}_k = \dim \bar{\mathcal{C}}_{G,k} \,, \tag{2.182}
$$

with $t_j = 1, 4, 2$ according to whether the representation $\mathcal{R}_{G,j}$ is of type I, II, III. Since the number of extended conjugacy classes is equal to the number of anti-unitary representations

$$
\sum_{j=1}^{\bar{n}} \frac{1}{t_j} \chi_{G,j}(\bar{\mathcal{C}}_{G,k}) \, \chi_{G,j}(\bar{\mathcal{C}}_{G,k'})^* = \frac{|G_0|}{\bar{d}_k} \, \delta_{k'k} \,. \tag{2.183}
$$

As a special case

$$
\sum_{j=1}^{\bar{n}} \frac{1}{t_j} \, (\dim \mathcal{R}_{G,j})^2 = |G_0| \,. \tag{2.184}
$$

As was described by Dyson[23] some problems can be circumvented by adopting a real form for the representation of G. Starting from a complex co-representation of G of dimension n satisfying (2.159), (2.160) then corresponding real matrices D_R are obtained by taking, for any $h \in G_0$, $a \in G_1$,

$$
\begin{aligned} D(h) = a(h) + i\, b(h) \in U(n) \quad &\to \quad D_R(h) = a(h)\, \mathbb{1}_2 + b(h)\, J \in O(2n, \mathbb{R}) \,, \\ D(a) = c(a) + i\, d(a) \in U(n) \quad &\to \quad D_R(a) = \big(c(a)\, \mathbb{1}_2 + d(a)\, J\big)\theta \in O(2n, \mathbb{R}) \,, \end{aligned} \tag{2.185}
$$

where

$$
J = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right) , \qquad \theta = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix} \right) , \qquad \theta J = -J\theta \,. \tag{2.186}
$$

$\{D_R(h), D_R(a)\}$ then form a standard representation of G and $2\, Re\, \chi(h) = \mathrm{tr}\big(D_R(h)\big)$ for $h \in G_0$. The conjugate representation $D(g)^* \to \theta D(g) \theta$ and so has an equivalent real representation. The real conjugacy classes for any $g \in G$ are then $\mathcal{C}_R(g) = \mathcal{C}_G(g)$ if $g^{-1} \in \mathcal{C}_G(g)$, $\mathcal{C}_G(g) \cup \mathcal{C}_G(g^{-1})$ otherwise.

[23] Freeman Dyson, 1923-2020, Britiah then American.

For the three cases described above

$$
\begin{aligned}
\mathrm{I}:&\quad D_R(h) = M_R(h)\,, \quad D_R(a) = N_R(a)\,, \quad n = d_0\,,\\
\mathrm{II}:&\quad D_R(h) = M_R(h)\otimes \mathbb{1}_2\,, \quad D_R(a) = N_R(a)\otimes J\,, \quad n = 2d_0\,,\\
\mathrm{III}:&\quad D_R(h) = M_R(h)\otimes P_+ + \bar M_R(h)\otimes P_-\,,\\
&\quad D_R(a) = N_R(a)\otimes P_+J - \bar N_R(a)\otimes P_-J\,, \quad P_\pm = \tfrac{1}{2}(\mathbb{1}_2 \pm \theta)\,, \quad n = 2d_0\,,
\end{aligned} \tag{2.187}
$$

where case I has been reduced to a single irreducible component and the real forms are obtained from (2.163) just as in (2.185). The representations in each case are characterised by the algebra of matrices commuting with $D_R(h)$, $D_R(a)$ for all h, a. For $\{M(h)\}$, $\{\bar M(h)\}$ forming irreducible representations of G_0 these have a real basis $\mathbb{1}_{d_0} e_\mu$ where

$$
\begin{aligned}
&\mathrm{I}:\ e_0 = \mathbb{1}_2\,, \qquad \mathrm{II}:\ e_0 = \mathbb{1}_2\otimes\mathbb{1}_2\,,\ e_1 = J\otimes\theta\,,\ e_2 = \mathbb{1}_2\otimes J\,,\ \ e_3 = J\otimes \theta J\,,\\
&\mathrm{III}:\ e_0 = \mathbb{1}_2\otimes\mathbb{1}_2\,,\ e_1 = J\otimes\theta\,,
\end{aligned} \tag{2.188}
$$

where ${e_0}^2 = e_0$ and in case II $e_i e_j = -\delta_{ij}\, e_0 + \epsilon_{ijk}\, e_k$ for $i,j,k = 1,2,3$, the algebra of quaternions, while for case III then ${e_1}^2 = -e_0$ the algebra corresponding to complex numbers. Only in case I is the usual form of Schur's lemma for irreducible representations applicable.[24] For case III although $\{M(h)\}$, $\{\bar M(h)\}$ define inequivalent representations of G_0 the corresponding real representations $\{M_R(h)\}$, $\{\bar M_R(h)\}$ may or may not be equivalent under real conjugation.

Commuting with $\mathbb{1}_{d_0} e_\mu$ does not uniquely characterise the representation matrices in (2.187) in each case. It is then necessary to augment the basic group G so that

$$
G \to \tilde G = \big\{ j^r G : r = 0,1,2,3,\ j^4 = e,\ hj = jh,\ h\in G_0,\ aj = j^3 a,\ a \in G_1\big\}\,, \quad |\tilde G| = 4|G|\,. \tag{2.189}
$$

Correspondingly $\tilde G_0 \simeq \mathbb{Z}_4 \times G_0$. The representation matrices for G are extended to $\tilde G$ by taking $D_R(j^r g) = \tilde J^r D_R(g)$ with $\tilde J = J$ or $J\otimes \mathbb{1}_2$ according to case I or cases II, III.

Just as in section 2.4 for two representations $\mathcal{R}, \mathcal{R}'$ with dimensions $d_\mathcal{R}, d_{\mathcal{R}'}$ respectively we may define in terms of the real representation matrices as in (2.187)

$$
S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = \frac{1}{|\tilde G|} \sum_{\tilde g\in\tilde G} D^{(\mathcal{R}')}_{R,rv}(\tilde g^{-1}) D^{(\mathcal{R})}_{R,us}(\tilde g)\,, \tag{2.190}
$$

so that as shown in (2.49) the two sets of real representation matrices for $\mathcal{R}, \mathcal{R}'$ are linked by $S^{(\mathcal{R}',\mathcal{R})}_{rt,uv} D^{(\mathcal{R})}_{R,ts}(\tilde g) = D^{(\mathcal{R}')}_{R,rt}(\tilde g)\, S^{(\mathcal{R}',\mathcal{R})}_{ts,uv}$ and $S^{(\mathcal{R}',\mathcal{R})}_{rs,uw} D^{(\mathcal{R}')}_{R,wv}(\tilde g) = D^{(\mathcal{R})}_{R,uw}(\tilde g)\, S^{(\mathcal{R}',\mathcal{R})}_{rs,wv}$. For $\mathcal{R}, \mathcal{R}'$ irreducible it is then necessary that $S^{(\mathcal{R}',\mathcal{R})}_{rs,uw} = 0$ for $\mathcal{R} \ne \mathcal{R}'$ but in this case instead of (2.50)

$$
S^{(\mathcal{R}',\mathcal{R})}_{rs,uv} = \delta_{\mathcal{R}\mathcal{R}'} \sum_{\mu,\nu} c_\mathcal{R}^{\mu\nu}\, (\mathbb{1}_{d_\mathcal{R}} e_\mu)_{rs}\, (\mathbb{1}_{d_\mathcal{R}} e_\nu)_{uv}\,. \tag{2.191}
$$

[24] The real matrices which commute with $D_R(h)$, $D_R(a)$ for all h, a form a real vector space which is closed under matrix multiplication. These may be restricted to form a division algebra where every non zero element has an inverse. A theorem due to Frobenius says that every such real division algebra has a basis in terms of real e_μ which satisfy the conditions as above for $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. This provides an alternative characterisation of the three possible anti-unitary representations.

It is easy to see that

$$S^{(\mathcal{R},\mathcal{R})}_{rs,uv}(\mathbb{1}_{d_\mathcal{R}}e_\omega)_{vu} = (\mathbb{1}_{d_\mathcal{R}}e_\omega)_{rs} \quad \Rightarrow \quad c_\mathcal{R}^{\mu\nu}\eta_{\nu\omega} = \frac{1}{t_\mathcal{R}\, d_\mathcal{R}}\delta^\mu{}_\omega \ \text{ for } \ \mathrm{tr}(e_\mu e_\nu) = c_R\,\eta_{\mu\nu}\,, \tag{2.192}$$

where, taking $c_\mathcal{R} = 2$ or $t_\mathcal{R} = 4$ according to whether the representation is of type I or type II,III, $\eta_{\mu\nu}$ is diagonal with $\eta_{00} = 1$, $\eta_{ij} = -\delta_{ij}$. Hence $c_\mathcal{R}^{\mu\nu} = \eta^{\mu\nu}/c_\mathcal{R}d_\mathcal{R}$ with $\eta^{\mu\nu}$ the inverse of $\eta_{\mu\nu}$. For orthogonal matrices $D^{(\mathcal{R})}_{R,rv}(\tilde g^{-1}) = D^{(\mathcal{R})}_{R,vr}(\tilde g)$ and the result (2.191) can be re-expressed by summing over $\tilde G/G$ as

$$\frac{1}{2|G|}\sum_{g\in G}\Big(D^{(\mathcal{R})}_{R,vr}(g)\,D^{(\mathcal{R})}_{R,us}(g) + (\tilde J D^{(\mathcal{R})}_{R}(g))_{vr}\,(\tilde J D^{(\mathcal{R})}_{R}(g))_{us}\Big) = \sum_{\mu,\nu} c_\mathcal{R}^{\mu\nu}\,(\mathbb{1}_{d_\mathcal{R}}e_\mu)_{rs}\,(\mathbb{1}_{d_\mathcal{R}}e_\nu)_{uv}\,. \tag{2.193}$$

By contracting indices

$$\begin{aligned}\frac{1}{2|G|}\sum_{g\in G}\Big(D^{(\mathcal{R})}_{R}(g)D^{(\mathcal{R})}_{R}(g) + \tilde J D^{(\mathcal{R})}_{R}(g)\tilde J D^{(\mathcal{R})}_{R}(g)\Big) &= \frac{1}{|G|}\sum_{a\in G_1} D^{(\mathcal{R})}_{R}(a^2)\\ &= \frac{1}{c_\mathcal{R}\,d_\mathcal{R}}\,\eta^{\mu\nu}\,\mathbb{1}_{d_\mathcal{R}}{e_\mu}^T e_\nu\,,\end{aligned} \tag{2.194}$$

since $D^{(\mathcal{R})}_{R}(g)\tilde J = \pm\tilde J D^{(\mathcal{R})}_{R}(g)$ for $g = h, a$. Using $\mathrm{tr}({e_\mu}^T e_\nu) = c_\mathcal{R}\,\delta_{\mu\nu}$

$$\bar{\mathcal{F}}_\mathcal{R} = \frac{1}{|G|}\sum_{a\in G_1}\mathrm{tr}\big(D^{(\mathcal{R})}_{R}(a^2)\big) = \frac{1}{|G_0|}\sum_{a\in G_1} Re\,\chi_\mathcal{R}(a^2) = \eta^{\mu\nu}\delta_{\mu\nu} = \begin{cases} 1\,, & \mathrm{I}\,,\\ -2\,, & \mathrm{II}\,,\\ 0\,, & \mathrm{III}\,.\end{cases} \tag{2.195}$$

This result extends the Frobenius-Schur indicator as in (2.63) to anti-unitary representations. The cases I, III, II may be labelled by $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ according to the different forms of the commutator algebra with the bases given by (2.188).

By taking traces in (2.193)

$$\frac{1}{|G_0|}\sum_{h\in G_0}\Big(\big(\mathrm{tr}\,a^{(\mathcal{R})}(h)\big)^2 + \big(\mathrm{tr}\,b^{(\mathcal{R})}(h)\big)^2\Big) = \frac{1}{|G_0|}\sum_{h\in G_0}\chi_\mathcal{R}(h)^*\chi_\mathcal{R}(h) = \eta^{\mu\nu}\eta_{\mu\nu}\,, \tag{2.196}$$

so that

$$\langle\chi_\mathcal{R},\chi_\mathcal{R}\rangle = t_\mathcal{R} = \begin{cases} 1\,, & \mathrm{I}\,,\\ 4\,, & \mathrm{II}\,,\\ 2\,, & \mathrm{III}\,.\end{cases} \tag{2.197}$$

This is as expected from the usual orthogonality relations for characters of irreducible representations of G_0 since for the three possible cases $\chi_{\mathcal{R},\mathrm{I}} = \chi_M$, $\chi_{\mathcal{R},\mathrm{II}} = 2\chi_M$ and $\chi_{\mathcal{R},\mathrm{III}} = \chi_M + \chi_{\bar M}$. The result is equivalent to (2.182). For any irreducible representation then

$$\bar{\mathcal{F}}_\mathcal{R} + \langle\chi_\mathcal{R},\chi_\mathcal{R}\rangle = \bar{\mathcal{F}}_\mathcal{R} + t_\mathcal{R} = 2\,. \tag{2.198}$$

Various different possibilities for anti-unitary representations can be illustrated by a range of different examples which are given in terms of tables where the characters for each extended conjugacy class are given for G_0. For $\{M(h)\}$ an irreducible complex representation of G_0 of dimension n there is an associated irreducible $2n$-dimensional epresentation of $G_0 \times \mathbb{Z}_2$ given by taking

$$D(h) = \begin{pmatrix} M(h) & 0 \\ 0 & M(h)^* \end{pmatrix}, \qquad D_T = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{2.199}$$

This corresponds to case III. If $\{\pm\mathbb{1}\} \subset \{M(h)\}$ then there is an anti-unitary representation for

$$D(h) = \begin{pmatrix} M(h) & 0 \\ 0 & M(h)^* \end{pmatrix}, \qquad D_T = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \tag{2.200}$$

corresponding to the central product $G = (G_0 \times \mathbb{Z}_4)/\mathbb{Z}_2$.

The results described below are all derived from the character tables in 2.7. We also list for each representation of G_0 whether it is real $\mathbb{R}$, pseudo-real $\mathbb{H}$ or complex $\mathbb{C}$.

For $G = \mathbb{Z}_n \times \mathbb{Z}_2$, $G_0 = \mathbb{Z}_n$,

$\mathbb{Z}_n$ n odd	$\bar{\mathcal{C}}_1$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{F}}$
$\mathcal{R}_1$, $\mathbb{R}$	1	1	1
$\mathcal{R}_{2,k}$, $\mathbb{C}$	2	$2\cos 2kr\pi/n$	0

$$k, r = 1, \dots, \tfrac{1}{2}(n-1)\,, \tag{2.201}$$

$\mathbb{Z}_n$ n even	$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$, $\mathbb{R}$	1	1	1	1
$\mathcal{R}_{2,k}$, $\mathbb{C}$	2	$2\cos 2kr\pi/n$	$2(-1)^k$	0
$\mathcal{R}_{1,2}$, $\mathbb{R}$	1	$(-1)^r$	$(-1)^{\frac{1}{2}n}$	1

$$k, r = 1, \dots, \tfrac{1}{2}n - 1\,, \tag{2.202}$$

For $G = \mathbb{Z}_{4n}$, $G_0 = \mathbb{Z}_{2n}$,

$\mathbb{Z}_{2n}$	$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$, $\mathbb{R}$	1	1	1	1
$\mathcal{R}_{2,k}$, $\mathbb{C}$	2	$2\cos kr\pi/n$	$2(-1)^k$	0
$\mathcal{R}_{1,2}$, $\mathbb{R}$	2	$2(-1)^r$	$2(-1)^n$	-2

$$k, r = 1, \dots, n-1\,. \tag{2.203}$$

For $G = D_n$, $G_0 = \mathbb{Z}_n$ the irreducible representations are all one dimensional,

$\mathbb{Z}_n$ n odd	$\bar{\mathcal{C}}_{1,r}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,k}$, $\begin{smallmatrix}\mathbb{R} & k=0 \\ \mathbb{C} & k\neq 0\end{smallmatrix}$	$\exp 2kr\pi\, i/n$	1

$$k, r = 0, \dots, n-1\,, \tag{2.204}$$

$\mathbb{Z}_n$ n even	$\bar{\mathcal{C}}_{1,r}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,k}$, $\begin{smallmatrix}\mathbb{R} & k=0,\frac{1}{2}n \\ \mathbb{C} & k\neq 0,\frac{1}{2}n\end{smallmatrix}$	$\exp 2kr\pi\, i/n$	1

$$k, r = 0, \dots, n-1\,. \tag{2.205}$$

For $G = D_{4n}$, $G_0 = D_{2n}$,

$\mathbb{D}_{2n}$	$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{2,n}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$, $\mathbb{R}$	1	1	1	1	1
$\mathcal{R}_{1,2}$, $\mathbb{R}$	1	1	1	-1	1
$\mathcal{R}_{2,k}$, $\mathbb{R}$	2	$2(-1)^k$	$2\cos kr\pi/n$	0	1
$\mathcal{R}_{2,n}$, $\mathbb{R}$	2	$2(-1)^n$	$2(-1)^r$	0	0

$$, \qquad k,r = 1,\dots,n-1\,. \tag{2.206}$$

For $G = \mathrm{Q}_{4n}$, $G_0 = \mathbb{Z}_{2n}$ the conjugacy classes for $\mathbb{Z}_{2n}$ are not extended and

$\mathbb{Z}_{2n}$	$\bar{\mathcal{C}}_{1,r}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,k}$, $\mathbb{C}$ $k\neq 0,\frac{1}{2}n$; $\mathbb{R}$ $k=0,\frac{1}{2}n$	$\exp 2kr\pi\, i/n$	1
$\mathcal{R}_{2,k}$, $\mathbb{C}$ $k\neq \frac{1}{2}(n-1)$; $\mathbb{R}$ $k=\frac{1}{2}(n-1)$	$2\exp(2k+1)r\pi\, i/n$	-2

$$, \qquad \begin{array}{c} r = 0,\dots,2n-1\,, \\ k = 0,\dots,n-1\,. \end{array} \tag{2.207}$$

In this case $\bar{\mathcal{F}}_{\mathcal{R}} = \chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,n})$.

For other examples we may consider assuming quaternionic groups for G_0. For each quaternion group G_{Q} described in 1.5 there is a corresponding faithful irreducible two dimensional representation $\{M(h)\}$ for all $h \in G_{\mathrm{Q}}$ obtained by using the representation (1.115) for the quaternions. There is then a four dimensional anti-unitary representation for $G = G_{\mathrm{Q}} \times \mathbb{Z}_2$, $G_0 = G_{\mathrm{Q}}$ obtained by taking

$$D(h) = \begin{pmatrix} M(h) & 0 \\ 0 & M(h) \end{pmatrix}, \qquad D_T = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}, \tag{2.208}$$

which generically corresponds to case II. If $\{\pm 1\} \subset G_{\mathrm{Q}}$ then for $G = (G_{\mathrm{Q}} \times \mathbb{Z}_4)/\mathbb{Z}_2$, $G_0 = G_{\mathrm{Q}}$, there is a two dimensional anti-unitary representation given by

$$D(h) = M(h)\,, \qquad D_T = J\,. \tag{2.209}$$

Clearly now ${D_T}^2 = -\mathbb{1}_2$.

For $G = \mathrm{Q}_{4n} \times \mathbb{Z}_2$ and $G_0 = \mathrm{Q}_{4n}$ we may use (2.208)

Q_{4n} n odd	$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{2n}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$, $\mathbb{R}$	1	1	1	1	1
$\mathcal{R}_{1,2}$, $\mathbb{R}$	1	1	1	-1	1
$\mathcal{R}_{2,n}$, $\mathbb{C}$	2	-2	$2(-1)^r$	0	0
$\mathcal{R}_{2,k}$, $\mathbb{R}$	2	2	$2\cos 2kr\pi/n$	0	1
$\mathcal{R}_{4,k}$, $\mathbb{H}$	4	-4	$4\cos(2k-1)r\pi/n$	0	-2

$$, \qquad \begin{array}{c} r = 1,\dots,n-1\,, \\ k = 1,\dots,\frac{1}{2}(n-1)\,, \end{array} \tag{2.210}$$

Q_{4n} n even		$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{n,1}$	$\bar{\mathcal{C}}_{n,2}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$,	$\mathbb{R}$	1	1	1	1	1	1
$\mathcal{R}_{1,2}$,	$\mathbb{R}$	1	1	1	-1	-1	1
$\mathcal{R}_{1,3}$,	$\mathbb{R}$	1	1	$(-1)^r$	1	-1	1
$\mathcal{R}_{1,4}$,	$\mathbb{R}$	1	1	$(-1)^r$	-1	1	1
$\mathcal{R}_{2,k}$,	$\mathbb{R}$	2	2	$2\cos 2kr\pi/n$	0	0	1
$\mathcal{R}_{4,k}$,	$\mathbb{H}$	4	-4	$4\cos(2k-1)r\pi/n$	0	0	-2

$$, \quad \begin{array}{c} r = 1, \dots, n-1\,, \\ k = 1, \dots, \tfrac{1}{2}(n-2)\,. \end{array} \tag{2.211}$$

For the n odd case $4n\,\bar{\mathcal{F}}_{\mathcal{R}} = 2\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,1}) + 2n\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,2}) + 4\sum_{r=1}^{\frac{1}{2}(n-1)} \chi_{\mathcal{R}}(\bar{\mathcal{C}}_{2,2r})$ while for n even $4n\,\bar{\mathcal{F}}_{\mathcal{R}} = 2\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,1}) + 2(n+1)\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,2}) + 4\sum_{r=1}^{\frac{1}{2}(n-2)} \chi_{\mathcal{R}}(\bar{\mathcal{C}}_{2,2r})$.

For $G = \mathrm{Q}_{4n} \times \mathbb{Z}_4/\mathbb{Z}_2$ and $G_0 = \mathrm{Q}_{4n}$ we may use (2.209)

Q_{4n} n odd		$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{2n}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$,	$\mathbb{R}$	1	1	1	1	1
$\mathcal{R}_{1,2}$,	$\mathbb{R}$	1	1	1	-1	1
$\mathcal{R}_{2,n}$,	$\mathbb{C}$	2	-2	$2(-1)^r$	0	0
$\mathcal{R}_{2,k}$,	$\mathbb{H}$ k odd $\mathbb{R}$ k even	2	$2(-1)^k$	$2\cos kr\pi/n$	0	1

$$, \quad k, r = 1, \dots, n-1\,, \tag{2.212}$$

Q_{4n} n even		$\bar{\mathcal{C}}_{1,1}$	$\bar{\mathcal{C}}_{1,2}$	$\bar{\mathcal{C}}_{2,r}$	$\bar{\mathcal{C}}_{n,1}$	$\bar{\mathcal{C}}_{n,2}$	$\bar{\mathcal{F}}$
$\mathcal{R}_{1,1}$,	$\mathbb{R}$	1	1	1	1	1	1
$\mathcal{R}_{1,2}$,	$\mathbb{R}$	1	1	1	-1	-1	1
$\mathcal{R}_{1,3}$,	$\mathbb{R}$	1	1	$(-1)^r$	1	-1	1
$\mathcal{R}_{1,4}$,	$\mathbb{R}$	1	1	$(-1)^r$	-1	1	1
$\mathcal{R}_{2,k}$,	$\mathbb{H}$ k odd $\mathbb{R}$ k even	2	$2(-1)^k$	$2\cos kr\pi/n$	0	0	1

$$, \quad k, r = 1, \dots, n-1\,. \tag{2.213}$$

For the n odd case $4n\,\bar{\mathcal{F}}_{\mathcal{R}} = 2n\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,1}) + 2\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,2}) + 4\sum_{r=1}^{\frac{1}{2}(n-1)} \chi_{\mathcal{R}}(\bar{\mathcal{C}}_{2,2r-1})$ while for n even $4n\,\bar{\mathcal{F}}_{\mathcal{R}} = 2(n+1)\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,1}) + 2\,\chi_{\mathcal{R}}(\bar{\mathcal{C}}_{1,2}) + 4\sum_{r=1}^{\frac{1}{2}(n-2)} \chi_{\mathcal{R}}(\bar{\mathcal{C}}_{2,2r-1})$.

These various examples illustrate some of the possibilities where the different cases, I, II and IIIi, IIIii, of anti-unitary representations for G can be combined with real, pseudo-real or complex representations of G_0. There are ten possibilities altogether (eight are given in the above tables) since case IIIii for G requires a complex representation for G_0.

3 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 $|\lambda|^2 = 1$.

3.1 Three Dimensional Rotations

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

$$R^T R = 1\!\!1_3 \quad \Rightarrow \quad (\det R)^2 = 1\,. \tag{3.1}$$

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

$$SRS^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}. \tag{3.2}$$

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

$$\operatorname{tr} R = 2\cos\theta + 1\,. \tag{3.3}$$

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

$$\mathbf{x} \underset{R}{\to} \mathbf{x}' = \mathbf{x}^R\,, \tag{3.4}$$

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

$$x'_i = R_{ij} x_j\,, \tag{3.5}$$

For $\det R = -1$ the transformation involves a reflection. Of course rotations preserve scalar products and vector products up to a sign

$$\mathbf{x}^R \cdot \mathbf{y}^R = \mathbf{x} \cdot \mathbf{y}\,, \qquad \mathbf{x}^R \times \mathbf{y}^R = \det R\, (\mathbf{x} \times \mathbf{y})^R\,. \tag{3.6}$$

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

$$R_{ij}(\theta, \mathbf{n}) = \cos\theta\, \delta_{ij} + (1 - \cos\theta) n_i n_j - \sin\theta\, \varepsilon_{ijk} n_k\,, \tag{3.7}$$

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

$$\mathbf{n} \in S^2\,, \quad 0 \le \theta \le \pi\,, \qquad (\pi, \mathbf{n}) \simeq (\pi, -\mathbf{n})\,, \tag{3.8}$$

with S^2 the two-dimensional unit sphere. The space corresponding to (3.8) is then the ball of radius π with opposite points on the surface identified. Equivalently the range of the angle θ can be extended to 0 to 2π if

$$R(\theta, \mathbf{n}) = R(2\pi - \theta, -\mathbf{n}) = R(\theta + 2\pi, \mathbf{n})\,. \tag{3.9}$$

In (3.4) this requires

$$\mathbf{x}^{R(\theta,\mathbf{n})} = \cos\theta\, \mathbf{x} + (1 - \cos\theta)\, \mathbf{n}\, \mathbf{n}\cdot\mathbf{x} + \sin\theta\, \mathbf{n}\times\mathbf{x}\,. \tag{3.10}$$

For an arbitrary $R \in SO(3)$ then taking $\mathbf{x} \to \mathbf{x}^{R^{-1}}$ gives

$$((\mathbf{x}^{R^{-1}})^{R(\theta,\mathbf{n})})^R = \mathbf{x}^{RR(\theta,\mathbf{n})R^{-1}} = \cos\theta\, \mathbf{x} + (1 - \cos\theta)\, \mathbf{n}^R\, \mathbf{n}^R\cdot\mathbf{x} + \sin\theta\, \mathbf{n}^R\times\mathbf{x}\,. \tag{3.11}$$

since $\mathbf{n}^R\cdot\mathbf{x} = \mathbf{n}\cdot\mathbf{x}^{R^{-1}}$ and $\mathbf{n}^R\times\mathbf{x} = (\mathbf{n}\times\mathbf{x}^{R^{-1}})^R$. Hence

$$R\,R(\theta, \mathbf{n})R^{-1} = R(\theta, \mathbf{n}^R)\,, \tag{3.12}$$

so that all rotations with the same θ belong to a single conjugacy class.

The additional transformation

$$\mathbf{x}' = \mathbf{x} - 2\,\mathbf{m}\,\mathbf{m}\cdot\mathbf{x}\,, \quad \mathbf{m}^2 = 1 \quad \Rightarrow \quad \mathbf{x}'^2 = \mathbf{x}^2\,, \tag{3.13}$$

corresponds to a reflection through the plane perpendicular to $\mathbf{m}$. The associated matrix given by

$$R_{ij} = \delta_{ij} - 2\,m_i\,m_j\,, \quad R_{ik}R_{kj} = \delta_{ij}\,, \tag{3.14}$$

has eigenvalues $1, 1, -1$ and belongs to $O(3)$. Arbitrary elements of $O(3)$ can be obtained by combining rotations and reflections. For two reflections defined by unit vectors $\mathbf{m}$ and then $\mathbf{l}$

$$\begin{aligned}\mathbf{x}' &= \mathbf{x} - 2\,\mathbf{l}\,\mathbf{l}\cdot\mathbf{x} - 2\,\mathbf{m}\;\mathbf{m}\cdot\mathbf{x} + 4\,\mathbf{l}\;\mathbf{l}\cdot\mathbf{m}\,\mathbf{m}\cdot\mathbf{x}\\ &= \big(2\,(\mathbf{l}\cdot\mathbf{m})^2 - 1\big)\mathbf{x} + 2\,(\mathbf{l}\times\mathbf{m})\,(\mathbf{l}\times\mathbf{m})\cdot\mathbf{x} + 2\,\mathbf{l}\cdot\mathbf{m}\,(\mathbf{l}\times\mathbf{m})\times\mathbf{x}\,,\end{aligned} \tag{3.15}$$

which, by comparing with (3.10), if $\mathbf{l}\cdot\mathbf{m} = \cos\phi$ corresponds to a rotation through an angle 2ϕ and about an axis $\mathbf{n}$ where $\mathbf{l}\times\mathbf{m} = \sin\phi\,\mathbf{n}$. Any rotation $R(2\phi, \mathbf{n})$ can be expressed as a product of two reflections by choosing a unit vector $\mathbf{m}$, $\mathbf{m}\cdot\mathbf{n} = 0$, and then defining $\mathbf{l} = \cos\phi\,\mathbf{m} + \sin\phi\,(\mathbf{m}\times\mathbf{n})$, $\mathbf{l}^2 = 1$.

For an infinitesimal rotation $R(\delta\theta, \mathbf{n})$ acting on a vector $\mathbf{x}$ and using standard vector notation we then have

$$\mathbf{x} \xrightarrow[R(\delta\theta,\mathbf{n})]{} \mathbf{x}' = \mathbf{x} + \delta\theta\,\mathbf{n}\times\mathbf{x}\,. \tag{3.16}$$

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

Although the group $SO(3)$ is infinite dimensional the notion of a sum over group elements for a finite group can be extended to an integration over the three dimensional unit ball. The crucial property (1.7) can be extended by requiring

$$\int \mathrm{d}\mu(\theta,\mathbf{n})\, f\big(R(\theta,\mathbf{n})\big) = \int \mathrm{d}\mu(\theta,\mathbf{n})\, f\big(R(\theta,\mathbf{n})R\big) \quad \text{for all} \quad R \in SO(3)\,. \tag{3.17}$$

This is satisfied, as shown later, by taking

$$\mathrm{d}\mu(\theta,\mathbf{n}) = \mathrm{d}\Omega_{\mathbf{n}}\, \mathrm{d}\theta \sin^2 \tfrac{1}{2}\theta, \qquad \mathbf{n} \in S^2\,,\ 0 \le \theta \le \pi\,. \tag{3.18}$$

The group volume $\mathcal{V}_{SO(3)} = \int_{SO(3)} \mathrm{d}\mu(\theta,\mathbf{n}) = 2\pi^2$.

3.2 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 $\{\mathbb{1},-\mathbb{1}\}$, is of crucial importance in understanding the role of spinors under rotations. To demonstrate this we introduce the standard *Pauli*[25] *matrices*, a set of three 2×2 matrices which have the explicit form

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.19}$$

These matrices satisfy the algebraic relations

$$\sigma_i \sigma_j = \delta_{ij}\, \mathbb{1}_2 + i\, \epsilon_{ijk} \sigma_k\,, \tag{3.20}$$

and also are traceless and hermitian. The matrices $(\mathbb{1}_2, i\sigma_3, i\sigma_2, i\sigma_1)$ provide a two dimensional complex representation of of the quaternion algebra (1.79) identical to (1.115). Adopting a vector notation $\boldsymbol{\sigma} = (\sigma_1,\sigma_2,\sigma_3)$, so that (3.20) is equivalent to $\mathbf{a}\cdot\boldsymbol{\sigma}\ \mathbf{b}\cdot\boldsymbol{\sigma} = \mathbf{a}\cdot\mathbf{b}\,\mathbb{1} + i\,\mathbf{a}\times\mathbf{b}\cdot\boldsymbol{\sigma}$, we have

$$\boldsymbol{\sigma}^\dagger = \boldsymbol{\sigma}\,, \qquad \mathrm{tr}(\boldsymbol{\sigma}) = 0\,. \tag{3.21}$$

Using (3.20) then gives

$$\mathrm{tr}(\sigma_i\sigma_j) = 2\delta_{ij}\,, \tag{3.22}$$

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

$$A = \tfrac{1}{2}\mathrm{tr}(A)\,\mathbb{1} + \tfrac{1}{2}\mathrm{tr}(\boldsymbol{\sigma}A)\cdot\boldsymbol{\sigma}\,, \tag{3.23}$$

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

$$\mathbf{x} \to \mathbf{x}\cdot\boldsymbol{\sigma} = (\mathbf{x}\cdot\boldsymbol{\sigma})^\dagger\,, \qquad \mathbf{x} = \tfrac{1}{2}\mathrm{tr}(\boldsymbol{\sigma}\,\mathbf{x}\cdot\boldsymbol{\sigma})\,, \tag{3.24}$$

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

$$(\mathbf{x}\cdot\boldsymbol{\sigma})^2 = \mathbf{x}^2\mathbb{1}\,. \tag{3.25}$$

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

From (3.25) and (3.21) the eigenvalues of $\mathbf{x}\cdot\boldsymbol{\sigma}$ must be $\pm\sqrt{\mathbf{x}^2}$ and in consequence we have

$$\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = -\mathbf{x}^2 \,. \tag{3.26}$$

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

$$\mathbf{x}'\cdot\boldsymbol{\sigma} = A\,\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger \,, \tag{3.27}$$

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

$$AA^\dagger = \mathbb{1} \,. \tag{3.28}$$

With, $\mathbf{x}'$ defined by (3.27) and using (3.26),

$$\mathbf{x}'^2 = -\det(\mathbf{x}'\cdot\boldsymbol{\sigma}) = -\det(A\,\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger) = -\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = \mathbf{x}^2 \,, \tag{3.29}$$

using $\det(XY) = \det X \ \det Y$ and from (3.28) $\det A \ \det A^\dagger = 1$. Hence, since this shows that $|\mathbf{x}'| = |\mathbf{x}|$,

$$x'_i = R_{ij}x_j \,, \tag{3.30}$$

with $[R_{ij}]$ an orthogonal matrix. Furthermore since as $A \to \mathbb{1}$, $R_{ij} \to \delta_{ij}$ we must have $\det[R_{ij}] = 1$. Explicitly from (3.27) and (3.22)

$$\sigma_i R_{ij} = A\sigma_j A^\dagger \quad \Rightarrow \quad R_{ij} = \tfrac{1}{2}\,\mathrm{tr}(\sigma_i A\sigma_j A^\dagger) \,. \tag{3.31}$$

To show the converse then from (3.31), using (note $\sigma_j\sigma_i\sigma_j = -\sigma_i$) $\sigma_j A^\dagger \sigma_j = 2\,\mathrm{tr}(A^\dagger)\mathbb{1} - A^\dagger$, we obtain

$$R_{jj} = |\mathrm{tr}(A)|^2 - 1 \,, \qquad \sigma_i R_{ij}\sigma_j = 2\,\mathrm{tr}(A^\dagger)A - \mathbb{1} \,. \tag{3.32}$$

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 (3.32) may be solved for $\mathrm{tr}(A)$ and then A,

$$A = \pm\frac{\mathbb{1} + \sigma_i R_{ij}\sigma_j}{2(1+R_{jj})^{\frac{1}{2}}} \,. \tag{3.33}$$

The arbitrary sign, which cancels in (3.31), ensures that in general $\pm A \leftrightarrow R_{ij}$. This ensures $SO(3) \simeq SU(2)/\mathbb{Z}_2$. Any arbitrary 2×2 matrix A can be expanded in a basis formed by the Pauli matrices and the unit matrix and then

$$A = x_0\,\mathbb{1} - i\,\mathbf{x}\cdot\boldsymbol{\sigma} \in SU(2) \quad \Leftrightarrow \quad {x_0}^2 + \mathbf{x}^2 = 1 \quad \Leftrightarrow \quad (x_0, \mathbf{x}) \in S^3 \,. \tag{3.34}$$

The additional transformation

$$\mathbf{x}'\cdot\boldsymbol{\sigma} = -\mathbf{m}\cdot\boldsymbol{\sigma}\,\mathbf{x}\cdot\boldsymbol{\sigma}\,\mathbf{m}\cdot\boldsymbol{\sigma} \,, \quad \mathbf{m}^2 = 1 \quad \Rightarrow \quad \mathbf{x}' = \mathbf{x} - 2\,\mathbf{m}\,\mathbf{m}\cdot\mathbf{x} \,, \tag{3.35}$$

corresponds to the reflection (3.35).

For a $SO(3)$ rotation R through an infinitesimal angle as in (3.16) then from (3.7)

$$R_{ij} = \delta_{ij} - \delta\theta\,\varepsilon_{ijk}n_k \,, \tag{3.36}$$

and it is easy to obtain for the associated $SU(2)$ matrix ,

$$A_R(\delta\theta,\mathbf{n}) = \mathbb{1} - \tfrac{1}{2}\delta\theta\, i\mathbf{n}\cdot\boldsymbol{\sigma}\,. \tag{3.37}$$

Note that since $\det(\mathbb{1}+X) = 1+\mathrm{tr}\,X$, to first order in X, for any matrix then the tracelessness of the Pauli matrices is necessary for (3.37) to be compatible with $\det A_R = 1$. For a finite rotation angle θ then, with (3.3), (3.32) gives $|\mathrm{tr}(A_R)| = 2|\cos\frac{1}{2}\theta|$ and the matrix A_R can be found by exponentiation, where corresponding to (3.7),

$$A_R(\theta,\mathbf{n}) = e^{-\frac{1}{2}i\theta\,\mathbf{n}\cdot\boldsymbol{\sigma}} = \cos\tfrac{1}{2}\theta\;\mathbb{1} - \sin\tfrac{1}{2}\theta\; i\mathbf{n}\cdot\boldsymbol{\sigma}\,,\;, \qquad \pm A_R(\theta,\mathbf{n}) \to R(\theta,\mathbf{n}) \tag{3.38}$$

The parameters $(\theta,\mathbf{n})$ cover all $SU(2)$ matrices for

$$\mathbf{n}\in S^2\,, \qquad 0\le\theta<2\pi\,, \tag{3.39}$$

in contrast to (3.8). For the matrices in (3.38) $A_R(2\pi,\mathbf{n}) = -\mathbb{1} \in \mathcal{Z}(SU(2))$. For the matrices $\{A_R(\theta,\mathbf{n})\}$ integration can be defined just as in (3.18) with the integration range on θ extended to be from 0 to 2π.

3.2.1 Non Compact Isomorphisms

The relation between the compact groups $SU(2)$ and $SO(3)$ can be extended to related non compact groups. If we define $\tilde{\boldsymbol{\sigma}} = (\sigma_1, i\sigma_2, \sigma_3)$ then the $\tilde{\sigma}_i$ matrices are all real and traceless. Hence for $y_0,\mathbf{y}$ real

$$A = y_0\,\mathbb{1} + \mathbf{y}\cdot\tilde{\boldsymbol{\sigma}} \in Sl(2,\mathbb{R}) \quad\Leftrightarrow\quad {y_0}^2 - {y_1}^2 + {y_2}^2 - {y_3}^2 = 1\,. \tag{3.40}$$

Clearly the parameters have an infinite range, it contains the subgroup corresponding to (1.123) by taking $y_0 = \cosh\theta$, $y_1 = \sinh\theta$, $y_2 = y_3 = 0$. Since

$$\det(\mathbf{x}\cdot\tilde{\boldsymbol{\sigma}}) = -{x_1}^2 + {x_2}^2 - {x_3}^2\,, \tag{3.41}$$

then

$$\mathbf{x}'\cdot\tilde{\boldsymbol{\sigma}} = A\,\mathbf{x}\cdot\tilde{\boldsymbol{\sigma}}A^{-1}\,,\quad A\in Sl(2,\mathbb{R}) \quad\Leftrightarrow\quad -{x_1}^2 + {x_2}^2 - {x_3}^2 = -{x'_1}^2 + {x'_2}^2 - {x'_3}^2\,. \tag{3.42}$$

An alternative non compact group is $SU(1,1)$ which is defined by

$$B^\dagger\sigma_3 B = \sigma_3\,, \qquad \det B = 1\,. \tag{3.43}$$

A basis of generators in terms of the Pauli matrices is $\hat{\boldsymbol{\sigma}} = (\sigma_1,\sigma_2,i\sigma_3)$ where $\sigma_3\,\hat{\boldsymbol{\sigma}} + \hat{\boldsymbol{\sigma}}^\dagger\sigma_3 = 0$. In consequence

$$B = y_0\,\mathbb{1} + \mathbf{y}\cdot\hat{\boldsymbol{\sigma}} \in SU(1,1) \quad\Leftrightarrow\quad {y_0}^2 - {y_1}^2 - {y_2}^2 + {y_3}^2 = 1\,. \tag{3.44}$$

In this case

$$\det(\mathbf{x}\cdot\hat{\boldsymbol{\sigma}}) = -{x_1}^2 - {x_2}^2 + {x_3}^2\,, \tag{3.45}$$

and

$$\mathbf{x}'\cdot\hat{\boldsymbol{\sigma}} = B\,\mathbf{x}\cdot\hat{\boldsymbol{\sigma}}B^{-1}\,,\quad B\in SU(1,1)\qquad\Leftrightarrow\qquad -{x_1}^2-{x_2}^2+{x_3}^2 = -{x'_1}^2-{x'_2}^2+{x'_3}^2\,. \tag{3.46}$$

There is an isomorphism of $Sl(2,\mathbb{R})$ and $SU(1,1)$ by rotating σ_2 into σ_3. This may be achieved for $A(\mathbf{y})\in Sl(2,\mathbb{R})$ then $e^{i\frac{1}{4}\pi\sigma_1}A(\mathbf{y})e^{-i\frac{1}{4}\pi\sigma_1} = B(\mathbf{y}')\in SU(1,1)$, with $\mathbf{y}' = (y_1,y_3,y_2)$. In each case, by similar arguments to previously, we then have

$$Sl(2,\mathbb{R})/\mathbb{Z}_2 \simeq SU(1,1)/\mathbb{Z}_2 \simeq SO(2,1)\,. \tag{3.47}$$

3.3 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 (3.16). 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 the commutator

$$R = {R_2}^{-1}{R_1}^{-1}R_2R_1 = \mathbb{1} + \mathrm{O}(\delta\theta_1\delta\theta_2)\,. \tag{3.48}$$

Acting on a vector $\mathbf{x}$ and using (3.16) and keeping only terms which are $\mathrm{O}(\delta\theta_1\delta\theta_2)$ we find

$$\begin{aligned}\mathbf{x}\underset{R}{\to}\mathbf{x}' &= \mathbf{x} + \delta\theta_1\delta\theta_2\big(\mathbf{n}_2\times(\mathbf{n}_1\times\mathbf{x}) - \mathbf{n}_1\times(\mathbf{n}_2\times\mathbf{x})\big)\\ &= \mathbf{x} + \delta\theta_1\delta\theta_2\,(\mathbf{n}_2\times\mathbf{n}_1)\times\mathbf{x}\,,\end{aligned} \tag{3.49}$$

using standard vector product identities.

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

$$U[R(\delta\theta,\mathbf{n})] = 1 - i\delta\theta\,\mathbf{n}\cdot\mathbf{J}\,, \tag{3.50}$$

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

$$\mathbf{J}^\dagger = \mathbf{J}\,, \tag{3.51}$$

or each J_i is hermitian. If we consider the combined rotations as in (3.48) in conjunction with (3.49) and (3.50) we find

$$\begin{aligned}U[R] &= 1 - i\delta\theta_1\delta\theta_2\,(\mathbf{n}_2\times\mathbf{n}_1)\cdot\mathbf{J}\\ &= U[R_2]^{-1}U[R_1]^{-1}U[R_2]U[R_1]\\ &= 1 - \delta\theta_1\delta\theta_2\,[\mathbf{n}_2\cdot\mathbf{J}\,,\,\mathbf{n}_1\cdot\mathbf{J}]\,,\end{aligned} \tag{3.52}$$

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

$$[\mathbf{n}_2\cdot\mathbf{J}\,,\,\mathbf{n}_1\cdot\mathbf{J}] = i\,(\mathbf{n}_2\times\mathbf{n}_1)\cdot\mathbf{J}\,, \tag{3.53}$$

or equivalently

$$[J_i,\,J_j] = i\,\varepsilon_{ijk}J_k\,. \tag{3.54}$$

Acting on functions of $\mathbf{x}$

$$\mathbf{J} \to \mathbf{L} = -i\,\mathbf{x} \times \nabla\,, \tag{3.55}$$

so that, neglecting $\delta\theta^2$,

$$(1 - i\delta\theta\,\mathbf{n}\cdot\mathbf{L})f(\mathbf{x}) = f(\mathbf{x} - \delta\theta\,\mathbf{n}\times\mathbf{x})\,, \tag{3.56}$$

corresponds to an infinitesimal rotation. It is straightforward to verify that L_i satisfies the commutation relations (3.54).

Although (3.50) expresses U in terms of $\mathbf{J}$ for infinitesimal rotations it can be extended to finite rotations since

$$U[R(\theta,\mathbf{n})] = \exp(-i\theta\,\mathbf{n}\cdot\mathbf{J}) = \lim_{N\to\infty}\left(1 - i\,\frac{\theta}{N}\,\mathbf{n}\cdot\mathbf{J}\right)^N. \tag{3.57}$$

Under rotations $\mathbf{J}$ is a vector. From (3.12), $U[R]U[R(\delta\theta,\mathbf{n})]U[R]^{-1} = U[R(\delta\theta,\mathbf{n}^R)]$ which in turn from (3.50) implies

$$U[R]J_iU[R]^{-1} = (R^{-1})_{ij}J_j\,. \tag{3.58}$$

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 (3.54) 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

$$U[R]HU[R]^{-1} = H \qquad \Leftrightarrow \qquad [\mathbf{J}, H] = 0\,. \tag{3.59}$$

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

3.4 Representations of Angular Momentum Commutation Relations

We here describe how the commutation relations (3.54) 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

$$J_\pm = J_1 \pm iJ_2\,, \tag{3.60}$$

and then (3.54) is equivalent to

$$[J_3, J_\pm] = \pm J_\pm\,, \tag{3.61a}$$

$$[J_+, J_-] = 2J_3\,. \tag{3.61b}$$

The hermeticity conditions (3.51) then become

$$J_+{}^\dagger = J_-\,, \qquad J_3{}^\dagger = J_3\,. \tag{3.62}$$

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

$$J_3|m\rangle = m|m\rangle\,. \tag{3.63}$$

Then from (3.61a) it is easy to see that

$$J_\pm|m\rangle \propto |m \pm 1\rangle \quad \text{or} \quad 0\,, \tag{3.64}$$

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

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

$$J_-|m\rangle = |m-1\rangle\,, \qquad J_+|m\rangle = \lambda_m|m+1\rangle\,, \tag{3.65}$$

and hence

$$J_+J_-|m\rangle = \lambda_{m-1}|m\rangle\,, \qquad J_-J_+|m\rangle = \lambda_m|m\rangle\,. \tag{3.66}$$

By considering $[J_+, J_-]|m\rangle$ we have from (3.61b), if $|m \pm 1\rangle$ are non zero,

$$\lambda_{m-1} - \lambda_m = 2m\,. \tag{3.67}$$

This can be solved for any m by

$$\lambda_m = j(j+1) - m(m+1)\,, \tag{3.68}$$

for some constant written as $j(j+1)$. For sufficiently large positive or negative m we clearly have $\lambda_m < 0$. The hermeticity conditions (3.62) 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

$$J_+|m_{\text{max}}\rangle = 0 \quad \Rightarrow \quad \lambda_{m_{\text{max}}} = (j - m_{\text{max}})(j + m_{\text{max}} + 1) = 0\,, \tag{3.69a}$$

$$J_-|m_{\text{min}}\rangle = 0 \quad \Rightarrow \quad \lambda_{m_{\text{min}}-1} = (j + m_{\text{min}})(j - m_{\text{min}} + 1) = 0\,, \tag{3.69b}$$

where also

$$m_{\text{max}} - m_{\text{min}} = 0, 1, 2, \dots\,. \tag{3.70}$$

Taking $j \geq 0$ the result (3.68) then requires

$$m_{\text{max}} = j\,, \qquad m_{\text{min}} = -j\,. \tag{3.71}$$

For this to be possible we must have

$$j \in \{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots\}\,, \tag{3.72}$$

and then for each value of j

$$m \in \{-j, -j+1, \dots j-1, j\}\,. \tag{3.73}$$

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

3.5 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 (3.72) and (3.73), satisfying

$$\langle j\,m|j\,m'\rangle = \delta_{mm'}\,. \tag{3.74}$$

These are eigenvectors of J_3 as before

$$J_3|j\,m\rangle = m|j\,m\rangle\,. \tag{3.75}$$

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

$$J_+|j\,j\rangle = 0\,. \tag{3.76}$$

A state satisfying both (3.75) and (3.76) is called a highest weight state. In this case the action of $J_\pm$ gives

$$J_\pm|j\,m\rangle = N^\pm_{jm}|j\,m{\pm}1\rangle\,, \tag{3.77}$$

where $N^\pm_{jm}$ are determined by requiring (3.74) to be satisfied. From (3.66) and (3.68) we must then have

$$|N^+_{jm}|^2 = \lambda_m = (j-m)(j+m+1)\,, \qquad |N^-_{jm}|^2 = \lambda_{m-1} = (j+m)(j-m+1)\,. \tag{3.78}$$

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

$$N^\pm_{jm} = \sqrt{(j\mp m)(j\pm m+1)}\,. \tag{3.79}$$

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

$$(J_-)^n|j\,j\rangle = \left(\tfrac{n!\,(2j)!}{(2j-n)!}\right)^{\frac{1}{2}}|j\,j-n\rangle\,, \quad n = 0,1,\dots,2j\,. \tag{3.80}$$

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

$$\mathbf{J}^2 = \begin{cases} J_-J_+ + J_3^2 + J_3\,, \\ J_+J_- + J_3^2 - J_3\,. \end{cases} \tag{3.81}$$

With the first form in (3.81) and using (3.76) we then get acting on the highest weight state

$$\mathbf{J}^2|j\,j\rangle = j(j+1)|j\,j\rangle\,. \tag{3.82}$$

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

$$[\mathbf{J}^2\,,\,J_i] = 0\,, \quad i = 1,2,3\,. \tag{3.83}$$

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

$$\mathbf{J}^2|j\,m\rangle = j(j+1)|j\,m\rangle\,, \tag{3.84}$$

as well as (3.75). Nevertheless we require (3.77), with (3.79), to determine the relative phases of all states.

3.5.1 Action of Time Reversal

The anti-unitary operator T acts on the angular momentum operators according to

$$T\,\mathbf{J}\,T^{-1} = -\mathbf{J}\,, \qquad T J_{\pm} T^{-1} = -J_{\mp}\,, \quad T J_3 T^{-1} = -J_3 \tag{3.85}$$

This is compatible with the commutation relations (3.54) or (3.61a), (3.61b) since T is anti-linear. Hence in terms of the $|jm\rangle$ basis if this is invariant under the action of T we have

$$T|jm\rangle = e^{i\phi}(-1)^{j-m}\,|j\,{-m}\rangle\,, \tag{3.86}$$

for some complex phase $e^{i\phi}$. Hence

$$T^2|jm\rangle = (-1)^{2j}\,|jm\rangle\,, \tag{3.87}$$

which generalises to $T^2 = (-1)^F$ with F the fermion number.

3.5.2 Representation Matrices

Using the $|j\,m\rangle$ basis it is straightforward to define corresponding representation matrices for each j belonging to (3.72). For the angular momentum operator

$$\mathbf{J}^{(j)}{}_{m'm} = \langle j\,m'|\mathbf{J}|j\,m\rangle \tag{3.88}$$

or alternatively

$$\mathbf{J}|j\,m\rangle = \sum_{m'} |j\,m'\rangle\,\mathbf{J}^{(j)}{}_{m'm}\,. \tag{3.89}$$

The $(2j+1)\times(2j+1)$ matrices $\mathbf{J}^{(j)} = [\mathbf{J}^{(j)}{}_{m'm}]$ then satisfy the angular momentum commutation relations (3.54). From (3.75) and (3.77)

$$J_3^{(j)}{}_{m'm} = m\,\delta_{m',m}\,, \qquad J_{\pm}^{(j)}{}_{m'm} = \sqrt{(j\mp m)(j\pm m+1)}\,\delta_{m',m\pm 1}\,. \tag{3.90}$$

For R a rotation then

$$D^{(j)}_{m'm}(R) = \langle j\,m'|U[R]|j\,m\rangle\,, \tag{3.91}$$

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 $\mathcal{V}_j$,

$$U[R]|jm\rangle = \sum_{m'} |jm'\rangle D^{(j)}_{m'm}(R)\,. \tag{3.92}$$

Note that $D^{(0)}(R) = 1$ is the trivial representation and for an infinitesimal rotation as in (3.16)

$$D^{(j)}\big(R(\delta\theta,\mathbf{n})\big) = \mathbb{1}_{2j+1} - i\delta\theta\,\mathbf{n}{\cdot}\mathbf{J}^{(j)}\,. \tag{3.93}$$

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

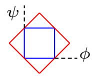
$$R_{\phi,\theta,\psi} = R(\phi,\mathbf{e}_3)R(\theta,\mathbf{e}_2)R(\psi,\mathbf{e}_3)\,, \quad \{R_{\phi,\theta,\psi}\} = SO(3) \text{ for } 0\le\theta\le\pi\,,\ 0\le\phi,\psi\le 2\pi\,, \tag{3.94}$$

where $\mathbf{e}_2, \mathbf{e}_3$ are unit vectors along the $2, 3$ directions. For the corresponding two dimensional matrix

$$A_{\phi,\theta,\psi} = e^{-\frac{1}{2}i\phi\,\sigma_3} e^{-\frac{1}{2}i\theta\,\sigma_2} e^{-\frac{1}{2}i\psi\,\sigma_3} = \begin{pmatrix} \cos\frac{1}{2}\theta\, e^{-\frac{1}{2}i(\phi+\psi)} & -\sin\frac{1}{2}\theta\, e^{-\frac{1}{2}i(\phi-\psi)} \\ \sin\frac{1}{2}\theta\, e^{\frac{1}{2}i(\phi-\psi)} & \cos\frac{1}{2}\theta\, e^{\frac{1}{2}i(\phi+\psi)} \end{pmatrix}, \tag{3.95}$$

with

$$\begin{aligned} \{A_{\phi,\theta,\psi}\} = SU(2) \quad \text{for} \quad & 0 \le \theta \le \pi\,, \quad -2\pi \le \phi - \psi \le 2\pi\,, \quad 0 \le \phi + \psi \le 4\pi\,, \\ \text{or} \quad & 0 \le \theta \le \pi\,, \quad 0 \le \phi \le 2\pi\,, \quad 0 \le \psi \le 4\pi\,. \end{aligned} \tag{3.96}$$

The allowed regions of (ϕ, ψ) in (3.94), (3.96) are enclosed in blue, red in  . Since

$$A_{\phi,\theta\pm 2\pi,\psi} = A_{\phi,\theta,\psi\pm 2\pi} = -A_{\phi,\theta,\psi}\,, \tag{3.97}$$

the region in (3.96) maps into (3.94) allowing for a change of sign.

In terms the angular momentum operators $\mathbf{J}$ the rotation operators in terms of Euler angles are then

$$U[R_{\phi,\theta,\psi}] = e^{-i\phi J_3}\, e^{-i\theta J_2}\, e^{-i\psi J_3}\,, \tag{3.98}$$

so that in (3.91)

$$D^{(j)}_{m'm}(R_{\phi,\theta,\psi}) = e^{-im'\phi - im\psi}\, d^{(j)}_{m'm}(\theta)\,, \qquad d^{(j)}_{m'm}(\theta) = \langle j\, m'|e^{-i\theta J_2}|j\, m\rangle\,. \tag{3.99}$$

The matrices $d^{(j)}(\theta) = [d^{(j)}_{m'm}(\theta)]$ satisfy $d^{(j)}(\theta)d^{(j)}(\theta') = d^{(j)}(\theta+\theta')$. For the special cases of $\theta = \pi, 2\pi$,

$$d^{(j)}_{m'm}(\pi) = (-1)^{j-m}\delta_{m',-m}\,, \qquad d^{(j)}_{m'm}(2\pi) = (-1)^{2j}\delta_{m',m}\,. \tag{3.100}$$

Since $iJ_2 = \frac{1}{2}(J_+ - J_-)$ then with the conventions (3.75) and (3.77) $d^{(j)}_{m'm}(\theta)$ are real.

In general $D^{(j)}(R(2\pi,\mathbf{n})) = (-1)^{2j}\mathbb{1}_{2j+1}$, 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 (3.72) is allowed since we require representations only up to a phase factor. From the result for $\theta = \pi$ we have

$$e^{-i\pi J_2}|j\, m\rangle = (-1)^{j-m}|j\, {-m}\rangle\,. \tag{3.101}$$

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 (3.99)

$$d^{(j)}_{m'm}(\theta) = d^{(j)}_{mm'}(-\theta) = (-1)^{m'-m} d^{(j)}_{-m'\,-m}(\theta) = (-1)^{m'-m} d^{(j)}_{m'm}(-\theta)\,. \tag{3.102}$$

Furthermore $d^{(j)}_{m'm'}(\pi-\theta) = \sum_{m''} d^{(j)}_{m'm'}(-\theta) d^{(j)}_{m''m}(\pi) = (-1)^{j+m'} d^{(j)}_{m'-m}(\theta)$.

For the simplest case $j = \frac{1}{2}$, it is easy to see from (3.90) that

$$J_+^{(\frac{1}{2})} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad J_-^{(\frac{1}{2})} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad J_3^{(\frac{1}{2})} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.103}$$

and hence we have

$$\mathbf{J}^{(\frac{1}{2})} = \tfrac{1}{2}\boldsymbol{\sigma}\,, \tag{3.104}$$

where σ_i, $i = 1,2,3$ are the Pauli matrices as given in (3.19). It is clear that $\frac{1}{2}\sigma_i$ must satisfy the commutation relations (3.54). The required commutation relations are a consequence of (3.20). For $j = \frac{1}{2}$ we also have

$$d^{(\frac{1}{2})}(\theta) = \begin{pmatrix} \cos\frac{1}{2}\theta & -\sin\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix}. \tag{3.105}$$

For $j = 1$

$$J_+^{(1)} = \left(\begin{smallmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{smallmatrix}\right), \qquad J_-^{(1)} = \left(\begin{smallmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{smallmatrix}\right), \qquad J_3^{(1)} = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{smallmatrix}\right), \tag{3.106}$$

and

$$d^{(1)}(\theta) = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & -\sqrt{2}\sin\theta & 1-\cos\theta \\ \sqrt{2}\sin\theta & 2\cos\theta & -\sqrt{2}\sin\theta \\ 1-\cos\theta & \sqrt{2}\sin\theta & 1+\cos\theta \end{pmatrix}. \tag{3.107}$$

For any integer j

$$d_{00}^{(j)}(\theta) = P_j(\cos\theta)\,, \tag{3.108}$$

with P_j the usual Legendre polynomials.

For general j there is an expression for $d_{m'm}^{(j)}(\theta)$ in terms of classical Jacobi polynomials. To obtain the associated differential equation we can start from (3.98) to obtain the differential relations

$$J_\pm\, U[R_{\phi,\theta,\psi}] = -\mathcal{J}_\pm\, U[R_{\phi,\theta,\psi}]\,, \qquad J_3\, U[R_{\phi,\theta,\psi}] = -\mathcal{J}_3\, U[R_{\phi,\theta,\psi}]\,, \tag{3.109}$$

where

$$\mathcal{J}_\pm = e^{\pm i\phi}\left(i\cot\theta\,\frac{\partial}{\partial\phi} \pm \frac{\partial}{\partial\theta} - i\csc\theta\,\frac{\partial}{\partial\psi}\right), \qquad \mathcal{J}_3 = -i\,\frac{\partial}{\partial\phi}\,. \tag{3.110}$$

This follows using $\csc\theta\, e^{-i\theta J_2} J_3 = (\cot\theta J_3 + J_1)e^{-i\theta J_2}$. It is straightforward to verify that $\mathcal{J}_\pm$, $\mathcal{J}_3$ satisfy the usual commutation relations (3.61a), (3.61b) and

$$\boldsymbol{\mathcal{J}}^2 = -\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2} - \cot\theta\,\frac{\partial}{\partial\theta} - \left(\cot\theta\,\frac{\partial}{\partial\phi} - \csc\theta\,\frac{\partial}{\partial\psi}\right)^2. \tag{3.111}$$

Using this with (3.109) to evaluate $\langle jm'|\,\mathbf{J}^2\, U[R_{\phi,\theta,\psi}]|jm\rangle$ leads to the differential equation

$$\left(-\frac{\partial^2}{\partial\theta^2} - \cot\theta\,\frac{\partial}{\partial\theta} + m'^{\,2} + \big(m'\cot\theta - m\csc\theta\big)^2\right) d_{m'm}^{(j)}(\theta) = j(j+1)\, d_{m'm}^{(j)}(\theta) \tag{3.112}$$

Writing, for $m \geq m'$,

$$d_{m'm}^{(j)}(\theta) = \big(\cos\tfrac{1}{2}\theta\big)^{m+m'}\big(\sin\tfrac{1}{2}\theta\big)^{m-m'} P(\cos\theta)\,, \tag{3.113}$$

the differential equation translates into a hypergeometric form

$$(1-t^2)P''(t) + 2\big(m' - (m+1)\,t\big)P'(t) + (j-m)(j+m+1)\,P(t) = 0 \tag{3.114}$$

The solutions are Jacobi polynomials giving

$$d^{(j)}_{m'm}(\theta) = \bigg(\frac{(j+m)!\,(j-m)!}{(j+m')!\,(j-m')!}\bigg)^{\frac{1}{2}} \big(\cos\tfrac{1}{2}\theta\big)^{m+m'}\big(\sin\tfrac{1}{2}\theta\big)^{m-m'}\,P^{(m-m',m+m')}_{j-m}(\cos\theta)\,. \tag{3.115}$$

For $m = m' = 0$ this reduces to (3.108). To check the normalisation for $\theta \to 0$ and $m \geq m'$

$$\begin{aligned} d^{(j)}_{m'm}(\theta) &\sim (\tfrac{1}{2}\theta)^{m-m'}\frac{1}{(m-m')!}\,\langle jm'|(J_-)^{m-m'}|jm\rangle \\ &= (\tfrac{1}{2}\theta)^{m-m'}\frac{1}{(m-m')!}\bigg(\frac{(j+m)!\,(j-m')!}{(j+m')!\,(j-m)!}\bigg)^{\frac{1}{2}}\,, \end{aligned} \tag{3.116}$$

and $P^{(m-m',m+m')}_{j-m}(1) = \frac{(j-m')!}{(j-m)!(m-m')!}$.

Defining

$$U^{(j)} = D^{(j)}\big(R_{\frac{\pi}{2},\frac{\pi}{2},-\frac{\pi}{2}}\big)\,, \tag{3.117}$$

then

$$\big(U^{(j)}d^{(j)}(\theta)\,U^{(j)\dagger}\big)_{m'm} = e^{im\theta}\delta_{m',m}\,. \tag{3.118}$$

As special cases

$$U^{(\frac{1}{2})} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \qquad U^{(1)} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2}\,i & -1 \\ \sqrt{2}\,i & 0 & \sqrt{2}\,i \\ -1 & \sqrt{2}\,i & 1 \end{pmatrix}. \tag{3.119}$$

3.5.3 Integration over $SO(3)$ and orthogonality relations

Corresponding to the sum over group elements for a finite group there is an integration over the group parameters, here the Euler angles θ, ϕ, ψ, for a continuous group. The crucial requirement is to respect the property (1.7). If $\mathrm{d}\mu_{\theta,\phi,\psi}$ this requires that $\mathrm{d}\mu_{\theta,\phi,\psi} = \mathrm{d}\mu_{\theta',\phi',\psi'}$ where the change of variables $(\phi,\theta,\psi) \to (\theta',\phi',\psi')$ is obtained by $R_{\theta,\phi,\psi}\,R = R_{\theta',\phi',\psi'}$ for any rotation R. Infinitesimally with the $j = \frac{1}{2}$ rotation matrices in (3.95) $A_{\theta,\phi,\psi}\,A_{\epsilon_\theta,\epsilon_\phi,\epsilon_\psi} = A_{\theta+\delta\theta,\phi+\delta\phi,\psi+\delta\psi}$ which gives, to first order in $\epsilon_\theta, \epsilon_\phi, \epsilon_\psi$,

$$\delta\theta = \cos\psi\,\epsilon_\theta\,, \qquad \delta\phi = \frac{\sin\psi}{\sin\theta}\,\epsilon_\theta\,, \qquad \delta\psi = \epsilon_\phi + \epsilon_\psi - \sin\psi\,\cot\theta\,\epsilon_\theta\,. \tag{3.120}$$

In consequence

$$\mathrm{d}(\delta\theta,\delta\phi,\delta\psi) = (\mathrm{d}\theta,\mathrm{d}\phi,\mathrm{d}\psi)J\,\epsilon_\theta\,, \quad J = \begin{pmatrix} 0 & -\sin\psi\,\frac{\cot\theta}{\sin\theta} & \frac{\sin\psi}{\sin^2\theta} \\ 0 & 0 & 0 \\ -\sin\psi & \frac{\cos\psi}{\sin\theta} & -\cos\psi\,\cot\theta \end{pmatrix}, \tag{3.121}$$

and hence

$$\delta(\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\psi\,\sin\theta) = \mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\psi\,(\sin\theta\,\mathrm{tr}J + \cos\theta\,\cos\psi)\epsilon_\theta = 0\,, \tag{3.122}$$

so that we may take for the integration measure in terms of Euler angles

$$\mathrm{d}\mu_{\theta,\phi,\psi} = \frac{1}{8\pi^2}\,\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\psi\,\sin\theta\,, \qquad \int_{SO(3)}\mathrm{d}\mu_{\theta,\phi,\psi} = 1\,, \tag{3.123}$$

with the integration ranges as in (3.94).

The orthogonality relation (2.50) now translates for the unitary rotation matrices into

$$\int_{SO(3)}\mathrm{d}\mu_{\theta,\phi,\psi}\; D^{(j_1)}_{m_1{}'m_1}(R_{\phi,\theta,\psi})\, D^{(j_2)}_{m_2{}'m_2}(R_{\phi,\theta,\psi})^* = \frac{1}{2j_1+1}\,\delta_{j_1j_2}\,\delta_{m_1{}'m_2{}'}\,\delta_{m_1m_2}\,. \tag{3.124}$$

With the decomposition (3.99) this is reducible to

$$\int_0^\pi \mathrm{d}\theta\; d^{(j_1)}_{m'm}(\theta)\, d^{(j_2)}_{m'm}(\theta) = \frac{2}{2j_1+1}\,\delta_{j_1j_2}\,. \tag{3.125}$$

The integration measure in (3.123) in terms of Euler angles can be transformed to other parameterisations. The overall angle of rotation Θ can be expressed in terms of Euler angles using, from (3.38) and (3.95),

$$\cos\tfrac{1}{2}\Theta = \tfrac{1}{2}\,\mathrm{tr}\big(A_{\phi,\theta,\psi}\big) = \cos\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}(\phi+\psi)\,. \tag{3.126}$$

For any function of Θ then

$$\int\mathrm{d}\mu_{\theta,\phi,\psi}\; f(\Theta) = \frac{1}{2}\int_0^{2\pi}\mathrm{d}\Theta\;\sin\tfrac{1}{2}\Theta\, f(\Theta)\int\mathrm{d}\mu_{\theta,\phi,\psi}\;\delta\big(\cos\tfrac{1}{2}\Theta - \cos\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}(\phi+\psi)\big)\,, \tag{3.127}$$

where if $\cos\frac{1}{2}\Theta > 0$ and taking $\phi_\pm = \phi \pm \psi$,

$$\begin{aligned}
&\int\mathrm{d}\mu_{\theta,\phi,\psi}\;\delta\big(\cos\tfrac{1}{2}\Theta - \cos\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}\phi_+\big)\\
&= \frac{1}{16\pi^2}\int_{-2\pi}^{2\pi}\mathrm{d}\phi_-\int_0^{2\pi}\mathrm{d}\phi_+\int_0^\pi\mathrm{d}\theta\,\sin\theta\,\delta\big(\cos\tfrac{1}{2}\Theta - \cos\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}\phi_+\big)\\
&= \frac{1}{\pi}\cos\tfrac{1}{2}\Theta\int_0^\Theta \mathrm{d}\phi_+\;\frac{1}{\cos^2\tfrac{1}{2}\phi_+} = \frac{2}{\pi}\,\sin\tfrac{1}{2}\Theta\,.
\end{aligned} \tag{3.128}$$

The final result remains if $\cos\frac{1}{2}\Theta < 0$ so that the measure then becomes

$$\frac{1}{\pi}\,\mathrm{d}\Theta\;\sin^2\tfrac{1}{2}\Theta\,, \qquad 0 \le \Theta \le 2\pi\,. \tag{3.129}$$

3.5.4 Characters for $SU(2)$

With the definition of characters in (2.51) the rotation group characters

$$\chi_j(\theta) = \mathrm{tr}\Big(D^{(j)}\big(R(\theta,\mathbf{n})\big)\Big)\,, \tag{3.130}$$

depend only on the rotation angle θ. Since

$$D^{(j)}_{m'm}\big(R(\theta,\mathbf{z})\big) = \delta_{m'm}\,e^{-im\theta}\,, \tag{3.131}$$

they may be easily calculated

$$\chi_j(\theta) = \sum_{m=-j}^{j} e^{-im\theta} = \frac{\sin(j+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\,. \tag{3.132}$$

The orthogonality properties corresponding to (2.56) follow from using (3.129) with $\Theta \to \theta$

$$\frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, \sin^2\tfrac{1}{2}\theta\, \chi_{j_1}(\theta)\chi_{j_2}(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\theta\, \big(\cos(j_1-j_2)\theta - \cos(j_1+j_2+1)\theta\big) = \delta_{j_1,j_2}\,,$$
$$2j_1,\, 2j_2 = 0,1,2,\dots\,. \tag{3.133}$$

Furthermore corresponding to (2.61)

$$\sum_{j=0,\frac{1}{2},1,\dots} \chi_j(\theta)\chi_j(\theta') = \frac{1}{4\,\sin^2\frac{1}{2}\theta}\sum_{n=-\infty}^{\infty}\big(e^{i\,\frac{1}{2}n(\theta-\theta')} - e^{i\,\frac{1}{2}n(\theta+\theta')}\big)$$
$$= \frac{\pi}{\sin^2\frac{1}{2}\theta}\sum_{n=-\infty}^{\infty}\big(\delta(\theta-\theta'-4n\pi) - \delta(\theta+\theta'-4n\pi)\big) = \frac{\pi}{\sin^2\frac{1}{2}\theta}\delta(\theta-\theta')\,,\ 0<\theta,\theta'<2\pi\,. \tag{3.134}$$

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 of (3.133) is halved. There is a corresponding modification in (3.134) and it is necessary to restrict $0<\theta,\theta'<\pi$ to get a single δ-function.

In addition since

$$\chi_j(2\theta) = \textstyle\sum_{m=-j}^{j} e^{i\,m\,2\theta} = \sum_{J=0}^{2j} (-1)^{2j-J}\chi_J(\theta)\,, \tag{3.135}$$

then

$$\frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, \sin^2\tfrac{1}{2}\theta\, \chi_j(2\theta) = (-1)^{2j}\,, \tag{3.136}$$

since only $\chi_0(\theta) = 1$ survives after integration. By virtue of (2.63) this shows that the representations are real for j integral, pseudo-real for j half integral.

3.6 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 (3.54) 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

$$|j_1\,m_1\rangle_1 |j_2\,m_2\rangle_2\,. \tag{3.137}$$

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 (3.54)

$$[J_{1,i}\,,\,J_{1,j}] = i\varepsilon_{ijk}J_{1,k}\,,$$
$$[J_{2,i}\,,\,J_{2,j}] = i\varepsilon_{ijk}J_{2,k}\,. \tag{3.138}$$

They may be extended to act on $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ since with the basis (3.137)

$$\begin{aligned} \mathbf{J}_1 &\equiv \mathbf{J}_1 \otimes 1_2 \,, \qquad \mathbf{J}_1\big(|j_1 m_1\rangle_1 \, |j_2\, m_2\rangle_2\big) = \mathbf{J}_1 |j_1 m_1\rangle_1 \, |j_2\, m_2\rangle_2 \,, \\ \mathbf{J}_2 &\equiv 1_1 \otimes \mathbf{J}_2 \,, \qquad \mathbf{J}_2\big(|j_1 m_1\rangle_1 \, |j_2\, m_2\rangle_2\big) = |j_1 m_1\rangle_1 \, \mathbf{J}_2 |j_2\, m_2\rangle_2 \,. \end{aligned} \tag{3.139}$$

With this definition it is clear that they commute

$$[J_{1,i} \,,\, J_{2,j}] = 0 \,. \tag{3.140}$$

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

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \,. \tag{3.141}$$

It is easy to see that this has the standard commutation relations (3.54).

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,

$$\begin{aligned} J_3|JM\rangle &= M|JM\rangle \,, \\ \mathbf{J}^2|JM\rangle &= J(J+1)|JM\rangle \,. \end{aligned} \tag{3.142}$$

These states are chosen to be orthonormal so that

$$\langle J'M'|JM\rangle = \delta_{J'J}\delta_{M'M} \,, \tag{3.143}$$

and satisfy (3.77). All states in $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$ must be linear combinations of the basis states (3.137) so that we may write

$$|JM\rangle = \sum_{m_1,m_2} |j_1 m_1\rangle_1 |j_2 m_2\rangle_2 \, \langle j_1 m_1 \, j_2 m_2|JM\rangle \,. \tag{3.144}$$

Here

$$\langle j_1 m_1 \, j_2 m_2|JM\rangle \,, \tag{3.145}$$

are *Clebsch-Gordan coefficients*[26].

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

$$J_3|JJ\rangle = J|JJ\rangle \,, \qquad J_+|JJ\rangle = 0 \,. \tag{3.146}$$

We may then determine the states $|JM\rangle$ by applying J_- as in (3.80). There is clearly a unique highest weight state with $J = j_1 + j_2$ given by

$$|j_1{+}j_2 \, j_1{+}j_2\rangle = |j_1 \, j_1\rangle_1 \, |j_2 \, j_2\rangle_2 \,, \tag{3.147}$$

so that

$$\langle j_1 j_1 \, j_2 j_2 | j_1 + j_2 \, j_1 + j_2\rangle = 1 \,. \tag{3.148}$$

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

From $|j_1{+}j_2\, j_1{+}j_2\rangle$ states $|j_1{+}j_2\, M\rangle$ for any M are obtained as in (3.80). Using $J_-{}^n = \sum_{r=0}^n \binom{n}{r} J_{1-}{}^r J_{2-}{}^{n-r}$ for $n = 1, \dots j_1 + j_2$ we may then derive

$$\langle j_1 m_1\, j_2 m_2 | J\, M\rangle = \left(\frac{(2j_1)!(2j_2)!}{(2J)!} \frac{(J-M)!(J+M)}{(j_1-m_1)!(j_1+m_1))!(j_2-m_2)!j_2+m_2))!} \right)^{\frac{1}{2}} ,$$
$$J = j_1 + j_2 \,, \qquad M = m_1 + m_2 \,. \tag{3.149}$$

Clearly as well as (3.148) $\langle j_1{-}j_1\, j_2{-}j_2 | j_1{+}j_2\, {-}(j_1{+}j_2)\rangle = 1$. We may then construct the states $|JM\rangle$ for $J = j_1 + j_2 - 1, \dots$ 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 (3.137) 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 $\dim \mathcal{V}^{(M)} = 2j_2 + 1$ for $M \le j_1 - j_2$. Assume all states $|J'M\rangle$ have been found as in (3.144) 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 $|JJ\rangle$. In constructing a normalised $|JJ\rangle$ in terms of a real linear combination of the states $|j_1 m_1\rangle_1 |j_2 m_2\rangle_2$, $J = m_1 + m_2$ there is an overall choice of sign, conventionally the coefficient for the largest m_1 is positive. 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

$$J \in \{ j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2| + 1, |j_1 - j_2| \} \,, \tag{3.150}$$

or

$$\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2} = \bigoplus_{J=|j_1-j_2|}^{j_1+j_2} \mathcal{V}_J \,. \tag{3.151}$$

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

$$\sum_{J=j_1-j_2}^{j_1+j_2} (2J+1) = \sum_{J=j_1-j_2}^{j_1+j_2} \big((J+1)^2 - J^2 \big)$$
$$= (j_1 + j_2 + 1)^2 - (j_1 - j_2)^2 = (2j_1 + 1)(2j_2 + 1) \,, \tag{3.152}$$

so that the basis $\{|JM\rangle\}$ has the correct dimension to span the vector space $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$. 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 (3.150). By convention they are chosen to be real and for each J there is a standard choice of the overall sign. With the common conventions

$$\langle j_1 m_1\, j_2 m_2 | JM\rangle = (-1)^{j_1+j_2-J} \langle j_2 m_2\, j_1 m_1 | JM\rangle \,,$$
$$= (-1)^{j_1+j_2-J} \langle j_1{-}m_1\, j_2{-}m_2 | J{-}M\rangle \,. \tag{3.153}$$

The first arises since interchanging j_1 and j_2 changes the overall sign whenever $j_1 + j_2 - J$ is odd, the second since construction of Clebsch-Gordan coefficients can equally be given starting from $|j_1{+}j_2\, {-}j_1{-}j_2\rangle = |j_1\, {-}j_1\rangle_1\, |j_2\, {-}j_2\rangle_2$ instead of (3.147) but the sign prescription changes as previously.

For $j_1 = j_2 = j$ the decomposition of the tensor product $\mathcal{V}_j \otimes \mathcal{V}_j$ in (3.151) can be separated into contributions which are symmetric or antisymmetric under interchange

$$
\begin{aligned}
&j \text{ integral}, \ \ \vee^2 \mathcal{V}_j = (\mathcal{V}_j \otimes \mathcal{V}_j)_{\text{sym}} = \bigoplus_{n=0}^{j} \mathcal{V}_{2n}\,, \qquad \wedge^2 \mathcal{V}_j = (\mathcal{V}_j \otimes \mathcal{V}_j)_{\text{antisym}} = \bigoplus_{n=0}^{j-1} \mathcal{V}_{2n+1}\,, \\
&j \text{ half integral}, \ \ \vee^2 \mathcal{V}_j = \bigoplus_{n=0}^{j-\frac{1}{2}} \mathcal{V}_{2n+1}\,, \qquad \wedge^2 \mathcal{V}_j = \bigoplus_{n=0}^{j-\frac{1}{2}} \mathcal{V}_{2n}\,, \qquad (3.154)
\end{aligned}
$$

As a check for j integral the total dimensions of the symmetric and antisymmetric subspaces in $\mathcal{V}_j \otimes \mathcal{V}_j$ are then $\sum_{n=0}^{j}(4n+1) = (j+1)(2j+1)$, $\sum_{n=0}^{j-1}(4n+3) = j(2j+1)$.

Since the original basis (3.137) and $\{|JM\rangle\}$ are both orthonormal we have the orthogonality/completeness conditions

$$
\begin{aligned}
\sum_{m_1,m_2} \langle j_1 m_1\, j_2 m_2 | JM\rangle \langle j_1 m_1\, j_2 m_2 | J'M'\rangle &= \delta_{JJ'}\, \delta_{MM'}\,, \\
\sum_{JM} \langle j_1 m_1\, j_2 m_2 | JM\rangle \langle j_1 m'_1\, j_2 m'_2 | JM\rangle &= \delta_{m_1 m'_1}\, \delta_{m_2 m'_2}\,. \qquad (3.155)
\end{aligned}
$$

Together with, from applying $J_\pm$ to (3.144),

$$
\begin{aligned}
N^\pm_{JM} \langle j_1 m_1\, j_2 m_2 | JM{\pm}1\rangle = N^\pm_{j_1 m_1 -1} \langle j_1 m_1{\mp}1\, j_2 m_2 | JM\rangle + N^\pm_{j_2 m_2 -1} \langle j_1 m_1\, j_2 m_2{\mp}1 | JM\rangle\,, \\
M = m_1 + m_2 \mp 1\,, \qquad (3.156)
\end{aligned}
$$

these determine all Clebsch-Gordan coefficients up to a choice of sign. Using (3.155) (3.144) can be inverted

$$
|j_1 m_1\rangle_1 |j_2 m_2\rangle_2 = \sum_{J,M} |JM\rangle\, \langle j_1 m_1\, j_2 m_2 | JM\rangle\,. \qquad (3.157)
$$

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 (3.144) to give the decomposition into irreducible representations for each J allowed by (3.150)

$$
\sum_{m_1{}',m_1} \sum_{m_2{}',m_2} D^{(j_1)}_{m_1{}' m_1}(R)\, D^{(j_2)}_{m_2{}' m_2}(R)\, \langle j_1 m_1{}'\, j_2 m_2{}' | J'M'\rangle \langle j_1 m_1\, j_2 m_2 | JM\rangle = \delta_{J'J}\, D^{(J)}_{M'M}(R)\,. \qquad (3.158)
$$

For rotation matrices expressed in terms of Euler angles as in (3.99) the dependence on ϕ, ψ factorises and the relation holds when $D^{(j)}_{m'm}(R) \to d^{(j)}_{m'm}(\theta)$. With the aid of the orthogonality relations (3.155) this can be rewritten as

$$
\sum_{m_1{}'+m_2{}'=m'} d^{(j_1)}_{m_1{}' m_1}(\theta)\, d^{(j_2)}_{m_2{}' m_2}(\theta)\, \langle j_1 m_1{}'\, j_2 m_2{}' | j m'\rangle = d^{(j)}_{m' m_1+m_2}(\theta)\, \langle j_1 m_1\, j_2 m_2 | j m_1{+}m_2\rangle\,. \qquad (3.159)
$$

3.7 Examples of the calculation of Clebsch-Gordan coefficients

A simple example where the Clebsch-Gordan coefficients can be quite easily calculated is when $j_1 = j$ and $j_2 = \frac{1}{2}$ so that there are $2(2j+1)$ states altogether. In this case the states can be combined using Clebsch-Gordan coefficients to form states with total angular momentum $J = j \pm \frac{1}{2}$ and the states $|JM\rangle$ are just linear combinations of $|j\, m\rangle_1$ and $|\frac{1}{2}\, \pm\frac{1}{2}\rangle_2$ with $m = M \mp \frac{1}{2}$.

In general then

$$|j{+}\tfrac{1}{2}\, m{+}\tfrac{1}{2}\rangle = a_m\, |j\, m\rangle_1\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle_2 + b_m\, |j\, m{+}1\rangle_1\, |\tfrac{1}{2}\, {-}\tfrac{1}{2}\rangle_2\,. \tag{3.160}$$

As a consequence of (3.147) we must have

$$a_j = 1\,, \qquad b_j = 0\,. \tag{3.161}$$

Applying $J_\pm$ to (3.160) we may directly obtain recurrence relations for a_{m-1} in tems of a_m and b_{m+1} in terms of b_m where using (3.77) and (3.79)

$$a_{m-1} = \left(\frac{j+m}{j+m+1}\right)^{\frac{1}{2}} a_m\,, \qquad b_{m+1} = \left(\frac{j-m-1}{j-m}\right)^{\frac{1}{2}} b_m\,. \tag{3.162}$$

These are easily solved

$$a_m = \left(\frac{j+m+1}{2j+1}\right)^{\frac{1}{2}}, \qquad b_m = \left(\frac{j-m}{2j+1}\right)^{\frac{1}{2}}, \tag{3.163}$$

where the normalisation of a_m is determined from (3.161). To check the normalisation of b_m it is sufficient to note that applying J_- to (3.160) also gives

$$b_{m-1} = \left(\frac{1}{(j+m+1)(j-m+1)}\right)^{\frac{1}{2}} a_m + \left(\frac{j-m}{j-m+1}\right)^{\frac{1}{2}} b_m\,, \tag{3.164}$$

which is satisfied by (3.163). Clearly $a_{-j-1} = 0$, $b_{-j-1} = 1$ so that $|j{+}\frac{1}{2}\, {-}j{-}\frac{1}{2}\rangle = |j\, {-}j\rangle_1\, |\frac{1}{2}\, {-}\frac{1}{2}\rangle_2$. Also ${a_m}^2 + {b_m}^2 = 1$ which is necessary for $|j{+}\frac{1}{2}\, m{+}\frac{1}{2}\rangle$ to be normalised.

The corresponding states with $J = j - \frac{1}{2}$ are orthogonal to the states defied in (3.160). In this case it is sufficient to take

$$|j{-}\tfrac{1}{2}\, m{+}\tfrac{1}{2}\rangle = -b_m\, |j\, m\rangle_1\, |\tfrac{1}{2}\, \tfrac{1}{2}\rangle_2 + a_m\, |j\, m{+}1\rangle_1\, |\tfrac{1}{2}\, {-}\tfrac{1}{2}\rangle_2\,, \tag{3.165}$$

where here $m = j-1, j-2, \dots, -j+1$. This result is unique to within an overall phase which we have taken in accordance with the so-called Condon and Shortley phase convention. We may directly verify that $J_+|j{-}\frac{1}{2}\, j{-}\frac{1}{2}\rangle = 0$ and that acting repeatedly with J_- respects the conventions relating $|j{-}\frac{1}{2}\, M{-}1\rangle$ to $|j{-}\frac{1}{2}\, M\rangle$.

In the end the Clebsch-Gordan coefficients are then

$$\begin{aligned} \langle jm\, \tfrac{1}{2}\tfrac{1}{2}|j{+}\tfrac{1}{2}\, m{+}\tfrac{1}{2}\rangle &= \sqrt{\tfrac{j+m+1}{2j+1}}\,, \qquad & \langle jm\, \tfrac{1}{2}{-}\tfrac{1}{2}|j{+}\tfrac{1}{2}\, m{-}\tfrac{1}{2}\rangle &= \sqrt{\tfrac{j-m+1}{2j+1}}\,, \\ \langle jm\, \tfrac{1}{2}\tfrac{1}{2}|j{-}\tfrac{1}{2}\, m{+}\tfrac{1}{2}\rangle &= -\sqrt{\tfrac{j-m}{2j+1}}\,, \qquad & \langle jm\, \tfrac{1}{2}{-}\tfrac{1}{2}|j{-}\tfrac{1}{2}\, m{-}\tfrac{1}{2}\rangle &= \sqrt{\tfrac{j+m}{2j+1}}\,, \end{aligned} \tag{3.166}$$

The Condon and Shortley convention requires taking $\langle jj\, \frac{1}{2}{-}\frac{1}{2}|j{-}\frac{1}{2}\, j{-}\frac{1}{2}\rangle = \sqrt{\frac{2j}{2j+1}} > 0$.

By combining the results in (3.166) with (3.159) for $j_1 = j - \frac{1}{2}$, $j_2 = \frac{1}{2}$ and using (3.105) we may obtain

$$(j \pm m)^{\frac{1}{2}}\, d^{(j)}_{m'm}(\theta) = (j \pm m')^{\frac{1}{2}} \cos\tfrac{1}{2}\theta\, d^{(j-\frac{1}{2})}_{m'\mp\frac{1}{2}\, m\mp\frac{1}{2}}(\theta) \pm (j \mp m')^{\frac{1}{2}} \sin\tfrac{1}{2}\theta\, d^{(j-\frac{1}{2})}_{m'\pm\frac{1}{2}\, m\mp\frac{1}{2}}(\theta)\,. \tag{3.167}$$

3.7.1 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 (3.150) this is only possible for $j_1 = j_2 = j$ and the singlet state must have the general form

$$|0\,0\rangle = \sum_m a_m \, |j\,m\rangle_1 |j\,{-m}\rangle_2 \,. \tag{3.168}$$

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

$$|0\,0\rangle = \frac{1}{\sqrt{2j+1}} \sum_{n=0}^{2j} (-1)^n |j\,j{-n}\rangle_1 |j\,{-j{+}n}\rangle_2 \quad \Rightarrow \quad \langle jm\,j{-m}|00\rangle = \frac{1}{\sqrt{2j+1}} (-1)^{j-m} \,. \tag{3.169}$$

Note that $|0\,0\rangle$ is symmetric, antisymmetric under $1 \leftrightarrow 2$ according to whether $2j$ is even, odd.

3.7.2 Construction of Highest Weight States

The construction of $|0\,0\rangle$ can be generalised to find all highest weight states $|J\,J\rangle$ contained in $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}$. These have the form

$$|J\,J\rangle = \sum_{m_1+m_2=J} |j_1\,m_1\rangle_1 |_2 j\,m_2\rangle_2 \, \langle j_1 m_1\, j_2 m_2 | JJ\rangle \,. \tag{3.170}$$

Requiring $J_+|J\,J\rangle = 0$ gives

$$N^+_{j_1 m_1-1} \, \langle j_1 m_1{-}1\, j_2 m_2 | JJ\rangle + N^+_{j_2 m_2-1} \, \langle j_1 m_1\, j_2 m_2{-}1 | JJ\rangle = 0 \,, \tag{3.171}$$

which leads to

$$\langle j_1 m_1\, j_2 m_2 | JJ\rangle = (-1)^{j_1-m_1} \left(\frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!} \right)^{\frac{1}{2}} A_{j_1 j_2 J} \,, \quad m_1 + m_2 = J \,. \tag{3.172}$$

For normalisation we require[27]

$$\begin{aligned} &\sum_{m_1=J-j_2}^{j_1} \frac{(j_1+m_1)!\,(j_2+J-m_1)!}{(j_1-m_1)!\,(j_2-J+m_1)!} \, {A_{j_1 j_2 J}}^2 \\ &\qquad = \frac{(j_1-j_2+J)!\,(j_2-j_1+J)!\,(j_1+j_2+J+1)!}{(j_1+j_2-J)!\,(2J+1)!} \, {A_{j_1 j_2 J}}^2 = 1 \,. \end{aligned} \tag{3.173}$$

where $A_{j_1 j_2 J} = A_{j_2 j_2 J}$ is given by the positive square root and $|j_1 - j_2| \le J \le j_1 + j_2$.

[27] The required summation is obtained by using, for $K \ge p$, $L \ge 0$, $\sum_{n=0}^{p} \frac{(K-n)!(L+n)!}{n!(p-n)!} = \frac{L!(K-p)!(K+L+1)!}{p!\,(K+L-p+1)!}$, here $K - p = j_2 - j_1 + J$, $L = j_1 - j_2 + J$, $p = j_1 + j_2 - J$.

3.7.3 Special Cases of Clebsch-Gordan Coefficients

A very similar discussion to that just given can be applied to obtain an expression for $\langle j_1 j_1\, j_2 m_2 | JM\rangle$. Applying J_+ to (3.157) gives a two term relation for this Clebsch-Gordan coefficient which requires

$$\langle j_1 j_1\, j_2 m_2 | JM\rangle = \left(\frac{(j_2-m_2)!\,(J+M)!}{(j_2+m_2)!\,(J-M)!}\right)^{\frac{1}{2}} B_{j_1 j_2 J}\,, \quad j_1 + m_2 = M\,. \tag{3.174}$$

$B_{j_1 j_2 J}$ can be determined in terms of $A_{j_1 j_2 J}$, as given by (3.173), by taking $M = J$ and comparing with (3.172) for $m_1 = j_1$. This gives

$$B_{j_1 j_2 J} = (2J+1)^{\frac{1}{2}} \left(\frac{(2j_1)!\,(j_2-j_1+J)!}{(j_1+j_2-J)!\,(j_1-j_2+J)!\,(j_1+j_2+J+1)!}\right)^{\frac{1}{2}}. \tag{3.175}$$

Results for Clebsch-Gordan coefficients may be derived by successively applying J_- to (3.170) with (3.172). The expressions thereby obtained in general can not be reduced to a single term. However applying ${J_-}^J$ to (3.170) and using from (3.80), for integer j and $m \geq 0$, ${J_-}^m |jm\rangle = \sqrt{\frac{(j+m)!}{(j-m)!}}\, |j0\rangle$ then for j_1, j_2, J integers

$$\begin{aligned}
\sqrt{(2J)!}\,\langle j_1 0\, j_2 0 | J0\rangle &= \sum_{m_1=0}^{J} \binom{J}{m_1} (-1)^{j_1-m_1} \frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!}\, A_{j_1 j_2 J} \\
&= (-1)^{j_1-J} \textstyle\prod_{r=1}^{J} (j_1-j_2-J-1+2r)(j_1+j_2-J+2r)\; A_{j_1 j_2 J} \\
&= \begin{cases} 0\,, & j_1+j_2+J \text{ odd} \\ (-1)^{\frac{1}{2}(j_1+j_2-J)} \frac{(j_1-j_2+J)!\,(j_2-j_1+J)!\,(\frac{1}{2}(j_1+j_2+J))!}{(\frac{1}{2}(j_1-j_2+J))!\,(\frac{1}{2}(j_2-j_1+J))!\,(\frac{1}{2}(j_1+j_2-J))!}\, A_{j_1 j_2 J}\,, & j_1+j_2+J \text{ even}\,. \end{cases}
\end{aligned} \tag{3.176}$$

3.8 $3j$ Symbol

Besides (3.153) there are further symmetry relations for Clebsch-Gordan coefficients. Less obviously

$$\langle j_1 m_1\, j_2 m_2 | JM\rangle = C_{j_1 j_2 J}\, (-1)^{j_1-m_1} \langle j m_1\, J{-}M | j_2{-}m_2\rangle\,, \quad m_1 + m_2 = M\,. \tag{3.177}$$

The dependence on m_1, m_2, M is dictated by the recurrence relations (3.156). Applying these to (3.177), both sides agree as a consequence of $N^{\pm}_{jm} = N^{\mp}_{j-m\pm1}$ and where $(-1)^{j_1-m_1}$ provides a necessary sign flip in one term when $m_1 \to m_1 \pm 1$. The overall coefficient $C_{j_1 j_2 J}$ is determined by setting $m_1 = j_1$ and using (3.174) so that

$$C_{j_1 j_2 J} = \frac{B_{j_1 j_2 J}}{B_{j_1 J j_2}} = \left(\frac{2J+1}{2j_2+1}\right)^{\frac{1}{2}}. \tag{3.178}$$

Defining a *3j symbol* by

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{j_1-j_2-m_3}}{\sqrt{2j_3+1}} \langle j_1 m_1\, j_2 m_2 | j_3\, {-m_3}\rangle\,, \quad m_1 + m_2 + m_3 = 0\,, \tag{3.179}$$

then (3.153) and (3.177) ensure

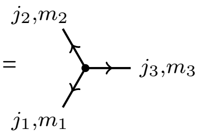
$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} . \tag{3.180}$$

These further entail that the $3j$ symbols are invariant under cyclic permutations and also for $m_i \to -m_i$ from (3.153) it follows that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} . \tag{3.181}$$

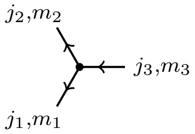
Hence, although equivalent to Clebsch-Gordan coefficients, $3j$ symbols are much nicer. They are non zero only if j_1, j_2, j_3 satisfy triangular inequalities $|j_1 - j_2| \le j_3 \le j_1 + j_2$ and $j_1 + j_2 + j_3$ is an integer. These symmetries are reflected by $A_{Jj_2j_1}/(2j_1+1)^{\frac{1}{2}} = B_{j_1j_2J}/(2J+1)^{\frac{1}{2}} = B_{j_1Jj_2}/(2j_2+1)^{\frac{1}{2}}$.

Many results for $3j$ symbols can be simply expressed diagrammatically

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} =$$ 

, (3.182)

where the diagram is invariant under rotation and for a non zero result $\sum_i m_i = 0$. For later convenience we also define

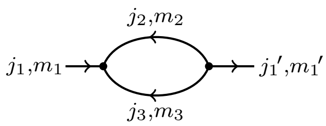


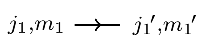
$$= (-1)^{j_3+m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} , \tag{3.183}$$

and similarly for other lines. With this convention from (3.181)  = .

The orthogonality conditions in (3.155), using (3.181), can then be reexpressed in the form

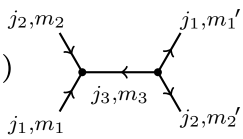
$$\sum_{m_2,m_3} (-1)^{\sum(j_i+m_i)} \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \begin{pmatrix} j_1{}' & j_2 & j_3 \\ m_1{}' & m_2 & m_3 \end{pmatrix} = \frac{1}{2j_1+1} \delta_{j_1j_1{}'} \delta_{m_1m_1{}'} ,$$

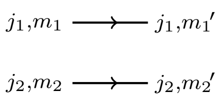
$\sum_{m_2,m_3}$ 

$= \frac{1}{2j_1+1}$ 

,

$$\sum_{j_3,m_3} (2j_3+1)(-1)^{\sum_i(j_i+m_i)} \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1{}' & m_2{}' & m_3 \end{pmatrix} = \delta_{m_1m_1{}'} \delta_{m_2m_2{}'} ,$$

$\sum_{j_3,m_3} (2j_3+1)$ 

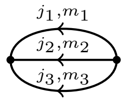
= 

. (3.184)

For each summation $\sum_m (-1)^{j+m} |j{-}m\rangle |jm\rangle$ generates an angular momentum singlet.

The essential normalisation of the $3j$ symbols is then diagrammatically expressed as

$$\textstyle\sum_{m_i}(-1)^{\sum_i(j_i+m_i)}\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \sum_{m_i} \qquad = 1\,. \tag{3.185}$$



As a special case from (3.173) and (3.176), for j_1, j_2, j_3 integers,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ 0 & 0 & 0 \end{pmatrix} = (-1)^J\left(\frac{(2k_1)!\,(2k_2)!\,(2k_3)!}{(2J+1)!}\right)^{\frac{1}{2}}\frac{J!}{k_1!\,k_2!\,k_3!}\,,\quad j_1{+}j_2{+}j_3 = 2J\,,\ k_i = J{-}j_i\,, \tag{3.186}$$

for J an integer.

If we consider the tensor product space $\mathcal{V}_{j_1}\otimes\mathcal{V}_{j_2}\otimes\mathcal{V}_{j_3}$, then so long as j_1, j_2, j_3 obey the required conditions, we may form a singlet state by

$$|0\,0\rangle = \sum_{m_1,m_2,m_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}|j_1\,m_1\rangle_1|j_2\,m_2\rangle_2|j_3\,m_3\rangle_3\,. \tag{3.187}$$

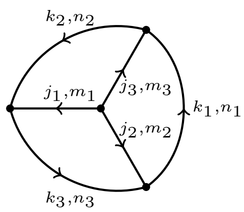
This is a singlet since, coupling first $|j_1\,m_1\rangle_1|j_2\,m_2\rangle_2$ to form a state $|j_3{-}m_3\rangle$, we have $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1-j_2+j_3}\langle j_1m_1\,j_2m_2|j_3\,{-}m_3\rangle\langle j_3{-}m_3\,j_3m_3|00\rangle$ using (3.169).

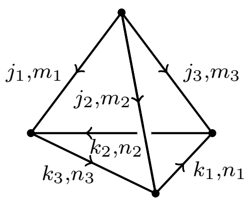
3.9 $6j$ Symbol

$3j$ symbols can be combined to form invariants which can be represented by vacuum graphs with trivalent vertices. The simplest defines the normalisation as in (3.185). The graph with 4 vertices defines the *6j symbol* as in

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} = \sum_{m_i,n_i}(-1)^{\sum_i(j_i+m_i+k_i+n_i)}\begin{pmatrix} j_1 & k_2 & k_3 \\ -m_1 & n_2 & -n_3 \end{pmatrix}\begin{pmatrix} k_1 & j_2 & k_3 \\ -n_1 & -m_2 & n_3 \end{pmatrix}\begin{pmatrix} k_1 & k_2 & j_3 \\ n_1 & -n_2 & -m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

$$= \qquad \text{or} \qquad , \tag{3.188}$$





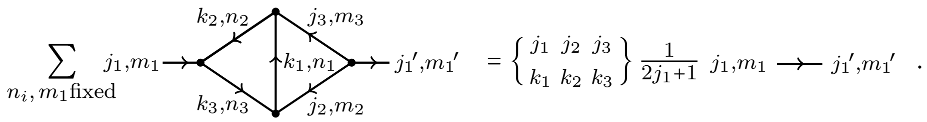
where the sums over n_i, m_i are constrained by the 4 linear conditions necessary for non vanishing $3j$ symbols. Due to these constraints there are just two non trivial summations. This is non zero so long as j_i, k_i satisfy the triangle inequalities corresponding to them being the edge lengths of the dual tetrahedron, obtained in the above picture by tasking $j_i \leftrightarrow k_i$ for each i, in Euclidean space and the sum of the three j_i, k_i incident at each vertex is an integer. Since $\sum_i m_i = 0$ in any non zero contribution the associated sign factor can be dropped. Using the symmetries in (3.180) and (3.181)

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_3 \\ k_2 & k_1 & k_3 \end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & j_2 \\ k_3 & k_1 & k_2 \end{Bmatrix} = \begin{Bmatrix} k_1 & k_2 & j_3 \\ j_1 & j_2 & k_3 \end{Bmatrix}\,. \tag{3.189}$$

so that the $6j$ symbol is invariant under permutations of columns and interchanging the upper and lower elements of any two columns. These correspond to any permutations of vertices in the figure in (3.188) so they generate the tetrahedral symmetry group $\mathcal{S}_4$.

Directly from (3.188)

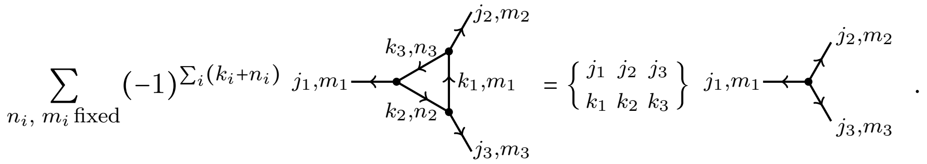
$$\sum_{n_i,\, m_1\,\text{fixed}} (-1)^{\sum_i (j_i+m_i+k_i+n_i)} \begin{pmatrix} j_1 & k_2 & k_3 \\ -m_1 & -n_2 & n_3 \end{pmatrix} \begin{pmatrix} k_1 & j_2 & k_3 \\ n_1 & -m_2 & -n_3 \end{pmatrix} \begin{pmatrix} k_1 & k_2 & j_3 \\ -n_1 & n_2 & -m_3 \end{pmatrix} \begin{pmatrix} j_1{}' & j_2 & j_3 \\ m_1{}' & m_2 & m_3 \end{pmatrix}$$
$$= \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} \frac{1}{2j_1+1}\, \delta_{j_1 j_1{}'} \delta_{m_1 m_1{}'} \, ,$$



(3.190)

Multiplying with $(2j_1+1)(-1)^{j_1+m_1} \begin{pmatrix} j_1{}' & j_2 & j_3 \\ m_1{}' & m_2{}' & m_3{}' \end{pmatrix}$ and summing $j_1{}', m_1{}'$, using (3.184), gives the vertex relation

$$\sum_{n_i,\, m_i\,\text{fixed}} (-1)^{\sum_i (k_i+n_i)} \begin{pmatrix} j_1 & k_2 & k_3 \\ m_1 & n_2 & -n_3 \end{pmatrix} \begin{pmatrix} k_1 & j_2 & k_3 \\ -n_1 & m_2 & n_3 \end{pmatrix} \begin{pmatrix} k_1 & k_2 & j_3 \\ n_1 & -n_2 & m_3 \end{pmatrix} = \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} ,$$

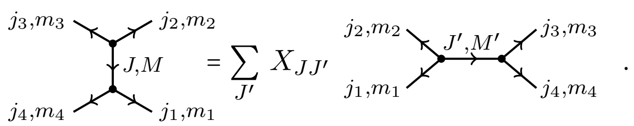


(3.191)

By using the orthogonality relations further this leads to the crossing relations

$$(-1)^{J+M} \begin{pmatrix} j_1 & j_4 & J \\ m_1 & m_4 & -M \end{pmatrix} \begin{pmatrix} j_3 & j_2 & J \\ m_3 & m_2 & M \end{pmatrix} = \sum_{J'} X_{JJ'}\, (-1)^{J'+M'} \begin{pmatrix} j_1 & j_2 & J' \\ m_1 & m_2 & M' \end{pmatrix} \begin{pmatrix} j_3 & j_4 & J' \\ m_3 & m_4 & -M' \end{pmatrix} ,$$
$$M = m_1 + m_4 \, , \ \ M' = m_3 + m_4 \, , \ \ \textstyle\sum_i m_i = 0 \, , \quad X_{JJ'} = (2J'+1)(-1)^{2j_4} \begin{Bmatrix} j_1 & j_2 & J \\ j_3 & j_4 & J' \end{Bmatrix} , \qquad (3.192)$$

where the summation over J' is constrained by j_i, J, J' forming a tetrahedron, $J' \geq |M'|$ and $\sum_{i=1}^{4} j_i$ is an integer. Diagrammatically

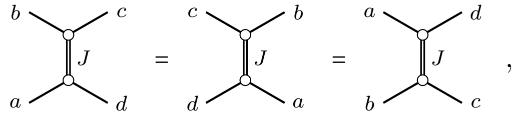


(3.193)

The $6j$ symbols play a crucial role in the decomposition of three spins. Tensor products are associative so that $\mathcal{V}_{j_1} \otimes (\mathcal{V}_{j_2} \otimes \mathcal{V}_{j_3}) \simeq (\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2}) \otimes \mathcal{V}_{j_3}$ but different bases for given total angular momentum j_4 are formed according to whether we decompose $\mathcal{V}_{j_2} \otimes \mathcal{V}_{j_3} = \bigoplus_J \mathcal{V}_J$ or $\mathcal{V}_{j_1} \otimes \mathcal{V}_{j_2} = \bigoplus_{J'} \mathcal{V}_{J'}$. The $6j$ symbols relate the two bases. Generalisations of $6j$ symbols can be obtained by considering further vacuum diagrams with trivalent vertices which are not reducible using $6j$ symbol relations.

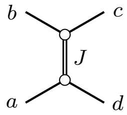
3.9.1 Crossing Relations

The relation (3.193) leads to identities for $6j$ symbols. We first consider such relations more generically where we consider couplings between four vectors, labelled by a, b, c, d, which are expressible diagrammatically, where here the lines do not carry any arrows, as

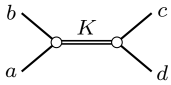


$$,\tag{3.194}$$

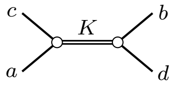
corresponding to a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry. The crossing equations then take the form



$$= \sum_K X^{abcd}_{JK}$$



$$= \sum_K Y^{abcd}_{JK}$$



$$.\tag{3.195}$$

Consistency with (3.194) requires

$$X^{abcd}_{JK} = X^{badc}_{JK} = X^{dcba}_{JK}\,, \qquad Y^{abcd}_{JK} = Y^{dcba}_{JK} = Y^{cdab}_{JK}\,. \tag{3.196}$$

Applying (3.195) repeatedly gives the crossing relations

$$\textstyle\sum_K X^{abcd}_{JK} X^{cbad}_{KL} = \delta_{JL}\,, \quad \sum_{K,L} Y^{abcd}_{JK} Y^{bcad}_{KL} Y^{cabd}_{LM} = \delta_{JM}\,, \tag{3.197}$$

with the condition

$$\textstyle\sum_{K,L} X^{abcd}_{JK}\, Y^{cbad}_{KL} Y^{bacd}_{LM} = \sum_K Y^{abcd}_{JK}\, X^{bcad}_{KM}\,. \tag{3.198}$$

These ensure X and Y generate $\mathcal{S}_3$ corresponding to permutations of a, b, c. For $\mathcal{S}_3$ defined abstractly by elements a, b such that $a^3 = b^2 = e,\ ba^2 = ab$ it is evident that these relations correspond to (3.197) and (3.198) with $Y \sim a,\ X \sim b$.

In terms of $6j$ symbols these results are applicable by taking

$$X^{j_1j_2j_3j_4}_{JK} = (2K+1)(-1)^{2j_4} \begin{Bmatrix} j_1 & j_2 & K \\ j_3 & j_4 & J \end{Bmatrix}, \quad Y^{j_1j_2j_3j_4}_{JK} = (2K+1)(-1)^{j_2+j_3+2j_4+J} \begin{Bmatrix} j_1 & j_3 & K \\ j_2 & j_4 & J \end{Bmatrix}. \tag{3.199}$$

(3.196) is modified to

$$\begin{aligned} X^{j_1j_2j_3j_4}_{JK} &= (-1)^{2K} X^{j_2j_1j_4j_3}_{JK} = (-1)^{2J} X^{j_4j_3j_2j_1}_{JK}\,, \\ Y^{j_1j_2j_3j_4}_{JK} &= (-1)^{2J} Y^{j_4j_3j_2j_1}_{JK} = (-1)^{2J+2K} Y^{j_3j_4j_1j_2}_{JK}\,, \end{aligned} \tag{3.200}$$

and (3.197) reduces to

$$\begin{aligned} &\sum_K (2J+1)(2K+1) \begin{Bmatrix} j_1 & j_2 & K \\ j_3 & j_4 & J \end{Bmatrix} \begin{Bmatrix} j_3 & j_2 & L \\ j_1 & j_4 & K \end{Bmatrix} = \delta_{JL}\,, \\ &\sum_{K,L} (-1)^{J+K+L} (2J+1)(2K+1)(2L+1) \begin{Bmatrix} j_1 & j_3 & K \\ j_2 & j_4 & J \end{Bmatrix} \begin{Bmatrix} j_2 & j_1 & L \\ j_3 & j_4 & K \end{Bmatrix} \begin{Bmatrix} j_3 & j_2 & M \\ j_1 & j_4 & L \end{Bmatrix} = \delta_{JM}\,. \end{aligned} \tag{3.201}$$

Subject to (3.189) these are equivalent to standard identities for $6j$ symbols. Furthermore

$$\sum_K Y^{j_1j_2j_3j_4}_{JK} X^{j_2j_2j_1j_4}_{KL} = (-1)^{j_2+j_3+J} \delta_{JL}\,. \tag{3.202}$$

3.10 Tensor Products and Characters

The decomposition of tensor products can equally be determined in terms of the characters given in (3.132)

$$\begin{aligned}\chi_{j_1}(\theta)\,\chi_{j_2}(\theta) = \chi_{j_1}(\theta)\sum_{m=-j_2}^{j_2} e^{-im\theta} &= \frac{1}{2i\,\sin\frac{1}{2}\theta}\sum_{m=-j_2}^{j_2}\big(e^{(j_1+m+\frac{1}{2})\theta} - e^{(-j_1+m+\frac{1}{2})\theta}\big)\\ &= \sum_{j=j_1-j_2}^{j_1+j_2}\chi_j(\theta) = \sum_{j=|j_1-j_2|}^{j_1+j_2}\chi_j(\theta)\,,\end{aligned} \tag{3.203}$$

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 (2.85) the result of this character calculation of course matches the tensor product decomposition given in (3.152).

For the symmetric and antisymmetric tensor products of the j representation from (2.92)

$$\begin{aligned}\chi_{\vee^2\mathcal{V}_j}(\theta) &= \tfrac{1}{2}\big(\chi_j(\theta)^2 + \chi_j(2\theta)\big) = \sum_{n=0}^{\lfloor j\rfloor}\chi_{2j-2n}(\theta)\,,\\ \chi_{\wedge^2\mathcal{V}_j}(\theta) &= \tfrac{1}{2}\big(\chi_j(\theta)^2 - \chi_j(2\theta)\big) = \sum_{n=0}^{\lfloor j\rfloor}\chi_{2j-1-2n}(\theta)\,,\end{aligned} \tag{3.204}$$

using (3.135). This of course agrees with (3.154).

The results can be extended to three fold tensor products

$$\begin{aligned}\chi_{j_1}(\theta)\,\chi_{j_2}(\theta)\,\chi_{j_2}(\theta) &= \sum_{J=j_1-j_2-j_3}^{j_1+j_2+j_3} N_J\,\chi_J(\theta)\,,\\ N_J &= \min\{2j_2+1, 2j_3+1, j_1+j_2+j_3-J+1, j_2+j_3-j_1+J+1\}\,.\end{aligned} \tag{3.205}$$

Although not manifestly symmetric cancellations arising from terms in the sum for $J < 0$ ensure it is so. For equal integer j

$$\chi_j(\theta)^3 = \sum_{J=0}^{j}(2J+1)\chi_J + \sum_{J=j+1}^{3j}(3j-J+1)\chi_J\,. \tag{3.206}$$

For j half integer a similar result obtains but with the first sum starting at $J = \frac{1}{2}$. From (2.92)

$$\left.\begin{matrix}\chi_{\vee^3\mathcal{V}_j}(\theta)\\ \chi_{\wedge^2\mathcal{V}^j}(\theta)\end{matrix}\right\} = \tfrac{1}{6}\big(\chi_{\mathcal{V}_j}(\theta)^3 \pm 3\,\chi_{\mathcal{V}_j}(\theta)\chi_{\mathcal{V}_j}(2\theta) + 2\,\chi_{\mathcal{V}_j}(3\theta)\big)\,, \tag{3.207}$$

giving, for j an integer,

$$\begin{aligned}\chi_{\vee^3\mathcal{V}_j} &= \sum_{k=0}^{2j}\big(1+\lfloor\tfrac{1}{6}k\rfloor\big)\chi_{3j-k} + \sum_{k=0}^{j-1}\big(1+\lfloor\tfrac{1}{3}k\rfloor\big)\chi_k - \sum_{k=0}^{\lfloor\frac{1}{3}(j-1)\rfloor}\chi_{3j-6k-1} - \sum_{k=0}^{\lfloor\frac{1}{2}(j-1)\rfloor}\chi_{j-2k-1}\,,\ j\ge 1\,,\\ \chi_{\wedge^3\mathcal{V}_j} &= \sum_{k=0}^{2j-2}\big(1+\lfloor\tfrac{1}{6}k\rfloor\big)\chi_{3j-k-3} + \sum_{k=0}^{j-2}\big(1+\lfloor\tfrac{1}{3}k\rfloor\big)\chi_k - \sum_{k=0}^{\lfloor\frac{1}{3}(j-2)\rfloor}\chi_{3j-6k-4} - \sum_{k=0}^{\lfloor\frac{1}{2}(j-2)\rfloor}\chi_{j-2k-2}\,,\ j\ge 2\,,\end{aligned} \tag{3.208}$$

while $\chi_{\wedge^3\mathcal{V}_1} = \chi_0$. For half integer j

$$\begin{aligned}
\chi_{\vee^3\mathcal{V}_j} &= \sum_{k=0}^{2j}\big(1+\lfloor\tfrac{1}{6}k\rfloor\big)\chi_{3j-k} + \sum_{k=\frac{3}{2}}^{j-1}\big(1+\lfloor\tfrac{1}{3}(k-1)\rfloor\big)\chi_k - \sum_{k=0}^{\lfloor\frac{1}{3}j\rfloor}\chi_{3j-6k-1}\,, \quad j\ge\tfrac{5}{2}\,,\\
\chi_{\wedge^3\mathcal{V}_j} &= \sum_{k=0}^{2j-3}\big(1+\lfloor\tfrac{1}{6}k\rfloor\big)\chi_{3j-k-3} + \sum_{k=\frac{3}{2}}^{j-1}\big(1+\lfloor\tfrac{1}{3}(k-1)\rfloor\big)\chi_k - \sum_{k=0}^{\lfloor\frac{1}{3}(j-2)\rfloor}\chi_{3j-6k-4}\,, \quad j\ge\tfrac{5}{2}\,,
\end{aligned} \tag{3.209}$$

with, for $j=\frac{3}{2}$, $\chi_{\vee^3\mathcal{V}_j} = \chi_{\frac{3}{2}} + \chi_{\frac{5}{2}} + \chi_{\frac{9}{2}}$, $\chi_{\wedge^3\mathcal{V}_j} = \chi_{\frac{3}{2}}$ and for $j=\frac{1}{2}$, $\chi_{\vee^3\mathcal{V}_j} = \chi_{\frac{3}{2}}$, $\chi_{\wedge^3\mathcal{V}_j} = 0$.

3.11 $SO(3)$ 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

$$V_i \underset{R}{\to} V'_i = R_{ij}V_j\,. \tag{3.210}$$

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

$$T_{i_1\dots i_n} \underset{R}{\to} T'_{i_1\dots i_n} = R_{i_1j_1}\dots R_{i_nj_n}T_{j_1\dots j_n}\,. \tag{3.211}$$

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_iV_j is a rank two tensor, and also that contraction of indices preserves tensorial properties essential because for any two vectors U_iV_i is a scalar and invariant under rotations, $U'_iV'_i = U_iV_i$. 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}

$$\delta'_{ij} = R_{ik}R_{jk} = \delta_{ij}\,, \tag{3.212}$$

as a consequence of the orthogonality property (3.1), and also the ε-symbol

$$\varepsilon'_{ijk} = R_{ij}R_{jm}R_{kn}\,\varepsilon_{lmn} = \det R\,\varepsilon_{ijk} = \varepsilon_{ijk}\,, \tag{3.213}$$

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

$$T_{ij} = S_{ij} + \varepsilon_{ijk}V_k + \tfrac{1}{3}\delta_{ij}\,T_{kk}\,, \tag{3.214}$$

for

$$S_{ij} = T_{(ij)} - \tfrac{1}{3}\delta_{ij}\,T_{kk}\,, \qquad V_i = \tfrac{1}{2}\varepsilon_{ijk}T_{jk}\,. \tag{3.215}$$

Each term in (3.214) 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

$$\delta_{i_r i_s}S_{i_1\dots i_n} = 0\,, \quad \varepsilon_{j i_r i_s}S_{i_1\dots i_n} = 0\,, \quad \text{for all} \quad r,s,\ 1 \le r < s \le n\,. \tag{3.216}$$

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

$$S_{i_1\dots i_n} = S_{(i_1\dots i_n)}\,, \tag{3.217}$$

and also traceless on any pair of indices. With the symmetry condition (3.217) it is sufficient to require just

$$S_{i_1\dots i_{n-2}jj} = 0\,. \tag{3.218}$$

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 (3.217), belonging to the symmetrised n-fold tensor product, $\mathrm{sym}(\mathcal{V}\otimes\dots\otimes\mathcal{V})$. 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

$$S_{\underbrace{1\dots1}_{r}\,\underbrace{2\dots2}_{s}\,\underbrace{3\dots3}_{t}} \quad \text{where} \quad r,s,t \ge 0\,,\ r+s+t = n\,. \tag{3.219}$$

Independent rank n symmetric tensors may then be counted by counting all r,s,t satisfying the conditions in (3.219), hence this gives

$$\dim\big(\mathrm{sym}(\underbrace{\mathcal{V}\otimes\dots\otimes\mathcal{V}}_{n})\big) = \tfrac{1}{2}(n+1)(n+2)\,. \tag{3.220}$$

To take the traceless conditions (3.218) 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

$$\dim\mathcal{V}_n = \tfrac{1}{2}(n+1)(n+2) - \tfrac{1}{2}(n-1)n = 2n+1\,. \tag{3.221}$$

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

For any rank n symmetric tensor $S_{i_1\dots i_n}$ there is a one to one correspondence with homogeneous polynomials of degree n in $\mathbf{x}$,

$$S_{i_1\dots i_n} \quad \leftrightarrow \quad S^{(n)}(\mathbf{x}) = S_{i_1\dots i_n} x_{i_1} \dots x_{i_n} \,, \tag{3.222}$$

and

$$S_{i_1\dots i_{n-2}jj} = 0 \quad \Leftrightarrow \quad \nabla^2 S^{(n)}(\mathbf{x}) = 0 \,, \tag{3.223}$$

where $S^{(n)}$ is then a *harmonic* function. As a particular case we have

$$S^{(n)}(\mathbf{x}) = (\mathbf{t}\cdot\mathbf{x})^n \quad \text{for} \quad \mathbf{t}^2 = 0 \,, \tag{3.224}$$

where $\mathbf{t}$ is any complex null vector. Since from (3.55)

$$\mathbf{L}^2 = -\mathbf{x}^2\nabla^2 + \mathbf{x}\cdot\nabla\big(\mathbf{x}\cdot\nabla + 1\big) \,. \tag{3.225}$$

and $\mathbf{x}\cdot\nabla S^{(n)}(\mathbf{x}) = n\, S^{(n)}(\mathbf{x})$ we have for harmonic polynomials

$$\mathbf{L}^2 S^{(n)}(\mathbf{x}) = n(n+1)\, S^{(n)}(\mathbf{x}) \,, \tag{3.226}$$

so that symmetric traceless tensors or harmonic polynomials of degree n correspond to angular momentum n for any integer n. Clearly $S^{(n)}(\mathbf{x}) = |\mathbf{x}|^n S^{(n)}(\hat{\mathbf{x}})$ for $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ and as $\mathbf{L}|\mathbf{x}| = 0$ (3.226) reduces to

$$\mathbf{L}^2 S^{(n)}(\hat{\mathbf{x}}) = n(n+1)\, S^{(n)}(\hat{\mathbf{x}}) \,, \tag{3.227}$$

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 (3.214) and (3.215). 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

$$\begin{aligned} & S_{1,(i_1\dots i_{n-r}\, j_1\dots j_r}\, S_{2,i_{n-r+1}\dots i_{n+m-2r})\, j_1\dots j_r} \,, && r = 0,\dots m \,, \\ & \varepsilon_{jk(i_1}\, S_{1,i_2\dots i_{n-r}\, j_1\dots j_r\, j}\, S_{2,i_{n-r+1}\dots i_{n+m-1-2r})\, j_1\dots j_r\, k} \,, && r = 0,\dots m-1 \,. \end{aligned} \tag{3.228}$$

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 (3.215) 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

$$U[R] V_i U[R]^{-1} = (R^{-1})_{ij} V_j \,, \tag{3.229}$$

as in (3.58), while for a rank n tensor operator

$$U[R] T_{i_1\dots i_n} U[R]^{-1} = (R^{-1})_{i_1 j_1} \dots (R^{-1})_{i_n j_n} T_{j_1\dots j_n} \,. \tag{3.230}$$

These may be decomposed into irreducible tensor operators as above. For infinitesimal rotations as in (3.16), with $U[R]$ correspondingly given by (3.50), then (3.229) gives

$$[J_i, V_j] = i\,\varepsilon_{ijk} V_k\,, \tag{3.231}$$

which is an alternative definition of a vector operator. From (3.230) we similarly get

$$[J_i, T_{j_1 j_2 \dots j_n}] = i\,\varepsilon_{ij_1 k} T_{k j_2 \dots j_n} + i\,\varepsilon_{ij_2 k} T_{j_1 k \dots j_n} + \dots + i\,\varepsilon_{ij_n k} T_{j_1 j_2 \dots k}\,. \tag{3.232}$$

The operators $\mathbf{x}, \mathbf{p}$ are examples of vector operators for the angular momentum operator given by $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ where $[x_i, p_j] = i\delta_{ij}$.

3.11.1 Spherical Harmonics

Rank n symmetric traceless tensors are directly related to *spherical harmonics*. If we choose an orthonormal basis for such tensors $S^{(n,m)}_{i_1 \dots i_n}$, labelled by m taking $2n{+}1$ values and satisfying $S^{(n,m)} \cdot S^{(n,m')} \propto \delta^{mm'}$, 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

$$Y_{nm}(\hat{\mathbf{x}}) = S^{(n,m)}_{i_1 \dots i_n}\, \hat{x}_{i_1} \dots \hat{x}_{i_n}\,. \tag{3.233}$$

For a standard basis we require m is an integer with $-n \le m \le n$ and

$$L_\pm Y_{nm}(\hat{\mathbf{x}}) = N^\pm_{nm}\, Y_{n\,m\pm1}(\hat{\mathbf{x}})\,, \qquad L_3 Y_{nm}(\hat{\mathbf{x}}) = m\, Y_{nm}(\hat{\mathbf{x}})\,, \tag{3.234}$$

where $L_\pm$, L_3 are the angular momentum operators acting on functions of $\hat{\mathbf{x}}$ and $N^\pm_{nm}$ is defined in (3.79). Defining

$$x = \hat{x}_1 + i\hat{x}_2, \quad \bar{x} = \hat{x}_1 - i\hat{x}_2, \quad z = \hat{x}_3\,, \qquad x\bar{x} + z^2 = 1\,, \tag{3.235}$$

then

$$L_3 = x\frac{\partial}{\partial x} - \bar{x}\frac{\partial}{\partial \bar{x}}\,, \qquad L_+ = -x\frac{\partial}{\partial z} + 2z\frac{\partial}{\partial \bar{x}} \qquad L_- = \bar{x}\frac{\partial}{\partial z} - 2z\frac{\partial}{\partial x}\,. \tag{3.236}$$

In terms of usual spherical polar coordinates $z = \cos\theta$, $x = \sin\theta\, e^{i\phi}$, $\bar{x} = \sin\theta\, e^{-i\phi}$.

Spherical harmonics can be expressed in the form

$$Y_{nm}(\hat{\mathbf{x}}) = \begin{cases} (-x)^m p_{nm}(z)\,, & m \ge 0\,, \\ \bar{x}^{|m|}\, p_{n|m|}(z)\,, & m < 0\,, \end{cases} \tag{3.237}$$

where $p_{nm}(z)$, $0 \le m \le n$, is a polynomial of degree $n - m$ and $Y_{nm}(\hat{\mathbf{x}})^* = (-1)^m Y_{n-m}(\hat{\mathbf{x}})$ This expression automatically satisfies the L_3 equation in (3.234) and from the $L_\pm$ equations

$$\frac{\mathrm{d}}{\mathrm{d}z} p_{nm}(z) = N^+_{nm}\, p_{nm+1}(z)\,, \quad -(1-z^2)\frac{\mathrm{d}}{\mathrm{d}z} p_{nm}(z) + 2m\, z\, p_{nm}(z) = N^-_{nm}\, p_{nm-1}(z)\,. \tag{3.238}$$

Hence, since $N^-_{nm}\, N^+_{nm-1} = (n+m)(n-m+1)$,

$$\Big(\frac{\mathrm{d}}{\mathrm{d}z}(1-z^2)\frac{\mathrm{d}}{\mathrm{d}z} - 2m\, z\frac{\mathrm{d}}{\mathrm{d}z} + (n-m)(n+m+1)\Big) p_{nm}(z) = 0\,. \tag{3.239}$$

Regular solutions are given in terms of Gegenbauer polynomials which may be defined in terms of the generating function

$$\big(1-2z\,r+r^2\big)^{-\lambda} = \textstyle\sum_{n=0}^{\infty} C_n{}^{\lambda}(z)\, r^n\,, \qquad C_n{}^{\lambda}(-z) = (-1)^n C_n{}^{\lambda}(z)\,. \tag{3.240}$$

Since $C_n{}^{\lambda\prime}(z) = 2\lambda\, C_{n-1}^{\lambda+1}(z)$, $((2m-1)!! = (2m)!/(2^m m!) = (2m-1)(2m-3)\dots 1)$,

$$p_{nm}(z) = a_n\,(2m-1)!!\sqrt{\tfrac{(n-m)!}{(n+m)!}}\; C_{n-m}^{m+\frac{1}{2}}(z)\,. \tag{3.241}$$

Spherical harmonics are also expressible in terms of a generating function

$$e^{\mathbf{t}\cdot\hat{\mathbf{x}}}\,, \qquad \mathbf{t}\cdot\hat{\mathbf{x}} = v\big(z - \tfrac{1}{2}\lambda\, x + \tfrac{1}{2\lambda}\,\bar{x}\big)\,, \tag{3.242}$$

with $\hat{\mathbf{x}}$ defined in terms of x, $\bar{x}$, z as in (3.235). With (3.236) we have

$$L_3\, e^{\mathbf{t}\cdot\hat{\mathbf{x}}} = \lambda\frac{\partial}{\partial\lambda}\, e^{\mathbf{t}\cdot\hat{\mathbf{x}}}\,, \quad L_+\, e^{\mathbf{t}\cdot\hat{\mathbf{x}}} = \frac{1}{\lambda}\bigg(v\frac{\partial}{\partial v} + \lambda\frac{\partial}{\partial\lambda}\bigg)e^{\mathbf{t}\cdot\hat{\mathbf{x}}}\,, \quad L_-\, e^{\mathbf{t}\cdot\hat{\mathbf{x}}} = \lambda\bigg(v\frac{\partial}{\partial v} - \lambda\frac{\partial}{\partial\lambda}\bigg)e^{\mathbf{t}\cdot\hat{\mathbf{x}}}\,, \tag{3.243}$$

then as $\mathbf{t}$, as defined by (3.242), is a null vector so that, for arbitrary v,λ, $(\mathbf{t}\cdot\hat{\mathbf{x}})^n$ is a harmonic function we can expand the exponential in the form

$$e^{\mathbf{t}\cdot\hat{\mathbf{x}}} = \textstyle\sum_{n=0}^{\infty}\sum_{m=-n}^{n} b_{nm}\, v^n \lambda^m\, Y_{nm}(\hat{\mathbf{x}})\,. \tag{3.244}$$

(3.234) is satisfied so long as

$$(n+m+1)b_{nm+1} = N^+_{nm}\, b_{nm}\,, \quad (n-m+1)b_{nm-1} = N^-_{nm}\, b_{nm}\,, \tag{3.245}$$

or

$$b_{nm} = \frac{b_n}{\sqrt{(n-m)!(n+m)!}}\,. \tag{3.246}$$

The coefficients b_n, or a_n in (3.241), depend on the normalisation of Y_{nm} for differing n. It is conventional to impose

$$\int_{S^2}\!\mathrm{d}\Omega\; Y_{nm}(\hat{\mathbf{x}})\, Y_{n'm'}(\hat{\mathbf{x}})^* = \delta_{nn'}\delta_{mm'}\,, \quad \text{or} \quad \int_{S^2}\!\mathrm{d}\Omega\; Y_{nm}(\hat{\mathbf{x}})\, Y_{n'm'}(\hat{\mathbf{x}}) = (-1)^m\delta_{nn'}\delta_{m-m'}\,. \tag{3.247}$$

On integration over S^2 $L_3{}^\dagger = L_3$, $L_+{}^\dagger = L_-$.

The required normalisation to ensure (3.247) can be derived by considering the three dimensional integrals

$$\int \mathrm{d}^3x\; e^{-\mathbf{x}^2+\mathbf{k}\cdot\mathbf{x}} = \pi^{\frac{3}{2}}\, e^{\frac{1}{4}\mathbf{k}^2} \quad \Rightarrow \quad \int \mathrm{d}^3x\; e^{-\mathbf{x}^2}(\mathbf{k}\cdot\mathbf{x})^{2n} = \pi^{\frac{3}{2}}\,\frac{(2n)!}{2^{2n}n!}\,(\mathbf{k}^2)^n\,. \tag{3.248}$$

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

$$\int_{S^2}\!\mathrm{d}\Omega\;(\mathbf{k}\cdot\hat{\mathbf{x}})^{2n} = 4\pi\,\frac{(2n)!}{2^{2n}n!}\,\frac{1}{(\frac{3}{2})_n}\,(\mathbf{k}^2)^n = \frac{4\pi}{2n+1}\,(\mathbf{k}^2)^n\,, \quad (\tfrac{3}{2})_n = \tfrac{3}{2}\cdot\tfrac{5}{2}\dots(\tfrac{1}{2}+n)\,. \tag{3.249}$$

If now $\mathbf{k} = \mathbf{t}_1 + \mathbf{t}_2$ with $\mathbf{t}_1{}^2 = \mathbf{t}_2{}^2 = 0$, so that $\mathbf{k}^2 = 2\,\mathbf{t}_1 \cdot \mathbf{t}_2$, then

$$\int_{S^2} \mathrm{d}\Omega\; (\mathbf{t}_1 \cdot \hat{\mathbf{x}})^n (\mathbf{t}_2 \cdot \hat{\mathbf{x}})^{n'} = \delta_{nn'}\, 4\pi\, \frac{n!^2}{(2n+1)!}\, (2\,\mathbf{t}_1 \cdot \mathbf{t}_2)^n \,. \tag{3.250}$$

For $\mathbf{t}_1$, $\mathbf{t}_2$ as defined in (3.242) and expanding $\int_{S^2} \mathrm{d}\Omega\; e^{\mathbf{t}_1 \cdot \hat{\mathbf{x}} + \mathbf{t}_2 \cdot \hat{\mathbf{x}}}$

$$\sum_{n,m,n',m'} b_{nm} b_{n'm'}\, v^n \lambda^m v'^{n'} \lambda'^{m'} \int_{S^2} \mathrm{d}\Omega\; Y_{nm}(\hat{\mathbf{x}})\, Y_{n'm'}(\hat{\mathbf{x}}) = \sum_n \frac{4\pi}{(2n+1)!}\, (2\,\mathbf{t}_1 \cdot \mathbf{t}_2)^n \,, \tag{3.251}$$

where

$$2\,\mathbf{t}_1 \cdot \mathbf{t}_2 = -\frac{v_1 v_2}{\lambda_1 \lambda_2} (\lambda_1 - \lambda_2)^2 \,, \tag{3.252}$$

so that

$$(2\,\mathbf{t}_1 \cdot \mathbf{t}_2)^n = (v_1 v_2)^n \sum_{m=-n}^{n} \frac{(2n)!}{(n-m)!(n+m)!} \bigg(-\frac{\lambda_1}{\lambda_2} \bigg)^m \,. \tag{3.253}$$

Hence ensuring (3.247) gives

$$b_{nm}\, b_{n-m} = \frac{4\pi}{2n+1}\, \frac{1}{(n-m)!(n+m)!} \quad \Rightarrow \quad b_n = \bigg(\frac{4\pi}{2n+1} \bigg)^{\frac{1}{2}} \,. \tag{3.254}$$

From (3.244) we have $b_{nn} Y_{nn}(\hat{\mathbf{x}}) = \frac{1}{n!}(-\frac{1}{2}x)^n$ while from the solution given by (3.237) and (3.241) $Y_{nn}(\hat{\mathbf{x}}) = a_n (2n-1)!!/\sqrt{(2n)!}\,(-x)^n$ which requires $a_n b_n = 1$. The expansion (3.244), with b_{nm} determined by (3.246) and (3.254), is ascribed to *Herglotz*.[28]

These results can be extended to the integral of three spherical harmonics. Taking $\mathbf{k} = \mathbf{t}_1 + \mathbf{t}_2 + \mathbf{t}_3$ with $\mathbf{t}_i$ three null vectors

$$\begin{aligned} \int_{S^2} \mathrm{d}\Omega\; & (\mathbf{t}_1 \cdot \hat{\mathbf{x}})^{n_1} (\mathbf{t}_2 \cdot \hat{\mathbf{x}})^{n_2} (\mathbf{t}_3 \cdot \hat{\mathbf{x}})^{n_3} \\ & = 4\pi\, \frac{n!}{(2n+1)!}\, \frac{n_1!\, n_2!\, n_3!}{k_1!\, k_2!\, k_3!}\, (2\,\mathbf{t}_1 \cdot \mathbf{t}_2)^{k_3} (2\,\mathbf{t}_2 \cdot \mathbf{t}_3)^{k_1} (2\,\mathbf{t}_1 \cdot \mathbf{t}_3)^{k_2} \,. \\ & n_1 + n_2 + n_3 = 2n \,, \qquad k_i = n - n_i \,. \end{aligned} \tag{3.255}$$

We then have

$$\begin{aligned} & b_{n_1 m_1} b_{n_2 m_2} b_{n_3 m_3} \int_{S^2} \mathrm{d}\Omega\; Y_{n_1 m_1}(\hat{\mathbf{x}})\, Y_{n_2 m_2}(\hat{\mathbf{x}})\, Y_{n_3 m_3}(\hat{\mathbf{x}}) \\ & = \frac{4\pi}{(2n+1)!}\, \frac{n!}{k_1!\, k_2!\, k_3!} \sum_{b_1, b_2, b_3} \prod_{i=1}^{3} \frac{(-1)^{b_i} (2k_i)!}{(k_i - b_i)!\, (k_i + b_i)!} \bigg|_{b_1 - b_2 = m_3,\, b_2 - b_3 = m_1,\, b_3 - b_1 = m_2} \,. \end{aligned} \tag{3.256}$$

This is non zero only when $\sum_i m_i = 0$. For $m_i = 0$ then $b_i = b$ and the sum can be evaluated using

$$\sum_{-\min k_i \le b \le \min k_i} (-1)^b \frac{1}{\prod_{i=1}^{3} (k_i - b)!\, (k_i + b)!} = \frac{(k_1 + k_2 + k_3)!}{(k_1 + k_2)!\, (k_2 + k_3)!\, (k_3 + k_1)!\, k_1!\, k_2!\, k_3!} \,. \tag{3.257}$$

[28] Gustav Herglotz, 1881-1953, German.

Hence, since $b_{n0} = \sqrt{\frac{4\pi}{2n+1}}\Big/ n!$, and with the result (3.186)

$$\int_{S^2} \mathrm{d}\Omega \, Y_{n_1 0}(\hat{\mathbf{x}}) \, Y_{n_2 0}(\hat{\mathbf{x}}) \, Y_{n_3 0}(\hat{\mathbf{x}}) = \left(\frac{(2n_1+1)(2n_2+1)(2n_3+1)}{4\pi} \right)^{\frac{1}{2}} \begin{pmatrix} n_1 & n_2 & n_3 \\ 0 & 0 & 0 \end{pmatrix}^2 . \tag{3.258}$$

For arbitrary m_i, $\sum_i m_i = 0$, it then follows that

$$\begin{aligned} &\int_{S^2} \mathrm{d}\Omega \, Y_{n_1 m_1}(\hat{\mathbf{x}}) \, Y_{n_2 m_2}(\hat{\mathbf{x}}) \, Y_{n_3 m_3}(\hat{\mathbf{x}}) \\ &\quad = \left(\frac{(2n_1+1)(2n_2+1)(2n_3+1)}{4\pi} \right)^{\frac{1}{2}} \begin{pmatrix} n_1 & n_2 & n_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} n_1 & n_2 & n_3 \\ m_1 & m_2 & m_3 \end{pmatrix} , \end{aligned} \tag{3.259}$$

since the dependence on m_i on both sides is identical from (3.234) and the recurrence identities for $3j$ symbols following from (3.156).

3.12 Molien Series for $SU(2)$

The Molien series in (2.100) can be readily extended to the continuous group $SU(2)$ by replacing the finite sum over group elements by the corresponding invariant integegration. Since the formula only involves a sum over conjugacy classes in this case it reduces, for the representation j of dimension $2j+1$, to

$$M_{SU(2)}(\mathbb{C}^{2j+1}, t) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\theta \; (1-\cos\theta) \frac{1}{\det\big(\mathbb{1} - t\, D(\theta, \mathbf{n})\big)} , \tag{3.260}$$

where the choice of $\mathbf{n}$ is irelevant so that from (3.131) we may use

$$\det\big(\mathbb{1} - t\, D(\theta, \mathbf{z})\big) = \textstyle\prod_{m=-j}^{j} \big(1 - t\, e^{im\theta}\big) . \tag{3.261}$$

For integer spin the result reduces to the series for $SO(3)$. In this case with $z = e^{i\theta}$

$$M_{SO(3)}(\mathbb{R}^{2j+1}, t) = \frac{1}{1-t} \frac{1}{4\pi i} \oint_{|z|=1} \mathrm{d}z \, z^{\frac{1}{2}(j-1)(j+2)} \frac{(1-z)(1-z^{-1})}{\prod_{m=1}^{j} (z^m - t)(1 - t\, z^m)} . \tag{3.262}$$

For $|t| < 1$ this can be evaluated by summing the residues of the poles arising at $z^m = t$ and also for $j = 0, 1$ at $z = 0$. This gives for $j = 0, 1, 2$

$$M_{SO(3)}(\mathbb{R}, t) = \frac{1}{1-t} , \quad M_{SO(3)}(\mathbb{R}^3, t) = \frac{1}{1-t^2} , \quad M_{SO(3)}(\mathbb{R}^5, t) = \frac{1}{(1-t^2)(1-t^3)} . \tag{3.263}$$

For $(x_1, x_2, x_3) \in \mathbb{R}^3$ the fundamental invariant is just $\mathbf{x}^2$. The representation for $j = 2$ acts on the five dimensional space formed by traceless symmetric tensors T_{ij} and the fundamental primary invariants are $\mathrm{tr}(T^2)$, $\mathrm{tr}(T^3)$. For larger integer j the results become complicated and there is no simple general formula.

For half integer spin it is covenient to extend the integration to be from 0 to 4π and the with $w = e^{i\frac{1}{2}\theta}$

$$M_{SU(2)}(\mathbb{C}^{2j+1}, t) = \frac{1}{4\pi i} \oint_{|w|=1} \mathrm{d}w \, w^{(j-\frac{1}{2})(j+\frac{3}{2})} \frac{(1-w^2)(1-w^{-2})}{\prod_{m=\frac{1}{2}}^{j} (w^{2m} - t)(1 - t\, w^{2m})} . \tag{3.264}$$

As special cases for $j = \frac{1}{2}, \frac{3}{2}$

$$M_{SU(2)}(\mathbb{C}^2, t) = 1\,, \qquad M_{SU(2)}(\mathbb{C}^4, t) = \frac{1}{1 - t^4}\,. \tag{3.265}$$

There are no invariants for the two dimensional spin-$\frac{1}{2}$ but there is one of order 4 for spin-$\frac{3}{2}$.

Quantum fields with half integer spin are anticommuting so it is more natural to use (2.112) to count invariants in this case. As above this can be reduced to a contour integral analogous to (3.263) but only poles at $w = 0$ are relevant. The lowest cases are

$$\tilde{M}_{SU(2)}(\mathbb{M}^2, t) = 1 + t^2\,, \qquad \tilde{M}_{SU(2)}(\mathbb{M}^4, t) = 1 + t^2 + t^4\,. \tag{3.266}$$

There is now a quadratic invariant since $(\mathcal{V}_{\frac{1}{2}} \otimes \mathcal{V}_{\frac{1}{2}})_{\text{antisym}} \supset \mathcal{V}_0$. For $j = \frac{9}{2}$ the obvious pattern does not extend since there are then two quartic and two sextic invariants. Extending to integer j

$$\tilde{M}_{SO(3)}(\mathbb{M}^3, t) = 1 + t^3\,, \qquad \tilde{M}_{SO(3)}(\mathbb{M}^5, t) = 1 + t^5\,, \tag{3.267}$$

where the first result reflects the existence of the invariant tensor ϵ_{ijk}.

3.13 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

$$k \in \left\{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots\right\}, \tag{3.268}$$

and

$$q \in \{-k, -k+1, \dots, k-1, k\}\,, \tag{3.269}$$

for each k in (3.268), constitute a *tensor operator* of rank k if they satisfy the commutation relations

$$\begin{aligned} [J_3\,,\, T_{kq}] &= q\, T_{kq}\,, \\ [J_\pm\,,\, T_{kq}] &= N^\pm_{kq}\, T_{kq\pm 1}\,, \end{aligned} \tag{3.270}$$

with $N^\pm_{kq}$ given by (3.79). This definition is of course modelled exactly on that for the $|j\,m\rangle$ states in (3.75) and (3.77) 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

$$V_{1\pm 1} = \mp\sqrt{\tfrac{1}{2}}(V_1 \pm iV_2)\,, \qquad V_{10} = V_3\,, \tag{3.271}$$

ensures that V_{1q} satisfy (3.270) for $k=1$ as a consequence of (3.231).

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

$$T_{2q} = \sum_{m,m'} V_{1m} U_{1m'} \langle 1m\, 1m' | 2q \rangle \,, \tag{3.272}$$

with $V_{1m}, U_{1m'}$ defined as in (3.271). This gives

$$\begin{aligned} T_{22} &= V_{11}U_{11}\,, \quad T_{21} = \sqrt{\tfrac{1}{2}}\big(V_{11}U_{10} + V_{10}U_{11}\big)\,, \\ T_{20} &= \sqrt{\tfrac{1}{6}}\big(V_{11}U_{1-1} + 2V_{10}U_{10} + V_{1-1}U_{11}\big)\,, \\ T_{2-1} &= \sqrt{\tfrac{1}{2}}\big(V_{10}U_{1-1} + V_{1-1}U_{10}\big)\,, \quad T_{2-2} = V_{1-1}U_{1-1}\,. \end{aligned} \tag{3.273}$$

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[29]

$$\langle \alpha', j'\, m' | T_{kq} | \alpha, j\, m\rangle = \langle jm\, kq | j'\, m'\rangle \, C\,, \tag{3.274}$$

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

$$C = \langle \alpha' j' \| T_k \| \alpha j\rangle\,, \tag{3.275}$$

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

$$\begin{aligned} \langle \alpha', j'\, m' | T_{00} | \alpha, j\, m\rangle &= \langle jm\, 00 | j' m'\rangle \, \langle \alpha' j' \| T_0 \| \alpha j\rangle \\ &= \delta_{jj'}\, \delta_{mm'}\ \langle \alpha' j' \| T_0 \| \alpha j\rangle\,, \end{aligned} \tag{3.276}$$

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

$$\sum_{q,m} T_{kq}|\alpha, j\, m\rangle \langle kq\, jm | JM\rangle = |J\, M\rangle \tag{3.277}$$

[29] Carl Henry Eckart, 1902-1973, American.

defines a set of states $\{|J\,M\rangle\}$ satisfying, by virtue of the definition of Clebsch-Gordan coefficients in (3.144),

$$J_3|J\,M\rangle = M|J\,M\rangle\,, \qquad J_\pm|J\,M\rangle = N^\pm_{J,M}|J\,M{\pm}1\rangle\,. \tag{3.278}$$

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

$$\langle\alpha', j'm'|J\,M\rangle = C_J\,\delta_{j'J}\delta_{m'M}\,, \tag{3.279}$$

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

$$\begin{aligned}\langle\alpha', JM|J\,M\rangle N^+_{JM-1} &= \langle\alpha', JM|J_+|J\,M{-}1\rangle \\ &= \langle\alpha', JM|J_-{}^\dagger|J\,M{-}1\rangle = \langle\alpha', JM{-}1|J\,M{-}1\rangle N^-_{JM}\,.\end{aligned} \tag{3.280}$$

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 $m' = M$, (3.279) is independent of M. Inverting (3.277)

$$T_{kq}|\alpha, j\,m\rangle = \sum_{JM}|J\,M\rangle\langle kq\,jm|JM\rangle\,, \tag{3.281}$$

and then taking the matrix element with $\langle\alpha', j'm'|$ gives the Wigner-Eckart theorem, using (3.279), with $C_{j'} = \langle\alpha' j'\|T_k\|\alpha j\rangle$.

3.14 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 (3.28), and consider a vector η belonging to the r-dimensional representation space for the fundamental representation and transforming as

$$\eta_\alpha \underset{A}{\to} \eta'_\alpha = A_\alpha{}^\beta\eta_\beta\,. \tag{3.282}$$

The extension to a tensor with n indices is straightforward

$$\psi_{\alpha_1\dots\alpha_n} \underset{A}{\to} \psi'_{\alpha_1\dots\alpha_n} = A_{\alpha_1}{}^{\beta_1}\dots A_{\alpha_n}{}^{\beta_n}\psi_{\beta_1\dots\beta_n}\,, \tag{3.283}$$

Since A is unitary

$$(A_\alpha{}^\beta)^* = (A^{-1})_\beta{}^\alpha\,. \tag{3.284}$$

The complex conjugation of (3.282) defines a transformation corresponding to the conjugate representation. If we define

$$\bar\eta^\alpha = (\eta_\alpha)^*\,, \tag{3.285}$$

then using (3.284) allows the conjugate transformation rule to be written as

$$\bar\eta^\alpha \underset{A}{\to} \bar\eta'^\alpha = \bar\eta^\beta(A^{-1})_\beta{}^\alpha\,. \tag{3.286}$$

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 (3.282), each lower one as (3.286).

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

$$\varepsilon'_{\alpha\beta} = A_\alpha{}^\gamma A_\beta{}^\delta \varepsilon_{\gamma\delta} = \det A \, \varepsilon_{\alpha\beta} = \varepsilon_{\alpha\beta} \quad \text{for} \quad A \in SU(2) \,, \tag{3.287}$$

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

$$\delta'_\alpha{}^\beta = A_\alpha{}^\gamma \delta_\gamma{}^\delta (A^{-1})_\delta{}^\beta = \delta_\alpha{}^\beta \,. \tag{3.288}$$

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

$$\varepsilon_{\alpha\beta}\,\varepsilon^{\gamma\delta} = -\delta_\alpha{}^\gamma \delta_\beta{}^\delta + \delta_\alpha{}^\delta \delta_\beta{}^\gamma \,, \qquad \varepsilon_{\alpha\gamma}\varepsilon^{\gamma\beta} = \delta_\alpha{}^\beta \,. \tag{3.289}$$

Rank n tensors as in (3.283) 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 $\phi_{\alpha_1\dots\alpha_n}$ it is sufficient to impose $\varepsilon^{\alpha_r\alpha_s}\phi_{\alpha_1\dots\alpha_n} = 0$ for all $r < s$. The irreducible tensors must then be totally symmetric $\phi_{\alpha_1\dots\alpha_n} = \phi_{(\alpha_1\dots\alpha_n)}$. To count these we may restrict to those of the form

$$\phi_{\underbrace{1\dots1}_{r}\,\underbrace{2\dots2}_{s}} \quad \text{where} \quad r = 0, \dots, n \,, \ r + s = n \,. \tag{3.290}$$

Hence there are $n+1$ independent symmetric tensors $\phi_{\alpha_1\dots\alpha_n}$ 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

$$\eta^\alpha = \varepsilon^{\alpha\beta}\,\eta_\beta \,, \tag{3.291}$$

which transforms as in (3.286) and correspondingly

$$\bar{\eta}_\alpha = \varepsilon_{\alpha\beta}\,\bar{\eta}^\beta \,, \tag{3.292}$$

As a consequence of (3.289) 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 (3.283), need be considered.

For an infinitesimal $SU(2)$ transformation, with A as in (3.37), the corresponding change in a spinor arising from the transformation (3.282) is

$$\delta\eta_\alpha = -i\delta\theta \, \tfrac{1}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})_\alpha{}^\beta \eta_\beta \,. \tag{3.293}$$

For a tensor then correspondingly from (3.283)

$$\delta\psi_{\alpha_1\dots\alpha_n} = -i\delta\theta \sum_{r=1}^{n} \tfrac{1}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})_{\alpha_r}{}^\beta \psi_{\alpha_1\dots\alpha_{r-1}\beta\alpha_{r+1}\dots\alpha_n} \,, \tag{3.294}$$

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

Making use of (3.289) we have

$$\varepsilon^{\alpha\gamma}\varepsilon_{\beta\delta}\,\boldsymbol{\sigma}_\gamma{}^\delta = \boldsymbol{\sigma}_\beta{}^\alpha\,, \tag{3.295}$$

since $\mathrm{tr}(\boldsymbol{\sigma}) = 0$. From (3.295) we get

$$\varepsilon^{\alpha\gamma}\boldsymbol{\sigma}_\gamma{}^\beta = \varepsilon^{\beta\gamma}\boldsymbol{\sigma}_\gamma{}^\alpha\,, \tag{3.296}$$

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

$$(\boldsymbol{\sigma}\varepsilon)_{\alpha\beta}\cdot(\varepsilon\boldsymbol{\sigma})^{\gamma\delta} = \delta_\alpha{}^\gamma\,\delta_\beta{}^\delta + \delta_\alpha{}^\delta\,\delta_\beta{}^\gamma\,, \qquad (\varepsilon\boldsymbol{\sigma})^{\alpha\beta}\cdot(\varepsilon\boldsymbol{\sigma})^{\gamma\delta} = -\varepsilon^{\alpha\gamma}\,\varepsilon^{\beta\delta} - \varepsilon^{\alpha\delta}\,\varepsilon^{\beta\gamma}\,. \tag{3.297}$$

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

$$T_{i_1\dots i_n} = (\varepsilon\sigma_{i_1})^{\alpha_1\beta_1}\dots(\varepsilon\sigma_{i_n})^{\alpha_n\beta_n}\,\psi_{\alpha_1\dots\alpha_n\beta_1\dots\beta_n}\,, \tag{3.298}$$

where it is easy to see that $T_{i_1\dots i_n}$ is symmetric and also zero on contraction of any pair of indices, as a consequence of (3.297). For an odd number of indices we may further define

$$T_{\alpha\, i_1\dots i_n} = (\varepsilon\sigma_{i_1})^{\alpha_1\beta_1}\dots(\varepsilon\sigma_{i_n})^{\alpha_n\beta_n}\,\psi_{\alpha\alpha_1\dots\alpha_n\beta_1\dots\beta_n}\,, \tag{3.299}$$

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

$$(\sigma_j)_\alpha{}^\beta T_{\beta\, i_1\dots i_{n-1}j} = 0\,. \tag{3.300}$$

For two symmetric spinorial tensors $\phi_{1,\alpha_1\dots\alpha_n}, \phi_{2,\beta_1\dots\beta_m}$, with $j_1 = \frac{1}{2}n$, $j_2 = \frac{1}{2}m$, their product can be decomposed into symmetric rank $(n+m-2r)$-tensors, for $r = 0,\dots m$ if $n \geq m$, where for each r,

$$\varepsilon^{\beta_1\gamma_1}\dots\varepsilon^{\beta_r\gamma_r}\phi_{1,(\alpha_1\dots\alpha_{n-r}\,\beta_1\dots\beta_r}\phi_{2,\alpha_{n-r+1}\dots\alpha_{n+m-2r})\gamma_1\dots\gamma_r}\,, \qquad r = 0,\dots,m\,. \tag{3.301}$$

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

$$\eta_{1\alpha}\,\eta_{2\beta} = \eta_{1(\alpha}\eta_{2\beta)} + \varepsilon_{\alpha\beta}\,\tfrac{1}{2}\eta_1{}^\gamma\eta_{2\gamma}\,, \tag{3.302}$$

where $\eta_{1(\alpha}\eta_{2\beta)}$ may be re-expressed as a vector using (3.298). 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.

For two symmetric spinorial tensors as above with $n = m$ there is a $SU(2)$ invariant

$$\varepsilon^{\alpha_1\beta_1}\dots\varepsilon^{\alpha_n\beta_n}\phi_{1,\alpha_1\dots\alpha_n}\phi_{2,\beta_1\dots\beta_n}\,, \tag{3.303}$$

which is clearly symmetric, antisymmetric under $\phi_1 \leftrightarrow \phi_2$ according to whether n is even, odd. For four $n = 4$ symmetric spinorial tensors there is an independent antisymmetric invariant

$$\varepsilon^{\alpha_1\beta_1}\varepsilon^{\alpha_2\gamma_1}\varepsilon^{\alpha_1\delta_1}\varepsilon^{\beta_2\gamma_2}\varepsilon^{\beta_3\delta_2}\varepsilon^{\gamma_3\delta_3}\phi_{1,\alpha_1\alpha_2\alpha_3}\,\phi_{2,\beta_1\beta_2\beta_3}\,\phi_{3,\gamma_1\gamma_2\gamma_3}\,\phi_{4,\delta_1\delta_2\delta_3}\,, \tag{3.304}$$

where the indices are contracted so as to correspond to the lines joining the vertices of a tetrahedron. Interchanging any two vertices then generates a minus sign since ε for the line joining the two vertices changes sign. The invariant in (3.304) corresponds to the t^4 term appearing in (3.266).

3.15 Spinor Representation for Angular Momentum

Finding explicit expressions for $3j$ and $6j$ symbols can be a rather involved combinatorial exercise. This can be made more algebraic by using spinor operators to provide representations of the angular momentum operators To this end we introduce operators $\hat{\phi}^\alpha$ satisfying the commutation relations[30]

$$\big[\hat{\phi}^\alpha, \hat{\phi}^\beta\big] = 2\,\varepsilon^{\alpha\beta}\,1\,. \tag{3.305}$$

The operators $\hat{\phi}^\alpha$ are arbitrary up to $\hat{\phi}^\alpha \to \hat{\phi}^\beta M_\beta{}^\alpha$ for $[M_\beta{}^\alpha] \in Sp(2,\mathbb{C})$. As a consequence of (3.305)

$$J_+ = \tfrac{1}{4}\,\hat{\phi}^1\hat{\phi}^1\,, \qquad J_- = -\tfrac{1}{4}\,\hat{\phi}^2\hat{\phi}^2\,, \qquad J_3 = \tfrac{1}{8}\big(\hat{\phi}^1\hat{\phi}^2 + \hat{\phi}^2\hat{\phi}^1\big)\,, \tag{3.306}$$

satisfy the standard commutation relations (3.61a), (3.61b) and

$$\begin{aligned}
&\big[J_+, e^{x\cdot\hat{\phi}}\big] = x_2\frac{\partial}{\partial x_1}\,e^{x\cdot\hat{\phi}}\,, \qquad \big[J_-, e^{x\cdot\hat{\phi}}\big] = x_1\frac{\partial}{\partial x_2}\,e^{x\cdot\hat{\phi}}\,,\\
&\big[J_3, e^{x\cdot\hat{\phi}}\big] = \frac{1}{2}\Big(x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2}\Big)\,e^{x\cdot\hat{\phi}}\,, \qquad x\cdot\hat{\phi} = x_\alpha\hat{\phi}^\alpha
\end{aligned} \tag{3.307}$$

Hence

$$e^{x\cdot\hat{\phi}} = \sum_{j=0,\frac{1}{2},1,\dots}\ \sum_{m=-j}^{j} \frac{1}{n_{jm}}\,{x_1}^{j+m}{x_2}^{j-m}\,O_{j\,m}\,, \qquad n_{jm} = \big((j+m)!(j-m)!\big)^{\frac{1}{2}}\,, \tag{3.308}$$

ensures that O_{jm} are irreducible tensor operators satisfying (3.270). The operators O_{jm} form a basis for the $2j+1$ symmetric tensor operators $\hat{\phi}^{(\alpha_1}\hat{\phi}^{\alpha_2}\dots\hat{\phi}^{\alpha_{2j})}$. By using (3.305) any arbitrary product $\hat{\phi}^{\alpha_1}\hat{\phi}^{\alpha_2}\dots\hat{\phi}^{\alpha_n}$ can be reduced to a sum over symmetrised products and hence O_{jm} for $2j \le n$.

For the product of two exponentials we have

$$\begin{aligned}
&e^{x\cdot\hat{\phi}}e^{y\cdot\hat{\phi}} = e^{(x+y)\cdot\hat{\phi}}e^{\varepsilon^{\beta\gamma}x_\beta y_\gamma}\\
&= \sum_{j=0,\frac{1}{2},1,\dots}\ \sum_{m=-j}^{j}\frac{1}{n_{jm}}\,(x_1+y_1)^{j+m}(x_2+y_2)^{j-m}\,O_{j\,m}\sum_{r,s\ge 0}\frac{1}{r!s!}(-1)^s\,(x_1y_2)^r(x_2y_1)^s\,.
\end{aligned} \tag{3.309}$$

Expanding $e^{x\cdot\hat{\phi}}$ in terms of $O_{j_1m_1}$ and $e^{y\cdot\hat{\phi}}$ in terms of $O_{j_2m_2}$ as in (3.308) and comparing coefficients gives

$$\begin{aligned}
O_{j_1\,m_1}\,O_{j_2\,m_2} &= n_{j_1m_1}n_{j_2m_2}\sum_{r,s\ge 0}\frac{(-1)^s}{r!s!(j_1+m_1-r)!(j_1-m_1-s)!(j_2+m_2-s)!(j_2-m_2-r)!}\\
&\qquad\qquad \times n_{j_1+j_2-r-s\;m_1+m_2}\,O_{j_1+j_2-r-s\;m_1+m_2}\\
&= (-1)^{j_1-j_2-m_3}\sum_{|j_1-j_2|\le j_3\le j_1+j_2} F\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}\,O_{j_3\,-m_3}\,, \quad m_3 = -m_1-m_2\,,
\end{aligned} \tag{3.310}$$

[30] A particular representations for such operators is unimportant for the considerations in this section. Necessarily such operators act on an infinite dimensional space. A particular choice, acting on functions $f(x)$, is provided by taking $\hat{\phi}^1 = 2x$, $\hat{\phi}^2 = -\frac{\mathrm{d}}{\mathrm{d}x}$ or in terms of creation and annihilation operators $\hat{\phi}^1 = \sqrt{2}\,a^\dagger$, $\hat{\phi}^2 = -\sqrt{2}\,a$.

for

$$F\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = n_{j_1m_1} n_{j_2m_2}\, n_{j_3m_3} \sum_s \frac{(-1)^{s+j_1-j_2-m_3}}{s!(j_1-m_1-s)!(j_2+m_2-s)!(j_3-j_2+m_1+s)!(j_3-j_1-m_2+s)!(j_1+j_2-j_3-s)!} \,. \tag{3.311}$$

In (3.311) the range of the s sum is dictated by the factorials in the denominator. For $j_3 = 0$, $j_1 = j_2$ and $m_2 = -m_1$ there is just one term when $s = j_1 - m_1$ for a non zero contribution giving

$$F\begin{pmatrix} j & j & 0 \\ m & -m & 0 \end{pmatrix} = (-1)^{j-m} \,. \tag{3.312}$$

Another such case arises for $j_3 = j_1 + j_2$ when we must take $s = 0$ so that

$$F\begin{pmatrix} j_1 & j_2 & j_1+j_2 \\ m_1 & m_2 & -m_1-m_2 \end{pmatrix} = (-1)^{j_1-j_2+m_1+m_2} \frac{n_{j_1+j_2\, m_1+m_2}}{n_{j_1m_1}\, n_{j_2m_2}} \,. \tag{3.313}$$

For $j_2 = \frac{1}{2}$ there remains one term and

$$F\begin{pmatrix} j & \frac{1}{2} & j+\frac{1}{2} \\ m & \pm\frac{1}{2} & -m\mp\frac{1}{2} \end{pmatrix} = \pm(j \pm m + 1)^{\frac{1}{2}} (-1)^{j+m} \,, \quad F\begin{pmatrix} j & \frac{1}{2} & j-\frac{1}{2} \\ m & \pm\frac{1}{2} & -m\mp\frac{1}{2} \end{pmatrix} = -(j \mp m)^{\frac{1}{2}} (-1)^{j+m} \,. \tag{3.314}$$

From (3.311) we may readily derive the symmetry properties by shifting $s \to j_1+j_2-j_3-s$ and $s \to j_2 - j_3 - m_1 + s$

$$F\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = F\begin{pmatrix} j_2 & j_1 & j_3 \\ -m_2 & -m_1 & -m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} F\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = F\begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} \,. \tag{3.315}$$

Detailed results are based on

$$\begin{aligned} \exp\big(- u\, \Delta_{xy}\big)\, e^{x\cdot\hat\phi} e^{z\cdot\hat\phi}\, e^{y\cdot\hat\phi}\Big|_{x=y=0} &= \sum_{j=0,\frac{1}{2},1,\dots} u^{2j} \sum_{m=-j}^{j} (-1)^{j+m}\, O_{j\,-m}\, e^{z\cdot\hat\phi}\, O_{j\,m} \\ &= \frac{1}{(1-u)^2}\, e^{u'\, z\cdot\hat\phi} \,, \qquad \Delta_{xy} = \varepsilon_{\alpha\beta} \frac{\partial^2}{\partial x_\alpha \partial y_\beta} \,, \qquad u' = \frac{1+u}{1-u} \,, \end{aligned} \tag{3.316}$$

where the calculation of the action of the derivatives is described later. This reduces to

$$\sum_{j=0,\frac{1}{2},1,\dots} u^{2j} \sum_{m=-j}^{j} (-1)^{j+m}\, O_{j\,-m}\, O_{kn}\, O_{j\,m} = \frac{1}{(1-u)^2}\, u'^{2k}\, O_{kn} \,. \tag{3.317}$$

As a special case, taking $k = 0$, this gives[31]

$$\sum_{m=-j}^{j} (-1)^{j+m}\, O_{j\,-m}\, O_{j\,m} = (2j+1)\, 1 \,, \tag{3.318}$$

[31] Alternatively starting from (3.310)

$$\sum_{m=-j}^{j} (-1)^{j+m}\, O_{j\,-m}\, O_{j\,m} = \sum_{j'=0}^{2j} j'!\, O_{2j'\,0} \sum_{s=0}^{2j-j'} \frac{1}{s!\,(2j-j'-s)!} \sum_{p=0}^{j'} (-1)^p \frac{(p+s)!\,(2j-s-p)!}{p!^2\,(j'-p)!^2} \,,$$

gives the same result.

and therefore, as a consequence of (3.318), from (3.310) and using (3.315)

$$\begin{aligned}&\sum_{m_1,m_2}(-1)^{j_1+j_2+m_1+m_2}\,O_{j_2\,-m_2}\,O_{j_1\,-m_1}\,O_{j_1\,m_1}\,O_{j_2\,m_2}=(2j_1+1)(2j_2+1)\,1\\&=\sum_{m_1,m_2}\sum_{j_3{}',j_3=|j_1-j_2|}^{j_1+j_2}F\begin{pmatrix}j_1&j_2&j_3{}'\\m_1&m_2&m_3\end{pmatrix}F\begin{pmatrix}j_1&j_2&j_3\\m_1&m_2&m_3\end{pmatrix}(-1)^{j_1+j_2-m_3}\,O_{j_3{}'\,m_3}O_{j_3\,-m_3}\,.\end{aligned}\tag{3.319}$$

For consistency it is necessary that

$$\sum_{m_1,m_2,m_1+m_2=-m_3}F\begin{pmatrix}j_1&j_2&j_3{}'\\m_1&m_2&m_3\end{pmatrix}F\begin{pmatrix}j_1&j_2&j_3\\m_1&m_2&m_3\end{pmatrix}=f(j_1,j_2,j_3)\,\frac{1}{2j_3+1}\,\delta_{j_3{}'j_3}\,,\tag{3.320}$$

so that

$$\sum_{m_i,\sum m_j=0}F\begin{pmatrix}j_1&j_2&j_3\\m_1&m_2&m_3\end{pmatrix}F\begin{pmatrix}j_1&j_2&j_3\\m_1&m_2&m_3\end{pmatrix}=f(j_1,j_2,j_3)\,,\tag{3.321}$$

where the symmetry conditions (3.315) requires $f(j_1,j_2,j_3)$ to be totally symmetric.

To obtain an explicit form for $f(j_1,j_2,j_3)$ we may use a generalisation of (3.316)

$$\begin{aligned}&\exp\big(-r\,\Delta_{zz'}-s\,\Delta_{ww'}\big)\,e^{z\cdot\hat\phi}e^{w\cdot\hat\phi}\otimes e^{w'\cdot\hat\phi}\,e^{z'\cdot\hat\phi}\Big|_{z=w=z'=w'=0}\\&=\sum_{j_1,j_2=0,\frac12,\dots}r^{2j_1}s^{2j_2}\sum_{j_3,j_3'=|j_1-j_2|}^{j_1+j_2}\\&\quad\times\sum_{m_1,m_2}(-1)^{j_1+m_1+j_2+m_2}F\begin{pmatrix}j_1&j_2&j_3\\-m_1&-m_2&m_3\end{pmatrix}F\begin{pmatrix}j_2&j_1&j_3{}'\\m_2&m_1&-m_3\end{pmatrix}O_{j_3-m_3}\otimes O_{j_3'm_3}\,.\end{aligned}\tag{3.322}$$

Hence, using (3.319) and (3.318), and also with u' as in (3.316) and similarly for v' in terms of v,

$$\begin{aligned}&\exp\big(-r\,\Delta_{x_2x_7}-s\,\Delta_{x_3x_6}-u\,\Delta_{x_1x_4}-v\,\Delta_{x_5x_8}\big)\,e^{x_1\cdot\hat\phi}e^{x_2\cdot\hat\phi}\dots e^{x_8\cdot\hat\phi}\Big|_{x_i=0}\\&=\frac{1}{(1-u)^2(1-v)^2}\sum_{j_1,j_2,j_3=0,\frac12,\dots}r^{2j_1}s^{2j_2}\,(u'v')^{2j_3}\,(-1)^{j_1+j_2-j_3}\,f(j_1,j_2,j_3)\,1\\&=\frac{1}{(1-u)^2(1-v)^2}\,\frac{1}{(1+r\,s-r\,u'v'-s\,u'v')^2}\,1\,.\end{aligned}\tag{3.323}$$

Since $\sum_{a,b,c\geq0}(-1)^c(a+b+c+1)!/(a!\,b!\,c!)\,x^ay^bz^c=(1-x-y+z)^{-2}$ we then have

$$f(j_1,j_2,j_3)=\frac{(j_1+j_2+j_3+1)!}{(j_1+j_2-j_3)!\,(j_2+j_3-j_1)!\,(j_3+j_1-j_2)!}\,.\tag{3.324}$$

From the generating function for $f(j_1,j_2,j_3)$ given by (3.323) this satisfies

$$\begin{aligned}\sum_{j_1,j_2=0,\frac12,\dots}r^{2j_1}s^{2j_2}\sum_{j_3=|j_1-j_2|}^{j_1+j_2}(-1)^{j_1+j_2-j_3}\,f(j_1,j_2,j_3)&=\frac{1}{(1-r)^2(1-s)^2}\\&=\sum_{j_1,j_2=0,\frac12,\dots}(2j_1+1)(2j_2+1)\,r^{2j_1}s^{2j_2}\,,\end{aligned}\tag{3.325}$$

as is necessary for (3.319) using (3.321).

Manifestly (3.310) is just the usual Clebsch-Gordan angular momentum decomposition as in (3.157) so that the $3j$ symbol $\left(\begin{smallmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{smallmatrix}\right)$ is determined up to a normalisation factor independent of m_1, m_2, m_3. The normalisation is determined by (3.320) so that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \Delta(j_1, j_2, j_3)\, F \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \qquad \Delta(j_1, j_2, j_3) = f(j_1, j_2, j_3)^{-\frac{1}{2}} . \tag{3.326}$$

These techniques can be extended to derive results for $6j$ symbols. Following a similar route as before

$$\begin{aligned}
&\exp\big(-u_1\,\Delta_{x_1x_4} - u_3\,\Delta_{x_5x_8} - u_2\,\Delta_{x_9x_{12}} - v_2\,\Delta_{x_3x_6} - v_1\,\Delta_{x_7x_{10}} - v_3\,\Delta_{x_2x_1}\big)\, e^{x_1\cdot\hat\phi} e^{x_2\cdot\hat\phi} \dots e^{x_{12}\cdot\hat\phi}\Big|_{x_i=0} \\
&= \exp\big(-u_1\,\Delta_{x_1x_4} - u_3\,\Delta_{x_5x_8} - u_2\,\Delta_{x_9x_{12}}\big) \prod_{i=1}^{3} \sum_{k_i=0,\frac{1}{2},\dots} {v_i}^{2k_i} \sum_{n_i=-k_i}^{k_i} (-1)^{k_i+n_i} \\
&\qquad \times e^{x_1\cdot\hat\phi}\, O_{k_3 -n_3}\, O_{k_2 -n_2}\, e^{x_4\cdot\hat\phi}\, e^{x_5\cdot\hat\phi}\, O_{k_2 n_2}\, O_{k_1 -n_1}\, e^{x_8\cdot\hat\phi}\, e^{x_9\cdot\hat\phi} O_{k_1 n_1}\, O_{k_3 n_3}\, e^{x_{12}\cdot\hat\phi}\Big|_{x_i=0} \\
&= \prod_{i=1}^{3} \frac{1}{(1-u_i)^2} \sum_{j_i,k_i=0,\frac{1}{2},\dots} {u_i'}^{2j_1} {v_i}^{2j_i}\ (-1)^{\sum_{i=1}^{3}(2j_i+2k_i)}\, F \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} 1 \\
&= \prod_{i=1}^{3} \frac{1}{(1-u_i)^2} \\
&\quad \times \frac{1}{\big(1 - u_2' u_3' v_1 - u_1' u_3' v_2 - u_1' u_2' v_3 - v_1 v_2 v_3 + u_1' u_2' v_1 v_2 + u_1' u_3' v_1 v_3 + u_2' u_3' v_2 v_3\big)^2}\, 1 ,
\end{aligned} \tag{3.327}$$

using that $j_1 + k_2 + k_3$, $j_2 + k_1 + k_3$ are integers to achieve a symmetric form for the sign factor and where we have required, similarly to (3.190),

$$\begin{aligned}
\sum_{n_i,\, m_1 \text{ fixed}} (-1)^{\sum_i (j_i+m_i+k_i+n_i)} F \begin{pmatrix} k_2 & k_3 & j_1 \\ n_2 & -n_3 & -m_1 \end{pmatrix} F \begin{pmatrix} k_3 & k_1 & j_2 \\ n_3 & -n_1 & -m_3 \end{pmatrix} F \begin{pmatrix} k_1 & k_2 & j_3 \\ n_1 & -n_2 & -m_3 \end{pmatrix} F \begin{pmatrix} j_2 & j_3 & j_1' \\ m_2 & m_3 & m_1 \end{pmatrix} \\
= F \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} \tfrac{1}{2j_1+1}\, \delta_{j_1 j_1'} .
\end{aligned} \tag{3.328}$$

Clearly from (3.326)

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix} = \Delta(k_2, k_3, j_1)\, \Delta(k_3, k_1, j_2)\, \Delta(k_1, k_2, j_3)\, \Delta(j_1, j_2, j_3)\, F \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix}, \tag{3.329}$$

and the symmetry properties in (3.189) follow directly from the generating function given by (3.328). By using

$$\frac{1}{\big(1 - \sum_{i=1}^4 x_i + \sum_{j=1}^3 y_j\big)^2} = \sum_{a_i,b_j\ge 0} \frac{(\sum_i a_i + \sum_j b_j + 1)!}{\prod_{i=1}^4 a_i!\, \prod_{j=1}^3 b_j!}\, (-1)^{\sum_j b_j} \prod_{i=1}^{4} x_i^{a_i} \prod_{j=1}^{3} y_j^{b_j} \tag{3.330}$$

to expand (3.323) with $x_1 = u_2' u_3' v_1$, $x_2 = u_1' u_3' v_2$, $x_3 = u_1' u_2' v_3$, $x_4 = v_1 v_2 v_3$, $y_1 = u_2' u_3' v_2 v_3$, $y_2 = u_1' u_3' v_1 v_3$, $y_3 = u_1' u_2' v_1 v_2$ and matching with $\prod_{i=1}^3 {u_i'}^{2j_1} {v_i}^{2j_i}$ requires taking

$$\begin{aligned}
& a_1 = s - j_1 - k_2 - k_3\,,\ a_2 = s - j_2 - k_3 - k_1\ \ a_3 = s - j_2 - k_1 - k_2,\ a_4 = s - j_1 - j_2 - j_3\,, \\
& b_1 = j_2 + j_3 + k_2 + k_3 - s\,, \quad b_2 = j_3 + j_1 + k_3 + k_1 - s\,, \quad b_3 = j_1 + j_2 + k_1 + k_2 - s\,.
\end{aligned} \tag{3.331}$$

with s an integer constrained to a finite range by requiring each a_i, b_j is a positive integer or zero and $\sum_i a_i + \sum_j b_j = s$. Hence

$$F\left\{\begin{smallmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{smallmatrix}\right\} = \sum_s (-1)^s \, \frac{(s+1)!}{a_1!\, a_2!\, a_3!\, a_4!\, b_1!\, b_2!\, b_3!}\,, \tag{3.332}$$

with a_i, b_j as in (3.331) and where the sign $(-1)^{\sum_j b_j} = (-1)^{\sum_j (2j_j + 2k_j) + s}$. As a special case for $k_1 = 0$ there is only one term in the sum with a non zero contribution requiring $k_2 = j_3$, $k_3 = j_2$ and

$$F\left\{\begin{smallmatrix} j_1 & j_2 & j_3 \\ 0 & j_3 & j_2 \end{smallmatrix}\right\} = (-1)^{\sum_i j_i} f(j_1, j_2, j_3) \quad \Rightarrow \quad \left\{\begin{smallmatrix} j_1 & j_2 & j_3 \\ 0 & j_3 & j_2 \end{smallmatrix}\right\} = \frac{(-1)^{\sum_i j_i}}{\sqrt{(2j_2+1)(2j_3+1)}}\,. \tag{3.333}$$

The generating functions obtained here were first determined by Schwinger.[32]

3.15.1 Calculation of Action of Derivatives

The calculations to obtain the results in (3.316), (3.322), (3.323) or (3.328) is complicated by the non commutavity of $\hat{\phi}^\alpha$. We consider in general

$$\exp\Big(- \sum_{i<j \in I} u_{i,j} \, \Delta_{x_i x_j} \Big) \, e^{x_1 \cdot \hat{\phi}} e^{x_2 \cdot \hat{\phi}} \dots e^{x_n \cdot \hat{\phi}} \Big|_{x_i = 0, i \in I}\,, \tag{3.334}$$

for $I \subset \{1, 2, \dots, n\}$ with $p = \dim I$ even. For $u_{i,j}$, $u_{i'j'}$ to be both non zero we require $i \neq i'$, $j \neq j'$. With our conventions $\varepsilon_{\alpha\beta} = -1$ and from (3.316),

$$\sum_{i<j \in I} u_{i,j} \, \Delta_{x_i x_j} = \sum_{i,j \in I} \partial_{x_{i,2}} \, U_{ij} \, \partial_{x_{j,1}}\,, \quad U_{ij} = \begin{cases} u_{i,j} & i<j \\ -u_{i,j} & i>j \\ 0 & \text{otherwise} \end{cases}\,. \tag{3.335}$$

$U = [U_{ij}]$ is then an antisymmetric $p \times p$ matrix with one non zero element in each row and column.

For $p = n$ (3.334) defines an invariant. To evaluate (3.334) we define a normal ordering whereby all operators $\hat{\phi}^1$ are moved to the left of $\hat{\phi}^2$. Thus

$$e^{x \cdot \hat{\phi}} = N\big(e^{x \cdot \hat{\phi}}\big) \, e^{-x_1 x_2}\,, \quad N\big(e^{x \cdot \hat{\phi}}\big) = e^{x_1 \cdot \hat{\phi}^1} e^{x_2 \cdot \hat{\phi}^2}\,, \tag{3.336}$$

and in general

$$\begin{gathered} e^{x_1 \cdot \hat{\phi}} \dots e^{x_n \cdot \hat{\phi}} = N\big(e^{x_1 \cdot \hat{\phi}} \dots e^{x_n \cdot \hat{\phi}}\big) \, \exp\Big(- \sum_{i,j} x_{i,1} V_{ij} \, x_{j,2} \Big)\,, \quad V_{ij} = \begin{cases} 0 & i<j \\ 1 & i=j \\ 2 & i>j \end{cases}\,, \\ N\big(e^{x_1 \cdot \hat{\phi}} \dots e^{x_n \cdot \hat{\phi}}\big) = \exp\big(\textstyle\sum_i x_{i,1} \hat{\phi}^1 \big) \, \exp\big(\textstyle\sum_i x_{i,2} \, \hat{\phi}^2 \big)\,. \end{gathered} \tag{3.337}$$

[32] Juliian Seymour Schwinger, 1918-1994, American. Nobel prize 1965.

Defining coherent states such that $\langle\phi^1|\hat{\phi}^1 = \phi^1\langle\phi^1|,\ \hat{\phi}^2|\phi^2\rangle = \phi^2\,|\phi^2\rangle$ then

$$\langle\phi^1|N\big(f(\hat{\phi}^1,\hat{\phi}^2)\big)|\phi^2\rangle = f(\phi^1,\phi^2)\,\langle\phi^1|\phi^2\rangle\,. \tag{3.338}$$

With these results the calculation of (3.334) can be reduced to considering

$$\begin{aligned}
&\exp\Big(-\sum_{i<j\in I} u_{i,j}\,\Delta_{x_ix_j}\Big)\langle\phi^1|\,e^{x_1\cdot\hat{\phi}}\dots e^{x_n\cdot\hat{\phi}}\,|\phi^2\rangle\Big|_{x_i=0,i\in I}\\
&= \exp\Big(-\sum_{i,j\in I}\partial_{x_{i,2}}\,U_{ij}\,\partial_{x_{j,1}}\Big)\,\exp\Big(\sum_{i=1}^n(x_{i,1}\,\phi^1+\phi^2\,x_{i,2})-\sum_{i,j=1}^n x_{i,1}V_{ij}\,x_{j,2}\Big)\Big|_{x_i=0,i\in I}\,\langle\phi^1|\phi^2\rangle\,,
\end{aligned} \tag{3.339}$$

where, with $I\cup\bar{I}=\{1,2,\dots,n\}$,

$$\begin{aligned}
&\sum_{i,j=1}^n x_{i,1}V_{ij}\,x_{j,2}\\
&= \sum_{i,j\in I} x_{i,1}V^{(p)}{}_{ij}\,x_{j,2} + \sum_{i,j\in\bar{I}} x_{i,1}V^{(n-p)}{}_{ij}\,x_{j,2} + 2\sum_{i\in I}\Big(x_{i,1}\sum_{j\in\bar{I},j>i}x_{j,2}+\sum_{j\in\bar{I},j<i}x_{j,1}\,x_{i,2}\Big),
\end{aligned} \tag{3.340}$$

with $V^{(p)}$ the $p\times p$ lower triangle matrix with 1's on the diagonal and $2's$ below the diagonal.

The evaluation of (3.339) follows from

$$\begin{aligned}
&\exp\big(-\partial_{\tilde{x}}\cdot U\cdot\partial_x\big)\,\exp\big(x\cdot\tilde{y}+y\cdot\tilde{x}-x\cdot V\cdot\tilde{x}\big)\Big|_{x=\tilde{x}=0} = \exp\big(-\partial_{\tilde{y}}\cdot V\cdot\partial_y\big)\,\exp\big(-y\cdot U\cdot\tilde{y}\big)\\
&\quad= \exp\big(-\partial_{\tilde{y}}\cdot V\cdot\partial_y\big)\det\big(U^{-1}\big)\int\mathrm{d}z\mathrm{d}\tilde{z}\ \exp\big(z\cdot U^{-1}\cdot\tilde{z}+z\cdot\tilde{y}+y\cdot\tilde{z}\big)\\
&\quad= \det\big(U^{-1}\big)\int\mathrm{d}z\mathrm{d}\tilde{z}\ \exp\big(z\cdot(U^{-1}-V)\cdot\tilde{z}+z\cdot\tilde{y}+y\cdot\tilde{z}\big)\\
&\quad= \frac{1}{\det(\mathbb{1}-V\cdot U)}\,\exp\big(-y\cdot(\mathbb{1}-V\cdot U)^{-1}\cdot U\cdot\tilde{y}\big)\,,
\end{aligned} \tag{3.341}$$

for $\tilde{x}$, $\tilde{y}$ column vectors, x, y row vectors and U, V non singular square matrices.

For application to (3.316)

$$x=\big(x_1\ y_1\big)\,,\quad \tilde{x}=\begin{pmatrix}x_2\\ y_2\end{pmatrix},\quad y=\big(\phi^2-2z_1\ \phi^2\big)\,,\quad \tilde{y}=\begin{pmatrix}\phi^1\\ \phi^1-2z_2\end{pmatrix},\quad U=\begin{pmatrix}0 & u\\ -u & 0\end{pmatrix},\quad V=\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}, \tag{3.342}$$

so that

$$\begin{aligned}
&\exp\big(-\partial_{\tilde{x}}\cdot U\cdot\partial_x\big)\,\exp\big(x\cdot\tilde{y}+y\cdot\tilde{x}-x\cdot V\cdot\tilde{x}\big)\Big|_{x=\tilde{x}=0}\times\exp\big(z_1\phi^1+\phi^2z_2-z_1z_2\big)\\
&= \tfrac{1}{(1-u)^2}\,\exp\big(u'(z_1\phi^1+\phi^2z_2)-u'^2z_1z_2\big)\,,\qquad u'=\tfrac{1+u}{1-u}\,.
\end{aligned} \tag{3.343}$$

which corresponds to the result in (3.316) after normal ordering. For (3.328) the matrix U becomes

$$U=\begin{pmatrix}
0&0&0&u_1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&v_3&0\\
0&0&0&0&0&v_2&0&0&0&0&0&0\\
-u_1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&u_3&0&0&0&0\\
0&0&-v_2&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&v_1&0&0\\
0&0&0&0&-u_3&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&u_2\\
0&0&0&0&0&0&-v_1&0&0&0&0&0\\
0&-v_3&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&-u_2&0&0&0
\end{pmatrix}, \tag{3.344}$$

where $1/\det(\mathbb{1}_{12} - V \cdot U)$ gives the final result for the generating function.

3.16 Isospin

The symmetry which played a significant role in the early days of nuclear and particle physics is isospin, was initially 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 sam1 mathematical properties as discussed in its applications to rotations, but with a very different physical interpretation. Results for Clebsch-Gordan coefficients are crucial in the applications of isospin symmetry. 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

$$q = \begin{pmatrix} u \\ d \end{pmatrix}, \qquad \bar{q} = \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}, \tag{3.345}$$

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

$$[I_3, I_\pm] = \pm I_\pm\,, \quad [I_+, I_-] = 2I_3 \quad \text{or} \quad [I_a, I_b] = i\,\varepsilon_{abc} I_c\,, \tag{3.346}$$

as in (3.61a),(3.61b) or (3.54), and also commute with the Hamiltonian

$$[I_a, H] = 0\,. \tag{3.347}$$

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

$$I_3|p\rangle = \tfrac{1}{2}|p\rangle\,, \quad I_3|n\rangle = -\tfrac{1}{2}|n\rangle\,, \qquad I_-|p\rangle = |n\rangle\,, \quad I_+|n\rangle = |p\rangle\,. \tag{3.348}$$

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

$$I_3(|\pi^+\rangle, |\pi^0\rangle, |\pi^-\rangle) = (|\pi^+\rangle, 0, -|\pi^-\rangle)\,, \qquad I_-|\pi^+\rangle = \sqrt{2}|\pi^0\rangle\,, \quad I_-|\pi^0\rangle = \sqrt{2}|\pi^-\rangle\,. \tag{3.349}$$

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, (3.347), 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.

$$\begin{aligned}\sum_{m'_3,m'_4,m'_1,m'_2} D^{(I_3)}_{m'_3 m_3}(R)^* D^{(I_4)}_{m'_4 m_4}(R)^* D^{(I_1)}_{m'_1 m_1}(R) D^{(I_2)}_{m'_2 m_2}(R)\, \langle I_3 m'_3, I_4 m'_4|T|I_1 m'_1, I_2 m'_2\rangle \\ = \langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle\,. \qquad (3.350)\end{aligned}$$

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

$$\begin{aligned}|I_1 m_1, I_2 m_2\rangle &= \sum_{I,M} |IM\rangle\, \langle I_1 m_1, I_2 m_2\, |IM\rangle\,, \\ \langle I_3 m_3, I_4 m_4| &= \sum_{I,M} \langle I_3 m_3, I_4 m_4|IM\rangle\langle IM|\,, \qquad (3.351)\end{aligned}$$

since then, as in (3.276),

$$\langle I'M'|T|IM\rangle = A_I\, \delta_{I'I}\delta_{M'M}\,, \qquad (3.352)$$

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

$$\langle I_3 m_3, I_4 m_4|T|I_1 m_1, I_2 m_2\rangle = \sum_I A_I\, \langle I_3 m_3, I_4 m_4|IM\rangle\langle I_1 m_1, I_2 m_2\, |IM\rangle\,. \qquad (3.353)$$

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

$$\begin{aligned}|\pi^+ p\rangle &= |\tfrac{3}{2}\,\tfrac{3}{2}\rangle\,, & |\pi^0 p\rangle &= \sqrt{\tfrac{2}{3}}|\tfrac{3}{2}\,\tfrac{1}{2}\rangle - \sqrt{\tfrac{1}{3}}|\tfrac{1}{2}\,\tfrac{1}{2}\rangle\,, \\ |\pi^0 n\rangle &= \sqrt{\tfrac{2}{3}}|\tfrac{3}{2}\,{-\tfrac{1}{2}}\rangle + \sqrt{\tfrac{1}{3}}|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle\,, & |\pi^- p\rangle &= \sqrt{\tfrac{1}{3}}|\tfrac{3}{2}\,{-\tfrac{1}{2}}\rangle - \sqrt{\tfrac{2}{3}}|\tfrac{1}{2}\,{-\tfrac{1}{2}}\rangle\,, \qquad (3.354)\end{aligned}$$

using the Clebsch-Gordan coefficients which have been calculated in (3.166) for $j = 1$. Hence we have the results for the scattering amplitudes

$$\begin{aligned}\langle \pi^+ p|T|\pi^+ p\rangle &= A_{\frac{3}{2}}\,, \\ \langle \pi^- p|T|\pi^- p\rangle &= \tfrac{1}{3}\, A_{\frac{3}{2}} + \tfrac{2}{3}\, A_{\frac{1}{2}}\,, \\ \langle \pi^0 n|T|\pi^- p\rangle &= \tfrac{\sqrt{2}}{3}\big(A_{\frac{3}{2}} - A_{\frac{1}{2}}\big)\,, \qquad (3.355)\end{aligned}$$

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

$$\sigma_{\pi^+p\to\pi^+p} = k\big|A_{\frac{3}{2}}\big|^2\,, \quad \sigma_{\pi^-p\to\pi^-p} = \tfrac{1}{9}k\big|A_{\frac{3}{2}} + 2A_{\frac{1}{2}}\big|^2\,, \quad \sigma_{\pi^-p\to\pi^0n} = \tfrac{2}{9}k\big|A_{\frac{3}{2}} - A_{\frac{1}{2}}\big|^2\,, \tag{3.356}$$

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 $|pp\rangle = |11\rangle$, $|pn\rangle = \frac{1}{\sqrt{2}}\big(|10\rangle + |00\rangle\big)$. 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.

3.16.1 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

$$\mathcal{C}|\pi^0\rangle = |\pi^0\rangle\,. \tag{3.357}$$

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

$$\mathcal{C}|\pi^\pm\rangle = -|\pi^\mp\rangle\,. \tag{3.358}$$

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

$$G = \mathcal{C}e^{-i\pi I_2}\,. \tag{3.359}$$

The action of $e^{-i\pi I_2}$ on an isospin multiplet is determined for any representation by (3.101). In this case we have

$$e^{-i\pi I_2}|\pi^+\rangle = |\pi^-\rangle\,, \quad e^{-i\pi I_2}|\pi^0\rangle = -|\pi^0\rangle\,, \quad e^{-i\pi I_2}|\pi^-\rangle = |\pi^+\rangle\,, \tag{3.360}$$

and hence on any pion state

$$G|\pi\rangle = -|\pi\rangle\,. \tag{3.361}$$

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[33] transformations, forming the Lorentz group, have a vital importance. Extending Lorentz transformations by translations, in space and time, generates the Poincaré[34] 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

$$\Lambda^T g \Lambda = g \,, \qquad \Lambda = [\Lambda^\mu{}_\nu] \,, \quad g = [g_{\mu\nu}] = \begin{pmatrix} 1 & 0 \\ 0 & -\mathbb{1}_3 \end{pmatrix} . \tag{4.5}$$

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

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

[33] Hendrik Antoon Lorentz, 1853-1928, Dutch. Nobel prize 1902.
[34] 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, defining a Killing vector, 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 $\Lambda^T\Lambda = \mathbb{1}$ as well as (4.5). In this case the Lorentz transform matrix has the form,

$$\Lambda_R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}, \qquad R^T R = \mathbb{1}_3\,, \tag{4.17}$$

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

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

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

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

Hence (4.5) requires

$$\mathcal{B}n = \cosh\alpha\, n \, , \qquad \mathcal{B}^2 - \sinh^2\alpha\, nn^T = \mathbb{1}_3 \, . \tag{4.21}$$

The solution is just

$$\mathcal{B} = \mathbb{1}_3 + (\cosh\alpha - 1) nn^T . \tag{4.22}$$

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

$$B(\alpha, \mathbf{n}) = \begin{pmatrix} \cosh\alpha & \sinh\alpha\, n^T \\ \sinh\alpha\, n & \mathbb{1}_3 + (\cosh\alpha - 1) nn^T \end{pmatrix} , \tag{4.23}$$

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

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

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

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 $\mathbf{n}$ fixed and α varying which is isomorphic to $SO(1,1)$ with matrices as in (1.123). With Λ_R as in (4.17) then for B as in (4.23)

$$\Lambda_R B(\alpha, \mathbf{n}) \Lambda_R{}^{-1} = B(\alpha, \mathbf{n}^R) \, , \tag{4.25}$$

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 $(\Lambda B^{-1})^T \Lambda B^{-1} = B^{-1}\Lambda^T\Lambda B^{-1} = \mathbb{1}$ 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 $U[\Lambda^{-1}] = 1 + i\,\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu}$ (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 (3.229) 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 (3.54) 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 (3.231). 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 (3.50).

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

$$\sigma_\mu = (1\!\!1_2, \boldsymbol{\sigma}) = \sigma_\mu{}^\dagger\,, \qquad \bar\sigma_\mu = (1\!\!1_2, -\boldsymbol{\sigma}) = \bar\sigma_\mu{}^\dagger\,. \tag{4.48}$$

Both σ_μ and $\bar\sigma_\mu$ form a complete set of hermitian 2×2 matrices. As a consequence of (3.20) we have

$$\sigma_\mu\,\bar\sigma_\nu + \sigma_\nu\,\bar\sigma_\mu = 2g_{\mu\nu}\,1\!\!1\,, \qquad \bar\sigma_\mu\,\sigma_\nu + \bar\sigma_\nu\,\sigma_\mu = 2g_{\mu\nu}\,1\!\!1\,, \tag{4.49}$$

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 (3.19)

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

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

$$\mathrm{x}\,\bar{\mathrm{x}} = x^2 1\!\!1\,, \qquad \bar{\mathrm{x}}\mathrm{x} = x^2 1\!\!1\,. \tag{4.55}$$

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

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

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

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

and hence, for $\operatorname{tr} A = e^{i\phi}|\operatorname{tr} A|$,

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

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 $\Lambda \in SO(3,1)$. All elements in $Sl(2,\mathbb{C})$ are continuously connected to the identity so (4.51) does not allow for spatial or time reflections.

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 (3.27) and (3.30). If $A^\dagger = A$ then Λ, given by (4.59), is symmetric so this is a boost. Taking

$$A_B(\alpha, \mathbf{n}) = \cosh \tfrac{1}{2}\alpha \; \mathbb{1} + \sinh \tfrac{1}{2}\alpha \; \mathbf{n}\cdot\boldsymbol{\sigma} \,, \quad -\infty < \alpha < \infty \,, \tag{4.62}$$

corresponds to the Lorentz boost in (4.23). Rotations remain in the form in (3.38).

For a general infinitesimal Lorentz transformation as in (4.14) then, using $\Lambda^\mu{}_\mu = 4$ to this order and $\sigma_\mu \bar\sigma^\mu = 4\,\mathbb{1}$, (4.61) gives

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

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

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

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

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

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

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

4.3.2 Spinors, Dotted and Undotted Indices

In a similar fashion to the discussion in section 3.14 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 $\det A = 1$[35],

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

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

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

---

[35] Using (3.289), $\varepsilon^{\alpha\gamma} A_\gamma{}^\delta\, \varepsilon_{\delta\beta} = \delta_\beta{}^\alpha \mathrm{tr}(A) - A_\beta{}^\alpha = (A^{-1})_\beta{}^\alpha$, since for any 2×2 matrix the characteristic equation requires $A^2 - \mathrm{tr}(A)\; A + \det A\, \mathbb{1} = 0$, 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*[36] *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

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2 g_{\mu\nu}\, \mathbb{1}_4 \,. \tag{4.81}$$

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,

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

[36] 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

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

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 $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations while the Dirac spinor corresponds to $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$. These representations are not unitary since there is no positive group invariant scalar product, for the simplest cases of a vector or a $(\frac{1}{2}, 0)$ 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

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

If Υ is irreducible, as in (4.87), corresponding to the $(\frac{1}{2}n, \frac{1}{2}n)$ real representation, then $T_{\mu_1 \dots \mu_n}$ 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 (3.20) we get

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

using $(\sigma_0 \bar\sigma_1 \sigma_2 \bar\sigma_3)^2 = \sigma_0 \bar\sigma_1 \sigma_2 \bar\sigma_3 \sigma_3 \bar\sigma_2 \sigma_1 \bar\sigma_0 = -\mathbb{1}$, and similarly $(\bar\sigma_0 \sigma_1 \bar\sigma_2 \sigma_3)^2 = -\mathbb{1}$, 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)

$$\tfrac{1}{24}\, \varepsilon^{\mu\nu\sigma\rho} \gamma_\mu \gamma_\nu \gamma_\sigma \gamma_\rho = \gamma_0 \gamma_1 \gamma_2 \gamma_3 = i \gamma_5 \,, \qquad \gamma_5 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} . \tag{4.92}$$

As a consequence of (4.91) we may further obtain[37]

$$\tfrac{1}{2}\, \varepsilon^{\mu\nu\sigma\rho} \sigma_\sigma \bar\sigma_\rho = -i\, \sigma^{[\mu} \bar\sigma^{\nu]} \,, \qquad \tfrac{1}{2}\, \varepsilon^{\mu\nu\sigma\rho} \bar\sigma_\sigma \sigma_\rho = i\, \bar\sigma^{[\mu} \sigma^{\nu]} \,. \tag{4.93}$$

[37] For a somewhat convoluted demonstration note, that since the indices only take four values, $\varepsilon^{\mu\nu\sigma\rho} \sigma_{[\mu} \bar\sigma_\nu \sigma_\sigma \bar\sigma_\rho \sigma_{\lambda]} = \frac{1}{5} \varepsilon^{\mu\nu\sigma\rho} (\sigma_\mu \bar\sigma_\nu \sigma_\sigma \bar\sigma_\rho \sigma_\lambda - \sigma_\mu \bar\sigma_\nu \sigma_\sigma \bar\sigma_\lambda \sigma_\rho + \sigma_\mu \bar\sigma_\nu \sigma_\lambda \bar\sigma_\sigma \sigma_\rho - \sigma_\mu \bar\sigma_\lambda \sigma_\nu \bar\sigma_\sigma \sigma_\rho + \sigma_\lambda \bar\sigma_\mu \sigma_\nu \bar\sigma_\sigma \sigma_\rho) = 0$. 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), $i\sigma_\lambda = -\frac{1}{6} \varepsilon_\lambda{}^{\nu\sigma\rho} \sigma_\nu \bar\sigma_\sigma \sigma_\rho$. Similarly $i\bar\sigma_\mu = \frac{1}{6} \varepsilon_\mu{}^{\nu\sigma\rho} \bar\sigma_\nu \sigma_\sigma \bar\sigma_\rho$. Using these results, $i(\sigma_\lambda \bar\sigma_\mu - \sigma_\mu \bar\sigma_\sigma) = -\frac{1}{6} \varepsilon_\lambda{}^{\nu\sigma\rho} (\sigma_\nu \bar\sigma_\sigma \sigma_\rho \bar\sigma_\mu + \sigma_\mu \bar\sigma_\nu \sigma_\sigma \bar\sigma_\rho)$. 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

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

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

$$U[\Lambda, a] = T[a] U[\Lambda]\,, \qquad U[\Lambda, 0] = U[\Lambda]\,, \quad T[a] = U[\mathbb{1}, a]\,, \tag{4.109}$$

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 2.2 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 as discussed in 1.3. 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 3.5, 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

$$U[\Lambda]\,|p, s\, s_3\rangle = \sum_{s'_3} |\Lambda p, s\, s'_3\rangle\, D^{(s)}_{s'_3\, s_3}\big(\mathcal{R}(p,\Lambda)\big)\,, \tag{4.127}$$

with $\mathcal{R}(\Lambda p, \Lambda')\mathcal{R}(p,\Lambda) = \mathcal{R}(p, \Lambda'\Lambda)$.

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 invariance under Lorentz transformations follows from

$$\mathrm{d}^4 p\, \delta(p^2 + m^2)\theta(p^0) = \frac{\mathrm{d}^3 p}{2p^0} \quad \text{where} \quad p^0 = (\mathbf{p}^2 + m^2)^{\frac{1}{2}}\,. \tag{4.130}$$

The precise definition of the representation depends on the choice of $L(p)$ satisfying (4.116). This does 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

$$L(p) = B(\alpha, \hat{\mathbf{p}}) \quad \text{for} \quad p^\mu = B(\alpha, \hat{\mathbf{p}})^\mu{}_\nu \mathring{p}^\nu = m(\cosh\alpha, \sinh\alpha\, \hat{\mathbf{p}})\,, \tag{4.131}$$

where $B(\alpha, \hat{\mathbf{p}})$ is the boost Lorentz transformation defined in (4.23). Then in (4.126)

$$U[L(p)] \to U[B(\alpha, \hat{\mathbf{p}})] = e^{i\alpha \hat{\mathbf{p}}\cdot\mathbf{K}} . \qquad (4.132)$$

Using (4.131) if we consider a rotation Λ_R then $L(\Lambda_R\, p) = B(\alpha, \hat{\mathbf{p}}^R)$ and, by virtue of (4.25),

$$B(\alpha, \hat{\mathbf{p}}^R)^{-1}\Lambda_R B(\alpha, \hat{\mathbf{p}}) = \Lambda_R , \qquad (4.133)$$

so that

$$U[\Lambda_R]\,|p, s\, s_3\rangle = U[L(\Lambda_R p)]\, U[\Lambda_R]\,|\mathring{p}, s\, s_3\rangle = \sum_{s'_3} |\Lambda_R p, s\, s'_3\rangle\, D^{(s)}_{s'_3\, s_3}(R) . \qquad (4.134)$$

The Wigner rotation given by (4.123) with this definition of $L(p)$ is then just $\mathcal{R}(p, \Lambda_R) = R$.

4.5.2 Helicity States

An alternative prescription for $L(p)$ in (4.117), giving a different but equivalent basis for $\mathcal{V}_{m,s}$, is to first boost along the 3-direction and then rotate so that

$$\mathbf{e}_3 \to \hat{\mathbf{p}}(\theta, \phi) = \cos\theta\, \mathbf{e}_3 + \sin\theta(\cos\phi\, \mathbf{e}_1 + \sin\phi\, \mathbf{e}_2) , \qquad (4.135)$$

which is just the radial unit vector in spherical polar coordinates. This rotation corresponds to $R_{\phi,\theta,-\phi}$, in terms of Euler angles, or equivalently $R(\Theta, \mathbf{n})$ with $\mathbf{n} = -\sin\phi\, \mathbf{e}_1 + \cos\phi\, \mathbf{e}_2$. Hence $p = L(p)\mathring{p}$, with $p^\mu = (p^0, |\mathbf{p}|\, \hat{\mathbf{p}})$ can be obtained by taking

$$L(p) = \Lambda_{R(\Theta,\mathbf{n})} B(\alpha, \mathbf{e}_3) = \begin{pmatrix} \cosh\alpha & 0 & 0 & \sinh\alpha \\ \cos\phi\sin\theta\sinh\alpha & \sin^2\!\phi{+}\cos^2\!\phi\cos\theta & \cos\phi\sin\phi(\cos\theta{-}1) & \cos\phi\sin\theta\cosh\alpha \\ \sin\phi\sin\theta\sinh\alpha & \cos\phi\sin\phi(\cos\theta{-}1) & \cos^2\!\phi{+}\sin^2\!\phi\cos\theta & \sin\phi\sin\theta\cosh\alpha \\ \cos\theta\sinh\alpha & -\cos\phi\sin\theta & -\sin\phi\sin\theta & \cos\theta\cosh\alpha \end{pmatrix} ,$$
$$p^\mu = m(\cosh\alpha\, , \sinh\alpha\, \hat{\mathbf{p}}(\theta, \phi)) . \qquad (4.136)$$

Helicity states are then defined by

$$|p, h\rangle = U[R(\Theta, \mathbf{n})]\, U[B(\alpha, \mathbf{e}_3)]\, |\mathring{p}, h\rangle , \quad U[R(\Theta, \mathbf{n})] = e^{-i\Theta\mathbf{n}\cdot\mathbf{J}} = e^{-i\phi J_3} e^{-i\theta J_2} e^{i\phi J_3} , \qquad (4.137)$$

suppressing the label s for the particle spin which is fixed. Since J_3 commutes with $U[B(\alpha, \mathbf{e}_3)]$ and

$$e^{-i\Theta\mathbf{n}\cdot\mathbf{J}} J_3\, e^{i\Theta\mathbf{n}\cdot\mathbf{J}} = \hat{\mathbf{p}} \cdot \mathbf{J} , \qquad (4.138)$$

we must have

$$\hat{\mathbf{p}} \cdot \mathbf{J}\, |p, h\rangle = h\, |p, h\rangle , \qquad (4.139)$$

so that h is the component of spin along the direction of motion, or helicity.

Wigner rotations are defined just as in (4.123). Helicity states transform very simply under rotations since for any rotation R

$$R\, R(\Theta, \mathbf{n}) = R(\Theta', \mathbf{n}')\, R(\chi, \mathbf{e}_3) = R_{\phi',\theta',-\phi'}\, R(\chi, \mathbf{e}_3) , \qquad \mathbf{n}' = -\sin\phi'\, \mathbf{e}_1 + \cos\phi'\, \mathbf{e}_2 , \qquad (4.140)$$

and

$$\Lambda_R L(p) = L(\Lambda_R p)\,, \qquad (\Lambda_R p)^\mu = \big(p^0, |\mathbf{p}|\,\hat{\mathbf{p}}(\theta',\phi')\big)\,, \tag{4.141}$$

so that

$$U[\Lambda_R]\,|p,h\rangle = |\Lambda_R p, h\rangle\, e^{-i\,h\chi}\,. \tag{4.142}$$

For $R = R(\gamma,\mathbf{e}_3)$ then $\chi = \gamma$ and $\theta' = \theta,\ \phi' = \phi + \gamma$. For $R = R(\eta,\mathbf{e}_2)$ and taking $\phi = 0$ then $\chi = \phi' = 0$ and $\theta' = \theta + \eta$.

When Lorentz boosts are involved the Wigner rotation mixes different helicity states. For a boost along the 3-direction

$$\mathcal{R}\big(p, B(\beta,\mathbf{e}_3)\big) = R_{\phi,\chi,-\phi}\,, \quad \cot\chi = \cot\theta\,\cosh\alpha + \coth\beta\,\mathrm{cosec}\,\theta\,\sinh\alpha\,, \tag{4.143}$$

so that

$$U[B(\beta,\mathbf{e}_3)]\,|p,h\rangle = \sum_{h'} |p',h'\rangle\, D^{(s)}_{h'h}(R_{\phi,\chi,-\phi}) \tag{4.144}$$

with $p' = B(\beta,\mathbf{e}_3)p$ given in terms of α',θ',ϕ as in (4.136) with $\cosh\alpha' = \cosh\beta\,\cosh\alpha + \sinh\beta\,\sinh\alpha\,\cos\theta,\ \cot\theta' = \cosh\beta\,\cot\theta + \sinh\beta\,\coth\alpha\,\mathrm{cosec}\,\theta$.

4.5.3 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

$$\mathring{p}^\mu = \mathring{\omega}(1,0,0,1)\,, \qquad \mathring{\omega} > 0\,, \tag{4.145}$$

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

$$\omega^\mu{}_\nu \mathring{p}^\nu = 0\,, \qquad \omega^{\mu\nu} = -\omega^{\nu\mu}\,. \tag{4.146}$$

This linear equation is easy to solve giving

$$\omega^0{}_3 = 0\,, \qquad \omega^1{}_0 = -\omega^1{}_3\,, \qquad \omega^2{}_0 = -\omega^2{}_3\,, \qquad \omega^3{}_0 = 0\,. \tag{4.147}$$

These reduce the six independent $\omega^{\mu\nu} = -\omega^{\nu\mu}$ to three so that

$$\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu} = \omega^{12}M_{12} + \omega^{01}(M_{01} + M_{31}) + \omega^{02}(M_{02} + M_{32})\,. \tag{4.148}$$

Identifying the operators

$$J_3 = M_{12}\,, \qquad E_1 = M_{01} + M_{31} = K_1 + J_2\,, \qquad E_2 = M_{02} + M_{32} = K_2 - J_1\,, \tag{4.149}$$

we find the commutators from (4.32), or from (3.54), (4.42) and (4.43),

$$[J_3, E_1] = iE_2\,, \qquad [J_3, E_2] = -iE_1\,, \qquad [E_1, E_2] = 0\,. \tag{4.150}$$

A unitary operator corresponding to finite group elements of $G_{\mathring{p}}$ is then

$$e^{-i(a_1 E_1 + a_2 E_2)} e^{-i\chi J_3}\,, \tag{4.151}$$

Noting that

$$e^{-i\chi J_3}(a_1 E_1 + a_2 E_2) e^{i\chi J_3} = a_1{}^\chi E_1 + a_2{}^\chi E_2\,, \quad \begin{pmatrix} a_1{}^\chi \\ a_2{}^\chi \end{pmatrix} = \begin{pmatrix} \cos\chi & -\sin\chi \\ \sin\chi & \cos\chi \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{4.152}$$

then if (4.151) corresponds to a group element (χ, a_1, a_2), with χ an angle with period 2π, we have the group multiplication rule

$$(\chi', a'_1, a'_2)(\chi, a_1, a_2) = (\chi' + \chi, a_1{}^{\chi'} + a'_1, a_2{}^{\chi'} + a'_2)\,. \tag{4.153}$$

The group multiplication rule (4.153) 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

$$G_{\mathring{p}} \simeq ISO(2) \simeq SO(2) \ltimes T_2\,. \tag{4.154}$$

This group is isomorphic to the group of rotations and translations in two dimensions. The representations of this group can be obtained in a very similar fashion to that of the Poincaré group. Define vectors $|b_1, b_2\rangle$ such that

$$(E_1, E_2)|b_1, b_2\rangle = (b_1, b_2)|b_1, b_2\rangle\,, \tag{4.155}$$

and then we assume, consistency with the group multiplication (4.153),

$$e^{-i\chi J_3}|b_1, b_2\rangle = e^{-ih\chi}|b_1{}^\chi, b_2{}^\chi\rangle\,, \tag{4.156}$$

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

$$E_1|h\rangle = E_2|h\rangle = 0\,, \qquad J_3|h\rangle = h|h\rangle\,, \tag{4.157}$$

so that the essential group action is

$$e^{-i\chi J_3}|h\rangle = e^{-ih\chi}|h\rangle\,. \tag{4.158}$$

For applications to representations of the Poincaré group $e^{-i\chi J_3}$ corresponds to a subgroup of the $SO(3)$ rotation group so it is necessary to require in (4.157) and (4.158)

$$h = 0, \pm\tfrac{1}{2}, \pm 1, \dots\,. \tag{4.159}$$

For the associated Lorentz transformations then a general element corresponding to the little group is $\Lambda_{(a_1,a_2)}\Lambda_\chi$ where

$$\Lambda_{(a_1,a_2)} = \begin{pmatrix} 1+\frac{1}{2}(a_1{}^2+a_2{}^2) & a_1 & a_2 & -\frac{1}{2}(a_1{}^2+a_2{}^2) \\ a_1 & 1 & 0 & -a_1 \\ a_2 & 0 & 1 & -a_2 \\ \frac{1}{2}(a_1{}^2+a_2{}^2) & a_1 & a_2 & 1-\frac{1}{2}(a_1{}^2+a_2{}^2) \end{pmatrix}, \qquad \Lambda_\chi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\chi & -\sin\chi & 0 \\ 0 & \sin\chi & \cos\chi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.160}$$

It is easy to see that $\Lambda_{(a_1,a_2)}\mathring{p} = \Lambda_\chi \mathring{p} = \mathring{p}$ with $\mathring{p}$ as in (4.145).

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

$$P^\mu |\mathring{p}, h\rangle = \mathring{p}^\mu |\mathring{p}, h\rangle \,, \qquad J_3 |\mathring{p}, h\rangle = h |\mathring{p}, h\rangle \,, \tag{4.161}$$

a basis $\{|p, h\rangle : p^2 = 0, p_0 > 0\}$, for $\mathcal{V}_h$ is formed by

$$|p, h\rangle = U[L(p)]\,|\mathring{p}, h\rangle \,, \quad \text{for} \quad p^\mu = L(p)^\mu{}_\nu \mathring{p}^\nu \,, \tag{4.162}$$

where $L(p)$ is assumed to be determined uniquely by p. Using

$$L(\Lambda p)^{-1}\Lambda L(p) = \Lambda_{(a_1,a_2)}\Lambda_\chi \in G_{\mathring{p}} \,, \quad \text{for} \quad a_{1,2}(p,\Lambda)\,,\ \chi(p,\Lambda)\,, \tag{4.163}$$

so that

$$U[\Lambda_{(a_1,a_2)}]U[\Lambda_\chi]\,|\mathring{p}, h\rangle = |\mathring{p}, h\rangle\, e^{-ih\chi} \,, \tag{4.164}$$

then, for any $\Lambda \in SO(3,1)^\uparrow$, the action of the corresponding unitary operator on $\mathcal{V}_h$ is given by

$$U[\Lambda]\,|p, h\rangle = |\Lambda p, h\rangle\, e^{-ih\chi(p,\Lambda)} \,. \tag{4.165}$$

Group multiplication requires $\chi(p,\Lambda) + \chi(\Lambda p, \Lambda') = \chi(p, \Lambda'\Lambda)$.

For $\mathring{p}$ as in (4.145), and

$$p^\mu = \omega(1, \hat{\mathbf{p}})\,, \qquad \omega > 0\,, \tag{4.166}$$

then $L(p)$, satisfying (4.116), is determined by assuming it is given by the expression (4.136) with now $e^\alpha = \omega/\mathring{\omega}$. By the same arguments as for massive helicity states

$$\hat{\mathbf{p}} \cdot \mathbf{J}\,|p, h\rangle = h\,|p, h\rangle \,, \tag{4.167}$$

so that the component of the angular momentum along the direction of motion, or *helicity*, is again h.

The irreducible representations of the Poincaré group for massless particles require only a single helicity h, with values as in (4.159). 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

$$\mathcal{P}\mathbf{J}\mathcal{P}^{-1} = \mathbf{J}\,, \quad \mathcal{P}\mathbf{K}\mathcal{P}^{-1} = -\mathbf{K}\,, \quad \mathcal{P}H\mathcal{P}^{-1} = H\,, \quad \mathcal{P}\mathbf{P}\mathcal{P}^{-1} = -\mathbf{P}\,. \tag{4.168}$$

In consequence $\mathcal{P}\mathbf{P}\cdot\mathbf{J}\mathcal{P}^{-1} = -\mathbf{P}\cdot\mathbf{J}$ so that, from (4.167), $\mathcal{P}|p, h\rangle$ 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.

4.5.4 Two Particle States and Angular Momentum Decomposition

Scattering experiments generally involve collisions of two particles. So called *in states* are formed from tensor products of free particle states in the distant past. A convenient basis can be formed starting from two particle states in the centre of mass frame. Initially we define

$$|\mathring{P}, p\mathbf{e}_3, h_1h_2\rangle = U_1[B(\alpha_1,\mathbf{e}_3)]\,|\mathring{p}_1,h_1\rangle_1\,U_2[B(\alpha_2,\mathbf{e}_3)]\,|\mathring{p}_2,-h_2\rangle_2\,, \tag{4.169}$$

with α_1,α_2 such that

$$\begin{aligned}&B(\alpha_1,\mathbf{e}_3)^\mu{}_\nu\,\mathring{p}_1{}^\nu = (E_1,p\,\mathbf{e}_3)\,, \qquad B(\alpha_2,\mathbf{e}_3)^\mu{}_\nu\,\mathring{p}_2{}^\nu = (E_2,-p\,\mathbf{e}_3)\,, \quad p>0\,,\\ &\mathring{P}^\mu = (E,\mathbf{0})\,, \quad E = E_1+E_2 \quad E_1 = (m_1{}^2+p^2)^{\frac12}\,, \quad E_2 = (m_2{}^2+p^2)^{\frac12}\,.\end{aligned} \tag{4.170}$$

A general centre of mass state is obtained by a rotation taking $\mathbf{e}_3 \to \hat{\mathbf{p}}$ as in (4.135)

$$\begin{aligned}&|\mathring{P}, p\hat{\mathbf{p}}, h_1h_2\rangle = U[R_{\phi,\theta,-\phi}]\,|\mathring{P}, p\mathbf{e}_3, h_1h_2\rangle\,,\\ &|\mathring{P}, p\hat{\mathbf{p}}, h_1h_2\rangle = |p_1,h_1\rangle_1\,|p_2,h_2\rangle_2\,, \quad p_1{}^\mu = (E_1,p\hat{\mathbf{p}})\,, \quad p_2{}^\mu = (E_2,-p\hat{\mathbf{p}})\,,\end{aligned} \tag{4.171}$$

where $U[R]$ is a rotation generated by the total angular momentum $\mathbf{J} = \mathbf{J}_1+\mathbf{J}_2$ and $\mathbf{p} = p\hat{\mathbf{p}}$ as in (4.135). It is easy to see that

$$\hat{\mathbf{p}}\cdot\mathbf{J}\,|\mathring{P},\mathbf{p},h_1h_2\rangle = h\,|\mathring{P},\mathbf{p},h_1h_2\rangle\,, \qquad h = h_1-h_2\,. \tag{4.172}$$

A basis for two particle states $|P,\mathbf{p},h_1h_2\rangle$ with arbitrary total 4-momentum P^μ can be obtained by acting on $|\mathring{P},\mathbf{p},h_1h_2\rangle$ with a boost taking $\mathring{P}^\mu \to P^\mu$. For any translation invariant operator A acting on two particle states the overall momentum conservation δ-function can factored off to defined a reduced matrix element for centre of mass states

$$\langle P',\mathbf{p}',h_1{}'h_2{}'|A|P,\mathbf{p},h_1h_2\rangle = (2\pi)^4\delta^4(P'-P)\,\frac{E}{p}\,\langle\hat{\mathbf{p}}',h_1{}'h_2{}'\|A\|\hat{\mathbf{p}},h_1h_2\rangle\,. \tag{4.173}$$

Since

$$(2\pi)^6 4p_1{}^0p_2{}^0\,\delta^3(\mathbf{p}_1{}'-\mathbf{p}_1)\,\delta^3(\mathbf{p}_2{}'-\mathbf{p}_2) = (2\pi)^4\delta^4(p_1{}'+p_2{}'-p_1-p_2)\,(4\pi)^2\frac{E}{p}\,\delta^2(\hat{\mathbf{p}}',\hat{\mathbf{p}})\,, \tag{4.174}$$

with $E,p,\hat{\mathbf{p}}',\hat{\mathbf{p}}$ defined by transforming p_1,p_2 to the centre of mass frame as in (4.171), then

$$\langle\hat{\mathbf{p}}',h_1{}'h_2{}'\|1\|\hat{\mathbf{p}},h_1h_2\rangle = (4\pi)^2\,\delta^2(\hat{\mathbf{p}}',\hat{\mathbf{p}})\,\delta_{h_1{}'h_1}\delta_{h_2{}'h_2}\,. \tag{4.175}$$

Here $\delta^2(\hat{\mathbf{p}}',\hat{\mathbf{p}})$ is the delta function on the unit sphere so that $\int \mathrm{d}\Omega_{\hat{\mathbf{p}}}\,\delta^2(\hat{\mathbf{p}}',\hat{\mathbf{p}})f(\hat{\mathbf{p}}) = f(\hat{\mathbf{p}}')$.

Using the orthogonality condition (3.124) the centre of mass states in (4.169) can be projected onto states of definite angular momentum by taking

$$\begin{aligned}|\mathring{P},JM,h_1h_2\rangle &= N_J\,\frac{1}{8\pi^2}\int_{SO(3)}\mathrm{d}\mu_{\theta,\phi,\psi}\;U[R_{\phi,\theta,\psi}]\,|\mathring{P},p\mathbf{e}_3,h_1h_2\rangle D^{(J)}_{Mh}(R_{\phi,\theta,\psi})^*\\ &= N_J\,\frac{1}{4\pi}\int_0^{2\pi}\!\mathrm{d}\phi\int_0^{2\pi}\!\mathrm{d}\theta\,\sin\theta\,|\mathring{P},p\hat{\mathbf{p}},h_1h_2\rangle D^{(J)}_{Mh}(R_{\phi,\theta,-\phi})^*\,,\end{aligned} \tag{4.176}$$

where for the ψ integration to be non zero requires $m = h = h_1 - h_2$. These states are then a basis for two particle states of total angular momentum J,

$$\begin{aligned}\langle J'M', h_1{}'h_2{}' \| 1 \| JM, h_1 h_2\rangle = {}& N_J{}^2 \int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\theta \, \sin\theta \, D^{(J')}_{M'h}(R_{\phi,\theta,-\phi}) D^{(J)}_{Mh}(R_{\phi,\theta,-\phi})^* \, \delta_{h_1{}'h_1} \delta_{h_2{}'h_2} \\ = {}& \delta_{J'J}\, \delta_{M'M}\, \delta_{h_1{}'h_1} \delta_{h_2{}'h_2} \,, \quad \text{for} \quad N_J{}^2 = \tfrac{2J+1}{4\pi} \,. \end{aligned} \tag{4.177}$$

Using the completeness relation

$$\frac{1}{4\pi} \sum_{J,M} (2J{+}1)\, D^{(J)}_{Mh}(R_{\phi',\theta',-\phi'})\, D^{(J)}_{Mh}(R_{\phi,\theta,-\phi})^* = \delta^2(\hat{\mathbf{p}}', \hat{\mathbf{p}}) = \delta(\cos\theta' {-} \cos\theta)\delta(\phi' {-} \phi) \,, \tag{4.178}$$

(4.176) can be inverted as an expansion over angular momentum states

$$|\mathring{P}, p\hat{\mathbf{p}}, h_1 h_2\rangle = \sum_{J,M} N_J \, |\mathring{P}, JM, h_1 h_2\rangle \, D^{(J)}_{M\, h_1 - h_2}(R_{\phi,\theta,-\phi}) \,. \tag{4.179}$$

For a rotationally scalar operator, which commutes with $\mathbf{J}$,

$$\langle J'M', h_1{}'h_2{}' \| A \| JM, h_1 h_2\rangle = A^{(J)}_{h_1{}'h_2{}', h_1 h_2} \, \delta_{J'J}\, \delta_{M'M} \,, \tag{4.180}$$

with $A^{(J)}_{h_1 h_2}$ independent of M by virtue of the Wigner-Eckart theorem. With the expansion (4.179)

$$\begin{aligned}\langle \hat{\mathbf{p}}', h_1{}'h_2{}' \| A \| \hat{\mathbf{p}}, h_1 h_2\rangle = {}& \frac{1}{4\pi} \sum_{J,M} (2J+1)\, A^{(J)}_{h_1{}'h_2{}', h_1 h_2} \, D^{(J)}_{Mh'}(R_{\theta',\phi',-\phi'})^* D^{(J)}_{Mh}(R_{\phi,\theta,-\phi}) \,, \\ & h' = h_1{}' - h_2{}' \,, \quad h = h_1 - h_2 \,, \end{aligned} \tag{4.181}$$

The two D-functions can be combined using the group property but more simply $\hat{\mathbf{p}}$ can be chosen to be along the 3-direction giving

$$\langle \hat{\mathbf{p}}, h_1{}'h_2{}' \| A \| \mathbf{e}_3, h_1 h_2\rangle = \frac{1}{4\pi} \sum_J (2J+1)\, A^{(J)}_{h_1{}'h_2{}', h_1 h_2} \, d^{(J)}_{hh'}(\theta)\, e^{i(h-h')\phi} \,, \tag{4.182}$$

for $\hat{\mathbf{p}}$ given in terms of θ, ϕ according (4.135).

This approach to the angular momentum decomposition of relativistic two particle states was first introduced by Jacob and Wick.[38] It avoids the complications of combining spins and then orbital angular momentum using Clebsch-Gordan coefficients. It was of course designed to apply to scattering amplitudes for spinning particles where taking $A \to S$, scattering operator, in (4.182) gives the partial wave decomposition. By virtue of (4.177) the unitarity condition for elastic scattering amplitudes $S^{(J)}_{h_1{}'h_2{}', h_1 h_2}$ reduces to $\sum_{h_1{}'', h_2{}''} S^{(J)}_{h_1{}'', h_2{}'', h_1{}'h_2{}'}{}^* S^{(J)}_{h_1{}''h_2{}'', h_1 h_2} = \delta_{h_1{}'h_1} \delta_{h_2{}'h_2}$. For the spinless case, using (3.108), (4.182) becomes the standard expansion in terms of Legendre polynomials.

[38] Maurice René Michel Jacob, 1933-2007, French. Gian Carlo Wick, 1909-1992, Italian.

4.5.5 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

$$p_{\alpha\dot\alpha} = p^\mu(\sigma_\mu)_{\alpha\dot\alpha}\,, \qquad [p_{\alpha\dot\alpha}] = \begin{pmatrix} p^0+p^3 & p^1-ip^2 \\ p^1+ip^2 & p^0-p^3 \end{pmatrix}\,, \tag{4.183}$$

similarly to (4.52), then for $p^2=0$

$$\det[p_{\alpha\dot\alpha}] = 0 \quad \Rightarrow \quad p_{\alpha\dot\alpha} = \lambda_\alpha \bar\lambda_{\dot\alpha}\,. \tag{4.184}$$

The spinor λ_α and its conjugate $\bar\lambda_{\dot\alpha}$ are arbitrary up to the $U(1)$ transformation given by $\lambda_\alpha \to \lambda_\alpha e^{-i\eta}$, $\bar\lambda_{\dot\alpha} \to \bar\lambda_{\dot\alpha} e^{i\eta}$. To determine λ_α precisely we choose the phase so that λ_1 is real. For $\mathring{p}$ given by (4.145) then, for simplicity choosing $2\mathring{\omega}=1$,

$$[\mathring{p}_{\alpha\dot\alpha}] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \mathring{\lambda} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,. \tag{4.185}$$

If $p_i{}^2 = p_j{}^2 = 0$ then

$$p_{i,\alpha\dot\alpha} = \lambda_{i,\alpha}\bar\lambda_{i,\dot\alpha}\,, \qquad p_{j,\alpha\dot\alpha} = \lambda_{j,\alpha}\bar\lambda_{j,\dot\alpha}\,, \tag{4.186}$$

and

$$2\,p_i\cdot p_j = p_{i,\alpha\dot\alpha}\,p_j{}^{\dot\alpha\alpha} = \langle ij\rangle[ij]\,, \qquad \langle ij\rangle = \varepsilon^{\alpha\beta}\lambda_{i,\alpha}\lambda_{j,\beta}\,, \quad [ij] = \varepsilon^{\dot\alpha\dot\beta}\bar\lambda_{i,\dot\beta}\bar\lambda_{j,\dot\beta}\,. \tag{4.187}$$

As a consequence of the map $SO(3,1)\to Sl(2,\mathbb{C})$ then for the massless case

$$L(p)\mathring{p} = p \quad \Rightarrow \quad A(\lambda_p)\mathring{\lambda} = \lambda_p\,, \tag{4.188}$$

defines $p\to\lambda_p$ uniquely, at least up to a sign, satisfying (4.184). For any Lorentz transformation $\Lambda\to A_\Lambda$ then (4.163) becomes equivalently

$$A(\lambda_{\Lambda p})^{-1}A_\Lambda A(\lambda_p) = A_{(a_1,a_2)}A_\chi\,, \tag{4.189}$$

where, from (4.160), we have correspondingly under $SO(3,1)\to Sl(2,\mathbb{C})$

$$\Lambda_{(a_1,a_2)} \to A_{(a_1,a_2)} = \begin{pmatrix} 1 & a_1-ia_2 \\ 0 & 1 \end{pmatrix}\,, \qquad \Lambda_\chi \to A_\chi = \begin{pmatrix} e^{-\frac12 i\chi} & 0 \\ 0 & e^{\frac12 i\chi} \end{pmatrix}\,. \tag{4.190}$$

With the definition (4.188) of λ_p, (4.189) implies

$$A_\Lambda \lambda_p = \lambda_{\Lambda p}\, e^{-\frac12 i\chi(p,\Lambda)}\,. \tag{4.191}$$

The spinorial formalism a convenient method of calculating $\chi(p,\Lambda)$. For helicity states with $L(p)$ as in (4.136) the corresponding $Sl(2,\mathbb{C})$ matrix is then

$$A_R(\Theta,\mathbf{n})A_B(\alpha,\mathbf{e}_3) = \begin{pmatrix} \cos\frac{1}{2}\theta\, e^{\frac{1}{2}\alpha} & -\sin\frac{1}{2}\theta\, e^{-i\phi-\frac{1}{2}\alpha} \\ \sin\frac{1}{2}\theta\, e^{i\phi+\frac{1}{2}\alpha} & \cos\frac{1}{2}\theta\, e^{-\frac{1}{2}\alpha} \end{pmatrix} . \tag{4.192}$$

In (4.188) we can then take

$$A(\lambda) = \begin{pmatrix} \lambda_1 & -\lambda_2{}^*/(\lambda_1{}^2+|\lambda_2|^2) \\ \lambda_2 & \lambda_1{}^*/(\lambda_1{}^2+|\lambda_2|^2) \end{pmatrix} , \quad \lambda_{p,1} = (p^0+p^3)^{\frac{1}{2}} , \quad \lambda_{p,2} = \frac{p^1+i\,p^2}{(p^0+p^3)^{\frac{1}{2}}} , \tag{4.193}$$

where the phase ambiguity in λ is resolved by requiring λ_1 to be real.

As a illustration we consider three examples

$$\begin{aligned}
&\Lambda = \Lambda_{R(\sigma,\mathbf{e}_3)} , \\
&\lambda_{\Lambda p,1} = \lambda_{p,1} , \ \lambda_{\Lambda p,2} = e^{i\sigma}\lambda_{p,1} , \quad \chi(p,\Lambda) = \sigma , \\
&\Lambda = \Lambda_{R(\rho,\mathbf{e}_2)} , \\
&\begin{pmatrix} \lambda_{\Lambda p,1} \\ \lambda_{\Lambda p,2} \end{pmatrix} = \begin{pmatrix} \cos\frac{1}{2}\rho & -\sin\frac{1}{2}\rho \\ \sin\frac{1}{2}\rho & \cos\frac{1}{2}\rho \end{pmatrix} \begin{pmatrix} \lambda_{p,1} \\ \lambda_{p,2} \end{pmatrix} e^{\frac{1}{2}i\eta} , \ \tan\tfrac{1}{2}\eta = \tan\tfrac{1}{2}\rho \, \frac{p^1+i\,p^2}{p^0+p^3} , \quad \chi(p,\Lambda) = \eta , \\
&\Lambda = B(\beta,\mathbf{e}_3) , \\
&\lambda_{\Lambda p,1} = e^{\frac{1}{2}\beta}\lambda_{p,1} , \ \lambda_{\Lambda p,2} = e^{-\frac{1}{2}\beta}\lambda_{p,1} , \quad \chi = 0 , \ a_1 - ia_2 = \frac{(e^{2\beta}-1)\lambda_{p,1}\lambda_{p,2}{}^*}{(\lambda_{p,1}{}^2+|\lambda_{p,2}|^2)(\lambda_{p,1}{}^2+e^{-2\beta}|\lambda_{p,2}|^2)} .
\end{aligned} \tag{4.194}$$

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 (3.83), 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*[39] *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

$$\tfrac{1}{4}M^{\mu\nu}M_{\mu\nu} , \qquad \tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho}M_{\mu\nu}M_{\sigma\rho} . \tag{4.195}$$

In terms of the generators $\mathbf{J},\mathbf{K}$, defined in (4.37),(4.41), and then $\mathbf{J}^\pm$, defined in (4.46),

$$\begin{aligned}
\tfrac{1}{4}M^{\mu\nu}M_{\mu\nu} &= \tfrac{1}{2}(\mathbf{J}^2-\mathbf{K}^2) = \mathbf{J}^{+2}+\mathbf{J}^{-2} , \\
\tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho}M_{\mu\nu}M_{\sigma\rho} &= \mathbf{J}\cdot\mathbf{K} = -i\big(\mathbf{J}^{+2}-\mathbf{J}^{-2}\big) .
\end{aligned} \tag{4.196}$$

[39]Hendrik Brugt Gerhard Casimir, 1909-2000, Dutch.

Since $\mathbf{J}^{\pm}$ both obey standard angular momentum commutation relations, as in (4.47), then for finite dimensional irreducible representations

$$\mathbf{J}^{+2} \to j(j+1)\,\mathbb{1}\,, \quad \mathbf{J}^{-2} \to \bar{\jmath}(\bar{\jmath}+1)\,\mathbb{1}\,, \qquad j,\bar{\jmath} = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots\,. \tag{4.197}$$

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

$$\begin{aligned} \tfrac{1}{4} s^{\mu\nu} s_{\mu\nu} &= -\tfrac{1}{32}\, \sigma^{\mu}\bar{\sigma}^{\nu}(\sigma_\mu\bar{\sigma}_\nu - \sigma_\nu\bar{\sigma}_\nu) = \tfrac{3}{4}\,\mathbb{1}\,, \\ \tfrac{1}{8}\varepsilon^{\mu\nu\sigma\rho} s_{\mu\nu} s_{\sigma\rho} &= \tfrac{1}{32}\, \varepsilon^{\mu\nu\sigma\rho}\sigma_\mu\bar{\sigma}_\nu\sigma_\sigma\bar{\sigma}_\rho = -\tfrac{3}{4} i\,\mathbb{1}\,, \end{aligned} \tag{4.198}$$

using (4.49) and (4.91). As expected this is in accord with (4.196) and (4.197) 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 (4.195) no longer provides Casimir operators because they fail to commute with P_μ. There is now only a single quadratic Casimir

$$P^2 = P^\mu P_\mu\,, \tag{4.199}$$

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*,

$$W^\mu = \tfrac{1}{2}\, \varepsilon^{\mu\nu\sigma\rho} P_\nu M_{\sigma\rho} = \tfrac{1}{2}\, \varepsilon^{\mu\nu\sigma\rho} M_{\sigma\rho} P_\nu\,. \tag{4.200}$$

Using $\varepsilon^{\mu\nu\sigma\rho} P_\nu P_\sigma = 0$ we have

$$[W^\mu, P_\nu] = 0\,. \tag{4.201}$$

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

$$\begin{aligned} [W^\mu, \tfrac{1}{2}\omega^{\sigma\rho} M_{\sigma\rho}] &= -\tfrac{1}{2} i\, \varepsilon^{\mu\nu\sigma\rho}\big(P_\lambda \omega^\lambda{}_\nu\, M_{\sigma\rho} + P_\nu\, M_{\lambda\rho}\omega^\lambda{}_\sigma + P_\nu\, M_{\sigma\lambda}\omega^\lambda{}_\sigma\big) \\ &= \tfrac{1}{2} i\, \omega^\mu{}_\lambda \varepsilon^{\lambda\nu\sigma\rho} P_\nu M_{\sigma\rho} = i\, \omega^\mu{}_\lambda W^\lambda\,, \end{aligned} \tag{4.202}$$

to obtain

$$[W^\mu, M_{\sigma\rho}] = i(\delta^\mu{}_\sigma W_\rho - \delta^\mu{}_\rho W_\sigma)\,. \tag{4.203}$$

With (4.201) and (4.203) we may then easily derive

$$[W^\mu, W^\nu] = i\, \varepsilon^{\mu\nu\sigma\rho} P_\sigma W_\rho\,. \tag{4.204}$$

It follows from (4.201) and (4.203) that

$$W_\mu W^\mu\,, \tag{4.205}$$

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),

$$W^0|\mathring{p}, s\, s_3\rangle = 0\,, \qquad W^i|\mathring{p}, s\, s_3\rangle = -m\,\varepsilon_{ijk}M_{jk}|\mathring{p}, s\, s_3\rangle = -m\, J_i|\mathring{p}, s\, s_3\rangle\,, \tag{4.206}$$

so that

$$W_\mu W^\mu|\mathring{p}, s\, s_3\rangle = -m^2\,\mathbf{J}^2|\mathring{p}, s\, s_3\rangle = -m^2 s(s+1)|\mathring{p}, s\, s_3\rangle\,. \tag{4.207}$$

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 (4.145),

$$\begin{aligned} &W^1|\mathring{p}, h\rangle = \mathring{\omega}\, E_2|\mathring{p}, h\rangle = 0\,, && W^2|\mathring{p}, h\rangle = -\mathring{\omega}\, E_1|\mathring{p}, h\rangle = 0\,,\\ &W^0|\mathring{p}, h\rangle = -\mathring{\omega}\, J_3|\mathring{p}, h\rangle = -\mathring{\omega}\, h|\mathring{p}, h\rangle\,, && W^3|\mathring{p}, h\rangle = -\mathring{\omega}\, J_3|\mathring{p}, h\rangle = -\mathring{\omega}\, h|\mathring{p}, h\rangle\,, \end{aligned} \tag{4.208}$$

using (4.157). Since W^μ, P^μ are both contravariant 4-vectors the result (4.208) requires

$$(W^\mu + h\, P^\mu)|p, h\rangle = 0\,, \tag{4.209}$$

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

$$U[\Lambda, a]\phi(x)U[\Lambda, a]^{-1} = \phi(\Lambda x + a)\,, \tag{4.210}$$

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

$$-\,i\big[\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu} - a^\mu P_\mu, \phi(x)\big] = (\omega^\mu{}_\nu x^\nu + a^\mu)\partial_\mu\phi(x)\,, \tag{4.211}$$

or

$$[M_{\mu\nu}, \phi(x)] = -L_{\mu\nu}\phi(x)\,, \quad L_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu)\,, \qquad [P_\mu, \phi(x)] = -i\partial_\mu\phi(x)\,. \tag{4.212}$$

$L_{\mu\nu}$ 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), $[\mathbf{P}, \phi] = i\,\nabla\phi$.

To describe particles with spin the quantum fields are required to transform according to a finite dimensional representation of the Lorentz group so that (4.210) is extended to

$$U[\Lambda, a]\phi(x)U[\Lambda, a]^{-1} = D(\Lambda)^{-1}\phi(\Lambda x + a)\,, \tag{4.213}$$

regarding ϕ now as a column vector and suppressing matrix indices. For an infinitesimal Lorentz transformation then assuming

$$D(\Lambda) = \mathbb{1} - i\,\tfrac{1}{2}\omega^{\mu\nu}S_{\mu\nu}\,, \qquad S_{\mu\nu} = -S_{\nu\mu}\,, \tag{4.214}$$

the commutator with $M_{\mu\nu}$ in (4.212) is extended to

$$[M_{\mu\nu}, \phi(x)] = -\big(L_{\mu\nu} + S_{\mu\nu}\big)\phi(x)\,. \tag{4.215}$$

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 $\mathcal{V}_{m,s}$,

$$\langle 0|\phi(x)|p, s\,s_3\rangle = u(p, s_3)\,e^{-ip\cdot x}\,, \qquad p^2 = m^2\,. \tag{4.216}$$

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 (4.216) is accord with translation invariance using (4.128). Using (4.213), for $a = 0$, $\Lambda \to \Lambda^{-1}$, with (4.127) we get

$$D(\Lambda)\,u(p, s_3) = \sum_{s'_3} u(\Lambda p, s'_3)\,D^{(s)}_{s'_3\,s_3}\big(\mathcal{R}(p,\Lambda)\big)\,, \tag{4.217}$$

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 (4.217) 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),

$$D(\Lambda_R)\,u(\mathring{p}, s_3) = \sum_{s'_3} u(\mathring{p}, s'_3)\,D^{(s)}_{s'_3\,s_3}(R)\,, \tag{4.218}$$

and then define, as in (4.126),

$$u(p, s_3) = D\big(L(p)\big)\,u(\mathring{p}, s_3)\,. \tag{4.219}$$

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 (4.218) 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 (4.216) we require

$$\langle 0|A^\mu(x)|p, 1\rangle = \epsilon^\mu(p)\,e^{-ip\cdot x}\,, \qquad p^2 = 0\,. \tag{4.220}$$

$\epsilon^\mu(p)$ 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 (4.145) then from the little group transformations as in (4.163) we require, for $h = 1$,

$$\Lambda_\chi \epsilon(\mathring{p}) = \epsilon(\mathring{p})\,e^{-i\chi}\,. \tag{4.221}$$

Using (4.160) this determines $\epsilon(\mathring{p})$ to be

$$\epsilon^\mu(\mathring{p}) = \tfrac{1}{\sqrt{2}}(0,1,i,0)\,, \tag{4.222}$$

with a normalisation $\epsilon^* \cdot \epsilon = -1$. Furthermore from the explicit form for $\Lambda_{(a_1,a_2)}$ also in (4.160) we then obtain

$$\Lambda_{(a_1,a_2)}\epsilon(\mathring{p}) = \epsilon(\mathring{p}) + c\,\mathring{p}\,, \qquad c = \tfrac{1}{\sqrt{2}}(a_1 + a_2)\,. \tag{4.223}$$

For general momentum $p = \omega(1,\mathbf{n})$, $p^2 = 0$, as in (4.166), we may define, for $L(p)$ given by (4.136),

$$\epsilon(p) = L(p)\,\epsilon(\mathring{p}) = \Lambda_{R(n)}\,\epsilon(\mathring{p})\,, \tag{4.224}$$

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 (4.136). With the definition (4.224)

$$p_\mu \epsilon^\mu(p) = \mathring{p}_\mu \epsilon^\mu(\mathring{p}) = 0\,. \tag{4.225}$$

For a general Lorentz transformation Λ then from (4.163) and (4.221),(4.222)

$$\Lambda\epsilon(p) = \big(\epsilon(\Lambda p) + c\,\Lambda p\big)\,e^{-i\chi(p,\Lambda)}\,, \tag{4.226}$$

for some c depending on p, Λ. This matches (4.165), for $h = 1$, save for the inhomogeneous term proportional to c (for $h = -1$ it is sufficient to take $\epsilon(p) \to \epsilon(p)^*$). (4.226) 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

$$\epsilon(p) \sim \epsilon(p) + c\,p\,, \quad \text{for arbitrary} \quad c\,. \tag{4.227}$$

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 (4.225), that scalar products of polarisation vectors depend only on their equivalence classes so that

$$\epsilon'(p)^* \cdot \epsilon'(p) = \epsilon(p)^* \cdot \epsilon(p) \quad \text{for} \quad \epsilon'(p) \sim \epsilon(p)\,. \tag{4.228}$$

The *gauge freedom* in (4.227) 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[40] groups is more accessible than that for finite discrete groups, the classification of the former was completed by Cartan[41] 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

$$g(0) = e \quad \Rightarrow \quad \varphi^r(0,a) = \varphi^r(a,0) = a^r \,, \tag{5.3}$$

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

$$\varphi^r(a,b) = a^r + b^r \,, \tag{5.6}$$

[40] Marius Sophus Lie, 1842-1899, Norwegian.
[41] Élie Joseph Cartan, 1869-1951, French.

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*[42] *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

$$\omega_j(x) \frac{\partial x^j}{\partial x'^i} = \omega'_i(x') . \tag{5.19}$$

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

$$(\mathrm{d}\omega)_{ij}(x) = \partial_i \omega_j(x) - \partial_j \omega_i(x) . \tag{5.23}$$

[42] Carl Gustav Jacob Jacobi, 1804-1851, German.

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 $U(1)$ corresponds 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 *derived* subalgebras $\mathfrak{g}^{(n)}$, $n = 1, 2, \dots$, forming the *derived series* 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 derived series 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 Matrix Lie Algebras and 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

$$D(\theta) = \mathbb{1} + \theta^a t_a\,, \tag{5.56}$$

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

$$A(\mathbf{u}) = u_0\,\mathbb{1} + i\,\mathbf{u}\cdot\boldsymbol{\sigma} \qquad u_0 = \sqrt{1 - \mathbf{u}^2}\,. \tag{5.61}$$

This parameterisation is valid for $u_0 \geq 0$. With, for infinitesimal $\boldsymbol{\theta}$, $A(\boldsymbol{\theta}) = \mathbb{1} + i\,\boldsymbol{\theta}\cdot\boldsymbol{\sigma}$ we get, using the standard results (3.20) 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 $\mathfrak{su}(2)$ are then

$$T_j(\mathbf{u}) = \mu_{ji}(\mathbf{u})\frac{\partial}{\partial u_i} \quad \Rightarrow \quad \mathbf{T} = u_0\,\nabla_{\mathbf{u}} + \mathbf{u}\times\nabla_{\mathbf{u}}\,. \tag{5.64}$$

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 Upper Triangular Matrices

The upper triangular and the strictly upper triangular matrices

$$\mathfrak{b} = \left\{\begin{pmatrix} x & x & x & . & x \\ 0 & x & x & . & x \\ 0 & 0 & x & . & x \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & . & x \end{pmatrix}\right\}, \qquad \mathfrak{n} = \left\{\begin{pmatrix} 0 & x & x & . & x \\ 0 & 0 & x & . & x \\ 0 & 0 & 0 & . & x \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & . & 0 \end{pmatrix}\right\}, \tag{5.67}$$

form Lie algebras with the commutator defined by usual matrix multiplication. It is easy to see that

$$\mathfrak{n} = [\mathfrak{b},\mathfrak{b}]\,, \tag{5.68}$$

and that the Lie algebras $\mathfrak{b}$ and hence also $\mathfrak{n}$ are solvable.

5.3.3 Representations and Lie Algebras

There is an intimate relation between representations of Lie algebras and Lie Groups. Just as described for groups in 2, 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,

$$D(g(\theta)) = \mathbb{1} + \theta^a t_a + \mathrm{O}(\theta^2)\,, \qquad D(g(\theta))^{-1} = \mathbb{1} - \theta^a t_a + \mathrm{O}(\theta^2)\,. \tag{5.69}$$

For unitary representations, as in (2.37), the matrix generators are then anti-hermitian,

$$t_a{}^\dagger = -t_a\,. \tag{5.70}$$

If the representation matrices have unit determinant, since $\det(\mathbb{1}+\epsilon X) = 1+\epsilon\,\mathrm{tr}(X)+\mathrm{O}(\epsilon^2)$, we must also have

$$\mathrm{tr}(t_a) = 0\,. \tag{5.71}$$

In a physics context it is commonplace to redefine the matrix generators so that $t_a = -i\hat{t}_a$ so that, instead of (5.70), the generators $\hat{t}_a$ are hermitian and satisfy the commutation relations $[\hat{t}_a,\hat{t}_b] = if^c{}_{ab}\hat{t}_c$.

Two representations of a Lie algebra $\{t'_a\}$ and $\{t_a\}$ are equivalent if, for some non singular S,

$$t'_a = S\,t_a S^{-1}\,. \tag{5.72}$$

For both representations to be unitary then S must be unitary. If the representation is irreducible then, by applying Schur's lemma,

$$t_a = S\,t_a S^{-1} \quad \text{or} \quad [S,t_a] = 0 \quad \Rightarrow \quad S \propto I\,. \tag{5.73}$$

The complex conjugate of a representation is also a representation, in general it is inequivalent. If it is equivalent then, for some C,

$$t_a{}^* = C\,t_a C^{-1}\,, \tag{5.74}$$

or for a unitary representation, assuming (5.70),

$$C\,t_a C^{-1} = -t_a{}^T\,. \tag{5.75}$$

Following the same argument as in 2.3.2, combining (5.75) with its transpose we get $C^{-1T}C\,t_a\,C^{-1}C^T = t_a$ so that for an irreducible representation

$$C^{-1\,T}C = c\,I \quad \Rightarrow \quad C = c\,C^T \quad \Rightarrow \quad c = \pm 1\,. \tag{5.76}$$

If $C = C^T$ then, by a transformation $C \to S^T CS$ 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 (1.108). The representation matrices then satisfy $D(g(\theta))^\dagger = -JD(g(\theta))J$, which is just as in (1.79). 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 3.14 is pseudo-real since

$$C\,\boldsymbol{\sigma}\,C^{-1} = -\boldsymbol{\sigma}^T \qquad \text{for} \quad C = i\sigma_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \tag{5.77}$$

which is equivalent to (3.296).

A corollary of (5.75) is that,for real or pseudo-real representations,

$$\mathrm{tr}\big(t_{(a_1}\dots t_{a_n)}\big) = 0 \qquad \text{for} \;\; n \;\; \text{odd}\,. \tag{5.78}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}s}\, a_s^r = \theta^a \mu_a{}^r(a_s)\,, \quad a_0^r = 0\,, \qquad \text{or} \quad \frac{\mathrm{d}}{\mathrm{d}s}\, g(a_s) = \theta^a T_a(a_s)\, g(a_s)\,. \tag{5.79}$$

To verify that this forms a subgroup consider $g(c) = g(a_t)g(a_s)$ where from (5.2)

$$c^r = \varphi^r(a_t, a_s)\,. \tag{5.80}$$

Using (5.79) with (5.31) we get

$$\frac{\partial}{\partial s}\, c^r = \theta^a \mu_a{}^u(a_s)\, \lambda_u{}^b(a_s)\, \mu_b{}^r(c) = \theta^b \mu_b{}^r(c)\,, \qquad c^r\big|_{s=0} = a_t^r\,. \tag{5.81}$$

The equation is then identical with (5.79), save for the initial condition at $s = 0$, and the solution then becomes

$$c^r = a_{s+t}^r \quad \Rightarrow \quad g(a_t)g(a_s) = g(a_{s+t})\,. \tag{5.82}$$

Since

$$g(a_s)^{-1} = g(a_{-s})\,, \tag{5.83}$$

then $\{g(a_s)\}$ forms an abelian subgroup of G depending on the parameter s. We may then define an *exponential map*

$$\exp : \mathfrak{g} \to G\,, \tag{5.84}$$

by

$$g(a_s) = \exp(s\, \theta^a T_a)\,. \tag{5.85}$$

For any representation we have

$$D(g(a_s)) = e^{s\, \theta^a t_a}\,, \tag{5.86}$$

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

$$\exp(tX)\,\exp(tY) = \exp(Z(t))\,, \qquad Z(t) \in \mathfrak{g}\,. \tag{5.87}$$

The *Baker Cambell Hausdorff*[43] *formula* gives an infinite series for $Z(t)$ in powers of t whose first terms are of the form

$$Z(t) = t(X+Y) + \tfrac{1}{2}t^2[X,Y] + \tfrac{1}{12}t^3\big([X,[X,Y]] - [Y,[X,Y]]\big) + \mathrm{O}(t^4)\,, \tag{5.88}$$

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

$$f(s) = e^{s(Z+\delta Z)}\, e^{-sZ}\,, \tag{5.89}$$

and then

$$\frac{\mathrm{d}}{\mathrm{d}s} f(s) = e^{s(Z+\delta Z)}\delta Z\, e^{-sZ} = e^{sZ}\delta Z\, e^{-sZ} + \mathrm{O}(\delta Z^2)\,. \tag{5.90}$$

Solving this equation

$$f(1) = \mathbb{1} + \int_0^1 \mathrm{d}s\; e^{sZ}\delta Z\, e^{-sZ} + \mathrm{O}(\delta Z^2)\,, \tag{5.91}$$

so that

$$e^{Z+\delta Z} - e^Z = \int_0^1 \mathrm{d}s\; e^{sZ}\delta Z\, e^{(1-s)Z} + \mathrm{O}(\delta Z^2)\,. \tag{5.92}$$

Hence for any $Z(t)$ we have the result for the derivative of its exponential

$$\frac{\mathrm{d}}{\mathrm{d}t}\, e^{Z(t)} = \int_0^1 \mathrm{d}s\; e^{sZ(t)}\, \frac{\mathrm{d}}{\mathrm{d}t} Z(t)\, e^{(1-s)Z(t)}\,. \tag{5.93}$$

If, instead of (5.87), we suppose,

$$e^{tX}\, e^{tY} = e^{Z(t)}\,, \tag{5.94}$$

then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{tX}\, e^{tY}\big)\, e^{-tY}\, e^{-tX} &= X + e^{tX}\, Y\, e^{-tX} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\, e^{Z(t)}\; e^{-Z(t)} = \int_0^1 \mathrm{d}s\; e^{sZ(t)}\, \frac{\mathrm{d}}{\mathrm{d}t}\, Z(t)\, e^{-sZ(t)}\,.\end{aligned} \tag{5.95}$$

[43] Henry Frederick Baker, 1866-1956, British, senior wrangler 1887. John Edward Cambell, 1862-1924, Irish. Felix Hausdorff, 1868-1942, German.

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

$$e^A B e^{-A} = B + [A, B] + \tfrac{1}{2}[A, [A, B]] + \dots , \tag{5.96}$$

(5.95) can be rewritten as an expansion in multiple commutators

$$X + e^{tX} Y e^{-tX} = \frac{\mathrm{d}}{\mathrm{d}t} Z(t) + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \underbrace{\big[Z(t), \dots \big[Z(t),}_{n} \frac{\mathrm{d}}{\mathrm{d}t} Z(t)\big] \dots \big], \tag{5.97}$$

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

$$f(X^{\mathrm{ad}})Y = \sum_{n=0}^{\infty} f_n \underbrace{\big[X, \dots \big[X,}_{n} Y\big] \dots \big] \qquad \text{for} \qquad f(x) = \sum_{n=0}^{\infty} f_n x^n , \tag{5.98}$$

so that (5.96) becomes $e^A B e^{-A} = e^{A^{\mathrm{ad}}} B$. Then, since $\int_0^1 \mathrm{d}s \, e^{sz} = (e^z - 1)/z$, (5.95) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} Z(t) = f\big(e^{Z(t)^{\mathrm{ad}}}\big)\big(X + e^{tX^{\mathrm{ad}}} Y\big) , \tag{5.99}$$

for, using the standard series expansion of $\ln(1 + x)$,

$$f(x) = \frac{\ln x}{x - 1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n + 1} (x - 1)^n . \tag{5.100}$$

Since

$$e^{Z(t)^{\mathrm{ad}}} U = e^{Z(t)} U e^{-Z(t)} = e^{tX} e^{tY} U e^{-tY} e^{-tX} = e^{tX^{\mathrm{ad}}} e^{tY^{\mathrm{ad}}} U , \tag{5.101}$$

we may replace $e^{Z(t)^{\mathrm{ad}}} \to e^{tX^{\mathrm{ad}}} e^{tY^{\mathrm{ad}}}$ on the right hand side of (5.99). With some intricate combinatorics (5.99) 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

$$e^{-tX} e^{-tY} e^{tX} e^{tY} = e^{t^2 [X,Y] + \mathrm{O}(t^3)} , \tag{5.102}$$

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

$$(p_{x_1 \to x_2} \circ p_{x_2 \to x_3})(s) = \begin{cases} p_{x_1 \to x_2}(2s) , & 0 \le s \le \frac{1}{2} , \\ p_{x_2 \to x_3}(2s - 1) , & \frac{1}{2} \le s \le 1 . \end{cases} \tag{5.103}$$

For any $p_{x_1\to x_2}$ the corresponding inverse, and also the trivial identity path, are defined by

$$p^{-1}_{x_2\to x_1}(s) = p_{x_1\to x_2}(1-s)\,, \qquad p^{\rm id}_{x\to x}(s) = x\,. \tag{5.104}$$

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 $x \in \mathcal{M}$,

$$\pi_1(\mathcal{M}) = \big\{[p_{x\to x}]\big\}\,. \tag{5.105}$$

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 (5.105). $\mathcal{M}$ 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 (3.8), $\mathcal{M}_{SO(3)}$ 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

$$\pi_1(SO(3))\simeq\mathbb{Z}_2\,. \tag{5.106}$$

As another example we may consider the group $U(1)$, as in (1.106), 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

$$\pi_1(U(1))\simeq\mathbb{Z}\,, \tag{5.107}$$

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

$$p_{i,e\to g}(s) = g_i(s)\,,\quad g_i(0)=e\,,\ \ g_i(1)=g\,,\qquad i=0,\dots,n-1\,, \tag{5.108}$$

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

$$g_i = \big(g, [p_{i,e\to g}]\big) \in \overline{G} \qquad \text{for all} \quad g \in G\,, \;\; i = 0, \dots, n-1\,, \tag{5.109}$$

with a corresponding group product

$$g_{1i}\; g_{2j} = g_k\,, \quad \text{for} \quad g = g_1 g_2\,, \;\; [p_{k,e\to g_1 g_2}] = [p_{i,e\to g_1} \circ g_1 p_{j,e\to g_2}]\,, \tag{5.110}$$

using the path composition as in (5.103) 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 (5.104),

$$g_i{}^{-1} = \big(g^{-1}, [g^{-1} p^{-1}_{i,g\to e}]\big)\,, \qquad e_0 = \big(e, [p_{0,e\to e}]\big)\,, \quad p_{0,e\to e} = p^{\rm id}_{e\to e}\,. \tag{5.111}$$

These definitions satisfy the group properties although associativity requires some care. $\overline{G}$ contains the normal subgroup given by

$$\big\{e_i : i = 0, \dots, n-1\big\} \simeq \pi_1(G)\,, \qquad e_i = \big(e, [p_{i,e\to e}]\big)\,. \tag{5.112}$$

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

$$G \simeq \overline{G}/\pi_1(G)\,, \qquad \pi_1(G) \subset \mathcal{Z}(\overline{G})\,. \tag{5.113}$$

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 (3.7) but allow the rotation angle range to be extended to $0 \to 2\pi$. Hence, instead of (3.8), we have

$$n \in S^2\,, \quad 0 \le \theta \le 2\pi\,, \qquad (\theta, n) \simeq (2\pi - \theta, -n)\,. \tag{5.114}$$

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 (5.114), by

$$p_{0,I\to R(\theta,n)}(s) = R(s\theta, n)\,, \quad p_{1,I\to R(\theta,n)}(s) = R(s(2\pi - \theta), -n)\,, \qquad 0 \le s \le 1\,, \tag{5.115}$$

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

$$R(\theta, n)_i\; R(\theta', n)_j = \begin{cases} R(\theta + \theta', n)_{\,i+j \bmod 2}\,, & 0 \le \theta + \theta' \le \pi\,, \\ R(\theta + \theta', n)_{\,i+j+1 \bmod 2}\,, & \pi \le \theta + \theta' \le 2\pi\,, \end{cases} \qquad 0 \le \theta, \theta' \le \pi\,. \tag{5.116}$$

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 (1.106) we may define

$$p_{n,1\to e^{i\theta}}(s) = e^{is(\theta+2n\pi)}\,, \qquad 0 \le s \le 1\,, \quad n \in \mathbb{Z}\,, \tag{5.117}$$

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

$$g_n\big(e^{i\theta}\big)\, g_{n'}\big(e^{i\theta'}\big) = \begin{cases} g_{n+n'}\big(e^{i(\theta+\theta')}\big)\,, & 0 \le \theta+\theta' \le 2\pi\,, \\ g_{n+n'+1}\big(e^{i(\theta+\theta')}\big)\,, & 2\pi \le \theta+\theta' \le 4\pi\,, \end{cases} \qquad 0 \le \theta,\theta' \le 2\pi\,. \tag{5.118}$$

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 (5.110). 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,

$$D(g_{1i})\, D(g_{2i}) = D(g_j) = D(g_j\, g_i{}^{-1})\, D(g_i) = D(e_k)\, D(g_i)\,, \tag{5.119}$$

where, by virtue of (5.112) and (5.113),

$$g_j\, g_i{}^{-1} = e_k \in \mathcal{Z}(\overline{G}) \quad \text{for some } k\,. \tag{5.120}$$

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,

$$D(e_k) = e^{i\gamma_k}\mathbb{1}\,, \tag{5.121}$$

where $\{e^{i\gamma_k} : k = 0,\dots,n-1\}$ form a one dimensional representation of $\pi_1(G)$. Combining (5.119) and (5.121) illustrates that $\{D(g_i)\}$, for i fixed, provide a projective representation of G as in (2.151).

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

$$e^{i\alpha\theta}\; e^{i\alpha\theta'} = \begin{cases} e^{i\alpha(\theta+\theta')}\,, & 0 \le \theta+\theta' \le 2\pi\,, \\ e^{2\pi i\alpha}\, e^{i\alpha(\theta+\theta'-2\pi)}\,, & 2\pi \le \theta+\theta' \le 4\pi\,. \end{cases} \tag{5.122}$$

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 (2.152) modulo trivial solutions of the form (2.153) 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 (2.151), 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 (2.152) 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

$$\gamma(c,\theta) + \gamma(a,b) = \gamma(a, b + \mathrm{d}b) + \gamma(b,\theta)\,. \tag{5.123}$$

Defining

$$\gamma_a(b) = \frac{\partial}{\partial\theta^a}\,\gamma(b,\theta)\bigg|_{\theta=0}\,, \tag{5.124}$$

and with (5.27) and the definition (5.32) then (5.123) becomes

$$T_a(b)\,\gamma(a,b) = \gamma_a(c) - \gamma_a(b)\,. \tag{5.125}$$

This differential equation for $\gamma(a,b)$ has integrability conditions obtained by considering

$$[T_a(b), T_b(b)]\,\gamma(a,b) = f^c{}_{ab} T_c(b)\,\gamma(a,b) \tag{5.126}$$

which applied to (5.125) 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

$$T_a(b)\,\gamma_b(b) - T_b(b)\,\gamma_a(b) - f^c{}_{ab}\,\gamma_c(b) = h_{ab} = -h_{ba}\,, \tag{5.127}$$

with h_{ab} a constant. Applying $T_c(b)$ and antisymmetrising the indices a, b, c gives, with (5.41),

$$0 = T_c\,h_{ab} + T_b\,h_{ca} + T_a\,h_{bc} = f^d{}_{ab}(T_d\gamma_c - T_c\gamma_d) + f^d{}_{bc}(T_d\gamma_a - T_a\gamma_d) + f^d{}_{ca}(T_d\gamma_b - T_b\gamma_d)\,, \tag{5.128}$$

and hence, with (5.127) and (5.43), there is then a constraint on h_{ab},

$$f^d{}_{ab}\,h_{dc} + f^d{}_{bc}\,h_{da} + f^d{}_{ca}\,h_{db} = 0\,. \tag{5.129}$$

As was discussed in 2.9 there are trivial solutions of the consistency conditions which are given by (2.153), 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 (5.124) gives

$$\gamma_a(b) = T_a(b)\,\alpha(b) - c_a\,, \qquad c_a = \frac{\partial}{\partial\theta^a}\,\alpha(\theta)\bigg|_{\theta=0}\,, \tag{5.130}$$

and then substituting in (5.127)

$$h_{ab} = f^c{}_{ab}\, c_c\,. \tag{5.131}$$

It is easy to verify that (5.130) and (5.131) satisfy (5.127) and (5.129)[44].

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 (2.151) requires

$$U(b)U(\theta) = e^{i\gamma_a(b)\theta^a} U(b+\mathrm{d}b)\,, \tag{5.132}$$

for infinitesimal θ^a. Assuming

$$U(\theta) = 1 - i\theta^a \hat{T}_a\,, \tag{5.133}$$

for hermitian operators $\hat{T}_a$, then, since $U(b+\mathrm{d}b) = U(b) + \theta^a T_a(b) U(b)$, we have

$$T_a(b)U(b) = -iU(b)\big(\hat{T}_a + \gamma_a(b)\big)\,. \tag{5.134}$$

By considering $[T_a, T_b]U(b)$ and using (5.127) then this requires that the hermitian operators $\{\hat{T}_a\}$ satisfy a modified Lie algebra

$$\big[\hat{T}_a, \hat{T}_b\big] = i f^c{}_{ab}\,\hat{T}_c - i\, h_{ab}\, 1\,. \tag{5.135}$$

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 (5.129) is necessary for $\{\hat{T}_a\}$ to satisfy the Jacobi identity, if (5.131) 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 (4.150), which has a central extension

$$[J_3, E_1] = iE_2\,, \qquad [J_3, E_2] = -iE_1\,, \qquad [E_1, E_2] = ic\,1\,. \tag{5.136}$$

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

$$\mathbf{x}' = R\mathbf{x} + \mathbf{a} + \mathbf{v}t\,, \qquad t' = t + b\,, \tag{5.137}$$

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 $c \to \infty$

[44] Alternatively, using the left invariant one forms in (5.48) and defining $h = \frac{1}{2} h_{ab}\,\omega^a \wedge \omega^b$, then (5.129) is equivalent, by virtue of (5.49), to $\mathrm{d}h = 0$, so that h is closed, while the trivial solution (5.131) 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 $\mathcal{M}_G$.

then the commutation relations from (4.42), (4.43) and (4.107), (4.108) become

$$\begin{aligned} &[J_i, J_j] = i\varepsilon_{ijk} J_k\,, \qquad [J_i, K_j] = i\varepsilon_{ijk} K_k\,, \qquad [K_i, K_j] = 0\,, \qquad [K_i, H] = iP_i\,, \\ &[K_i, P_j] = i\delta_{ij}\, M\,, \qquad [\mathbf{J}, M] = [\mathbf{K}, M] = [\mathbf{P}, M] = [H, M] = 0\,. \end{aligned} \tag{5.138}$$

When the Lie algebra is calculated just from the transformations in (5.137) 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

$$U[\mathbf{v}, \mathbf{a}] = e^{-i\mathbf{a}\cdot\mathbf{P}}\, e^{i\mathbf{v}\cdot\mathbf{K}}\,, \tag{5.139}$$

then a straightforward calculation shows that

$$U[\mathbf{v}', \mathbf{a}']\, U[\mathbf{v}, \mathbf{a}] = e^{iM\,\mathbf{v}'\cdot\mathbf{a}}\, U[\mathbf{v}' + \mathbf{v}, \mathbf{a}' + \mathbf{a}]\,. \tag{5.140}$$

For comparison with the preceding general discussion we should take $T_a \to (\nabla_{\mathbf{v}}, \nabla_{\mathbf{a}})$ and $\hat{T}_a \to (-\mathbf{K}, \mathbf{P})$. From (5.140) then $\gamma_a \to M(\mathbf{0}, \mathbf{v})$ and from (5.127) $h_{ab} \to M\big(\begin{smallmatrix} 0 & I_3 \\ -I_3 & 0 \end{smallmatrix}\big)$.

For representations of the Galilean group in quantum mechanics the central extension plays an essential role. Using (5.138)

$$e^{-i\mathbf{v}\cdot\mathbf{K}} \mathbf{P} e^{i\mathbf{v}\cdot\mathbf{K}} = \mathbf{P} + M\mathbf{v}\,, \qquad e^{i\mathbf{v}\cdot\mathbf{K}} H e^{i\mathbf{v}\cdot\mathbf{K}} = H + \mathbf{P}\cdot\mathbf{v} + \tfrac{1}{2} M \mathbf{v}^2\,. \tag{5.141}$$

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

$$|\mathbf{p}\rangle = e^{i\mathbf{v}\cdot\mathbf{K}} |\mathbf{0}\rangle\,, \qquad \mathbf{p} = M\mathbf{v}\,, \tag{5.142}$$

so that as a consequence of (5.142)

$$\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle\,, \qquad H|\mathbf{p}\rangle = \big(E_0 + \tfrac{\mathbf{p}^2}{2M}\big)|\mathbf{p}\rangle\,. \tag{5.143}$$

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 (1.106) 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,

$$\int_0^{2\pi} \mathrm{d}\theta \, f(\theta) = \int_0^{2\pi} \mathrm{d}\theta \, f(\theta' + \theta) \, . \tag{5.144}$$

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

$$\int_G \mathrm{d}\rho(b) \, f(g(b)) = \int_G \mathrm{d}\rho(b) \, f(g(c)) \quad \text{for} \quad g(c) = g(a)g(b) \, , \tag{5.145}$$

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

$$\mathrm{d}^n c = |J| \, \mathrm{d}^n b \, , \qquad J = \det\left[\frac{\partial c^r}{\partial b^s}\right], \tag{5.146}$$

and then require, to satisfy (5.145),

$$\mathrm{d}\rho(b) = \mathrm{d}\rho(c) \quad \Rightarrow \quad \rho(b) = |J| \, \rho(c) \, . \tag{5.147}$$

For a Lie group the fundamental result (5.31), with (5.30), ensures that

$$J = \det\big[\lambda(b)\big] \, \det\big[\mu(c)\big] = \frac{\det\big[\mu(c)\big]}{\det\big[\mu(b)\big]} \, . \tag{5.148}$$

Comparing (5.146) and (5.148) with (5.147) show that the invariant integration measure over a general Lie group G is obtained by taking

$$\mathrm{d}\rho(b) = \frac{C}{\big|\det\big[\mu(b)\big]\big|} \, \mathrm{d}^n b \, . \tag{5.149}$$

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,

$$\int_G \mathrm{d}\rho(b) = |G| < \infty \, , \tag{5.150}$$

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 ${u_0}^2 + \mathbf{u}^2 = 1$,

$$\det[\mu_{ji}(\mathbf{u})] = u_0 \,. \tag{5.151}$$

Hence (5.149) requires

$$\mathrm{d}\rho(\mathbf{u}) = \frac{1}{|u_0|}\mathrm{d}^3 u\,, \qquad -1 \le u_0 \le 1\,, \quad |\mathbf{u}| \le 1\,. \tag{5.152}$$

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 (3.38)

$$u_0 = \cos\tfrac{1}{2}\theta\,, \quad \mathbf{u} = -\sin\tfrac{1}{2}\theta\,\mathbf{n}\,, \qquad \mathrm{d}^3 u = |\mathbf{u}|^2 \mathrm{d}|\mathbf{u}|\,\mathrm{d}\Omega_{\mathbf{n}}\,, \tag{5.153}$$

so that

$$\mathrm{d}\rho(\theta,\mathbf{n}) = \tfrac{1}{2}\sin^2\tfrac{1}{2}\theta\,\mathrm{d}\theta\,\mathrm{d}\Omega_{\mathbf{n}}\,, \qquad 0 \le \theta \le 2\pi\,. \tag{5.154}$$

Since $\int_{S^2}\mathrm{d}\Omega_{\mathbf{n}} = 4\pi$ the group volume is easily found

$$\int_{SU(2)}\mathrm{d}\rho(\theta,\mathbf{n}) = 2\pi^2\,. \tag{5.155}$$

These results verify the integration measure in (3.18) for $SO3)$, where the range of θ is halved.

For the parameterisation of $SU(2)$ in terms of Euler angles ϕ, θ, ψ as in (3.96) the

$$\begin{aligned} u_0 &= \cos\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}(\phi+\psi)\,, \qquad & u_3 &= -\cos\tfrac{1}{2}\theta\,\sin\tfrac{1}{2}(\phi+\psi)\,,\\ u_1 &= \sin\tfrac{1}{2}\theta\,\sin\tfrac{1}{2}(\phi-\psi)\,, \qquad & u_2 &= \sin\tfrac{1}{2}\theta\,\cos\tfrac{1}{2}(\phi-\psi)\,. \end{aligned} \tag{5.156}$$

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

$$\mathrm{d}\rho(\phi,\theta,\psi) = \tfrac{1}{8}\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\psi\,, \tag{5.157}$$

reproducing (3.123).

For $SO(3)$, since $SU(2)$ is a double cover, the group volume is halved. In terms of the parameterisation $(\theta,\mathbf{n})$ used in (5.154) we should take $0 \le \theta \le \pi$ or in terms of the Euler angles modify (3.96) to $0 \le \psi \le 2\pi$.

For compact Lie groups the orthogonality relations for representations (2.50) or characters (2.56) remain valid if the summation is replaced by invariant integration over the group and $|G|$ by the group volume as in (5.150). For $SU(2)$ this corresponds to the results give in (3.133) and (3.134).

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 (3.19) a general real 2×2 matrix may be expressed as in (3.40)

$$A = v_0 + v_1\sigma_1 + v_2\, i\sigma_2 + v_3\sigma_3\,, \tag{5.158}$$

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 (3.20),

$$\begin{pmatrix} \mathrm{d}v_1 & \mathrm{d}v_2 & \mathrm{d}v_3 \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 \end{pmatrix} \begin{pmatrix} v_0 & v_3 & v_2 \\ v_3 & v_0 & -v_1 \\ -v_2 & -v_1 & v_0 \end{pmatrix}. \tag{5.161}$$

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

$$\begin{gathered} v_0 = \cosh\alpha\, \cos\beta\,,\ v_2 = \cosh\alpha\, \sin\beta\,,\ v_1 = \sinh\alpha\, \cos\gamma\,,\ v_3 = \sinh\alpha\, \sin\gamma\,, \\ \alpha \geq 0\,, \qquad 0 \leq \beta, \gamma \leq 2\pi\,. \end{gathered} \tag{5.164}$$

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 (1.123).

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

$$Y \underset{X}{\to} [X, Y] = X^{\mathrm{ad}} Y \,, \tag{5.166}$$

defines the linear map $X^{\mathrm{ad}} : \mathfrak{g} \to \mathfrak{g}$. 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^{ad} is defined by

$$Y \xrightarrow[\exp(X)]{} D^{\mathrm{ad}}\big(\exp(sX)\big)Y = e^{sX^{\mathrm{ad}}} Y = \sum_{n=0}^{\infty} \frac{s^n}{n!}\, \underbrace{\big[X, \dots \big[X}_{n}, Y\big] \dots \big] \,, \tag{5.167}$$

with similar notation to (5.98). To verify that (5.166) provides a representation of the Lie algebra the Jacobi identity is essential since from

$$Z^{\mathrm{ad}} X^{\mathrm{ad}} Y = [Z, [X, Y]] \,, \tag{5.168}$$

we obtain for the adjoint commutator, using (5.17),

$$\big[Z^{\mathrm{ad}}, X^{\mathrm{ad}}\big] Y = [Z, [X, Y]] - [X, [Z, Y]] = [[Z, X], Y] = [Z, Y]^{\mathrm{ad}} Y \,, \tag{5.169}$$

and hence in general

$$\big[Z^{\mathrm{ad}}, X^{\mathrm{ad}}\big] = [Z, Y]^{\mathrm{ad}} \,. \tag{5.170}$$

Explicit adjoint representation matrices are obtained by choosing a basis for $\mathfrak{g}$, $\{T_a\}$ so that for any $X \in \mathfrak{g}$ (5.166) becomes

$$[X, T_a] = T_b\, (X^{\mathrm{ad}})^b{}_a \,. \tag{5.171}$$

For $X \to T_a$ the corresponding adjoint representation matrices are then given by

$$[T_a, T_b] = T_c\, (T_a^{\mathrm{ad}})^c{}_b \qquad \Rightarrow \qquad (T_a^{\mathrm{ad}})^c{}_b = f^c{}_{ab} \,, \tag{5.172}$$

using (5.41). The commutator

$$\big[T_a^{\mathrm{ad}}, T_b^{\mathrm{ad}}\big] = f^c{}_{ab}\, T_c^{\mathrm{ad}} \,, \tag{5.173}$$

is directly equivalent to the Jacobi identity (5.42). The group representation matrices $D^{\mathrm{ad}}(\exp X) = e^{X^{\mathrm{ad}}}$, with $X^{\mathrm{ad}} = T_a^{\mathrm{ad}} X^a$, are then obtained using the matrix exponential. Close to the identity, in accord with (5.69),

$$D^{\mathrm{ad}}(\exp X) = I + X^{\mathrm{ad}} + \mathrm{O}(X^2) \,. \tag{5.174}$$

If the Lie algebra is abelian then clearly $X^{\mathrm{ad}} = 0$ for all X so the adjoint representation is trivial.

For $\mathfrak{su}(2)$ with the standard hermitian generators

$$[J_i, J_j] = i\varepsilon_{ijk}\, J_k \qquad \Rightarrow \qquad (J_i^{\rm ad})_{jk} = -i\varepsilon_{ijk}\,, \tag{5.175}$$

where $\mathbf{J}^{\rm ad}$ are three 3×3 hermitian matrices. If $\mathbf{n}$ is a unit vector $(\mathbf{n}\cdot\mathbf{J}^{\rm ad})^2 = I - n\,n^T$ from which we may deduce that $\mathbf{n}\cdot\mathbf{J}^{\rm ad}$ has eigenvalues $\pm 1, 0$ so that this is the spin 1 representation. For the the Lie algebra $\mathfrak{iso}(2)$, as given in (4.150), we have

$$E_1^{\rm ad} = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_2^{\rm ad} = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad J_3^{\rm ad} = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.176}$$

5.8.1 Killing Form

The *Killing*[45] *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

$$\kappa(X,Y) = \operatorname{tr}\big(X^{\rm ad}Y^{\rm ad}\big) \qquad \text{for all} \quad X, Y \in \mathfrak{g}\,, \tag{5.177}$$

or in terms of a basis as in (5.172)

$$\kappa_{ab} = \kappa(T_a, T_b) = f^c{}_{ad}\, f^d{}_{bc}\,, \tag{5.178}$$

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

$$\kappa([Z,X],Y) + \kappa(X,[Z,Y]) = 0\,. \tag{5.179}$$

The verification of this is simple since, from (5.170),

$$\kappa([Z,X],Y) = \operatorname{tr}\big([Z,X]^{\rm ad}Y^{\rm ad}\big) = \operatorname{tr}\big([Z^{\rm ad}, X^{\rm ad}]\,Y^{\rm ad}\big)\,, \tag{5.180}$$

and then (5.179) follows from $\operatorname{tr}\big([Z^{\rm ad}, X^{\rm ad}]\,Y^{\rm ad}\big) + \operatorname{tr}\big(X^{\rm ad}\,[Z^{\rm ad}, Y^{\rm ad}]\big) = 0$, using cyclic symmetry of the matrix trace. The result (5.179) also shows that the Killing form is invariant under the action of the corresponding Lie group G since

$$\kappa\big(e^{sZ^{\rm ad}}X,\, e^{sZ^{\rm ad}}Y\big) = \kappa\big(X,Y\big)\,, \tag{5.181}$$

which follows from (5.167) and differentiating with respect to s and then using (5.179).

Alternatively (5.179) may be expressed in terms of components using

$$\kappa\big([T_c, T_a], T_b\big) = f^d{}_{ca}\kappa(T_d, T_b) = f^d{}_{ca}\,\kappa_{db} \equiv f_{cab}\,, \tag{5.182}$$

in a form expressing κ_{ab} as an invariant tensor for the adjoint representation

$$\kappa_{db}\, f^d{}_{ca} + \kappa_{ad}\, f^d{}_{cb} = 0 \qquad \Leftrightarrow \qquad f_{cab} + f_{cba} = 0\,. \tag{5.183}$$

[45] Wilhelm Karl Joseph Killing, 1847-1923, German.

Since, from (5.39), $f_{cab} + f_{cba} = 0$ this implies

$$f_{abc} = f_{[abc]} \,. \tag{5.184}$$

If the Lie algebra $\mathfrak{g}$ contains an invariant subalgebra $\mathfrak{h}$ then in an appropriate basis we may write

$$T_a = (T_i, T_r)\,, \quad T_i \in \mathfrak{h} \qquad [T_i, T_j] = f^k{}_{ij} T_k\,, \quad [T_r, T_i] = f^j{}_{ri} T_j\,, \tag{5.185}$$

so that the Killing form restricted to $\mathfrak{h}$ is just

$$\kappa_{ij} = f^k{}_{il}\, f^l{}_{jk} = \mathrm{tr}_{\mathfrak{h}}(T_i^{\mathrm{ad}} T_j^{\mathrm{ad}})\,. \tag{5.186}$$

The crucial property of the Killing form is the invariance condition (5.179). If g_{ab} also defines an invariant bilinear form on the Lie algebra, as in (5.179), so that

$$g_{ab}\big([Z,X]^a Y^b + X^a [Z,Y]^b\big) = 0\,, \tag{5.187}$$

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 (5.187), 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^{ad} may be chosen to be unitary so that in (5.174) the adjoint Lie algebra generators are anti-hermitian, as in (5.70),

$$X^{\mathrm{ad}\dagger} = -X^{\mathrm{ad}}\,. \tag{5.188}$$

In this case

$$\kappa(X,X) \le 0\,, \qquad \kappa(X,X) = 0 \quad \Leftrightarrow \quad X^{\mathrm{ad}} = 0\,. \tag{5.189}$$

For $\mathfrak{su}(2)$ using the hermitian adjoint generators in (5.175) the Killing form is positive

$$\kappa_{ij} = \mathrm{tr}(J_i^{\mathrm{ad}}\, J_j^{\mathrm{ad}}) = i^2 \varepsilon_{ikl}\, \varepsilon_{jlk} = 2\,\delta_{ij}\,. \tag{5.190}$$

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 (5.176)

$$[\kappa_{ab}] = 2\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.191}$$

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

$$T_a = (T_i, T_r) \qquad \Rightarrow \qquad [T_i, T_j] = 0\,, \quad [T_r, T_i] = f^j{}_{ri} T_j\,. \tag{5.192}$$

Then from (5.178)

$$\kappa_{ai} = f^c{}_{ad}\, f^d{}_{ic} = f^r{}_{aj}\, f^j{}_{ir} = 0\,, \qquad \text{since} \quad f^r{}_{sj} = f^r{}_{kj} = 0\,, \tag{5.193}$$

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 (5.190) and (5.191) 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 (5.189) since if $X^{\text{ad}} = 0$, 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

$$G \simeq G_{\text{semi-simple}} \otimes U(1) \otimes \cdots \otimes U(1)/F\,, \tag{5.194}$$

with a $U(1)$ factor for each independent Lie algebra element with $X^{\text{ad}} = 0$ 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

$$\mathfrak{h}_\perp = \{X : \kappa(X,Y) = 0,\, Y \in \mathfrak{h}\}\,. \tag{5.195}$$

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 (5.195), 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

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}_\perp\,. \tag{5.196}$$

This decomposition may be continued to give until there are no remaining invariant spaces

$$\mathfrak{g} = \bigoplus_i \mathfrak{g}_i\,, \qquad \mathfrak{g}_i \;\; \text{simple}\,. \tag{5.197}$$

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

$$X = \sum_i X_i \,, \quad Y = \sum_i Y_i \,, \qquad \kappa(X,Y) = \sum_i \mathrm{tr}_{\mathfrak{g}_i}\big(X_i^{\mathrm{ad}} Y_i^{\mathrm{ad}}\big) \,, \tag{5.198}$$

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 (5.183) becomes $\kappa\, T_a^{\mathrm{ad}} + T_a^{\mathrm{ad}\,T} \kappa = 0$ so that from $[T_a^{\mathrm{ad}}, \kappa^{-1}\kappa] = 0$ we obtain $T_a^{\mathrm{ad}} \kappa^{-1} + \kappa^{-1} T_a^{\mathrm{ad}\,T} = 0$ or

$$f^b{}_{ad}\, \kappa^{dc} + f^c{}_{ad}\, \kappa^{bd} = 0 \,, \tag{5.199}$$

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

$$[t_a, \kappa^{bc}\, t_b t_c] = \kappa^{bc}\big(f^d{}_{ab}\, t_d t_c + f^d{}_{ac}\, t_b t_d\big) = \big(\kappa^{be} f^d{}_{ab} + \kappa^{dc} f^e{}_{ac}\big)\, t_d t_e = 0 \,. \tag{5.200}$$

In consequence $\kappa^{ab}\, t_a t_b$ is a *quadratic Casimir operator.*

To discuss central extensions we rewrite the fundamental consistency condition (5.129) in the form

$$h_{ae}\, f^e{}_{cd} = -h_{de}\, f^e{}_{ac} - h_{ce}\, f^e{}_{da} \,. \tag{5.201}$$

Then using (5.199)

$$h_{ae}\, f^e{}_{cd}\, f^c{}_{bg}\, \kappa^{gd} = -h_{ae}\, f^g{}_{cd}\, f^c{}_{bg}\, \kappa^{de} = h_{ae}\, \kappa_{db}\, \kappa^{ed} = h_{ab} \,, \tag{5.202}$$

and also, with (5.199) again,

$$\begin{aligned} \big(h_{de}\, f^e{}_{ac} + h_{ce}\, f^e{}_{da}\big) f^c{}_{bg}\, \kappa^{gd} &= \big(h_{de}\, f^e{}_{ac}\, f^c{}_{bg} + h_{ec}\, f^c{}_{bg}\, f^e{}_{ad}\big)\kappa^{gd} \\ &= h_{de}\, f^e{}_{ac}\, f^c{}_{bg}\, \kappa^{gd} - h_{ec}\, f^c{}_{bg}\, f^g{}_{ad}\, \kappa^{de} \end{aligned} \tag{5.203}$$

we may obtain from (5.201), re-expressing (5.203) as a matrix trace,

$$h_{ab} = -\mathrm{tr}\big(h\, [T_a^{\mathrm{ad}}, T_b^{\mathrm{ad}}]\kappa^{-1}\big) = -\mathrm{tr}\big(h\, T_c^{\mathrm{ad}} \kappa^{-1}\big) f^c{}_{ab} \,. \tag{5.204}$$

Hence h_{ab} is of the form given in (5.131) 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.

6 Lie Algebras for Matrix Groups

Here we obtain the Lie algebras $\mathfrak{g}$ corresponding to the various continuous matrix groups G described in section 1.6 by considering matrices close to the identity

$$M = \mathbb{1} + X + \mathrm{O}(X^2)\,, \tag{6.1}$$

with suitable conditions on X depending on the particular group.

6.1 Unitary Groups

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,

$$R^i{}_j = \begin{array}{c c} & \begin{array}{cccccc} & & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & & & 1 & & 0 \\ \vdots & & & & & \vdots \\ 0 & & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array}\,, \qquad i,j = 1,\dots,n\,, \qquad (R^i{}_j)^\dagger = R^j{}_i\,. \tag{6.2}$$

These matrices satisfy

$$\big[R^i{}_j, R^k{}_l\big] = \delta^k{}_j\, R^i{}_l - \delta^i{}_l\, R^k{}_j\,, \tag{6.3}$$

and

$$\mathrm{tr}\big(R^i{}_j\, R^k{}_l\big) = \delta^k{}_j\, \delta^i{}_l\,. \tag{6.4}$$

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

$$[X, R^i{}_j] = X^i{}_k\, R^k{}_j - R^i{}_k\, X^k{}_j \quad \Rightarrow \quad (X^{\mathrm{ad}})^l{}_{k,}{}^i{}_j = X^i{}_k\, \delta^l{}_j - X^l{}_j\, \delta^i{}_k\,. \tag{6.5}$$

Hence, for $X = R^i{}_j X^j{}_i, Y = R^i{}_j Y^j{}_i$,

$$\kappa(X,Y) = \mathrm{tr}\big(X^{\mathrm{ad}} Y^{\mathrm{ad}}\big) = 2\big(n \textstyle\sum_{i,j} X^j{}_i\, Y^i{}_j - \sum_i X^i{}_i \sum_j Y^j{}_j\big)\,. \tag{6.6}$$

Restricting to $\mathfrak{u}(n)$

$$\kappa(X,X) = -2n \textstyle\sum_{i,j} |\hat{X}^j{}_i|^2\,, \qquad \hat{X}^j{}_i = X^j{}_i - \frac{1}{n}\, \delta^j{}_i \sum_k X^k{}_k\,. \tag{6.7}$$

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 $\kappa(X,X) = 2n\,\mathrm{tr}(X^2) < 0$.

A basis of $n^2 - 1$ anti-hermitian generators for $\mathfrak{su}(n)$, forming the fundamental representation of this Lie algebra, is provided by taking

$$\big\{t_a\big\} = \big\{i(R^i{}_j + R^j{}_i), (R^i{}_j - R^j{}_i) : 1 \le i < j \le n\big\} \cup \big\{i(R^i{}_i - R^{i+1}{}_{i+1}); 1 \le i \le n-1\big\}\,. \tag{6.8}$$

These satisfy

$$\mathrm{tr}(t_a) = 0\,, \quad \mathrm{tr}(t_a t_b) = -2\,\delta_{ab}\,, \quad t_a t_a = -\tfrac{n^2-1}{n}\,\mathbb{1}\,, \qquad [t_a,\, t_b] = f^c{}_{ab}\, t_c\,, \tag{6.9}$$

and the completeness condition

$$X = t_a X_a \qquad X_a = \tfrac{1}{2}\,\mathrm{tr}(t_a X) \ \text{ for any } \ X \in \mathfrak{su}(n)\,. \tag{6.10}$$

Since

$$\sum_{i=1}^{r} \mathrm{tr}\big(t_{a_1}\dots[t_c, t_{a_i}]\dots t_{a_r}\big) = 0\,, \tag{6.11}$$

then $\mathrm{tr}\big(t_{a_1} t_{a_2}\dots t_{a_r}\big)$ is an invariant tensor for $\mathfrak{su}(n)$ in the basis provided by $\{t_a\}$. Invariant tensors can be chosen to be totally symmetric since otherwise the commutation relations (6.9) allow the trace over r generators to be reduced to a sum over traces with $r' < r$. A potential basis for invariant tensors is then provided by

$$\mathrm{tr}\big(t_{(a_1} t_{a_2}\dots t_{a_r)}\big)\,, \quad r = 2, \dots n\,, \tag{6.12}$$

since for an $n \times n$ matrix A, $\mathrm{tr}(A^r)$ for $r > n$ is reducible to products of traces $\mathrm{tr}(A^m)$ with $m \le n$. The invariant tensors are not unique since any symmetric rank r invariant tensor $d_{r,a_1\dots a_r}$ is invariant up symmetrised products of lower rank r_i tensors with $\sum r_i = r$.

Any representation of the $\mathfrak{su}(n)$ Lie algebra has a basis $\{T_a\}$ where $[T_a,\, T_b] = f^c{}_{ab}\,T_c$ and then for invariant tensors $d_{r,a_1\dots a_r}$

$$C_r = d_{r,a_1\dots a_r} T_{a_1}\dots T_{a_r} \quad \text{satisfies} \quad [T_a, C_r] = 0\,, \tag{6.13}$$

and so C_r is a Casimir operator. For $\mathfrak{su}(n)$ there are thus $n-1$ Casimirs. For an irreducible representation $C_r = c_r\,\mathbb{1}$.

6.2 Symplectic Groups

For $\mathfrak{sp}(2n,\mathbb{R})$ or $\mathfrak{sp}(2n,\mathbb{C})$ the condition (1.107) translates into

$$JX = -X^T J = (JX)^T\,, \tag{6.14}$$

where J is the standard antisymmetric matrix given in (1.108). It can be represented by

$$J_{ij} = -J_{ji} = -(-1)^i\,\delta_{ij'}\,, \qquad j' = j - (-1)^j\,. \tag{6.15}$$

It is convenient to define also

$$J^{ij} = -J^{ji}\,, \qquad J^{ik} J_{kj} = \delta^i{}_j\,, \tag{6.16}$$

where J_{ij}, J_{ij} can be used to raise and lower indices. A basis for $\mathfrak{sp}(2n,\mathbb{R})$, or $\mathfrak{sp}(2n,\mathbb{C})$, satisfying (6.14) is provided in terms of the $2n \times 2n$ matrices $\{T^i{}_j\}$

$$(T^i{}_j)^k{}_l = \delta^i{}_l\,\delta^k{}_j + J^{ik}\,J_{jl}\,, \qquad JT^i{}_j = (JT^i{}_j)^T\,. \tag{6.17}$$

With the choice (6.15) these have the explicit form

$$T^i{}_j = \begin{array}{cc} & \begin{array}{ccccccc} & & i' & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j' \\ \\ \\ \end{array} & \begin{pmatrix} 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & & 0 & & 1 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & & -(-1)^{i+j} & & 0 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array}, \qquad 1 \le i < j \le 2n\,. \tag{6.18}$$

A linearly independent basis is given by $\{T^i{}_j,\ 1 \le i < j \le 2n; T^{2i-1}{}_{2i-1},\ T^{2i}{}_{2i-1},\ 1 \le i \le n\}$. For $n = 1$ this prescription gives

$$T^1{}_1 = -T^2{}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad T^1{}_2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \qquad T^2{}_1 = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}. \tag{6.19}$$

From (6.17)

$$T^i{}_j = J^{il}\, J_{jk}\, T^k{}_l\,, \tag{6.20}$$

and the matrices $\{T^i{}_j\}$ satisfy the Lie algebra

$$\big[T^i{}_j, T^k{}_l\big] = \delta^i{}_l\, T^k{}_j - \delta^k{}_j\, T^i{}_l - J^{ik}\, J_{jm} T^m{}_l + J_{jl}\, J^{im} T^k{}_m\,, \tag{6.21}$$

and also

$$\operatorname{tr}\big(T^i{}_j\, T^k{}_l\big) = 2\big(\delta^i{}_l\, \delta^k{}_j + J^{ik}\, J_{jl}\big)\,. \tag{6.22}$$

For any $X \in \mathfrak{sp}(2n,\mathbb{R})$, or $\mathfrak{sp}(2n,\mathbb{C})$ then $X = \frac{1}{2} X^j{}_i T^i{}_j$. As a consequence of the symmetry of $T^i{}_j$ $X^j{}_i$ should be restricted so that $X^j{}_i = J^{jl} J_{ik} X^k{}_l$. Using (6.21)

$$[X, T^i{}_j] = X^k{}_j\, T^i{}_k - X^i{}_k\, T^k{}_j\,, \tag{6.23}$$

so that from (5.171) $[X, T^i{}_j] = \frac{1}{2}\, T^l{}_k (X^{\mathrm{ad}})^k{}_l,{}^i{}_j$ giving

$$(X^{\mathrm{ad}})^k{}_l,{}^i{}_j = \delta^i{}_l\, X^k{}_j - \delta^k{}_j\, X^i{}_l + J^{ki}\, J_{lm} X^m{}_j - J_{lj}\, J^{km} X^i{}_m\,, \tag{6.24}$$

and hence

$$\kappa(X,Y) = \tfrac{1}{4} (X^{\mathrm{ad}})^k{}_l,{}^i{}_j\, (Y^{\mathrm{ad}})^j{}_i,{}^l{}_k = (n+2)\, X^i{}_j Y^j{}_i\,. \tag{6.25}$$

An alternative basis for the Lie algebra is obtained by taking $T_{ij} = J_{ik}\, T^k{}_j = T_{ji}$ with $\operatorname{tr}(T_{ij}\, T_{kl}) = -2(J_{il} J_{jk} + J_{ik} J_{jl})$ and the Lie algebra becoming

$$\big[T_{ij}, T_{kl}\big] = J_{ik}\, T_{jl} + J_{il}\, T_{jk} + J_{jk}\, T_{il} + J_{jl}\, T_{ik}\,, \tag{6.26}$$

The generators here correspond to symmetric $2n$ dimensional matrices so the dimension is clearly $\frac{1}{2}\, 2n(2n+1)$.

For the corresponding compact group $Sp(n) = Sp(2n,\mathbb{C}) \cap SU(2n)$ we impose, as well as (6.14),

$$X^\dagger = -X \qquad \text{or} \qquad X^j{}_i{}^* = -X^i{}_j\,. \tag{6.27}$$

Then (6.25) gives

$$\kappa(X,X) = -(n+2)\,\textstyle\sum_{i,j}\big|X^j{}_i\big|^2\,. \tag{6.28}$$

From (6.14)

$$JXJ^{-1} = -X^T\,, \tag{6.29}$$

so that, following the discussion in section 5.3.3, the fundamental representation of compact $Sp(n)$ is pseudo-real.

In the alternative basis in (6.26) a general $X \in \mathfrak{sp}(2n,\mathbb{C})$ can be expressed as $X = \frac{1}{2}\,X^{ij}T_{ij}$ with X^{ij} symmetric and (6.27) becomes

$$X_{ij}{}^* = J^{ik}\,J^{jl}\,X_{kl}\,. \tag{6.30}$$

6.3 Orthogonal and Spin Groups

For $\mathfrak{o}(n)$ or $\mathfrak{so}(n)$ then in (6.1) 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

$$S_{ij} = -S_{ji} = \begin{array}{c} \\ \\ \\ i \\ \\ j \\ \\ \\ \end{array}\begin{array}{c}\begin{array}{ccccccc} & & i & & j & & \end{array}\\ \begin{pmatrix} 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & & 0 & & 1 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & & -1 & & 0 & & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}\end{array}, \qquad i\neq j = 1,\dots,n\,. \tag{6.31}$$

These satisfy

$$\big[S_{ij}, S_{kl}\big] = \delta_{jk}\,S_{il} - \delta_{ik}\,S_{jl} - \delta_{jl}\,S_{ik} + \delta_{il}\,S_{jk}\,, \tag{6.32}$$

and

$$\mathrm{tr}\big(S_{ij}\,S_{kl}\big) = 2\big(\delta_{il}\,\delta_{jk} - \delta_{ik}\,\delta_{jl}\big)\,. \tag{6.33}$$

For arbitrary $X \in \mathfrak{so}(2n)$ then $X = \frac{1}{2}X_{ij}S_{ij}$, where $X_{ij} = -X_{ji}$ is real. From (6.32)

$$[X, S_{ij}] = X_{ki}\,S_{kj} - X_{kj}\,S_{ki} \quad \Rightarrow \quad X^{\mathrm{ad}}_{kl,ij} = X_{ki}\,\delta_{lj} - X_{kj}\,\delta_{li} - X_{li}\,\delta_{kj} + X_{lj}\,\delta_{ki}\,, \tag{6.34}$$

and hence

$$\kappa(X,Y) = \tfrac{1}{4}X^{\mathrm{ad}}_{kl,ij}\,Y^{\mathrm{ad}}_{ij,kl} = (n-2)\,X_{ij}Y_{ji}\,. \tag{6.35}$$

The matrices (6.31) 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.

6.3.1 Spin Groups and Gamma Matrices

The relation $SO(3)$ and $SU(2)$, which is described in section 3.2, and also the introduction of spinorial representations, described in section 3.14, 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, $i = 1,\dots,n$ which form the basis for a Clifford algebra the *Clifford*[46] *algebra*,

$$\gamma_i\gamma_j + \gamma_j\gamma_i = 2\delta_{ij}\,\mathbb{1}\,, \qquad \gamma_i{}^\dagger = \gamma_i\,. \tag{6.36}$$

The algebra may be extended to pseudo-orthogonal groups such as the Lorentz group, which involve a metric g_{ij} as in (1.124), by taking $\delta_{ij} \to g_{ij}$ on the right hand side of (6.36). 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 (1.124) 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 (6.36), 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 (6.36) 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 (6.36)

$$(\gamma\cdot x)^2 = x^2\,\mathbb{1}, \qquad x \in \mathbb{R}^n\,. \tag{6.37}$$

This the primary definition of a Clifford algebra where there a product for two vectors belonging to a vector space V which is proportional to the unit operator on V.

To show the connection with $SO(n)$ we first define

$$s_{ij} = \tfrac{1}{2}\,\gamma_{[i}\,\gamma_{j]} = -s_{ij}{}^\dagger\,. \tag{6.38}$$

Using just (6.36) it is easy to obtain

$$[s_{ij},\gamma_k] = \delta_{jk}\,\gamma_i - \delta_{ik}\,\gamma_j\,, \tag{6.39}$$

and hence

$$\big[s_{ij},s_{kl}\big] = \delta_{jk}\,s_{il} - \delta_{ik}\,s_{jl} - \delta_{jl}\,s_{ik} + \delta_{il}\,s_{jk}\,. \tag{6.40}$$

This is identical with (6.32), the Lie algebra $\mathfrak{so}(n)$. Moreover for finite transformations, which involve the matrix exponential of $\frac{1}{2}\,\omega_{ij}s_{ij}$, $\omega_{ij} = -\omega_{ji}$,

$$e^{-\frac{1}{2}\,\omega_{ij}s_{ij}}\,\gamma\cdot x\;e^{\frac{1}{2}\,\omega_{ij}s_{ij}} = \gamma\cdot x'\,, \qquad x' = Rx\,, \quad R = e^{-\frac{1}{2}\,\omega_{ij}S_{ij}} \in SO(n)\,, \tag{6.41}$$

with $S_{ij} \in \mathfrak{so}(n)$ as in (6.31). It is easy to see that $x'^2 = x^2$, as required for rotations, as a consequence of (6.37). To show the converse we note that $\gamma'_i = \gamma_j R_{ji}$ also satisfies (6.36) 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

$$\mathrm{Spin}(n) = \big\{ e^{-\frac{1}{2}\,\omega_{ij}s_{ij}} : \omega_{ij} = -\omega_{ji} \in \mathbb{R}\big\}\,. \tag{6.42}$$

[46] William Kingdon Clifford, 1845-1879, English, second wrangler 1867.

Clearly 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 Spin$(3) \simeq SU(2)$. In general, since $\pm\mathbb{1} \in$ Spin(n) are mapped to $\mathbb{1} \in SO(n)$, we have $SO(n) \simeq \mathrm{Spin}(n)/\mathbb{Z}_2$.

Unlike $SO(n)$, Spin(n) is simply connected and is the covering group for $SO(n)$. For further analysis we define

$$\Gamma = \gamma_1\gamma_2\dots\gamma_n = (-1)^{\frac{1}{2}n(n-1)}\,\Gamma^\dagger\,, \qquad \Gamma^\dagger = \gamma_n\gamma_{n-1}\dots\gamma_1\,, \tag{6.43}$$

so that

$$\Gamma^2 = (-1)^{\frac{1}{2}n(n-1)}\,\mathbb{1}\,. \tag{6.44}$$

Directly from (6.36)

$$[\Gamma,\gamma_i] = 0\,, \quad n \text{ odd}\,, \quad \Gamma\,\gamma_i + \gamma_i\,\Gamma = 0\,, \quad n \text{ even}\,, \qquad i = 1,\dots n\,. \tag{6.45}$$

Using, similarly to (3.38),

$$e^{\alpha\, s_{ij}} = \cos\tfrac{1}{2}\alpha\,\mathbb{1} + \sin\tfrac{1}{2}\alpha\, 2s_{ij}\,, \tag{6.46}$$

then

$$e^{\pi\sum_{i=1}^m s_{2i-1\,2i}} = \Gamma\,, \qquad e^{-\pi\sum_{i=1}^m s_{2i-1\,2i}} = (-1)^m\Gamma\,, \qquad \text{for} \;\; n = 2m \;\; \text{even}\,. \tag{6.47}$$

This allows the identification of the centres of the spin groups

$$\begin{aligned} \mathcal{Z}(\mathrm{Spin}(n)) &= \{\mathbb{1},-\mathbb{1},\Gamma,-\Gamma\} \simeq \begin{cases} \mathbb{Z}_2\times\mathbb{Z}_2\,, & n = 4m\,, \\ \mathbb{Z}_4\,, & n = 4m+2\,, \end{cases} \\ \mathcal{Z}(\mathrm{Spin}(n)) &= \{\mathbb{1},-\mathbb{1}\} \simeq \mathbb{Z}_2\,, \qquad\qquad\qquad n = 2m+1\,. \end{aligned} \tag{6.48}$$

Spinors for general rotational groups are defined as belonging to the fundamental representation space $\mathcal{S}$ for Spin(n), so they form projective representations, up to a sign, of $SO(n)$.

6.3.2 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 (6.36), and also $\gamma_i^2 = \mathbb{1}$, 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 $i_1,\dots i_r$ all different indices,

$$\Gamma_{i_1\dots i_r} = \gamma_{[i_1}\dots\gamma_{i_r]} = (-1)^{\frac{1}{2}r(r-1)}\,\Gamma_{i_1\dots i_r}{}^\dagger\,, \quad \Gamma_{i_1\dots i_r}{}^\dagger = \Gamma_{i_r\dots i_1}\,, \qquad r = 1,\dots n\,, \tag{6.49}$$

where $\Gamma_{i_1\dots i_r}{}^2 = (-1)^{\frac{1}{2}r(r-1)}\,\mathbb{1}$ and with the usual summation convention

$$\tfrac{1}{r!}\,\Gamma_{i_1\dots i_r}\,\Gamma_{i_r\dots i_1} = \tbinom{n}{r}\,\mathbb{1}\,. \tag{6.50}$$

From the definition (6.43)

$$\Gamma_{i_1\dots i_n} = \varepsilon_{i_1\dots i_n}\,\Gamma\,, \qquad n = 1,2,\dots\,. \tag{6.51}$$

We also have the relations

$$\Gamma_{i_1\dots i_r} = (-1)^{\frac{1}{2}n(n-1)+\frac{1}{2}r(r-1)}\,\tfrac{1}{s!}\,\varepsilon_{i_1\dots i_r\,j_1\dots j_s}\Gamma_{j_1\dots j_s}\Gamma\,, \qquad r+s=n\,. \tag{6.52}$$

A basis for these products of γ-matrices is given by $\mathcal{C}_r = \{\Gamma_{i_1\dots i_r} : 1 \le i_1 < i_2 < \dots < i_r \le n\}$, with $\dim\mathcal{C}_r = \binom{n}{r}$, $\mathcal{C}_n = \{\Gamma\}$. It is easy to see that $\mathcal{C}^{(n)} = \{\pm\mathbb{1}, \pm\mathcal{C}_1, \dots, \pm\mathcal{C}_{n-1}, \pm\Gamma\}$ 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 $\{\mathbb{1}, \mathcal{C}_1, \dots, \mathcal{C}_n\}$ 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 (6.45) Γ commutes with all elements in $\mathcal{C}^{(n)}$ and so for an irreducible representation we must have $\Gamma \propto \mathbb{1}$. Taking into account (6.44)

$$\Gamma = \begin{cases} \pm\mathbb{1}\,, & n = 4m+1 \\ \pm i\,\mathbb{1}\,, & n = 4m+3 \end{cases}. \tag{6.53}$$

The $\pm$ signs correspond to inequivalent representations, linked by taking $\gamma_i \to -\gamma_i$. For a linearly independent basis then, as a consequence of (6.52), it is necessary to recognise that the products of gamma matrices are no longer independent if $r > \frac{1}{2}n$. The matrix groups formed from the irreducible representations for n odd are then, for the two cases in (6.53), $\mathcal{C}^{(4m+1)} = \{\pm\mathbb{1}, \pm\mathcal{C}_1, \dots, \pm\mathcal{C}_{2m}\}$, $\dim\mathcal{C}^{(4m+1)} = 2^{4m+1}$, and $\mathcal{C}^{(4m+3)} = \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm\mathcal{C}_1, \pm i\,\mathcal{C}_1, \dots, \pm i\,\mathcal{C}_{2m+1}\}$, $\dim\mathcal{C}^{(4m+3)} = 2^{4m+4}$. Thus for $n = 3$ there is the Pauli group of order 16 formed from the 2×2 matrices $\{\pm\mathbb{1}_2, \pm i\mathbb{1}_2, \pm\sigma_i, \pm i\sigma_i\}$, with σ_i the usual Pauli matrices or $\mathrm{Q}_8 \cup i\,\mathrm{Q}_8$. This group corresponds to $\mathrm{Q}_8 \times \mathbb{Z}_4/\mathbb{Z}_2$ with $\mathbb{Z}_4 = \{\pm 1, \pm i\}$ and $\mathbb{Z}_2 = \{\pm 1\}$. The Pauli group has a D_4 subgroup formed by $\{\pm\mathbb{1}_2, \pm\sigma_1, \pm\sigma_2, \pm i\sigma_3\}$.

For n even $\mathcal{C}^{(n)}$ does not contain any elements commuting with all γ_i but

$$[\Gamma, s_{ij}] = 0\,. \tag{6.54}$$

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 (6.44),

$$\Gamma = \begin{cases} \left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{smallmatrix}\right), & n = 4m\,, \\ i\left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right), & n = 4m+2\,, \end{cases} \qquad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ \bar\sigma_i & 0 \end{pmatrix}, \qquad s_{ij} = \begin{pmatrix} s_{+ij} & 0 \\ 0 & s_{-ij} \end{pmatrix}. \tag{6.55}$$

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 (6.38) we have

$$s_{\pm ij}{}^\dagger = -s_{\pm ij}\,. \tag{6.56}$$

With the decomposition in (6.55) the Clifford algebra (6.36) is equivalent to

$$\sigma_i\,\bar\sigma_j + \sigma_j\,\bar\sigma_i = 2\delta_{ij}\,\mathbb{1}\,, \qquad \bar\sigma_i\,\sigma_j + \bar\sigma_j\,\sigma_i = 2\delta_{ij}\,\mathbb{1}\,. \tag{6.57}$$

For traces of gamma matrices and their products we first note that from (6.36)

$$\operatorname{tr}\big(\gamma_j(\gamma_i\gamma_j+\gamma_j\gamma_i)\big) = 2\operatorname{tr}(\gamma_j\gamma_j\gamma_i) = 2\operatorname{tr}(\gamma_i) = 0\,, \quad j\neq i\,, \quad \text{no sum on } j\,. \tag{6.58}$$

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

$$\operatorname{tr}(\Gamma_{i_1\dots i_r}) = 0\,, \qquad \text{except when } r=n,\ n \text{ odd}\,. \tag{6.59}$$

Hence in general, for $r,s=0,\dots,n$ for n even, or with $r,s<\frac{1}{2}n$ for n odd,

$$\operatorname{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{j_s\dots j_1}\big) = \delta_{rs}\,d_n\,(A_r)_{i_1\dots i_r,j_1\dots j_r}\,, \tag{6.60}$$

where

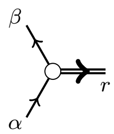
$$\begin{aligned} &d_n = \operatorname{tr}(\mathbb{1})\,,\\ &(A_r)_{i_1\dots i_r,j_1\dots j_r} = r!\,\delta_{[i_1|j_1|}\delta_{i_2|j_2|}\dots\delta_{i_r]j_r} = \begin{cases} \pm1\,, & (j_1,\dots,j_r) \text{ an even/odd} \\ & \ \text{permutation of } (i_1,\dots,i_r) \\ 0, & \text{otherwise}\end{cases}\,, \end{aligned} \tag{6.61}$$

so that $A_r/r!$ is a projection operator for antisymmetric rank r tensors.

With d_n-dimensional spinorial indices $\alpha,\beta,\dots$ the Γ-matrices can be represented diagrammatically

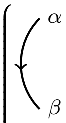
$(\Gamma_{i_1\dots i_r})_\alpha{}^\beta$ =

$r=0$

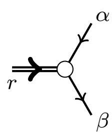


$r\geq1$,

$(\Gamma_{i_r\dots i_1})_\alpha{}^\beta$ =

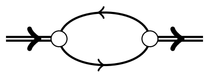


$r=0$

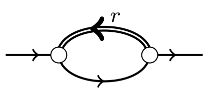


$r\geq1$, (6.62)

with  $= d_n$,  $=0$. (6.60) is then equivalent to

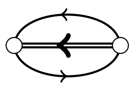
r  s $= d_n\,\delta_{rs}$  , (6.63)

The identity (6.50) corresponds to



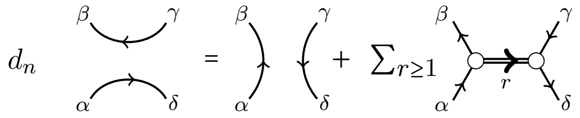
$= \binom{n}{r}$. (6.64)

These matrices then have a norm

$\frac{1}{r!}\operatorname{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{i_r\dots i_1}\big) =$  $= d_n\binom{n}{r}$. (6.65)

In general these products of gamma matrices form a complete set so that

$$d_n\, \delta_\alpha{}^\delta \delta_\gamma{}^\beta = \delta_\alpha{}^\beta \delta_\gamma{}^\delta + \sum_{r\geq 1} \tfrac{1}{r!}\, (\Gamma_{i_1\dots i_r})_\alpha{}^\beta\, (\Gamma_{i_r\dots i_1})_\gamma{}^\delta\,,$$



(6.66)

with $r = 1,\dots,n, \mathrm{d}_n = 2^{\frac{1}{2}n}$ for n even, and $r = 1,\dots,\frac{1}{2}(n-1),\ d_n = 2^{\frac{1}{2}(n-1)}$ for n odd.

6.3.3 Construction of Representations of the Clifford Algebra

For $n = 2m$ an easy way to construct the γ-matrices satisfying the Clifford algebra (6.36) explicitly is to define

$$a_r = \tfrac{1}{2}(\gamma_{2r-1} + i\gamma_{2r})\,, \qquad a_r{}^\dagger = \tfrac{1}{2}(\gamma_{2r-1} - i\gamma_{2r})\,, \qquad r = 1,\dots,m\,. \tag{6.67}$$

Then (6.36) becomes

$$a_r\, a_s + a_s\, a_r = 0\,, \qquad a_r\, a_s{}^\dagger + a_s\, a_r{}^\dagger = \delta_{rs}\, \mathbb{1}\,, \tag{6.68}$$

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 $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$,

$$a = \sigma_+ = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right), \qquad a^\dagger = \sigma_- = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right), \qquad \gamma_1\gamma_2 = i\,\sigma_3 = i\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right). \tag{6.69}$$

The general case is obtained using tensor products

$$a_r = \underbrace{\mathbb{1}\otimes\cdots\otimes\mathbb{1}}_{r-1} \otimes \sigma_+ \otimes \underbrace{\sigma_3\otimes\cdots\otimes\sigma_3}_{m-r}\,, \qquad a_r{}^\dagger = \underbrace{\mathbb{1}\otimes\cdots\otimes\mathbb{1}}_{r-1} \otimes \sigma_- \otimes \underbrace{\sigma_3\otimes\cdots\otimes\sigma_3}_{m-r}\,. \tag{6.70}$$

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 (6.70) $\gamma_{2r-1}\gamma_{2r} = i\,\mathbb{1}\otimes\cdots\otimes\mathbb{1}\otimes\sigma_3\otimes\mathbb{1}\cdots\otimes\mathbb{1}$ so that

$$\Gamma = i^m \underbrace{\sigma_3\otimes\sigma_3\otimes\cdots\otimes\sigma_3}_{m}\,. \tag{6.71}$$

These results are equivalent to defining the gamma matrices for increasing n, where $\gamma_i{}^{(n)}\gamma_j{}^{(n)} + \gamma_j{}^{(n)}\gamma_i{}^{(n)} = 2\delta_{ij}\,\mathbb{1}^{(n)}$, recursively in terms of the Pauli matrices by

$$\begin{aligned}
&\gamma_i{}^{(2m+2)} = \gamma_i{}^{(2m)} \otimes \sigma_3\,, \qquad i = 1,\dots,2m\,,\\
&\gamma_{2m+1}^{(2m+2)} = \mathbb{1}^{(2m)} \otimes \sigma_1\,, \qquad \gamma_{2m+2}^{(2m+2)} = \mathbb{1}^{(2m)} \otimes \sigma_2\,,\\
&\Gamma^{(2m+2)} = i\,\Gamma^{(2m)} \otimes \sigma_3\,.
\end{aligned} \tag{6.72}$$

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

$$\gamma_i{}^{(2m+1)} = \gamma_i{}^{(2m)}\,, \quad i = 1,\dots,2m\,, \quad \gamma^{(2m+1)}_{2m+1} = c_m\,\Gamma^{(2m)}\,, \ c_m = \begin{cases} \pm 1\,, & m \text{ even}\,, \\ \pm i\,, & m \text{ odd}\,, \end{cases} \tag{6.73}$$

where the $\pm$ signs correspond to inequivalent representations. Thus $\gamma_i{}^{(3)} = (\sigma_1, \sigma_2, \mp\sigma_3)$.

6.3.4 Conjugation Matrix for Gamma Matrices

It is easy to see that γ_i^T also obeys the Clifford algebra in (6.36) so that for an irreducible representation we must have

$$\begin{aligned} & C\gamma_i\,C^{-1} = -\,\gamma_i^T && \Rightarrow && C\,\Gamma\,C^{-1} = (-1)^{\frac12 n(n+1)}\,\Gamma^T \\ \text{or} \quad & C\gamma_i\,C^{-1} = \gamma_i^T && \Rightarrow && C\,\Gamma C^{-1} = (-1)^{\frac12 n(n-1)}\,\Gamma^T\,. \end{aligned} \tag{6.74}$$

When n is even then, by taking $C \to C\Gamma$, the two cases are equivalent. When n is odd, and we require (6.53), 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 (6.38)

$$C\,s_{ij}\,C^{-1} = -s_{ij}{}^T\,, \tag{6.75}$$

so that for the matrices defining Spin(n)

$$e^{-\frac12 \omega_{ij}s_{ij}}\,C\left(e^{-\frac12 \omega_{ij}s_{ij}}\right)^T = C\,. \tag{6.76}$$

With the recursive construction of the gamma matrices $\gamma_i^{(n)}$ in (6.72) we may also construct in a similar fashion $C^{(n)}$ iteratively since, using (5.77),

$$\begin{aligned} & C^{(n)}\,\gamma_i^{(n)}\,C^{(n)\,-1} = \gamma_i^{(n)\,T} \\ \Rightarrow \quad & C^{(n+2)} = C^{(n)} \otimes i\sigma_2 \quad \text{ensures} \quad C^{(n+2)}\,\gamma_i^{(n+2)}\,C^{(n+2)\,-1} = -\gamma_i^{(n+2)\,T}\,, \end{aligned} \tag{6.77}$$

and, using $\sigma_1\sigma_i\sigma_1 = \sigma_i^T$, $i = 1,2$, $\sigma_1\sigma_3\sigma_1 = -\sigma_3{}^T$,

$$\begin{aligned} & C^{(n)}\,\gamma_i^{(n)}\,C^{(n)\,-1} = -\gamma_i^{(n)\,T} \\ \Rightarrow \quad & C^{(n+2)} = C^{(n)} \otimes \sigma_1 \quad \text{ensures} \quad C^{(n+2)}\,\gamma_i^{(n+2)}\,C^{(n+2)\,-1} = \gamma_i^{(n+2)\,T}\,. \end{aligned} \tag{6.78}$$

Starting from $n = 0$, or $n = 2$, this construction gives (note that $(X \otimes Y)^T = X^T \otimes Y^T$),

$$\begin{array}{llll} C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \Gamma^T\,, & C = C^T\,, & n = 8k\,, \\ C\,\gamma_i\,C^{-1} = -\,\gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = -\,\Gamma^T\,, & C = -\,C^T\,, & n = 8k+2\,, \\ C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = \Gamma^T\,, & C = -\,C^T\,, & n = 8k+4\,, \\ C\,\gamma_i\,C^{-1} = -\,\gamma_i{}^T\,, & C\,\Gamma\,C^{-1} = -\,\Gamma^T\,, & C = C^T\,, & n = 8k+6\,. \end{array} \tag{6.79}$$

In each case we have $C s_{ij} C^{-1} = -s_{ij}{}^T$. Starting from (6.79) and with the construction in (6.73) for odd n,

$$\begin{aligned}
&C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, && C = C^T\,, && n = 8k+1\,,\\
&C\,\gamma_i\,C^{-1} = -\gamma_i{}^T\,, && C = -C^T\,, && n = 8k+3\,,\\
&C\,\gamma_i\,C^{-1} = \gamma_i{}^T\,, && C = -C^T\,, && n = 8k+5\,,\\
&C\,\gamma_i\,C^{-1} = -\gamma_i{}^T\,, && C = C^T\,, && n = 8k+7\,.
\end{aligned} \tag{6.80}$$

The definition of C for $n = 2m+1$ remains the same as in (6.79) for $n = 2m$ since in each n odd case we have $C\gamma_1\gamma_2\dots\gamma_n C^{-1} = (\gamma_1\gamma_2\dots\gamma_n)^T$.

If we consider a basis in which Γ is diagonal, as in (6.55), 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 $C \to S^T C S$ with $S\Gamma S^{-1} = \Gamma$ we may choose with the basis in (6.55),

$$\begin{aligned}
&C = \begin{pmatrix} 1\!\!1 & 0 \\ 0 & 1\!\!1 \end{pmatrix}, && \bar\sigma_i = \sigma_i{}^T, && s_{\pm ij} = -\,s_{\pm ij}{}^T, && n = 8k\,,\\
&C = \begin{pmatrix} 0 & 1\!\!1 \\ -1\!\!1 & 0 \end{pmatrix}, && \sigma_i = \sigma_i{}^T,\ \bar\sigma_i = \bar\sigma_i{}^T, && s_{\pm ij} = -\,s_{\mp ij}{}^T, && n = 8k+2\,,\\
&C = \begin{pmatrix} J & 0 \\ 0 & J \end{pmatrix}, && J = -\,J^T,\ J\bar\sigma_i = -(J\sigma_i)^T, && J s_{\pm ij} = (J s_{\pm ij})^T, && n = 8k+4\,,\\
&C = \begin{pmatrix} 0 & 1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}, && \sigma_i = -\,\sigma_i{}^T,\ \bar\sigma_i = -\bar\sigma_i{}^T, && s_{\pm ij} = -\,s_{\mp ij}{}^T, && n = 8k+6\,.
\end{aligned} \tag{6.81}$$

Here the antisymmetric matrix J can be taken to be of the standard form as in (1.108). For $n = 8k$ the matrices are real.

Since the generators of the two fundamental spinor representations satisfy (6.56) then as a consequence of the discussion in section 5.3.3 we have for these representations of $\mathrm{Spin}(n)$, for n even, from (6.81)

$$\begin{aligned}
&\mathrm{Spin}(8k):\ \text{real}\,, \qquad \mathrm{Spin}(8k+4):\ \text{pseudo-real}\,,\\
&\mathrm{Spin}(8k+2)\,,\ \mathrm{Spin}(8k+6):\ \text{complex}\,.
\end{aligned} \tag{6.82}$$

Furthermore for n odd the single spinor representation, from (6.80), satisfies

$$\mathrm{Spin}(8k+1)\,,\mathrm{Spin}(8k+7):\ \text{real}\,, \qquad \mathrm{Spin}(8k+3)\,,\mathrm{Spin}(8k+5):\ \text{pseudo-real}\,. \tag{6.83}$$

6.3.5 Special Cases

When $n = 2$ we may take

$$\sigma_i = (1, -i)\,, \qquad \bar\sigma_i = (1, i)\,, \tag{6.84}$$

while for $n = 4$ we may express $\sigma_i, \bar\sigma_i$ in terms of unit quaternions

$$\sigma_i = (1, -i, -j, -k)\,, \qquad \bar\sigma_i = (1, i, j, k)\,. \tag{6.85}$$

For low n results for γ-matrices may be used to identify $\mathrm{Spin}(n)$ with other groups. Thus

$$\begin{aligned} \mathrm{Spin}(3) &\simeq SU(2)\,, & \mathrm{Spin}(4) &\simeq SU(2)\times SU(2)\,, \\ \mathrm{Spin}(5) &\simeq Sp(2)\,, & \mathrm{Spin}(6) &\simeq SU(4)\,. \end{aligned} \tag{6.86}$$

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 (6.52), with the decomposition in (6.55), we have

$$s_{\pm ij} = \pm\tfrac{1}{2}\,\varepsilon_{ijkl}\,s_{\pm kl}\,, \tag{6.87}$$

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 (6.76) 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 (6.48), $\mathcal{Z}(\mathrm{Spin}(6)) \simeq \mathbb{Z}_4 \simeq \mathcal{Z}(SU(4))$. Using (6.81) with (6.55), the transformation (6.41) can be rewritten just in terms of the $SU(4)$ matrix

$$e^{-\frac{1}{2}\omega_{ij}s_{+ij}}\,\sigma\cdot x\,\big(e^{-\frac{1}{2}\omega_{ij}s_{+ij}}\big)^T = \sigma\cdot x'\,, \tag{6.88}$$

which is analogous to (3.27). The result that the transformation $x \to x'$ satisfies $x^2 = x'^2$ also follows in a similar fashion to (3.29), but in this case using the Pfaffian (1.109) 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 $n = 6$, $\det(\sigma\cdot x) = (x^2)^2$).

6.3.6 Fierz Identities

Fierz[47] identities which depend on the completeness properties of γ-matrices play a crucial role in many calculations. To derive Fierz identities it is convenient to map the γ-matrices to operator fermi fields $\hat\psi_i$ so that

$$\gamma_i \to \hat\psi_i\,, \qquad \{\hat\psi_i,\,\hat\psi_j\} = 2\,\delta_{ij}\,\mathbb{1}\,, \tag{6.89}$$

with i a n-dimensional index and where $\hat\psi_i$ is decomposed into fermionic creation and annihilation operators

$$\hat\psi_i = b_i + b_i{}^\dagger\,, \qquad \{b_i,\,b_j\} = 0\,, \qquad \{b_i,\,b_j{}^\dagger\} = \delta_{ij}\,\mathbb{1}\,. \tag{6.90}$$

Then from (6.49)

$$\Gamma_{i_1\ldots i_r} \to\, :\hat\psi_{i_1}\ldots\hat\psi_{i_r}:\,, \tag{6.91}$$

where $:\cdots:$ denotes normal ordering, where annihilation operators are moved to the right of creation operators taking into account anticommutation signs as required but dropping non zero terms from (6.90), thus $:\hat\psi_i\hat\psi_i: = 0$ while $\hat\psi_i\hat\psi_i = n\,\mathbb{1}$.

[47] Markus Fierz, 1912-2006, Swiss. A student of and assistant to Pauli.

The normal ordered products are conveniently represented in terms of a generating function

$$e^{\theta_i \hat\psi_i} = \textstyle\sum_{s=0}^\infty \frac{1}{s!}\, \theta_{i_s}\dots\theta_{i_1}\ :\hat\psi_{i_1}\dots\hat\psi_{i_s}: = \sum_{s=0}^\infty \frac{1}{s!}(-1)^s\ :\hat\psi_{i_s}\dots\hat\psi_{i_1}:\ \theta_{i_1}\dots\theta_{i_s}\,, \tag{6.92}$$

where $\{\theta_i\}$ are arbitrary Grassmannian, or anticommuting, variables. These ensure that only antisymmetrised products of $\hat\psi_i$ arise in the expansion and therefore the products are normal ordered. For n an integer the number of terms in the expansion is finite since $:\hat\psi_{i_1}\dots\hat\psi_{i_s}: = 0$ for $s > n$. As an application we may use

$$e^{\theta_i\hat\psi_i}\, e^{\tilde\theta_j\hat\psi_j} = e^{(\theta+\tilde\theta)_i\hat\psi_i} e^{\tilde\theta_j\theta_j}\,, \tag{6.93}$$

and expanding to $\mathrm{O}(\theta^r,\tilde\theta^s)$ gives

$$\begin{aligned}&\tilde\theta_{j_s}\dots\tilde\theta_{j_1}\,\theta_{i_r}\dots\theta_{i_1}:\hat\psi_{i_1}\dots\hat\psi_{i_r}::\hat\psi_{j_1}\dots\hat\psi_{j_s}:\\ &= \textstyle\sum_{t=0}^{\min(r,s)}\frac{r!\ s!}{(r-t)!\,(s-t)!\,t!}\,\tilde\theta_{j_{s-t}}\dots\tilde\theta_{j_1}\tilde\theta_{k_t}\dots\tilde\theta_{k_1}\theta_{k_1}\dots\theta_{k_t}\theta_{i_{r-t}}\dots\theta_{i_1}:\hat\psi_{i_1}\dots\hat\psi_{i_{r-t}}\,\hat\psi_{j_1}\dots\hat\psi_{j_{s-t}}:.\end{aligned} \tag{6.94}$$

This implies

$$\begin{aligned}&\Gamma_{i_1\dots i_r}\,\Gamma_{j_1\dots j_s}\\ &\quad= \textstyle\sum_{t=0}^{\min(r,s)}\frac{r!\ s!}{(r-t)!\,(s-t)!\,t!^2}\,(-1)^{\frac12 t(2s-t-1)}\,\Gamma_{[i_1\dots i_{r-t}|[j_1\dots j_{s-t}}\,(A_t)_{j_{s-t+1}\dots j_s],i_{r-t+1}\dots i_r]}\,,\end{aligned} \tag{6.95}$$

where $(A_s)_{j_1\dots j_s,i_1\dots i_s}$ is defined in (6.61) and acts as the δ-function for antisymmetric rank s tensors so that $\frac{1}{s!}(A_s)_{j_1\dots j_s,i_1\dots i_s}\theta_{1_1}\dots\theta_{i_s} = \theta_{j_1}\dots\theta_{j_s}$. In (6.95) the i-indices and j-indices are separately antisymmetrised.

In a similar fashion starting from $e^{\phi_i\hat\psi_i}\,e^{\theta_j\hat\psi_j}\,e^{\tilde\phi_k\hat\psi_k}$ we may obtain

$$\begin{aligned}&\tilde\phi_{k_t}\dots\tilde\phi_{k_1}\,\theta_{j_s}\dots\theta_{j_1}\,\phi_{i_r}\dots\phi_{i_1}\,\mathrm{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{j_1\dots j_s}\,\Gamma_{k_1\dots k_t}\big) = d_n\,\tfrac{r!\,s!\,t!}{a!\,b!\,c!}\,(\tilde\phi_i\theta_i)^a(\tilde\phi_j\phi_j)^b(\theta_k\phi_k)^c\,,\\ &\qquad a = \tfrac12(s+t-r)\,,\quad b = \tfrac12(t+r-s)\,,\quad c = \tfrac12(r+s-t)\,.\end{aligned} \tag{6.96}$$

This is non zero for $|r-s| \le t \le r+s$, $r+s+t$ even and implies

$$\begin{aligned}&\mathrm{tr}\big(\Gamma_{i_1\dots i_r}\,\Gamma_{j_1\dots j_s}\,\Gamma_{k_1\dots k_t}\big)\\ &\quad= d_n\,\tfrac{r!\,s!\,t!}{(a!\,b!\,c!)^2}\,(-1)^{sb}\,(A_a)_{[k_t\dots k_{b+1},[j_s\dots j_{c+1}}\,(A_b)_{k_b\dots k_1],[i_r\dots i_{c+1}}\,(A_c)_{j_c\dots j_1],i_c\dots i_1]}\,.\end{aligned} \tag{6.97}$$

Fierz identities are obtained by considering

$$\begin{aligned}&e^{u\,\partial^2/\partial\phi_l\partial\tilde\phi_l}\,e^{\phi_j\hat\psi_j}\,e^{\theta_i\hat\psi_i}\,e^{\tilde\phi_k\hat\psi_k}\big|_{\phi=\tilde\phi=0} = \textstyle\sum_{r=0}^\infty\frac{1}{r!}\,u^r\ :\hat\psi_{j_r}\dots\hat\psi_{j_1}:\ e^{\theta_i\hat\psi_i}:\hat\psi_{j_1}\dots\hat\psi_{j_r}:\\ &\quad= \textstyle\sum_{s=0}^\infty\frac{1}{s!}\ \theta_{i_s}\dots\theta_{i_1}\ \sum_{r=0}^\infty\frac{1}{r!}\,u^r(-1)^{rs}\ :\hat\psi_{j_r}\dots\hat\psi_{j_1}::\hat\psi_{i_1}\dots\hat\psi_{i_s}::\hat\psi_{j_1}\dots\hat\psi_{j_r}:,\end{aligned} \tag{6.98}$$

where $\{\phi_j,\,\tilde\phi_j\}$ are further independent Grassmannian variables and $\frac{\partial}{\partial\phi_l}\phi_j + \phi_j\frac{\partial}{\partial\phi_l} = \delta_j^{\,k}$.

This can be evaluated using[48]

$$e^{u\,\partial^2/\partial\phi_l\partial\tilde\phi_l}\, e^{\phi_j\hat\psi_j}\, e^{\theta_i\hat\psi_i}\, e^{\tilde\phi_k\hat\psi_k}\Big|_{\phi=\tilde\phi=0} = e^{u\,\partial^2/\partial\phi_l\partial\tilde\phi_l}\, e^{\phi_j\hat\psi_j+\tilde\phi_j(\hat\psi_j+2\theta_j)-\phi_j\tilde\phi_j}\Big|_{\phi=\tilde\phi=0} e^{\theta_i\hat\psi_i}$$
$$= (1+u)^n\, e^{\frac{1-u}{1+u}\,\theta_i\hat\psi_i} = \textstyle\sum_{s=0}^\infty \frac{1}{s!}(1+u)^{n-s}(1-u)^s\,\theta_{i_s}\dots\theta_{i_1}\,:\hat\psi_{i_1}\dots\hat\psi_{i_s}:\,. \qquad (6.99)$$

Thus

$$\tfrac{1}{r!}\,:\hat\psi_{j_r}\dots\hat\psi_{j_1}::\hat\psi_{i_1}\dots\hat\psi_{i_s}::\hat\psi_{j_1}\dots\hat\psi_{j_r}: = \Omega_{rs}\,:\hat\psi_{i_1}\dots\hat\psi_{i_s}:\,, \qquad s=1,2,\dots\,,$$
$$\tfrac{1}{r!}\,:\hat\psi_{j_r}\dots\hat\psi_{j_1}::\hat\psi_{j_1}\dots\hat\psi_{j_r}: = \Omega_{r0}\,\mathbb{1}\,, \qquad (6.100)$$

for

$$\begin{aligned}\Omega_{rs} &= (-1)^{rs}\,\tfrac{1}{r!}\,\tfrac{\mathrm{d}^r}{\mathrm{d}u^r}(1+u)^{n-s}(1-u)^s\Big|_{u=0}\\ &= (-1)^{rs}\textstyle\sum_{t=0}^{\min(r,s)}(-1)^t\,\binom{s}{t}\binom{n-s}{r-t}\\ &= (-1)^{rs}\binom{n-s}{r}F(-r,-s;n-r-s+1;-1)\\ &= (-1)^{rs}\binom{n}{r}F(-r,-s;-n;2)\,,\end{aligned} \qquad (6.101)$$

where F is a hypergeometric function and the last line follows by the ralation between hypergeometric functions with argument x and $1-x$. For integer n it is necessary that $r,s\le n$. Manifestly $r!(n-r)!\,\Omega_{rs} = s!(n-s)!\,\Omega_{sr}$ and, using the properties of hypergeometric functions, $\Omega_{rs} = (-1)^{r(n-1)}\Omega_{r\,n-s} = (-1)^{s(n-1)}\Omega_{n-r\,s}$. In general $\Omega_{r0} = \binom{n}{r}$, $\Omega_{0s}=1$, $\Omega_{1s} = (-1)^s(n-2s)$ and also

$$\textstyle\sum_r\Omega_{rs} = \exp\big((-1)^s\frac{\mathrm{d}}{\mathrm{d}u}\big)(1+u)^{n-s}(1-u)^s\Big|_{u=0} = \begin{cases}(2+u)^{n-s}(-u)^s\big|_{u=0}\,, & s \text{ even}\\ u^{n-s}(2-u)^s\big|_{u=0}\,, & s\text{ odd}\end{cases}$$
$$= \begin{cases}2^n\,, & s=0\,,\ s=n\,,\ n\text{ odd}\\ 0\,, & \text{otherwise}\end{cases}\,. \qquad (6.102)$$

As particular cases for even n

$$\Omega\big|_{n=2} = \begin{pmatrix}1&1&1\\2&0&-2\\1&-1&1\end{pmatrix}, \qquad \Omega\big|_{n=4} = \begin{pmatrix}1&1&1&1&1\\4&-2&0&2&-4\\6&0&-2&0&6\\4&2&0&-2&-4\\1&-1&1&-1&1\end{pmatrix},$$
$$\Omega\big|_{n=6} = \begin{pmatrix}1&1&1&1&1&1&1\\6&-4&2&0&-2&4&-6\\15&5&-1&-3&-1&5&15\\20&0&-4&0&4&0&-20\\15&-5&-1&3&-1&-5&15\\6&4&2&0&-2&-4&-6\\1&-1&1&-1&1&-1&1\end{pmatrix}, \qquad (6.103)$$

[48]When the derivatives act on a normal ordered function of $\hat\psi$ then, in a similar fashion to the discussion in 3.15.1, the operator anticommutation relations are irrelevant and for evaluation we can let $\hat\psi\to\psi$, an ordinary Grassmann variable. The necessary identity may then be obtained, for $\tilde\phi$, ϕ, $\tilde\psi$, ψ n dimensional Grassmann variables and U, V $n\times n$ matrices, by using

$$e^{-\partial/\partial\tilde\phi\,U\,\partial/\partial\phi}\,e^{-\phi\,V\,\tilde\phi}\,e^{\phi\tilde\psi+\psi\tilde\phi}\Big|_{\phi=\tilde\phi=0} = e^{\partial/\partial\tilde\psi\,V\,\partial/\partial\psi}\,e^{\psi\,U\,\tilde\psi} = e^{\partial/\partial\tilde\psi\,V\,\partial/\partial\psi}\det U\int \mathrm{d}^n\eta\,\mathrm{d}^n\tilde\eta\; e^{-\eta\,U^{-1}\,\tilde\eta+\eta\tilde\psi+\psi\tilde\eta}$$
$$= \det U\int\mathrm{d}^n\eta\,\mathrm{d}^n\tilde\eta\;e^{-\eta\,(U^{-1}+V)\,\tilde\eta+\eta\tilde\psi+\psi\tilde\eta} = \det(1+U\,V)\,e^{\psi\,(1+UV)^{-1}U\,\tilde\psi}\,,$$

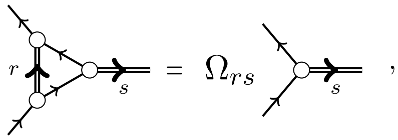
where the crucial step is to express $e^{\psi\,U\,\tilde\psi}$ in terms of an integral over Grassmann variables η, $\tilde\eta$ and with the conventions here $e^{\eta\tilde\psi+\psi\tilde\eta}$ is an eigenvector for $\partial/\partial\tilde\psi,\partial/\partial\psi$ with eigenvalues $-\eta$, $\tilde\eta$. In application to (6.99) $\phi\to\phi+\tilde\phi$, $\tilde\psi\to\psi$, $\psi\to-2\theta$ and $U\to u\,\mathbb{1}$, $V\to\mathbb{1}$ so that $\mathbb{1}-2(1+UV)^{-1}U\to\frac{1-u}{1+u}$.

and for n odd (6.101) gives

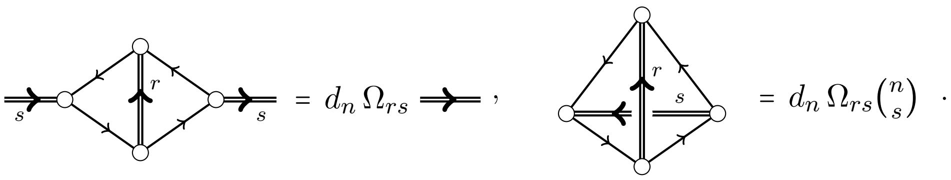
$$\Omega|_{n=3} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3 & -1 & -1 & 3 \\ 3 & -1 & -1 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad \Omega|_{n=5} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & -3 & 1 & 1 & -3 & 5 \\ 10 & 2 & -2 & -2 & 2 & 10 \\ 10 & 2 & -2 & -2 & 2 & 10 \\ 5 & -3 & 1 & 1 & -3 & 5 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{6.104}$$

By virtue of (6.91), (6.100) is equivalent to the γ-matrix identity

$$\tfrac{1}{r!}\,\Gamma_{j_r\dots j_1}\,\Gamma_{i_1\dots i_s}\,\Gamma_{j_1\dots j_r} = \Omega_{rs}\,\Gamma_{i_1\dots i_s}\,, \tag{6.105}$$



from which it follows that

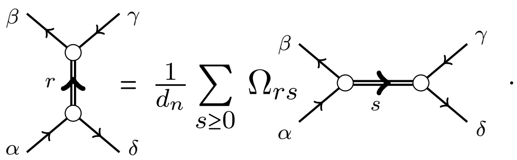


(6.106)

Clearly it is necessary that $\Omega_{rs}\binom{n}{s}$ is symmetric under $r \leftrightarrow s$ which is evident from (6.101).

With the completeness relation (6.66) this implies a crossing relation as in (3.194)

$$\tfrac{1}{r!}\,(\Gamma_{j_r\dots j_1})_\alpha{}^\delta\;(\Gamma_{j_1\dots j_r})_\gamma{}^\beta = \tfrac{1}{d_n}\textstyle\sum_{s\ge 0}\Omega_{rs}\,\tfrac{1}{s!}\,(\Gamma_{i_s\dots i_1})_\alpha{}^\beta\;(\Gamma_{i_1\dots i_s})_\gamma{}^\delta\,, \tag{6.107}$$



These relations are what is usually referred to as the Fierz identity. For consistency it is necessary that

$$\textstyle\sum_{t\ge 0}\Omega_{rt}\,\Omega_{ts} = d_n{}^2\,\delta_{rs}\,, \quad \text{or} \quad \Omega^2 = d_n{}^2\,\mathbb{1}_{n+1}\,, \tag{6.108}$$

which requires, except when n is an odd integer,

$$d_n = 2^{\frac{1}{2}n}\,. \tag{6.109}$$

As a check we may use the first line of (6.101) to obtain via a Taylor expansion

$$\textstyle\sum_{t\ge 0} x^t\,\Omega_{ts} = \big(1 + (-1)^s x\big)^{n-s}\big(1 - (-1)^s x\big)^s\,, \tag{6.110}$$

and then use this in (6.108) to obtain

$$\begin{aligned} \textstyle\sum_{t\ge 0}\Omega_{rt}\,\Omega_{ts} &= \tfrac{1}{r!}\,\tfrac{\mathrm{d}^r}{\mathrm{d}u^r}\big(1 + u + (-1)^{r+s}(1-u)\big)^{n-s}\big(1 + u - (-1)^{r+s}(1-u)\big)^s\Big|_{u=0} \\ &= \tfrac{1}{r!}\,2^n \begin{cases} \tfrac{\mathrm{d}^r}{\mathrm{d}u^r}\,u^s\big|_{u=0}\,, & r+s \ \text{even} \\ \tfrac{\mathrm{d}^r}{\mathrm{d}u^r}\,u^{n-s}\big|_{u=0}\,, & r+s \ \text{odd} \end{cases} = 2^n \begin{cases} \delta_{rs}\,, & r+s \ \text{even} \\ 1\,, & r+s = n \ \text{odd}\,. \end{cases} \end{aligned} \tag{6.111}$$

For $n = 2m+1$ odd $\{\mathbb{1}, \Gamma_{i_1\dots i_r} : 1 \le r \le m\}$ form an independent basis and we have, as a consequence of the symmetry properties of Ω,

$$\Omega = \begin{pmatrix} \hat{\Omega} & \hat{\Omega}\mathcal{C} \\ \mathcal{C}\hat{\Omega} & \mathcal{C}\hat{\Omega}\mathcal{C} \end{pmatrix}, \qquad \mathcal{C}_{rs} = \delta_{r+s,m+1}, \quad \mathcal{C}^2 = \mathbb{1}, \tag{6.112}$$

where $\hat{\Omega}$, $\mathcal{C}$ are $(m+1)\times(m+1)$ matrices, $\mathcal{C}$ is anti-diagonal. Clearly in this case from (6.104) $\det \Omega = 0$. The result (6.111) for $n = 2m+1$ requires $\Omega^2 = 2\begin{pmatrix} \hat{\Omega}^2 & \mathcal{C}\hat{\Omega}^2 \\ \hat{\Omega}^2\mathcal{C} & \mathcal{C}\hat{\Omega}^2\mathcal{C} \end{pmatrix} = 2^{2m+1}\begin{pmatrix} \mathbb{1} & \mathcal{C} \\ \mathcal{C} & \mathbb{1} \end{pmatrix}$ and is then satisfied for

$$\hat{\Omega}^2 = 2^{2m}\,\mathbb{1}_{m+1}. \tag{6.113}$$

Hence in (6.105) and (6.107) we may restrict $0 \le r, s \le m$ and take $\Omega \to \hat{\Omega}$, $d_n \to 2^m$. For $n = 3, 5$ the 2×2, 3×3 reduced matrices forming $\hat{\Omega}$ in these cases are given by the corresponding top left hand sub matrices in (6.104). For $n = 7$, $m = 3$

$$\hat{\Omega}|_{n=7} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 7 & -5 & 3 & -1 \\ 21 & 9 & 1 & -3 \\ 35 & -5 & -5 & 3 \end{pmatrix}. \tag{6.114}$$

In general Ω_{rs} is a polynomial in n of degree r and the result (6.101) with (6.109) may be extended to any complex n. In this case $\{\mathbb{1}, \Gamma_{i_1\dots i_r} : r = 1, 2, \dots\}$ span an infinite dimensional space although the norm in (6.65) is no longer positive definite when $r > n$. There is an infinite sum in (6.108) which is convergent for $\mathrm{Re}\, n > r + s$ and may then defined by analytic continuation in n.

7 $SU(3)$ 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 crucial breakthrough, leading to the realisation that quarks are the fundamental constituents of hadrons, 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 (6.3) where, since the corresponding matrices in (6.2) 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 $SU(2)$.

7.1 Recap of $\mathfrak{su}(2)$

For the basic generators of $\mathfrak{su}(2)$ we define in terms of 2×2 matrices as in (6.2)

$$e_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad e_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{7.1}$$

which satisfy the Lie algebra

$$[e_+, e_-] = h\,, \qquad [h, e_\pm] = \pm 2\, e_\pm\,. \tag{7.2}$$

These matrices satisfy

$$e_+{}^\dagger = e_-\,, \qquad h^\dagger = h\,. \tag{7.3}$$

Under interchange of the rows and columns

$$b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \Rightarrow \qquad b\big\{e_+, e_-, h\big\}b^{-1} = \big\{e_-, e_+, -h\big\}\,. \tag{7.4}$$

Clearly $b^2 = \mathbb{1}$ and $\big\{e_+, e_-, h\big\}$, $\{e_-, e_+, -h\big\}$ must satisfy the same commutation relations as in (7.2) so b generates an automorphism.

For representations of the $\mathfrak{su}(2)$ Lie algebra then we require operators

$$\mathfrak{l} = \{E_+, E_-, H\}\,, \qquad [E_+, E_-] = H\,, \quad [H, E_\pm] = \pm 2\, E_\pm\,. \tag{7.5}$$

It is easy to see that the commutation relations are identical with (3.61a) and (3.61a), and also the hermeticity conditions with (3.62), by taking $J_\pm \to E_\pm$, $2J_3 \to H$. Indeed the representation matrices in (7.1) then correspond exactly with (3.103).

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*[49] *group*. For $\mathfrak{su}(2)$ the relevant basis is given by (7.5) and then from (7.4) there is just one non trivial automorphism

$$\mathfrak{l} \underset{b}{\to} \mathfrak{l}_R = \{E_-, E_+, -H\}\,. \tag{7.6}$$

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 (7.5) and subsequently require there is a scalar product so that the operators satisfy the hermeticity conditions in (7.3). A basis for a representation space for $\mathfrak{su}(2)$ is given by $\{|r\rangle\}$ where

$$H|r\rangle = r|r\rangle\,. \tag{7.7}$$

The eigenvalue r is termed the *weight*. It is easy to see from (7.5) that

$$E_\pm|r\rangle \propto |r\pm 2\rangle \qquad \text{unless} \quad E_+|r\rangle = 0 \quad \text{or} \quad E_-|r\rangle = 0\,. \tag{7.8}$$

We consider representations where there is a *highest weight*, $r_{\max} = n$, and hence a highest weight vector $|n\rangle_{\rm hw}$ satisfying

$$E_+|n\rangle_{\rm hw} = 0\,. \tag{7.9}$$

The representation space V_n is then spanned by

$$\big\{ {E_-}^r|n\rangle_{\rm hw} : r = 0,1,\dots \big\}\,. \tag{7.10}$$

On this basis

$$H {E_-}^r|n\rangle_{\rm hw} = (n-2r)\, {E_-}^r|n\rangle_{\rm hw}\,, \tag{7.11}$$

and using

$$\big[E_+, {E_-}^r\big] = \sum_{s=0}^{r-1} {E_-}^{r-s-1}[E_+,E_-]{E_-}^s = {E_-}^{r-1}\sum_{s=0}^{r-1}\big(H-2s\big) = {E_-}^{r-1}\, r(H-r+1)\,, \tag{7.12}$$

then from (7.9),

$$E_+\, {E_-}^r|n\rangle_{\rm hw} = r(n-r+1)\, {E_-}^{r-1}|n\rangle_{\rm hw}\,. \tag{7.13}$$

(7.11) and (7.13) ensure that the commutation relations (7.5) are realised on V_n.

If $n \in \mathbb{N}_0$, or $n = 0,1,2,\dots$, then from (7.13)

$$|-n-2\rangle_{\rm hw} = {E_-}^{n+1}|n\rangle_{\rm hw} \in V_n\,, \tag{7.14}$$

is also a highest weight vector, satisfying (7.9). From $|-n-2\rangle_{\rm hw}$ we may construct, just as in (7.10), a basis for an associated invariant subspace

$$V_{-n-2} \subset V_n\,. \tag{7.15}$$

[49]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

$$\mathcal{V}_n = V_n \big/ V_{-n-2} \,. \tag{7.16}$$

In general for a vector space V with a subspace U the quotient V/U is defined by

$$V/U = \left\{ |v\rangle/\!\sim : |v\rangle \sim |v'\rangle \text{ if } |v\rangle - |v'\rangle \in U \right\}. \tag{7.17}$$

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

$$U \underset{X}{\to} U \quad \Rightarrow \quad \{X|v\rangle/\sim\} = \{X|v'\rangle/\sim\} \;\text{ if }\; |v\rangle \sim |v'\rangle \quad \Rightarrow \quad X : V/U \to V/U \,. \tag{7.18}$$

Thus, if $U \subset V$ is an invariant subspace under X, then X has a well defined action on V/U. Furthermore for traces

$$\operatorname{tr}_{V/U}(X) = \operatorname{tr}_V(X) - \operatorname{tr}_U(X) \,. \tag{7.19}$$

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 (7.16). On $\mathcal{V}_n$ this ensures

$$E_-{}^{n+1}|n\rangle_{\text{hw}} = 0 \,, \tag{7.20}$$

so that there is a finite basis $\{E_-{}^r|n\rangle_{\text{hw}} : r = 0, \dots, n\}$. In terms of the angular momentum representations constructed in section 3, $n = 2j$. 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 (7.4) of the $\mathfrak{su}(2)$ Lie algebra.

If we define a formal trace over all vectors belonging to V_n then

$$C_n(t) = \widetilde{\operatorname{tr}}_{V_n}\big(t^H\big) = \sum_{r=0}^{\infty} t^{n-2r} = \frac{t^{n+2}}{t^2 - 1} \,, \tag{7.21}$$

where convergence of the sum requires $|t| > 1$. Then for the irreducible representation defined on the quotient $\mathcal{V}_n$, by virtue of (7.19), the character is

$$\chi_n(t) = \operatorname{tr}_{\mathcal{V}_n}\big(t^H\big) = C_n(t) - C_{-n-2}(t) = \frac{t^{n+2} - t^{-n}}{t^2 - 1} = \frac{t^{n+1} - t^{-n-1}}{t - t^{-1}} \,. \tag{7.22}$$

This is just the same as (3.132) with $t \to e^{i\frac{1}{2}\theta}$ and $n \to 2j$. It is easy to see that

$$\chi_n(1) = \dim \mathcal{V}_n = n + 1 \,. \tag{7.23}$$

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

$$\chi_n(t) = -\chi_{-n-2}(t) \,, \tag{7.24}$$

as follows directly from (7.22). Clearly $\chi_{-1}(t) = 0$.

The $\mathfrak{su}(2)$ Casimir operator in this basis

$$C = E_+E_- + E_-E_+ + \tfrac{1}{2}H^2 = 2E_-E_+ + \tfrac{1}{2}H^2 + H\,, \tag{7.25}$$

and it is easy to see that

$$C|n\rangle_{\rm hw} = c_n|n\rangle_{\rm hw} \qquad \text{for} \qquad c_n = \tfrac{1}{2}n(n+2)\,. \tag{7.26}$$

Note that $c_{-n-2} = c_n$ as required from (7.14) as all vectors belonging to V_n must have the same eigenvalue for C.

7.2 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 (6.2). Thus we define

$$e_{1+} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad e_{2+} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad e_{3+} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{7.27}$$

and their conjugates

$$e_{i-} = e_{i+}{}^{\dagger}\,, \quad i = 1,2,3\,, \tag{7.28}$$

together with the hermitian traceless diagonal matrices

$$h_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad h_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{7.29}$$

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 $(1\,2\,3)$

$$b = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad a = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{7.30}$$

then

$$\begin{aligned} b\{h_1,h_2\}b^{-1} &= \{-h_1, h_1+h_2\}\,, \qquad & b\{e_{1\pm}, e_{2\pm}, e_{3\pm}\}b^{-1} &= \{e_{1\mp}, e_{3\pm}, e_{2\pm}\}\,, \\ a\{h_1,h_2\}a^{-1} &= \{h_2, -h_1-h_2\}\,, \qquad & a\{e_{1\pm}, e_{2\pm}, e_{3\pm}\}a^{-1} &= \{e_{2\pm}, e_{3\mp}, e_{1\mp}\}\,. \end{aligned} \tag{7.31}$$

The matrices in (7.30) satisfy

$$b^2 = \mathbb{1}\,, \qquad a^3 = \mathbb{1}\,, \qquad ab = ba^2\,, \tag{7.32}$$

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

$$\{E_{1\pm}, E_{2\pm}, E_{3\pm}, H_1, H_2\} \;\to\; \big[\hat{R}^i{}_j\big] = \begin{pmatrix} \frac{1}{3}(2H_1+H_2) & E_{1+} & E_{3+} \\ E_{1-} & \frac{1}{3}(-H_1+H_2) & E_{2+} \\ E_{3-} & E_{2-} & -\frac{1}{3}(H_1+2H_2) \end{pmatrix}, \tag{7.33}$$

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

$$\big[\hat{R}^i{}_j,\, \hat{R}^k{}_l\big] = \delta^k{}_j\, \hat{R}^i{}_l - \delta^i{}_l\, \hat{R}^k{}_j\,, \tag{7.34}$$

since, for X, Y appropriate matrices, (7.34) requires

$$\big[\mathrm{tr}(X\hat{R}), \mathrm{tr}(Y\hat{R})\big] = \mathrm{tr}\big([X,Y]\hat{R}\big)\,, \tag{7.35}$$

and with the definitions (7.27) and (7.29) we have, from (7.33), $\mathrm{tr}(e_{i\pm}\hat{R}) = E_{i\pm}$, $i = 1,2,3$ 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 (7.33). Crucially there are two commuting generators H_1, H_2 so that

$$[H_1, H_2] = 0\,. \tag{7.36}$$

For E_{i+} the commutation relations are

$$[E_{1+}, E_{2+}] = E_{3+}\,, \qquad [E_{1+}, E_{3+}] = [E_{2+}, E_{3+}] = 0\,. \tag{7.37}$$

while under commutation with H_1, H_2

$$\begin{aligned}\big[H_1, \{E_{1\pm}, E_{2\pm}, E_{3\pm}\}\big] &= \pm\{2E_{1\pm}, -E_{2\pm}, E_{3\pm}\}\,,\\ \big[H_2, \{E_{1\pm}, E_{2\pm}, E_{3\pm}\}\big] &= \pm\{-E_{1\pm}, 2E_{2\pm}, E_{3\pm}\}\,.\end{aligned} \tag{7.38}$$

The remaining commutators involving $E_{i\pm}$ are

$$\begin{aligned}&[E_{1+}, E_{1-}] = H_1\,, &&[E_{1+}, E_{2-}] = 0\,, &&[E_{2+}, E_{2-}] = H_2\,,\\ &[E_{3+}, E_{1-}] = -E_{2+}\,, &&[E_{3+}, E_{2-}] = E_{1+}\,, &&[E_{3+}, E_{3-}] = H_1 + H_2\,,\end{aligned} \tag{7.39}$$

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 (7.33) can be decomposed into three $\mathfrak{su}(2)$ Lie algebras,

$$\mathfrak{l}_1 = \{E_{1+}, E_{1-}, H_1\}\,, \qquad \mathfrak{l}_2 = \{E_{2+}, E_{2-}, H_2\}\,, \qquad \mathfrak{l}_3 = \{E_{3+}, E_{3-}, H_1 + H_2\}\,, \tag{7.40}$$

where each $\mathfrak{l}_i$ satisfies (7.5). From (7.31) the automorphism symmetries of the privileged basis in (7.33) are generated by

$$\mathfrak{l}_1 \underset{b}{\to} \mathfrak{l}_{1R}\,, \quad \mathfrak{l}_2 \underset{b}{\to} \mathfrak{l}_3\,, \quad \mathfrak{l}_3 \underset{b}{\to} \mathfrak{l}_2\,, \qquad \mathfrak{l}_1 \underset{a}{\to} \mathfrak{l}_2\,, \quad \mathfrak{l}_2 \underset{a}{\to} \mathfrak{l}_{3R}\,, \quad \mathfrak{l}_3 \underset{a}{\to} \mathfrak{l}_{1R}\,, \tag{7.41}$$

with the reflected $\mathfrak{su}(2)$ Lie algebra defined by (7.6). The corresponding Weyl group, defined in terms of transformations a, b satisfying (7.32), $W(\mathfrak{su}(3)) \simeq S_3$.

If we define

$$H_\perp = \tfrac{1}{\sqrt{3}}(H_1 + 2H_2)\,, \tag{7.42}$$

then the automorphism symmetries become

$$\big(H_1, H_\perp\big) \underset{b}{\to} \big(-H_1, H_\perp\big)\,, \qquad \big(H_1, H_\perp\big) \underset{a}{\to} \big(-\tfrac{1}{2}H_1 + \tfrac{\sqrt{3}}{2}H_\perp, -\tfrac{\sqrt{3}}{2}H_1 - \tfrac{1}{2}H_\perp\big)\,. \tag{7.43}$$

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 $2\pi/3$.

7.3 Highest Weight Representations for $\mathfrak{su}(3)$

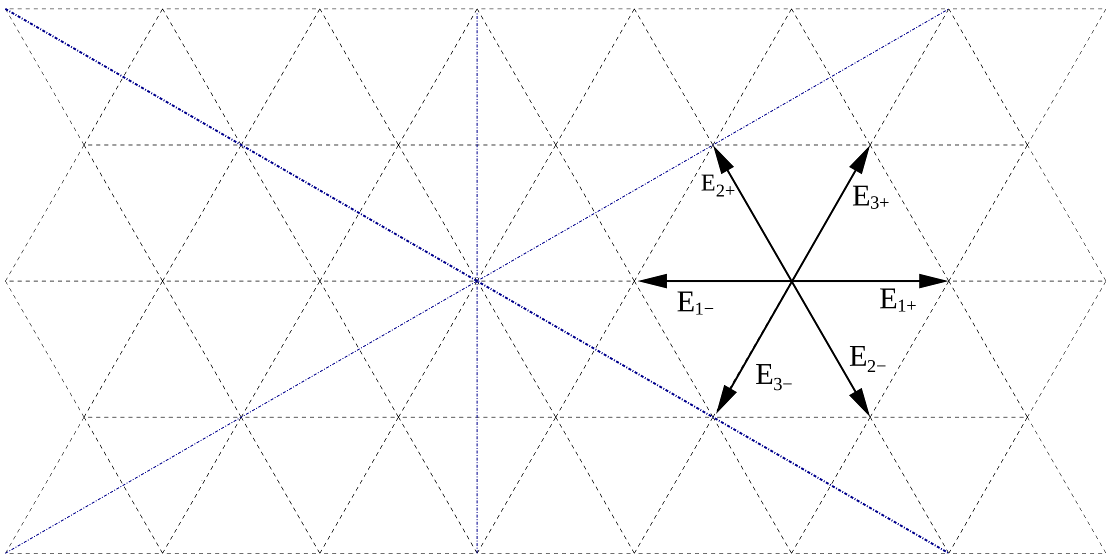
H_1, H_2 commute, (7.36), and a standard basis for the representation space for $\mathfrak{su}(3)$ is given by their simultaneous eigenvectors $|r_1, r_2\rangle$ where

$$H_1|r_1,r_2\rangle = r_1|r_1,r_2\rangle\,, \qquad H_2|r_1,r_2\rangle = r_2|r_1,r_2\rangle\,. \tag{7.44}$$

As a consequence of (7.38) we must then have

$$\begin{aligned} E_{1\pm}|r_1,r_2\rangle &\propto |r_1 \pm 2, r_2 \mp 1\rangle\,, \\ E_{2\pm}|r_1,r_2\rangle &\propto |r_1 \mp 1, r_2 \pm 2\rangle\,, \\ E_{3\pm}|r_1,r_2\rangle &\propto |r_1 \pm 1, r_2 \pm 1\rangle\,, \end{aligned} \tag{7.45}$$

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 (7.45), 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

$$H_i \underset{\sigma}{\to} H'_i\,, \quad H'_i|r_1,r_2\rangle = r'_i|r_1,r_2\rangle\,, \; i = 1,2 \quad \Rightarrow \quad [r'_1, r'_2] = \sigma[r_1,r_2]\,. \tag{7.46}$$

From (7.41) this is given by

$$\begin{aligned} &b[r_1,r_2] = [-r_1, r_1 + r_2]\,, \quad ab[r_1,r_2] = [r_1 + r_2, -r_2]\,, \quad a^2b[r_1,r_2] = [-r_2, -r_1]\,, \\ &a[r_1,r_2] = [r_2, -r_1 - r_2]\,, \qquad a^2[r_1,r_2] = [-r_1 - r_2, r_1]\,. \end{aligned} \tag{7.47}$$

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

$$E_{1+}|n_1,n_2\rangle_{\rm hw} = E_{2+}|n_1,n_2\rangle_{\rm hw} = 0 \quad \Rightarrow \quad E_{3+}|n_1,n_2\rangle_{\rm hw} = 0\,. \tag{7.48}$$

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 and may be defined by

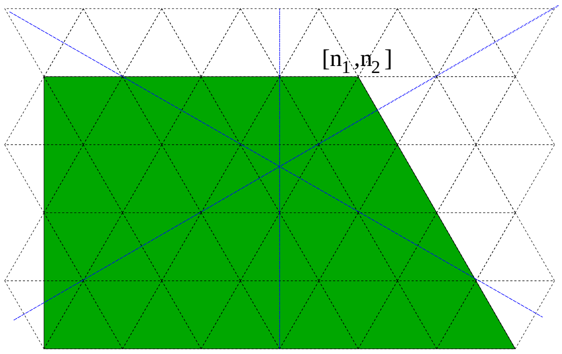
$$V_{[n_1,n_2]} = \operatorname{span}\big\{ {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} : r,s,t = 0,1,\dots \big\}\,. \tag{7.49}$$

The ordering of E_{1-},E_{2-},E_{3-} in the basis assumed in (7.49) 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 chosen basis in (7.49) using the commutation relations given by the conjugate of (7.37).

For these basis vectors

$$\begin{aligned} H_1\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} &= (n_1-2r+s-t)\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw}\,, \\ H_2\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} &= (n_2+r-2s-t)\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw}\,, \end{aligned} \tag{7.50}$$

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.



The representation of $\mathfrak{su}(3)$ is determined then in terms of the action of $E_{i\pm}$ on the basis (7.49). For the lowering operators it is easy to see that

$$\begin{aligned} E_{3-}\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} &= {E_{3-}}^{t+1} {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw}\,, \\ E_{2-}\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} &= {E_{3-}}^t {E_{2-}}^{s+1} {E_{1-}}^r |n_1,n_2\rangle_{\rm hw}\,, \\ E_{1-}\, {E_{3-}}^t {E_{2-}}^s {E_{1-}}^r |n_1,n_2\rangle_{\rm hw} &= {E_{3-}}^t {E_{2-}}^s {E_{1-}}^{r+1} |n_1,n_2\rangle_{\rm hw} \\ &\quad - s\, {E_{3-}}^{t+1} {E_{2-}}^{s-1} {E_{1-}}^r |n_1,n_2\rangle_{\rm hw}\,, \end{aligned} \tag{7.51}$$

using $\big[E_{1-},{E_{2-}}^s\big] = -s\,E_{3-}{E_{2-}}^{s-1}$.

The action of E_{i+} on the basis (7.49) may then be determined by using the basic commutation relations (7.39), with (7.38) and (7.37), and then applying (7.48). Just as in (7.12) we may obtain

$$\big[E_{1+},{E_{1-}}^r\big] = {E_{1-}}^{r-1}\, r(H_1-r+1)\,, \quad \big[E_{1+},{E_{2-}}^s\big] = 0\,, \quad \big[E_{1+},{E_{3-}}^t\big] = -t\,{E_{3-}}^{t-1}E_{2-}\,, \tag{7.52}$$

so that

$$E_{1+}\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1, n_2\rangle_{\rm hw}$$
$$= r(n_1 - r + 1)\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^{r-1} |n_1, n_2\rangle_{\rm hw} - t\, E_{3-}{}^{t-1} E_{2-}{}^{s+1} E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} \,. \qquad (7.53)$$

Similarly

$$\begin{aligned} \big[E_{2+}, E_{3-}{}^t\big] &= t\, E_{3-}{}^{t-1} E_{1-}\,, & \big[E_{1-}, E_{2-}{}^s\big] &= -s\, E_{3-} E_{2-}{}^{s-1}\,, \\ \big[E_{2+}, E_{2-}{}^s\big] &= E_{2-}{}^{s-1}\, s(H_2 - s + 1)\,, & \big[E_{2+}, E_{1-}{}^r\big] &= 0\,, \end{aligned} \qquad (7.54)$$

which leads to

$$E_{2+}\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1, n_2\rangle_{\rm hw}$$
$$= s(n_2 + r - s - t + 1)\, E_{3-}{}^t E_{2-}{}^{s-1} E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} + t\, E_{3-}{}^{t-1} E_{2-}{}^s E_{1-}{}^{r+1} |n_1, n_2\rangle_{\rm hw} \,. \qquad (7.55)$$

Furthermore

$$\begin{aligned} E_{3+}\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} &= [E_{1+}, E_{2+}]\, E_{3-}{}^t E_{2-}{}^s E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} \\ &= t(n_1 + n_2 - r - s - t + 1)\, E_{3-}{}^{t-1} E_{2-}{}^s E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} \\ &\quad + rs(n_1 - r + 1)\, E_{3-}{}^t E_{2-}{}^{s-1} E_{1-}{}^{r-1} |n_1, n_2\rangle_{\rm hw} \,. \end{aligned} \qquad (7.56)$$

The results (7.50), (7.51) with (7.53), (7.55) and (7.56) demonstrate how $V_{[n_1,n_2]}$ forms a representation space for $\mathfrak{su}(3)$ which is in general infinite dimensional.

The space $V_{[n_1,n_2]}$ defines a reducible representation of $\mathfrak{su}(3)$ when it contains vectors which satisfy the highest weight condition (7.48) since these generate invariant subspaces. Highest weight vectors may be constructed in $V_{[n_1,n_2]}$ in a similar fashion to the discussion for $\mathfrak{su}(2)$. Using, as a special case of (7.53) and (7.55),

$$\begin{aligned} E_{1+} E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} &= r(n_1 - r + 1) E_{1-}{}^{r-1} |n_1, n_2\rangle_{\rm hw}\,, & E_{2+} E_{1-}{}^r |n_1, n_2\rangle_{\rm hw} &= 0\,, \\ E_{2+} E_{2-}{}^s |n_1, n_2\rangle_{\rm hw} &= s(n_1 - s + 1) E_{2-}{}^{s-1} |n_1, n_2\rangle_{\rm hw}\,, & E_{1+} E_{2-}{}^s |n_1, n_2\rangle_{\rm hw} &= 0\,, \end{aligned} \qquad (7.57)$$

then for n_1, n_2 positive integers

$$\begin{aligned} |-n_1 - 2, n_1 + n_2 + 1\rangle_{\rm hw} &= E_{1-}{}^{n_1+1} |n_1, n_2\rangle_{\rm hw}\,, \\ |n_1 + n_2 + 1, -n_2 - 2\rangle_{\rm hw} &= E_{2-}{}^{n_2+1} |n_1, n_2\rangle_{\rm hw}\,, \end{aligned} \qquad (7.58)$$

satisfy the necessary conditions (7.48). Using the highest weight vectors obtained in (7.58) we may further obtain, for n_1, n_2 positive integers, two more highest weight vectors

$$\begin{aligned} |n_2, -n_1 - n_2 - 3\rangle_{\rm hw} &= E_{2-}{}^{n_2+n_1+2} |-n_1 - 2, n_1 + n_2 + 1\rangle_{\rm hw}\,, \\ |-n_1 - n_2 - 3, n_1\rangle_{\rm hw} &= E_{1-}{}^{n_2+n_1+2} |n_1 + n_2 + 1, -n_2 - 2\rangle_{\rm hw}\,. \end{aligned} \qquad (7.59)$$

This construction may also be applied to the highest weight vectors in (7.59) giving one further highest weight vector

$$|-n_2 - 2, -n_1 - 2\rangle_{\rm hw} = E_{1-}{}^{n_2+1} |n_2, -n_1 - n_2 - 3\rangle_{\rm hw} = E_{2-}{}^{n_1+1} |-n_1 - n_2 - 3, n_1\rangle_{\rm hw}\,. \qquad (7.60)$$

The highest weight vectors determined in (7.58), (7.59), (7.60) are non degenerate, in (7.60) this depends on the identity[50]

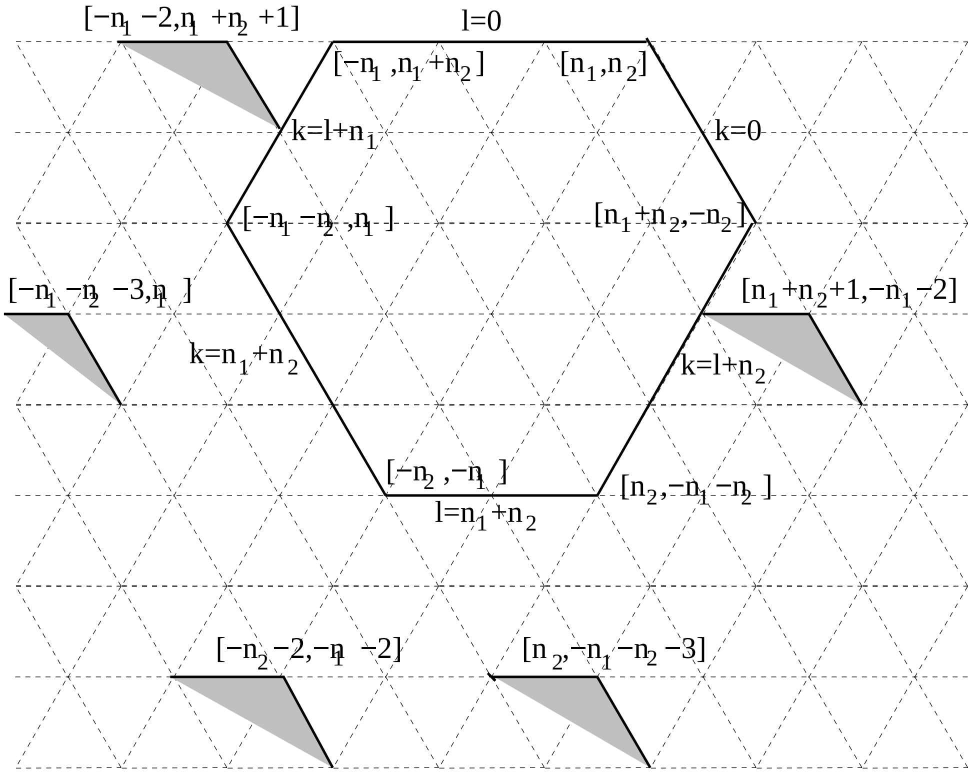
$$E_{1-}{}^{n_2+1} E_{2-}{}^{n_1+n_2+2} E_{1-}{}^{n_1+1} = E_{2-}{}^{n_1+1} E_{1-}{}^{n_1+n_2+2} E_{2-}{}^{n_2+1} \, . \tag{7.61}$$

Clearly this construction of highest weight vectors terminates with (7.60).

For each of highest weight vectors given in (7.58), (7.59) and (7.60), $|n'_1, n'_2\rangle_{\rm hw}$, there are associated invariant, under the action of $\mathfrak{su}(3)$, subspaces $V_{[n'_1,n'_2]}$, constructed as in (7.49), and contained in $V_{[n_1,n_2]}$. In particular

$$\begin{aligned}
& V_{[-n_1-2,\, n_1+n_2+1]} \, , \ V_{[n_1+n_2+1,-n_2-2]} \subset V_{[n_1,n_2]} \, , \\
& V_{[n_2,-n_1-n_2-3]} \subset V_{[-n_1-2,\, n_1+n_2+1]} \, , \ V_{[-n_1-n_2-3,\, n_1]} \subset V_{[n_1+n_2+1,-n_2-2]} \, , \\
& V_{[-n_2-2,-n_1-2]} \subset V_{[n_2,-n_1-n_2-3]} \cap V_{[-n_1-n_2-3,\, n_1]} \, .
\end{aligned} \tag{7.62}$$

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 of $V_{[n_1,n_2]}$ to an irreducible representation space becomes less trivial than that given by (7.16) for $\mathfrak{su}(2)$ due to this nested structure of invariant subspaces. Using the

[50]This may be shown using the identity $E_{1-}{}^r E_{2-}{}^s = \sum_{t=0}^{r,s} (-1)^t \binom{r}{t} \frac{s!}{(s-t)!} \, E_{3-}{}^t E_{2-}{}^{s-t} E_{1-}{}^{r-t}$, both sides of (7.61) give rise to the same expansion in $E_{3-}{}^t E_{2-}{}^{n_1+n_2+2-t} E_{1-}{}^{n_1+n_2+2-t}$.

same definition of the quotient of a vector space by a subspace as in (7.16) we may define

$$\begin{aligned}\mathcal{V}^{(2)}_{[n_1,n_2]} &= \big(V_{[n_2,-n_1-n_2-3]} \oplus V_{[-n_1-n_2-3,\,n_1]}\big)\big/\, V_{[-n_2-2,-n_1-2]}\,,\\ \mathcal{V}^{(1)}_{[n_1,n_2]} &= \big(V_{[-n_1-2,\,n_1+n_2+1]} \oplus V_{[n_1+n_2+1,-n_2-2]}\big)\big/\,\mathcal{V}^{(2)}_{[n_1,n_2]}\,,\\ \mathcal{V}_{[n_1,n_2]} &= V_{[n_1,n_2]}\big/\,\mathcal{V}^{(1)}_{[n_1,n_2]}\,. \end{aligned} \tag{7.63}$$

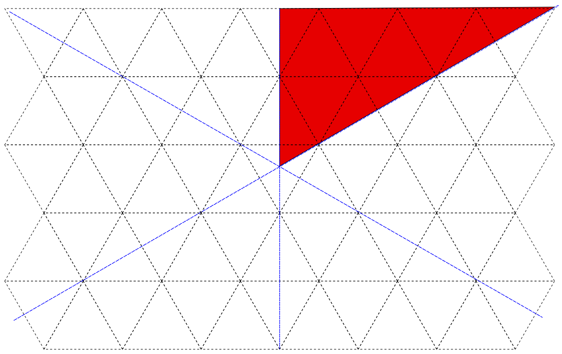
In $\mathcal{V}_{[n_1,n_2]}$ there then are no highest weight vectors other than $|n_1,n_2\rangle_{\rm hw}$ so invariant subspaces are absent and $\mathcal{V}_{[n_1,n_2]}$ is a representation space for an irreducible representation of $\mathfrak{su}(3)$. Although it remains to be demonstrated the representation space is then finite-dimensional and the corresponding weight diagram has vertices with weights

$$[n_1,n_2]\,,\quad [-n_1,n_1+n_2]\,,\quad [n_1+n_2,-n_2]\,,\quad [-n_1-n_2,n_1]\,,\quad [n_2,-n_1-n_2]\,,\quad [-n_2,-n_1]\,, \tag{7.64}$$

which are related by the transformations of the Weyl group as in (7.47). The sector of the weight diagram corresponding to highest weight states forming finite dimensional representations is then

$$\mathcal{W} = \{[m,n] : m,n \in \mathbb{N}_0\}\,, \tag{7.65}$$

which is illustrated by



7.3.1 Analysis of the Weight Diagram

It is clear that the construction (7.49) 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 $[r_1,r_2]$, (7.49) there is a finite dimensional subspace contained in $V_{[n_1,n_2]}$ given by

$$V^{(k,l)}_{[n_1,n_2]} = \operatorname{span}\big\{ {E_{3-}}^{t} {E_{2-}}^{l-t} {E_{1-}}^{k-t} |n_1,n_2\rangle_{\rm hw} : 0 \le t \le k,l \big\}\,, \tag{7.66}$$

where

$$k = \tfrac{1}{3}\big(2n_1+n_2-2r_1-r_2\big)\,, \qquad l = \tfrac{1}{3}\big(n_1+2n_2-r_1-2r_2\big)\,. \tag{7.67}$$

Clearly

$$\dim V^{(k,l)}_{[n_1,n_2]} = \begin{cases} k+1\,, & k\le l\,,\\ l+1\,, & l\le k\,. \end{cases} \tag{7.68}$$

To show how (7.63) leads to a finite-dimensional representation we consider how it applies to for the vectors corresponding to particular individual weights $[r_1,r_2]$. In a similar fashion to (7.66) we may define

$$\begin{aligned}
&V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]}\,, \qquad V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]}\,,\\
&V^{(k-n_1-1,l-n_1-n_2-2)}_{[n_2,-n_1-n_2-3]}\,, \qquad V^{(k-n_1-n_2-2,l-n_2-1)}_{[-n_1-n_2-3,n_1]}\,, \qquad V^{(k-n_1-n_2-2,l-n_1-n_2-2)}_{[-n_2-2,-n_1-2]}\,,
\end{aligned} \tag{7.69}$$

which form nested subspaces, just as in (7.62), and whose dimensions are given by the obvious extension of (7.68).

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 (7.64) 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 $V^{(k,l)}_{[n_1,n_2]}$. Taking into account the contributions from $V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]}$ and $V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]}$ gives

$$\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} = \begin{cases} 0 & \text{if} \quad k \ge l+n_1+1,\ l \ge 0\,,\\ 1 & \text{if} \quad k = l+n_1,\ l\ge 0\,, \end{cases} \tag{7.70}$$

and

$$\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} = \begin{cases} 0 & \text{if} \quad l \ge k+n_2+1,\ k \ge 0\,,\\ 1 & \text{if} \quad l = k+n_2,\ k\ge 0\,. \end{cases} \tag{7.71}$$

Furthermore

$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]}\\
&= \begin{cases} l+1-n_2-(l-n_2) = 1\,, & k = n_1+n_2,\ n_2 \le l \le n_1+n_2\,,\\ k+1-(k-n_1)-n_1 = 1\,, & l = n_1+n_2,\ n_1 \le k \le n_1+n_2\,. \end{cases}
\end{aligned} \tag{7.72}$$

The remaining contributions, when present, give rise to a complete cancellation so that the representation space given by (7.63) is finite dimensional. When $l \ge n_2$, $k \ge n_1+n_2+1$,

$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} + \dim V^{(k-n_1-n_2-2,l-n_2-1)}_{[-n_1-n_2-3,n_1]}\\
&= \begin{cases} (l+1)-(l+1)-(l-n_2)+(l-n_2)\,, & k \ge l+n_1+1\\ (l+1)-(k-n_1)-(l-n_2)+(k-n_1-n_2-1)\,, & l \le k \le l+n_1+1 \end{cases}\\
&= 0\,,
\end{aligned} \tag{7.73}$$

and for $k,l \ge n_1+n_2+1$ in an analogous fashion

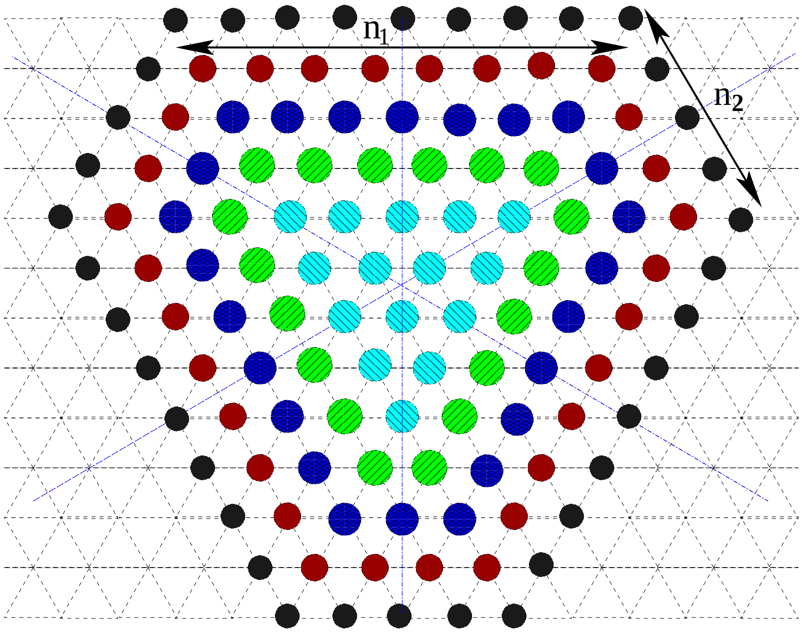
$$\begin{aligned}
&\dim V^{(k,l)}_{[n_1,n_2]} - \dim V^{(k-n_1-1,l)}_{[-n_1-2,n_1+n_2+1]} - \dim V^{(k,l-n_2-1)}_{[n_1+n_2+1,-n_2-2]} + \dim V^{(k-n_1-1,l-n_1-n_2-2)}_{[n_2,-n_1-n_2-3]}\\
&+ \dim V^{(k-n_1-n_2-2,l-n_2-1)}_{[-n_1-n_2-3,n_1]} - \dim V^{(k-n_1-n_2-2,l-n_1-n_2-2)}_{[-n_2-2,-n_1-2]} = 0\,.
\end{aligned} \tag{7.74}$$

For the finite representation space $\mathcal{V}_{[n_1,n_2]}$ then at each vertex of the weight diagram as in (7.64) there are associated vectors which satisfy analogous conditions to (7.48), in particular

$$\begin{aligned}
&(E_{1-}, E_{3+})\,|{-n_1}, n_1+n_2\rangle = 0\,, \qquad && (E_{2-}, E_{3+})\,|n_1+n_2, -n_2\rangle\,,\\
&(E_{2+}, E_{3-})\,|{-n_1}-n_2, n_1\rangle = 0\,, \qquad && (E_{1+}, E_{3-})\,|n_2, -n_1-n_2\rangle = 0\,,\\
&\quad (E_{1-}, E_{2-})\,|{-n_2}, -n_1\rangle = 0\,.
\end{aligned} \tag{7.75}$$

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 $n_1 \geq n_2$ 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.



7.3.2 $SU(3)$ 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 (7.21) and (7.22). For the highest weight representation space $V_{[n_1,n_2]}$ we then define in terms of the basis (7.49)

$$C_{[n_1,n_2]}(t_1,t_2) = \widetilde{\mathrm{tr}}_{V_{[n_1,n_2]}}\big({t_1}^{H_1}\,{t_2}^{H_2}\big) = \sum_{r,s,t\geq 0} {t_1}^{n_1-2r+s-t}\,{t_2}^{n_2+r-2s-t}$$
$$= {t_1}^{n_1}\,{t_2}^{n_2} \sum_{r,s,t\geq 0} \big(t_2/{t_1}^2\big)^r \big(t_1/{t_2}^2\big)^s \big(1/t_1t_2\big)^t\,. \tag{7.76}$$

For a succinct final expression it is more convenient to use the variables

$$u = (u_1,u_2,u_3)\,, \quad u_1 = t_1\,, \quad u_3 = 1/t_2\,, \quad u_1u_2u_3 = 1\,, \tag{7.77}$$

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

$$C_{[n_1,n_2]}(u) = \frac{{u_1}^{n_1+n_2+2}\,{u_2}^{n_2+1}}{(u_1-u_2)(u_2-u_3)(u_1-u_3)}\,. \tag{7.78}$$

Following (7.63) the character for the irreducible representation of $\mathfrak{su}(3)$ obtained from the highest weight vector $|n_1,n_2\rangle_{\mathrm{hw}}$ is then

$$\begin{aligned}\chi_{[n_1,n_2]}(u) &= C_{[n_1,n_2]}(u) - C_{[-n_1-2,n_1+n_2+1]}(u) - C_{[n_1+n_2+1,-n_2-2]}(u)\\ &\quad + C_{[-n_1-n_2-3,n_1]}(u) + C_{[n_2,-n_1-n_2-3]}(u) - C_{[-n_2-2,-n_1-2]}(u)\\ &= \frac{1}{(u_1-u_2)(u_2-u_3)(u_1-u_3)}\\ &\qquad \times \big({u_1}^{n_1+n_2+2}\,{u_2}^{n_2+1} - {u_2}^{n_1+n_2+2}\,{u_1}^{n_2+1} - {u_1}^{n_1+n_2+2}\,{u_3}^{n_2+1}\\ &\qquad\quad + {u_2}^{n_1+n_2+2}\,{u_3}^{n_2+1} - {u_3}^{n_1+n_2+2}\,{u_2}^{n_2+1} + {u_3}^{n_1+n_2+2}\,{u_1}^{n_2+1}\big)\,.\end{aligned} \tag{7.79}$$

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

$$\chi_{[n_1,n_2]}(q,1,q^{-1}) = \frac{1-q^{n_1+1}}{1-q}\,\frac{1-q^{n_2+1}}{1-q}\,\frac{1-q^{-n_1-n_2-2}}{1-q^{-2}}\,, \tag{7.80}$$

and hence it is then easy to calculate

$$\dim \mathcal{V}_{[n_1,n_2]} = \chi_{[n_1,n_2]}(1,1,1) = \tfrac{1}{2}(n_1+1)(n_2+1)(n_1+n_2+2)\,. \tag{7.81}$$

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

$$[r_1,r_2]^\sigma = \sigma[r_1+1,r_2+1] - [1,1]\,. \tag{7.82}$$

Directly from (7.47) we easily obtain

$$\begin{aligned} &[r_1,r_2]^b = [-r_1-2,r_1+r_2+1]\,, \qquad [r_1,r_2]^{ab} = [r_1+r_2+1,-r_2-2]\,,\\ &[r_1,r_2]^a = [r_2,-r_1-r_2-3]\,, \qquad [r_1,r_2]^{a^2} = [-r_1-r_2-3,r_1]\,,\\ &[r_1,r_2]^{a^2b} = [-r_2-2,-r_1-2]\,.\end{aligned} \tag{7.83}$$

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 (7.58), (7.59) and (7.60). Thus (7.79) may be written more concisely as

$$\chi_{[n_1,n_2]}(u) = \sum_{\sigma\in S_3} P_\sigma\, C_{[n_1,n_2]^\sigma}(u) = \sum_{\sigma\in S_3} C_{[n_1,n_2]}(\sigma u)\,, \tag{7.84}$$

with, for $\sigma \in S_3$,

$$P_\sigma = \begin{cases} -1\,, & \sigma \text{ odd permutation}\,, \\ 1\,, & \sigma \text{ even permutation}\,, \end{cases} \tag{7.85}$$

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

$$\chi_{[n_1,n_2]^\sigma}(u) = P_\sigma \chi_{[n_1,n_2]}(u)\,. \tag{7.86}$$

Since $[-1,r]^b = [-1,r]$, $[r,-1]^{ab} = [r,-1]$ and $[r,-r-2]^{a^2b} = [r,-r-2]$ we must then have

$$\chi_{[-1,r]}(u) = \chi_{[r,-1]}(u) = \chi_{[r,-r-2]}(u) = 0\,. \tag{7.87}$$

This shows the necessity of the three factors in the dimension formula (7.81). It is important to note that for any $[n_1,n_2]$

$$n_1,n_2 \neq -1,\ n_1+n_2 \neq -2\,, \qquad [n_1,n_2]^\sigma \in \mathcal{W} \quad \text{for a unique} \quad \sigma \in S_3\,, \tag{7.88}$$

where $\mathcal{W}$ is defined in (7.65).

7.3.3 Casimir operator

For the basis in (7.33) the $\mathfrak{su}(3)$ quadratic Casimir operator is given by

$$\begin{aligned} C = \hat{R}^i{}_j\, \hat{R}^j{}_i &= \textstyle\sum_{i=1}^3 \big(E_{i+}E_{i-} + E_{i-}E_{i+}\big) + \tfrac{2}{3}\big({H_1}^2 + {H_2}^2 + H_1H_2\big) \\ &= \textstyle\sum_{i=1}^3 E_{i-}E_{i+} + \tfrac{2}{3}\big({H_1}^2 + {H_2}^2 + H_1H_2\big) + 2\big(H_1+H_2\big)\,. \end{aligned} \tag{7.89}$$

Acting on a highest weight vector

$$C|n_1,n_2\rangle_{\rm hw} = c_{[n_1,n_2]}|n_1,n_2\rangle_{\rm hw}\,, \tag{7.90}$$

where, from the explicit form in (7.89),

$$c_{[n_1,n_2]} = \tfrac{2}{3}({n_1}^2 + {n_2}^2 + n_1n_2) + 2(n_1+n_2)\,. \tag{7.91}$$

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 $V_{[n_1,n_2]}$.

7.3.4 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 (7.63) 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 (7.40) 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 (7.20) the highest weight vector must satisfy

$$E_{1-}{}^{n_1+1}|n_1,0\rangle_{\rm hw} = 0\,, \qquad E_{2-}|n_1,0\rangle_{\rm hw} = 0\,. \tag{7.92}$$

Furthermore, using $[E_{3+}, E_{1-}{}^r] = -r E_{1-}{}^{r-1}E_{2+}$,

$$E_{3+}\,E_{1-}{}^r|n_1,0\rangle_{\rm hw} = 0\,, \qquad (H_1+H_2)\,E_{1-}{}^r|n_1,0\rangle_{\rm hw} = (n_1-r)\,E_{1-}{}^r|n_1,0\rangle_{\rm hw}\,, \tag{7.93}$$

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 (7.20) again

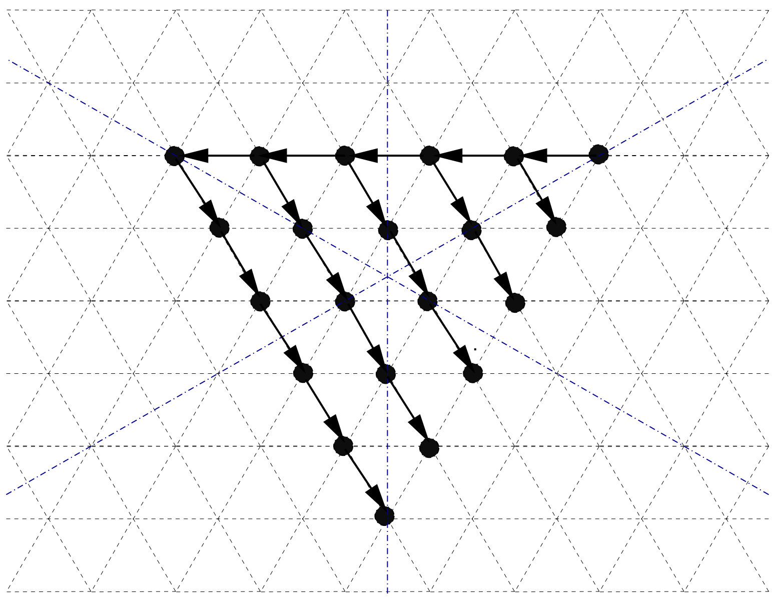
$$E_{3-}{}^{n_1-r+1}\,E_{1-}{}^r|n_1,0\rangle_{\rm hw} = 0\,. \tag{7.94}$$

Hence a finite dimensional basis for $\mathcal{V}_{[n_1,0]}$ is given by

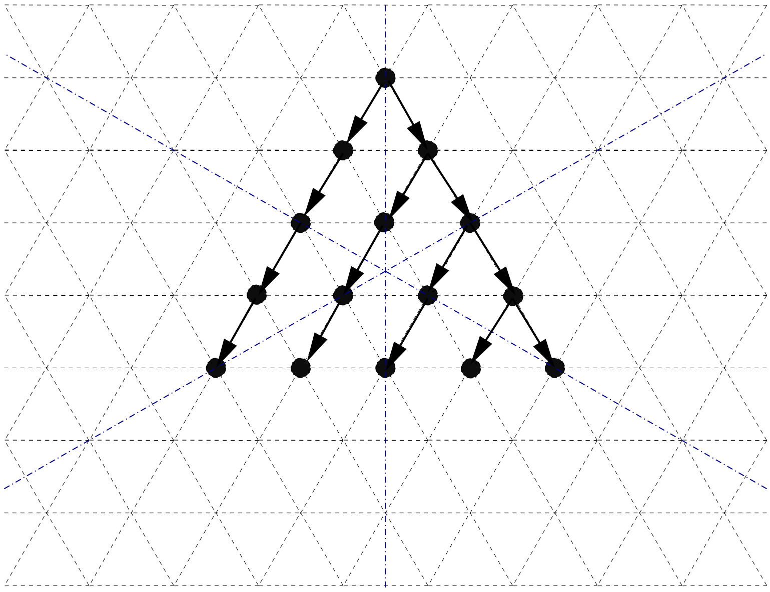
$$E_{3-}{}^tE_{1-}{}^r|n_1,0\rangle_{\rm hw}\,, \qquad t=0,\dots n_1-r\,,\ r=0,\dots,n_1\,, \tag{7.95}$$

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 (7.20) for the highest weight state require

$$ {E_{1-}}^2|1,1\rangle_{\rm hw} = {E_{2-}}^2|1,1\rangle_{\rm hw} = {E_{3-}}^3|1,1\rangle_{\rm hw} = 0\,. \tag{7.96} $$

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 (7.49) are then restricted to just

$$ \begin{aligned} &[-1,2]: \ E_{1-}|1,1\rangle_{\rm hw}\,, \quad [2,-1]: \ E_{2-}|1,1\rangle_{\rm hw}\,, \quad [0,0]: \ E_{3-}|1,1\rangle_{\rm hw}\,, \ E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,, \\ &[-2,1]: \ E_{3-}E_{1-}|1,1\rangle_{\rm hw}\,, \qquad [1,0]: \ E_{3-}E_{2-}|1,1\rangle_{\rm hw}\,, \ {E_{2-}}^2E_{1-}|1,1\rangle_{\rm hw}\,, \\ &[-1,-1]: {E_{3-}}^2|1,1\rangle_{\rm hw}\,, \ E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,. \end{aligned} \tag{7.97} $$

However (7.96) requires further relations since

$$ {E_{2-}}^2E_{1-}|1,1\rangle_{\rm hw} = \big(E_{2-}E_{1-}E_{2-} + E_{3-}E_{2-}\big)|1,1\rangle_{\rm hw} = 2\,E_{3-}E_{2-}|1,1\rangle_{\rm hw}\,, \tag{7.98} $$

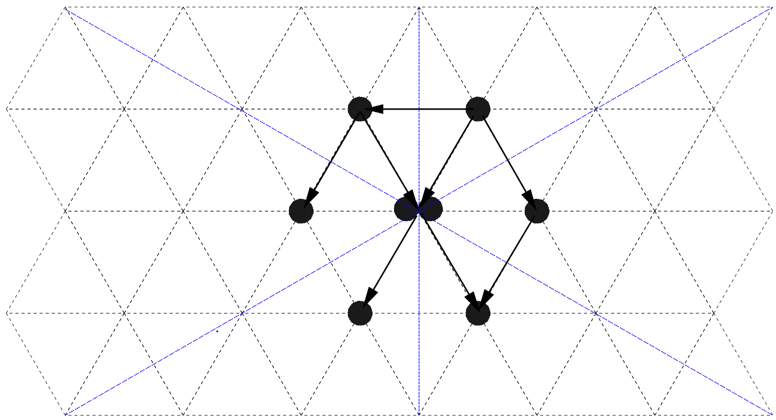
which then entails

$$ \begin{aligned} &E_{1-}{E_{2-}}^2E_{1-}|1,1\rangle_{\rm hw} = -2\,E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw} \\ &= 2\,E_{1-}E_{3-}E_{2-}|1,1\rangle_{\rm hw} = 2\big(E_{3-}E_{2-}E_{1-} - {E_{3-}}^2\big)|1,1\rangle_{\rm hw}\,, \end{aligned} \tag{7.99} $$

so that furthermore

$$E_{3-}{}^2|1,1\rangle_{\rm hw} = 2\,E_{3-}E_{2-}E_{1-}|1,1\rangle_{\rm hw}\,. \tag{7.100}$$

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.



7.4 $SU(3)$ 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

$$q^i \to A^i{}_j\, q^j\,, \qquad \bar{q}_i \to \bar{q}_j\,(A^{-1})^i{}_j\,, \qquad \big[A^j{}_i\big] \in SU(3)\,. \tag{7.101}$$

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

$$T^{i_1\dots i_r}_{j_1\dots j_s} \to A^{i_1}{}_{k_1}\dots A^{i_r}{}_{k_r}\,T^{k_1\dots k_r}_{l_1\dots l_s}\,(A^{-1})^{l_1}{}_{j_1}\dots(A^{-1})^{l_s}{}_{j_s}\,. \tag{7.102}$$

The conjugate of a (r,s)-tensor is a (s,r)-tensor

$$\bar{T}^{j_1\dots j_s}_{i_1\dots i_r} = \big(T^{i_1\dots i_r}_{j_1\dots j_s}\big)^*\,. \tag{7.103}$$

The invariant tensors are a natural extension of those for $SU(2)$, as exhibited in (3.287) and (3.288). 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,

$$\varepsilon^{ijk}\,, \qquad \varepsilon_{ijk}\,, \qquad \delta^i_j\,. \tag{7.104}$$

That ε^{ijk} and ε_{ijk} are invariant tensors is a consequence of the transformation matrix A satisfying $\det A = 1$. The transformation rules (7.102) guarantee that the contraction of an upper and lower index maintains the tensorial transformation properties. In consequence from a tensor $T^{i_1\ldots i_r}_{j_1\ldots j_s}$ then contracting with $\varepsilon^{ij_m j_n}$ or $\varepsilon_{ji_m i_n}$, 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,

$$S^{i_1\ldots i_r}_{j_1\ldots j_s} = S^{(i_1\ldots i_r)}_{(j_1\ldots j_s)}\,, \qquad S^{i_1\ldots i_{r-1}\,i}_{j_1\ldots j_{s-1}\,i} = 0\,. \tag{7.105}$$

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 (3.220) 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

$$\begin{aligned}\dim \mathcal{V}_{[r,s]} &= \tfrac{1}{2}(r+1)(r+2)\,\tfrac{1}{2}(s+1)(s+2) - \tfrac{1}{2}r(r+1)\,\tfrac{1}{2}s(s+1)\\ &= \tfrac{1}{2}(r+1)(s+1)(r+s+2)\,.\end{aligned} \tag{7.106}$$

This is of course identical to (7.81). 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.

7.4.1 $\mathfrak{su}(3)$ 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 (3.19), the Gell-Mann λ-matrices λ_a, $a = 1,\ldots,8$,

$$\begin{gathered}\lambda_1 = \begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}, \quad \lambda_2 = \begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix}, \quad \lambda_3 = \begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix},\\ \lambda_4 = \begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}, \quad \lambda_5 = \begin{pmatrix}0&0&-i\\0&0&0\\i&0&0\end{pmatrix}, \quad \lambda_6 = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}, \quad \lambda_7 = \begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix}.\end{gathered} \tag{7.107}$$

These satisfy

$$\mathrm{tr}(\lambda_a \lambda_b) = 2\,\delta_{ab}\,, \tag{7.108}$$

and

$$[\lambda_a, \lambda_b] = 2i\,f_{abc}\lambda_c\,, \tag{7.109}$$

for totally antisymmetric structure constants, f_{abc}. In terms of the matrices defined in (7.27) and (7.29) 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 (3.38). (3.23) needs only straightforward modification while instead of (3.20) we now have

$$\lambda_a \lambda_b = \tfrac{2}{3}\,\mathbb{1} + d_{abc}\,\lambda_c + i f_{abc}\,\lambda_c\,, \tag{7.110}$$

with d_{abc} totally symmetric and satisfying $d_{abb} = 0$.

7.5 $SU(3)$ and Physics

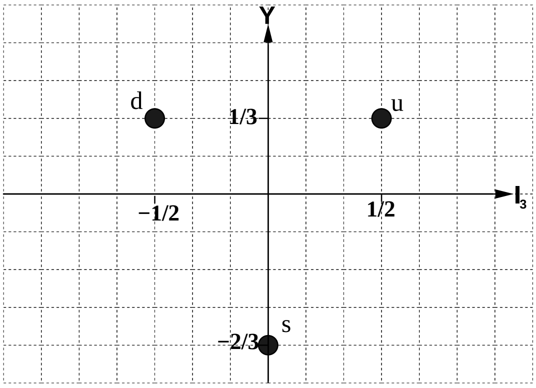
Besides its virtues in terms of understanding more general Lie groups a major motivation in studying $SU(3)$ is in terms of its role in physics. Historically $SU(3)$ was introduced, as a generalisation of the isospin $SU(2)_I$, to be an approximate symmetry group for strong interactions, in current terminology 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.

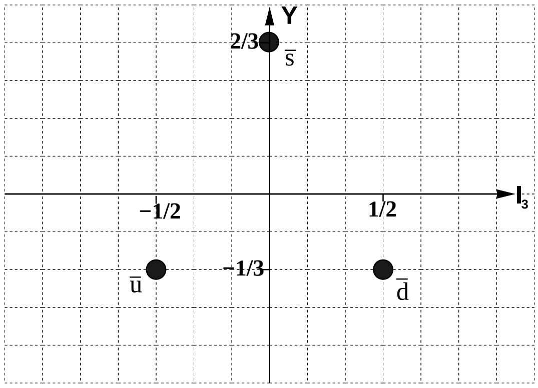
For $SU(3)_F$ the two commuting generators are identified with I_3, belonging to $SU(2)_I$, and also the hypercharge Y, where $[I_i, Y] = 0$ so that Y takes the same value for any isospin multiplet. Y is related to strangeness S, a quantum number invented to explain why the newly discovered, in the 1940's, 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 (7.107) is adopted then these are F_a, $a = 1, \ldots, 8$, where F_a are hermitian, and

$$[F_a, F_b] = i f_{abc} F_c\,, \qquad F_i = I_i\,, \quad i = 1,2,3\,, \qquad F_8 = \tfrac{1}{\sqrt{3}} Y\,. \tag{7.111}$$

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

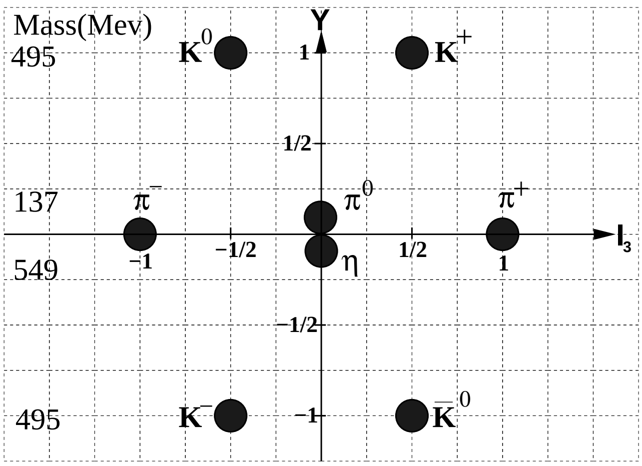




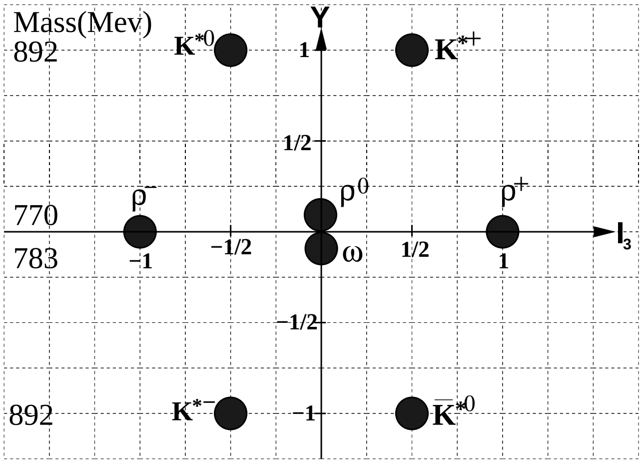
The charges of quarks are dictated by the requirement $Q = I_3 + \frac{1}{2}Y$ and so for q are fractional, $\frac{2}{3}$ and $-\frac{1}{3}$, while for $\bar{q}$ they are the opposite sign. We may further interpret the quantum numbers in terms of the numbers of particular quarks minus their anti-quarks, 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 $SU(3)/\mathbb{Z}_3$.

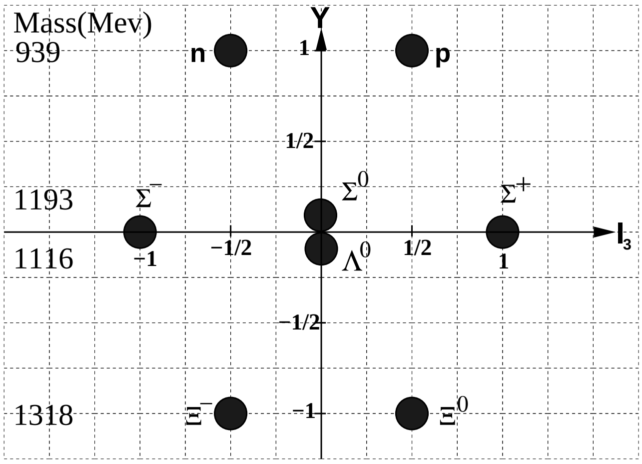
For the mesons we have self-conjugate octets belonging to the $[1,1]$, or $\mathbf{8}$, $SU(3)$ representations. The weight diagram for the lightest spin-0 negative parity mesons is



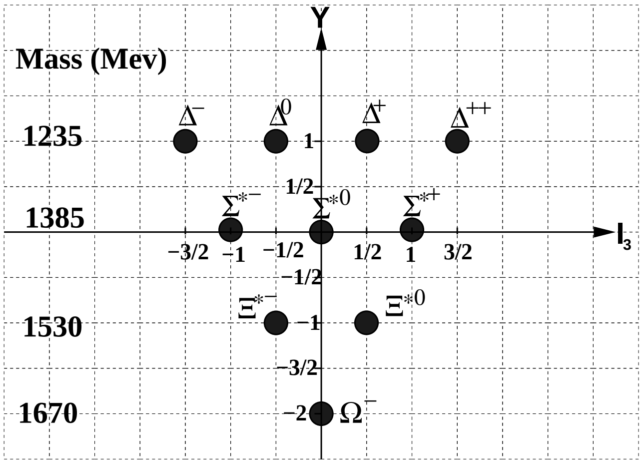
Here the kaons K^+, K^0 and $\bar{K}^0, K^-$ are $I = \frac{1}{2}$ strange particles with $S = 1$ and $S = -1$. A similar pattern emerges for the next lightest spin one negative parity mesons.



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.

7.5.1 $SU(3)_F$ 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

$$\begin{aligned}\mathcal{L}_m &= -m_u\,\bar{u}u - m_d\,\bar{d}d - m_s\,\bar{s}s \\ &= -\bar{m}\,\bar{q}q - \tfrac{1}{2}(m_u - m_d)\,\bar{q}\lambda_3 q - \tfrac{1}{2\sqrt{3}}(m_u + m_d - 2m_s)\,\bar{q}\lambda_8 q\,,\end{aligned} \tag{7.112}$$

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

$$H = H_0 + T_8\,, \tag{7.113}$$

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 (7.111), we have the commutation relations $[F_a, H_0] = 0$ and $[F_a, T_b] = if_{abc}T_c$. The Hamiltonian in (7.113) 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

$$m_{I,Y} = m_0 + \langle II_3, Y|T_8|II_3, Y\rangle\,. \tag{7.114}$$

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 3.13, 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 (7.33),

$$\big[\hat{R}^i{}_j, T^k{}_l\big] = \delta^k{}_j T^i{}_l - \delta^i_l T^k{}_j \,, \qquad T_8 = \tfrac{1}{3}\big(T^1{}_1 + T^2{}_2 - 2T^3{}_3\big) \,. \tag{7.115}$$

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 (7.115) requires

$$T_8 = \tfrac{1}{3}\big(H_1 + 2H_2\big) = Y \,, \tag{7.116}$$

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

$$T_8 = \tfrac{1}{4}\big(\hat{R}^1{}_k\hat{R}^k{}_1 + \hat{R}^k{}_1\hat{R}^1{}_k + \hat{R}^2{}_k\hat{R}^k{}_2 + \hat{R}^k{}_2\hat{R}^2{}_k - \hat{R}^3{}_k\hat{R}^k{}_3 - \hat{R}^k{}_3\hat{R}^3{}_k\big) - \tfrac{1}{6}\,C \,, \tag{7.117}$$

where C is the $SU(3)$ Casimir operator defined in (7.89). Using (7.33) then

$$\begin{aligned} T_8 &= \tfrac{1}{2}\big(E_{1+}E_{1-} + E_{1-}E_{1+} + \tfrac{1}{2}{H_1}^2\big) - \tfrac{1}{36}\big(H_1 + 2H_2\big)^2 - \tfrac{1}{6}\,C \\ &= I_iI_i - \tfrac{1}{4}\,Y^2 - \tfrac{1}{6}\,C \,, \end{aligned} \tag{7.118}$$

with I_i the isospin operators and (7.25) 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 (7.116) and (7.118). 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

$$m_{IY} = m_0 + aY + b\big(I(I+1) - \tfrac{1}{4}Y^2\big) \,, \tag{7.119}$$

with a, b undetermined coefficients.

For the baryon octet (7.119) 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 $4m_K^2 = 3m_\pi^2 + m_\eta^2$.

7.5.2 $SU(3)$ 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 $SU(3)_{\text{colour}}$ 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

$$\bar{q}_{1r} q_2{}^r \,, \qquad \varepsilon_{rst}\, q_1{}^r q_2{}^s q_3{}^t \,. \tag{7.120}$$

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.

7.6 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 7.4 with irreducible representations characterised by symmetric traceless tensors as in (7.105). Thus for the product of two fundamental representations it is sufficient to express it in terms of its symmetric and antisymmetric parts

$$q_1{}^i\, q_2{}^j = S^{ij} + \varepsilon^{ijk}\bar{q}_k \,, \qquad S^{ij} = q_1{}^{(i} q_2{}^{j)} \,, \quad \bar{q}_k = \tfrac{1}{2}\varepsilon_{kij} q_1{}^i q_2{}^j \,. \tag{7.121}$$

while for the product of the fundamental and its conjugate it is only necessary to separate out the trace

$$\bar{q}_i q^j = M_i{}^j + \delta_i^j S \,, \qquad M_i{}^j = \bar{q}_i q^j - \tfrac{1}{3}\delta_i^j\, \bar{q}_k q^k \,, \quad S = \tfrac{1}{3}\, \bar{q}_i q^i \,. \tag{7.122}$$

These correspond respectively to

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \mathbf{3}^* \,, \qquad \mathbf{3}^* \otimes \mathbf{3} = \mathbf{8} \oplus \mathbf{1} \,. \tag{7.123}$$

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

$$\begin{aligned} q_1{}^i\, q_2{}^j\, q_3{}^k &= D^{ijk} + \varepsilon^{ikl} B_l{}^j + \varepsilon^{jkl} B_l{}^i + \varepsilon^{ijl} B'_l{}^k + \varepsilon^{ijk} S \,, \\ D^{ijk} &= q_1{}^{(i}\, q_2{}^j\, q_3{}^{k)} \,, \qquad\qquad S = \tfrac{1}{6}\varepsilon_{ijk}\, q_1{}^i\, q_2{}^j\, q_3{}^k \,, \\ B_l{}^i &= \tfrac{1}{3}\varepsilon_{jkl}\, q_1{}^{(i}\, q_2{}^{j)}\, q_3{}^k \,, \quad B'_l{}^k = \tfrac{1}{2}\varepsilon_{ijl}\, q_1{}^i\, q_2{}^j\, q_3{}^k - \delta_l^k\, S \,. \end{aligned} \tag{7.124}$$

To verify that this is complete it is necessary to recognise, since the indices take only three values, that

$$\varepsilon^{ijl} B_l{}^k + \varepsilon^{kil} B_l{}^j + \varepsilon^{jkl} B_l{}^i = \varepsilon^{ijk} B_l{}^l = 0 \,, \tag{7.125}$$

for any B^i_j belonging to the **8** representation. (7.124) then corresponds to

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}\,. \tag{7.126}$$

These of course are the baryon representations for $SU(3)_F$.

In general it is only necessary to use the invariant tensors in (7.104) 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

$$B^i_j\, B'^k_l \to B^{(i}_{(j} B'^{k)}_{l)}\,,\ \varepsilon^{jl(m} B^i_j B'^{k)}_l\,,\ \ \varepsilon_{ik(m} B^i_j B'^k_{l)}\,,\ \ B^i_j\, B'^j_l\,,\ \ B^i_j\, B'^j_i\,,\ \ B^i_j\, B'^j_i\,. \tag{7.127}$$

and then subtracting the required terms to cancel all traces formed by contracting upper and lower indices, as in (7.122). This gives the decomposition

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{10} \oplus \mathbf{10}^* \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}\,. \tag{7.128}$$

7.6.1 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 (2.85). For $\mathfrak{su}(3)$, characters are given by (7.79). In general these have an expansion in terms of a sum over the weights in the associated weight diagram

$$\chi_{\underline{\Lambda}}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda},\underline{\lambda}}\, {u_1}^{r_1+r_2+2} {u_2}^{r_2+1}\,, \qquad \underline{\Lambda} = [n_1, n_2]\,, \quad \underline{\lambda} = [r_1, r_2]\,, \tag{7.129}$$

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

$$n_{\underline{\Lambda},\underline{\lambda}} = n_{\underline{\Lambda},\sigma\underline{\lambda}}\,. \tag{7.130}$$

Using (7.78) it is easy to see that

$$C_{\underline{\Lambda}}(u)\, \chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\, C_{\underline{\Lambda}+\underline{\lambda}}(u)\,, \tag{7.131}$$

and since, for the weights $\{\underline{\lambda}\}$ corresponding to the representation with highest weight $\underline{\Lambda}$,

$$\{\underline{\lambda}\} = \{\sigma\underline{\lambda}\}\,, \qquad (\underline{\Lambda} + \underline{\lambda})^\sigma = \underline{\Lambda}^\sigma + \sigma\underline{\lambda}\,, \tag{7.132}$$

then, with (7.130), we may use (7.84) to obtain

$$\chi_{\underline{\Lambda}}(u)\, \chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\, \chi_{\underline{\Lambda}+\underline{\lambda}}(u)\,. \tag{7.133}$$

However in general $\underline{\Lambda} + \underline{\lambda} \notin \mathcal{W}$, as defined in (7.65). In this case (7.86) may be used to rewrite (7.133) as

$$\chi_{\underline{\Lambda}}(u)\, \chi_{\underline{\Lambda}'}(u) = \sum_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\, P_\sigma\, \chi_{(\underline{\Lambda}+\underline{\lambda})^\sigma}(u)\,, \qquad (\underline{\Lambda} + \underline{\lambda})^\sigma \in \mathcal{W}\,, \tag{7.134}$$

dropping all terms where $\underline{\Lambda}+\underline{\lambda}$ satisfies any of the conditions in (7.87) ensuring $\chi_{\underline{\Lambda}+\underline{\lambda}}(u)=0$, so that, by virtue of (7.88), σ in (7.134) is then unique. Since in (7.134) 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 (7.134) 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,

$$\mathcal{V}_{\underline{\Lambda}} = \bigoplus_{\underline{\lambda}} \mathcal{V}_{\underline{\Lambda}}^{(\underline{\lambda})} , \qquad \dim \mathcal{V}_{\underline{\Lambda}}^{(\underline{\lambda})} = n_{\underline{\Lambda},\underline{\lambda}} , \tag{7.135}$$

and then (7.134) is equivalent to

$$\mathcal{V}_{\underline{\Lambda}} \otimes \mathcal{V}_{\underline{\Lambda}'} \simeq \bigoplus_{\underline{\lambda}} n_{\underline{\Lambda}',\underline{\lambda}}\, P_\sigma\, \mathcal{V}_{(\underline{\Lambda}+\underline{\lambda})^\sigma} , \qquad (\underline{\Lambda}+\underline{\lambda})^\sigma \in \mathcal{W} . \tag{7.136}$$

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

$$\mathcal{V}_{[1,0]} \to [1,0] ,\ [-1,1] ,\ [0,-1] , \tag{7.137}$$

and then

$$\begin{aligned} \mathcal{V}_{[n_1,n_2]} \otimes \mathcal{V}_{[1,0]} &\simeq \mathcal{V}_{[n_1+1,n_2]} \oplus \mathcal{V}_{[n_1-1,n_2+1]} \oplus \mathcal{V}_{[n_1,n_2-1]} \\ &= \begin{cases} \mathcal{V}_{[1,n_2]} \oplus \mathcal{V}_{[0,n_2-1]} , & n_1 = 0 , \\ \mathcal{V}_{[n_1+1,0]} \oplus \mathcal{V}_{[n_1-1,1]} , & n_2 = 0 . \end{cases} \end{aligned} \tag{7.138}$$

It is easy to see that this is in accord with the results in (7.126). For an octet

$$\mathcal{V}_{[1,1]} \to [1,1] ,\ [2,-1] ,\ [-1,2] ,\ [0,0]^2 ,\ [1,-2] ,\ [-2,1] ,\ [-1,-1] , \tag{7.139}$$

so that, for $n_1, n_2 \geq 2$,

$$\begin{aligned} \mathcal{V}_{[n_1,n_2]} \otimes \mathcal{V}_{[1,1]} \simeq{}& \mathcal{V}_{[n_1+1,n_2+1]} \oplus \mathcal{V}_{[n_1+2,n_2-1]} \oplus \mathcal{V}_{[n_1-1,n_2+2]} \oplus \mathcal{V}_{[n_1,n_2]} \\ &\oplus \mathcal{V}_{[n_1,n_2]} \oplus \mathcal{V}_{[n_1+1,n_2-2]} \oplus \mathcal{V}_{[n_1-2,n_2+1]} \oplus \mathcal{V}_{[n_1-1,n_2-1]} , \end{aligned} \tag{7.140}$$

with special cases

$$\mathcal{V}_{[1,1]} \otimes \mathcal{V}_{[1,1]} \simeq \mathcal{V}_{[2,2]} \oplus \mathcal{V}_{[3,0]} \oplus \mathcal{V}_{[0,3]} \oplus \mathcal{V}_{[1,1]} \oplus \mathcal{V}_{[1,1]} \oplus \mathcal{V}_{[0,0]} , \tag{7.141}$$

which is in accord with (7.128), and

$$\mathcal{V}_{[3,0]} \otimes \mathcal{V}_{[1,1]} \simeq \mathcal{V}_{[4,1]} \oplus \mathcal{V}_{[2,2]} \oplus \mathcal{V}_{[3,0]} \oplus \mathcal{V}_{[1,1]} , \tag{7.142}$$

using $\mathcal{V}_{[4,-2]} \simeq -\mathcal{V}_{[3,0]}$. Equivalently, labelling the representations by their dimensions

$$\mathbf{10} \otimes \mathbf{8} = \mathbf{35} \oplus \mathbf{27} \oplus \mathbf{10} \oplus \mathbf{8} . \tag{7.143}$$

8 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

$$A_\mu \in \mathcal{A} \, . \tag{8.1}$$

In a formal sense, the gauge group $\mathcal{G}$ is defined by

$$\mathcal{G} \simeq \bigotimes_x G_x \, , \tag{8.2}$$

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

$$A_\mu(x) \underset{g(x)}{\longrightarrow} A_\mu{}^{g(x)}(x) \sim A_\mu(x) \, . \tag{8.3}$$

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

$$\mathcal{A}/\mathcal{G} \, . \tag{8.4}$$

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[51],

Theory:	QED	Weinberg-Salam model	QCD,
Gauge Group:	$U(1)$	$SU(2) \otimes U(1)$	$SU(3)$.

[51] 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.

8.1 Abelian Gauge Theories

The simplest example arises for $G = U(1)$, which is the gauge group for Maxwell[52] 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

$$\phi \underset{e^{i\alpha}}{\longrightarrow} e^{iq\alpha}\phi = \phi' . \tag{8.5}$$

If the field ϕ forms a non projective representation we must have

$$q \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\} . \tag{8.6}$$

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 (8.6). 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 (8.5) for *global transformations*, where α is independent of x. Restricting to first derivatives this requires

$$\mathcal{L}_\phi(\phi, \partial_\mu\phi) = \mathcal{L}_\phi(\phi', \partial_\mu\phi') , \tag{8.7}$$

and an obvious solution, which defines a Lorentz invariant theory for complex scalars ϕ, is then

$$\mathcal{L}_\phi(\phi, \partial_\mu\phi) = \partial^\mu\phi^\star\phi_\mu\phi - V(\phi^*\phi) . \tag{8.8}$$

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

$$\partial_\mu\phi' = e^{iq\alpha}\big(\partial_\mu\phi + iq\partial_\mu\alpha\,\phi\big) , \tag{8.9}$$

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

$$D_\mu\phi = \partial_\mu\phi - iqA_\mu\phi . \tag{8.10}$$

If under a local $U(1)$ gauge transformation, as in (8.5), the gauge field transforms as

$$A_\mu \underset{e^{i\alpha}}{\longrightarrow} A_\mu + \partial_\mu\alpha = A'_\mu , \tag{8.11}$$

so that

$$D'_\mu\phi' = e^{i\alpha}D_\mu\phi , \tag{8.12}$$

[52] James Clerk Maxwell, 1831-79, Scottish, second wrangler 1854.

and then it is easy to see that, for any globally invariant lagrangian satisfying (8.7),

$$\mathcal{L}_\phi\big(\phi, D_\mu\phi\big) = \mathcal{L}_\phi\big(\phi', D'_\mu\phi'\big)\,. \tag{8.13}$$

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 (4.227) 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

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = F'_{\mu\nu}\,, \tag{8.14}$$

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

$$[D_\mu, D_\nu]\phi = -iqF_{\mu\nu}\phi\,. \tag{8.15}$$

The simplest Lorentz invariant, gauge invariant, lagrangian is then

$$\mathcal{L} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_\phi\big(\phi, D_\mu\phi\big)\,, \qquad \mathcal{L}_{\text{gauge}} = -\frac{1}{4e^2}\,F^{\mu\nu}F_{\mu\nu}\,, \tag{8.16}$$

with e an arbitrary parameter, unimportant classically. It is commonplace to rescale the fields so that

$$A_\mu \to eA_\mu\,, \qquad D_\mu\phi = \partial_\mu\phi - ieqA_\mu\phi\,, \tag{8.17}$$

so that e disappears from the gauge field term in (8.16). The dynamical equations of motion which flow from (8.16) are, for the gauge field,

$$\frac{1}{e^2}\,\partial^\mu F_{\mu\nu} = j_\nu = -\frac{\partial}{\partial A^\nu}\mathcal{L}_\phi\big(\phi, D_\mu\phi\big)\,, \tag{8.18}$$

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 (8.14),

$$\partial_\omega F_{\mu\nu} + \partial_\nu F_{\omega\mu} + \partial_\mu F_{\nu\omega} = 0\,. \tag{8.19}$$

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 $\mathrm{d}F = \mathrm{d}^2A = 0$.

8.2 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[53] so they are often referred to, for the particular gauge invariant lagrangian generalising the abelian lagrangian given in (8.16) and obtained below, as Yang-Mills theories. Nevertheless the same theory was also developed, but not published, by R. Shaw[54] (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

$$\phi(x) \underset{g(x)}{\longrightarrow} g(x)\phi(x) = \phi'(x) \,, \tag{8.20}$$

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

$$\partial_\mu\phi(x) \underset{g(x)}{\longrightarrow} g(x)\big(\partial_\mu\phi(x) + g(x)^{-1}\partial_\mu g(x)\,\phi(x)\big) = \partial_\mu\phi'(x) \,, \tag{8.21}$$

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,

$$A_\mu(x) = A^a{}_\mu(x)\, t_a \,, \tag{8.22}$$

and then

$$D_\mu\phi = (\partial_\mu + A_\mu)\phi \,. \tag{8.23}$$

Requiring

$$D'_\mu\phi' = gD_\mu\phi \,, \tag{8.24}$$

or

$$g^{-1}A'_\mu\, g + g^{-1}\partial_\mu g = A_\mu \,, \tag{8.25}$$

then the gauge field must transform under a gauge transformation as

$$A_\mu \underset{g}{\longrightarrow} A'_\mu = gA_\mu\, g^{-1} - \partial_\mu g\, g^{-1} = gA_\mu\, g^{-1} + g\partial_\mu g^{-1} \,. \tag{8.26}$$

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 (8.26).

It is also useful to note, since G is a Lie group, the associated infinitesimal transformations when

$$g = \mathbb{1} + \lambda \,, \qquad \lambda = \lambda^a t_a \,. \tag{8.27}$$

[53] Chen-Ning Franklin Yang, 1922-, Chinese then American, Nobel prize 1957. Robert L. Mills, 1927-99, American.

[54] Ron Shaw, 1929-2016, English.

Then from (8.20) and (8.24), for arbitrary $\lambda^a(x)$,

$$\delta\phi = \lambda\phi\,, \qquad \delta D_\mu\phi = \lambda D_\mu\phi\,, \tag{8.28}$$

and from (8.26)

$$\delta A_\mu = [\lambda, A_\mu] - \partial_\mu\lambda \qquad \Rightarrow \qquad \delta A^a{}_\mu = -f^a{}_{bc}A^b{}_\mu\lambda^c - \partial_\mu\lambda^a\,. \tag{8.29}$$

The associated curvature is obtained from the commutator of two covariant derivatives, as in the abelian case in (8.15), which gives

$$[D_\mu, D_\nu]\phi = F_{\mu\nu}\phi\,, \qquad F_{\mu\nu} = F^a{}_{\mu\nu}t_a\,, \tag{8.30}$$

so that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\,, \tag{8.31}$$

or

$$F^a{}_{\mu\nu} = \partial_\mu A^a{}_\nu - \partial_\nu A^a{}_\mu + f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\,. \tag{8.32}$$

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

$$F = \mathrm{d}A + A\wedge A\,, \qquad A = A_\mu\,\mathrm{d}x^\mu\,, \quad A\wedge A = \tfrac{1}{2}[A_\mu, A_\nu]\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\,. \tag{8.33}$$

For a gauge transformation as in (8.26)

$$F_{\mu\nu} \xrightarrow[g]{} F'_{\mu\nu} = gF_{\mu\nu}\,g^{-1}\,, \tag{8.34}$$

or, infinitesimally,

$$\delta F_{\mu\nu} = [\lambda, F_{\mu\nu}] \qquad \Rightarrow \qquad \delta F^a{}_{\mu\nu} = -f^a{}_{bc}F^b{}_{\mu\nu}\lambda^c\,, \tag{8.35}$$

which are homogeneous.

As a consistency check we verify the result (8.35) for $\delta F^a{}_{\mu\nu}$ from the expression (8.32) using (8.29) for $\delta A^a{}_\mu$. First

$$\delta\big(\partial_\mu A^a{}_\nu - \partial_\nu A^a{}_\mu\big) = -f^a{}_{bc}\big(\partial_\mu A^b{}_\nu - \partial_\nu A^b{}_\mu\big)\lambda^c - f^a{}_{bc}\big(A^b{}_\nu\partial_\mu\lambda^c - A^b{}_\mu\partial_\nu\lambda^c\big)\,. \tag{8.36}$$

Then

$$\delta\big(f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\big)\big|_{\partial\lambda} = -f^a{}_{bc}\big(\partial_\mu\lambda^b A^c{}_\nu + A^b{}_\mu\partial_\nu\lambda^c\big)\,, \tag{8.37}$$

which cancels, using (5.39), the $\partial\lambda$ terms in (8.36). Furthermore

$$\begin{aligned}\delta\big(f^a{}_{bc}\,A^b{}_\mu A^c{}_\nu\big)\big|_\lambda &= -\,f^a{}_{bc}\big(f^b{}_{de}A^d{}_\mu\lambda^e\,A^c{}_\nu + A^b{}_\mu\,f^c{}_{de}A^d{}_\nu\lambda^e\big)\\ &= -\,\big(f^a{}_{fd}f^f{}_{be} + f^a{}_{cf}f^f{}_{be}\big)A^b{}_\mu A^d{}_\nu\lambda^e = -f^a{}_{fe}\,f^f{}_{bd}A^b{}_\mu A^d{}_\nu\,\lambda^e\,,\end{aligned} \tag{8.38}$$

by virtue of the Jacobi identity in the form (5.43). Combining (8.36), (8.37) and (8.38) demonstrates (8.35) 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

$$D_\mu \Phi = \partial_\mu \Phi + [A_\mu, \Phi] \qquad \Rightarrow \qquad (D_\mu \Phi)^a = \partial_\mu \Phi^a + f^a{}_{bc} \, A^b{}_\mu \Phi^c \, . \tag{8.39}$$

This is in accord with the general form given by (8.23), with (8.22), using (5.172) for the adjoint representation generators. Note that (8.29) can be written as $\delta A^a{}_\mu = -(D_\mu \lambda)^a$ and for an arbitrary variation $\delta A^a{}_\mu$ from (8.32),

$$\delta F^a{}_{\mu\nu} = (D_\mu \delta A_\nu)^a - (D_\nu \delta A_\mu)^a \, . \tag{8.40}$$

From the identity

$$\big([D_\omega, [D_\mu, D_\nu]] + [D_\nu, [D_\omega, D_\mu]] + [D_\mu, [D_\nu, D_\omega]]\big)\phi = 0 \, , \tag{8.41}$$

for any representation, we have the non abelian Bianchi identity, generalising (8.19),

$$D_\omega F_{\mu\nu} + D_\nu F_{\omega\mu} + D_\mu F_{\nu\omega} = 0 \, , \tag{8.42}$$

where the adjoint covariant derivatives are as defined in (8.39). Alternatively with the notation in (8.33)

$$\mathrm{d}F + A \wedge F - F \wedge A = 0 \, . \tag{8.43}$$

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 (5.187) or equivalently

$$g_{db} f^d{}_{ca} + g_{ad} f^d{}_{cb} = 0 \, , \tag{8.44}$$

which also implies, for finite group transformations g and with X, Y belonging to the associated Lie algebra,

$$g_{ab} \, (g X g^{-1})^a \, (g Y g^{-1})^b = g_{ab} \, X^a Y^b \, . \tag{8.45}$$

If X, Y are then adjoint representation fields the definition of the adjoint covariant derivative in (8.39) gives

$$\partial_\mu \big(g_{ab} \, X^a Y^b \big) = g_{ab} \big((D_\mu X)^a Y^b + X^a (D_\mu Y)^b \big) \, , \tag{8.46}$$

in a similar fashion to covariant derivatives in general relativity.

The simplest gauge invariant lagrangian, extending the abelian result in (8.16), is then, as a result of the transformation properties (8.34) or (8.35), just the obvious extension of that proposed by Yang and Mills for $SU(2)$

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4} \, g_{ab} F^{a\mu\nu} F^b{}_{\mu\nu} \, . \tag{8.47}$$

It is essential that the metric be non degenerate $\det[g_{ab}] \neq 0$, and then using (8.40) requiring the action to be stationary gives the gauge covariant dynamical equations

$$(D^\mu F_{\mu\nu})^a = 0 \, . \tag{8.48}$$

These equations, as well as (8.42) 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 (8.26) and hence the time evolution of A_μ is arbitrary up to this extent, only gauge equivalence classes, belonging to (8.4), 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 (5.194). Each $U(1)$ factor corresponds to a simple abelian gauge theory as described in 8.1. 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

$$g_{ab} = \frac{1}{g^2}\,\delta_{ab}\,, \tag{8.49}$$

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 (5.136). Choosing $T_a = (E_1, E_2, J_3, 1)$ then it is straightforward to verify that

$$[g_{ab}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & c \\ 0 & 0 & c & 0 \end{pmatrix}, \qquad \beta \text{ arbitrary}\,, \tag{8.50}$$

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 (8.46) is given by the conservation of the gauge invariant energy momentum tensor defined by

$$T^{\mu\nu} = g_{ab}\big(F^{a\mu\sigma}F^{b\nu}{}_\sigma - \tfrac{1}{4}g^{\mu\nu}F^{a\sigma\rho}F^b{}_{\sigma\rho}\big)\,. \tag{8.51}$$

Then

$$\begin{aligned} \partial_\mu T^\mu{}_\nu &= g_{ab}\big((D_\mu F^{\mu\sigma})^a F^b{}_{\nu\sigma} + F^{a\mu\sigma}(D_\mu F_{\sigma\nu})^b - \tfrac{1}{2}F^{a\sigma\rho}(D_\nu F_{\sigma\rho})^b\big) \\ &= g_{ab}(D_\mu F^{\mu\sigma})^a - \tfrac{1}{2}g_{ab}F^{a\sigma\rho}\big((D_\rho F_{\nu\sigma})^b - (D_\sigma F_{\nu\rho})^b + (D_\nu F_{\sigma\rho})^b\big) \\ &= g_{ab}(D_\mu F^{\mu\sigma})^a\,, \end{aligned} \tag{8.52}$$

using the Bianchi identity (8.42). Clearly this is conserved subject to the dynamical equation (8.48).

8.2.1 Chern-Simons Theory

The standard gauge invariant lagrangian is provided by (8.47). 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[55] theories.

First we note that in four dimensions the Bianchi identity (8.42) may be alternatively be written using the four dimensional antisymmetric symbol as

$$\varepsilon^{\mu\nu\sigma\rho} D_\nu F_{\sigma\rho} = 0 \,. \tag{8.53}$$

Apart from (8.47) there is then another similar gauge invariant and Lorentz invariant

$$\frac{1}{4}\,\varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} \,, \tag{8.54}$$

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 (8.40) and (8.53) to give

$$\delta\frac{1}{4}\,\varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} = \varepsilon^{\mu\nu\sigma\rho} g_{ab} (D_\mu \delta A_\nu)^a F^b{}_{\sigma\rho} = \partial_\mu\big(\varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu\, F^b{}_{\sigma\rho}\big) \,. \tag{8.55}$$

This allows us to write

$$\frac{1}{4}\,\varepsilon^{\mu\nu\sigma\rho} g_{ab} F^a{}_{\mu\nu} F^b{}_{\sigma\rho} = \partial_\mu \omega^\mu \,, \tag{8.56}$$

where

$$\omega^\mu = \varepsilon^{\mu\nu\sigma\rho} g_{ab}\big(A^a{}_\nu \partial_\sigma A^b{}_\rho + \tfrac{1}{3} f^b{}_{cd} A^a{}_\nu A^c{}_\sigma A^d{}_\rho\big) \,, \tag{8.57}$$

since this has the variation

$$\begin{aligned} \delta\omega^\mu &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\big(\delta A^a{}_\nu \partial_\sigma A^b{}_\rho + A^a{}_\nu \partial_\sigma \delta A^b{}_\rho + f^b{}_{cd}\, \delta A^a{}_\nu\, A^c{}_\sigma A^d{}_\rho\big) \\ &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \partial_\sigma\big(A^a{}_\nu \delta A^b{}_\rho\big) + \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu \big(2\partial_\sigma A^b{}_\rho + f^b{}_{cd} A^c{}_\sigma A^d{}_\rho\big) \\ &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \partial_\nu\big(\delta A^a{}_\sigma A^b{}_\rho\big) + \varepsilon^{\mu\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu F^b{}_{\sigma\rho} \,, \end{aligned} \tag{8.58}$$

using that $g_{ab} f^b{}_{cd}$ is totally antisymmetric as a consequence of (8.44). The result is then in agreement with (8.55).

If the variation is a gauge transformation so that

$$A^a{}_\mu \underset{g}{\to} A'^a{}_\mu \qquad \Rightarrow \qquad \omega^\mu \underset{g}{\to} \omega'^\mu \,, \tag{8.59}$$

then since (8.56) is gauge invariant we must require

$$\partial_\mu \omega^\mu = \partial_\mu \omega'^\mu \,. \tag{8.60}$$

[55]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 (8.58) which then gives, using the Bianchi identity (8.53) again,

$$\begin{aligned}\delta\omega^\mu &= -\,\varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big((D_\sigma\lambda)^a A^b{}_\rho\big) - \varepsilon^{\mu\nu\sigma\rho} g_{ab}\,(D_\nu\lambda)^a F^b{}_{\sigma\rho} \\ &= \varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a (D_\sigma A)^b{}_\rho - \lambda^a F^b{}_{\sigma\rho}\big) \\ &= -\,\varepsilon^{\mu\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a \partial_\sigma A^b{}_\rho\big)\,. \end{aligned} \tag{8.61}$$

Hence it is evident that the result in (8.61) satisfies

$$\partial_\mu \delta\omega^\mu = 0\,. \tag{8.62}$$

In three dimensions the identities for ω^μ may be applied to

$$\mathcal{L}_{CS} = \varepsilon^{\nu\sigma\rho} g_{ab}\big(A^a{}_\nu \partial_\sigma A^b{}_\rho + \tfrac{1}{3} f^b{}_{cd} A^a{}_\nu A^c{}_\sigma A^d{}_\rho\big)\,, \tag{8.63}$$

which defines the Chern-Simons lagrangian for gauge fields. For an infinitesimal gauge transformation, by virtue of (8.61), $\mathcal{L}_{CS}$ becomes a total derivative since

$$\delta A^a{}_\nu = -(D_\nu\lambda)^a \quad \Rightarrow \quad \delta\mathcal{L}_{CS} = -\varepsilon^{\nu\sigma\rho} g_{ab}\,\partial_\nu\big(\lambda^a \partial_\sigma A^b{}_\rho\big)\,, \tag{8.64}$$

so that the corresponding action is invariant. Under a general variation

$$\delta \int \mathrm{d}^3x\, \mathcal{L}_{CS} = \int \mathrm{d}^3 x\, \varepsilon^{\nu\sigma\rho} g_{ab}\, \delta A^a{}_\nu F^b{}_{\sigma\rho}\,, \tag{8.65}$$

so that the dynamical equations are

$$F^a{}_{\mu\nu} = 0\,, \tag{8.66}$$

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

$$\begin{aligned}\mathrm{tr}\big(\underbrace{A\wedge\dots\wedge A}_{n}\big) = \mathrm{tr}\big((\underbrace{A\wedge\dots\wedge A}_{n-1})\wedge A\big) &= (-)^{n-1}\mathrm{tr}\big(A\wedge(\underbrace{A\wedge\dots\wedge A}_{n-1})\big) \\ &= 0 \qquad \text{for } \; n \; \text{ even}\,. \end{aligned} \tag{8.67}$$

The Chern-Simons theory is then defined in terms of the three-form

$$\omega = \mathrm{tr}\big(A \wedge \mathrm{d}A + \tfrac{2}{3}\, A \wedge A \wedge A\big) = \mathrm{tr}\big(A \wedge F - \tfrac{1}{3}\, A \wedge A \wedge A\big)\,, \tag{8.68}$$

with the two-form curvature F as in (8.33). It is easy to see that

$$\mathrm{d}\omega = \mathrm{tr}\big(\mathrm{d}A \wedge \mathrm{d}A + 2\,\mathrm{d}A \wedge A \wedge A\big) = \mathrm{tr}\big(F \wedge F\big)\,, \tag{8.69}$$

which is equivalent to (8.56) and (8.57). For a finite gauge transformation

$$A' = gAg^{-1} + g\mathrm{d}g^{-1}\,, \qquad F' = gFg^{-1}\,, \tag{8.70}$$

so that, from (8.68),

$$\begin{aligned}\omega' = \omega + \mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge (F - A \wedge A)\big) - \mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g \wedge A\big)\\ - \tfrac{1}{3}\,\mathrm{tr}\big(\mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g \wedge \mathrm{d}g^{-1}g\big)\,.\end{aligned} \tag{8.71}$$

Using

$$\mathrm{d}g^{-1}g = -g^{-1}\mathrm{d}g\,, \qquad \mathrm{d}\big(g^{-1}\mathrm{d}g\big) = -g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\,, \tag{8.72}$$

we get

$$\omega' = \omega + \mathrm{d}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge A\big) + \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big)\,. \tag{8.73}$$

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 (8.72),

$$\mathrm{d}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big) = -\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big) = 0\,, \tag{8.74}$$

by virtue (8.67), we have

$$\mathrm{d}\omega' = \mathrm{d}\omega\,, \tag{8.75}$$

which is equivalent to (8.60). 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

$$\int_{\mathcal{M}_3} \omega' \neq \int_{\mathcal{M}_3} \omega\,, \tag{8.76}$$

for some $g(x)$. The Cherns-Simons action is not then gauge invariant for such gauge transformations g.

To discuss

$$I = \int_{\mathcal{M}_3} \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\big)\,, \tag{8.77}$$

we note that for a variation of g, since

$$\delta(g^{-1}\mathrm{d}g) = g^{-1}\mathrm{d}(\delta g\, g^{-1})\, g\,, \tag{8.78}$$

then

$$\delta\,\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = \mathrm{tr}\big(\mathrm{d}(\delta g\, g^{-1})\wedge \mathrm{d}g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\big)$$
$$= \mathrm{d}\,\mathrm{tr}\big(\delta g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\wedge \mathrm{d}g\, g^{-1}\big)\,, \qquad (8.79)$$

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,

$$\delta I = 0\,, \qquad (8.80)$$

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

$$\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = \rho(\theta)\,\mathrm{d}^3\theta\,, \qquad (8.81)$$

The integration measure in (8.81) is defined in terms of the left invariant Lie algebra one forms so that for $g(\theta') = g_0 g(\theta)$ we have

$$\rho(\theta')\,\mathrm{d}^3\theta' = \rho(\theta)\,\mathrm{d}^3\theta\,. \qquad (8.82)$$

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

$$\tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) \approx \tfrac{1}{3}i^3\,\mathrm{tr}\big(\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\wedge\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\wedge\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{\theta}\big)$$
$$= \tfrac{2}{3}\,\varepsilon_{ijk}\,\mathrm{d}\theta^i\wedge\mathrm{d}\theta^j\wedge\mathrm{d}\theta^k = 4\,\mathrm{d}^3\theta\,, \qquad (8.83)$$

assuming (5.21) and standard formulae for the Pauli matrices in (3.20) with (3.22). Thus $\rho(0) = 4$ and the results for the group integration volume for $SU(2)$ in (5.155) then imply, integrating over $\mathcal{M}_{SU(2)}\simeq S^3$,

$$\int_{\mathcal{M}_{SU(2)}} \tfrac{1}{3}\,\mathrm{tr}\big(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\big) = 8\pi^2\,. \qquad (8.84)$$

In general the topological invariant defined by (8.77), 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 (8.84) then requires that in general

$$I = 8\pi^2 n \qquad \text{for} \qquad n\in\mathbb{Z}\,. \qquad (8.85)$$

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

$$S_{\mathrm{CS}} = \frac{k}{4\pi}\int_{\mathcal{M}_3}\mathrm{tr}\big(A\wedge\mathrm{d}A + \tfrac{2}{3}\,A\wedge A\wedge A\big)\,, \qquad k\in\mathbb{Z}\,, \qquad (8.86)$$

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.

8.3 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, (8.14). 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 (8.34). 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,

$$\operatorname{tr}\big(D_{\alpha_{11}} \dots D_{\alpha_{1r_1}} F_{\mu_1\nu_1}\, D_{\alpha_{21}} \dots D_{\alpha_{2r_2}} F_{\mu_2\nu_2} \, \dots \, D_{\alpha_{s1}} \dots D_{\alpha_{sr_s}} F_{\mu_s\nu_s}\big) . \tag{8.87}$$

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 (8.51). In general there are also derivative relations since

$$\partial_\mu \operatorname{tr}(X_1 \dots X_s) = \sum_{i=1}^{s} \operatorname{tr}(X_1 \dots \, D_\mu X_i \, \dots X_s) . \tag{8.88}$$

However, depending on the gauge group, the traces in (8.87) are not independent for arbitrary products of F's, even when no derivatives are involved. To show this we may consider the identity

$$\det(\mathbb{1} - X) = e^{\operatorname{tr} \ln(\mathbb{1} - X)} , \tag{8.89}$$

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

$$\begin{aligned} &\det(\mathbb{1} - X) = e^{-\sum_{r\geq 1} \operatorname{tr}(X^r)/r} \\ &= 1 - \operatorname{tr}(X) + \tfrac{1}{2}\big(\operatorname{tr}(X)^2 - \operatorname{tr}(X^2)\big) - \tfrac{1}{6}\big(\operatorname{tr}(X)^3 - 3\operatorname{tr}(X)\operatorname{tr}(X^2) + 2\operatorname{tr}(X^3)\big) + \dots . \end{aligned} \tag{8.90}$$

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[56]. If $N = 2$ this gives the relation

$$\operatorname{tr}(X^3) = \tfrac{3}{2}\operatorname{tr}(X)\operatorname{tr}(X^2) - \tfrac{1}{2}\operatorname{tr}(X)^3\,, \tag{8.91}$$

and if $N = 3$, and also we require $\operatorname{tr}(X) = 0$, the relevant identity becomes

$$\operatorname{tr}(X^4) = \tfrac{1}{2}\operatorname{tr}(X^2)^2\,. \tag{8.92}$$

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 (8.87) since the results for $\operatorname{tr}(X^n)$ are equivalent to relations for $\operatorname{tr}(t_{(a_1}\dots t_{a_N)})$.

8.3.1 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 (8.2), 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 (8.5) 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 (8.5), this is defined in terms of the differential equation

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t} - iq\,\dot{x}^\mu(t)A_\mu\big(x(t)\big)\Big)P(t,t') = 0\,, \qquad P(t,t) = 1\,, \quad \dot{x}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\,, \tag{8.93}$$

which has a solution,

$$P(t,t') = e^{iq\int_{t'}^{t}\mathrm{d}\tau\,\dot{x}^\mu(\tau)A_\mu(x(\tau))}\,. \tag{8.94}$$

We then require

$$P(\Gamma_{x,y}) = P(1,0) = e^{iq\int_{\Gamma_{x,y}}\mathrm{d}x^\mu A_\mu(x)} \in U(1)\,, \tag{8.95}$$

[56]Equivalently if $F(z) = \det(\mathbb{1} - zX) = 1 + \sum_{r=1}^{N} a_r(X)\,z^r$ then

$$-\frac{F'(z)}{F(z)} = \operatorname{tr}\big(X(\mathbb{1} - zX)^{-1}\big) = \sum_{r=0}^{\infty} z^r\operatorname{tr}(X^{r+1})\,,$$

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 (8.11)

$$P(\Gamma_{x,y}) \underset{e^{i\alpha}}{\longrightarrow} P(\Gamma_{x,y})\, e^{iq\int_{\Gamma_{x,y}} \mathrm{d}x^\mu \partial_\mu \alpha(x)} = e^{iq\,\alpha(x)}\, P(\Gamma_{x,y})\, e^{-iq\,\alpha(y)}\,, \tag{8.96}$$

demonstrating that, for ϕ transforming under gauge transformations as in (8.5),

$$P(\Gamma_{x,y})\phi(y) \underset{e^{i\alpha}}{\longrightarrow} e^{iq\alpha(x)} P(\Gamma_{x,y})\phi(y)\,. \tag{8.97}$$

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 (8.96) that $P(\Gamma)$ is gauge invariant. In this abelian case $P(\Gamma)$ may be expressed just in terms of the gauge invariant curvature in (8.14) using Stokes' theorem

$$P(\Gamma) = e^{iq\oint_\Gamma \mathrm{d}x^\mu A_\mu(x)} = e^{\frac{1}{2}iq\int_S \mathrm{d}S^{\mu\nu} F_{\mu\nu}(x)}\,, \tag{8.98}$$

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 $\oint_\Gamma \mathrm{d}\mathbf{x}\cdot\mathbf{A} = \int_S \mathrm{d}\mathbf{S}\cdot\mathbf{B}$, $\mathbf{B} = \nabla\times\mathbf{A}$ with $\mathrm{d}S_i = \frac{1}{2}\varepsilon_{ijk}\mathrm{d}S^{jk}$).

For the non abelian case (8.93) generalises to a matrix equation

$$\Big(\mathbb{1}\,\frac{\mathrm{d}}{\mathrm{d}t} + A(t)\Big)P(t,t') = 0\,, \qquad A(t) = \dot{x}^\mu A_\mu(x(t))\,, \quad P(t,t) = \mathbb{1}\,, \tag{8.99}$$

where $A(t)$ is a matrix belonging to the Lie algebra for a representation $\mathcal{R}$ of G. (8.99) may also be expressed in an equivalent integral form

$$P(t,t') = \mathbb{1} - \int_{t'}^{t} \mathrm{d}\tau\, A(\tau)P(\tau,t')\,. \tag{8.100}$$

Solving this iteratively gives

$$\begin{aligned} P(t,t') &= \mathbb{1} + \sum_{n\ge1}(-1)^n \int_{t'}^{t}\mathrm{d}t_1 \int_{t'}^{t_1}\mathrm{d}t_2 \cdots \int_{t'}^{t_{n-1}}\mathrm{d}t_n\, A(t_1)A(t_2)\dots A(t_n) \\ &= \mathbb{1} + \sum_{n\ge1}(-1)^n \frac{1}{n!}\prod_{r=1}^{n}\int_{t'}^{t}\mathrm{d}t_r\, \mathcal{T}\big\{A(t_1)A(t_2)\dots A(t_n)\big\}\,. \end{aligned} \tag{8.101}$$

where $\mathcal{T}$ denotes that the non commuting, for differing t, $A(t)$ are t-ordered so that

$$\mathcal{T}\big\{A(t)A(t')\big\} = \begin{cases} A(t)A(t')\,, & t \ge t'\,, \\ A(t')A(t)\,, & t < t'\,. \end{cases} \tag{8.102}$$

The final expression can be simply written as a $\mathcal{T}$-ordered exponential

$$P(t,t') = \mathcal{T}\Big\{e^{-\int_{t'}^{t}\mathrm{d}\tau\, A(\tau)}\Big\}\,. \tag{8.103}$$

The corresponding non abelian generalisation of (8.95) is then

$$P(\Gamma_{x,y}) = P(1,0) = \mathcal{P}\Big\{e^{-\int_{\Gamma_{x,y}}\mathrm{d}x^\mu A_\mu(x)}\Big\} \in \mathcal{R}\,, \tag{8.104}$$

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

$$P(\Gamma_{x,y})\, P(\Gamma_{y,z}) = P(\Gamma_{x,y} \circ \Gamma_{y,z})\,, \tag{8.105}$$

where $\Gamma_{x,y} \circ \Gamma_{y,z}$ denotes path composition, and, if $\mathcal{R}$ is a unitary representation

$$P(\Gamma_{x,y})^{-1} = P(\Gamma^{-1}_{y,x}) = P(\Gamma_{x,y})^\dagger\,, \tag{8.106}$$

with $\Gamma^{-1}_{y,x}$ the inverse path to $\Gamma_{x,y}$.

For a gauge transformation as in (8.26), $g(x) \in \mathcal{R}$, then in (8.99)

$$A(t) \underset{g}{\to} g(t)A(t)g(t)^{-1} - \dot g(t)g(t)^{-1}\,,\ g(t) = g(x(t)) \quad \Rightarrow \quad P(t) \underset{g}{\to} g(t)P(t,t')g(t')^{-1}\,, \tag{8.107}$$

and hence

$$P(\Gamma_{x,y}) \underset{g}{\longrightarrow} g(x)P(\Gamma_{x,y})\, g(y)^{-1}\,. \tag{8.108}$$

For $\Gamma = \Gamma_{x,x}$ a closed path then we may obtain a gauge invariant by taking the trace

$$W(\Gamma) = \mathrm{tr}\big(P(\Gamma_{x,x})\big)\,. \tag{8.109}$$

$W(\Gamma)$ is a *Wilson*[57] *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 (8.104), may also be used to construct gauge invariants involving local gauge dependent fields at different points. For the field ϕ, transforming as in (8.5), $\phi(x)^\dagger P(\Gamma_{x,y})\phi(y)$ is such a gauge invariant, assuming the gauge transformation g is unitary so that (8.5) 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 (8.87), at this point. As an illustration we consider a rectangular closed path with the associated Wilson loop

$$W(\square) = \mathrm{tr}\big(P(\Gamma_{x,x+be_j})\, P(\Gamma_{x+be_j,x+ae_i+be_j})\, P(\Gamma_{x+ae_i+be_j,x+ae_i})\, P(\Gamma_{x+ae_i,x})\big)\,, \tag{8.110}$$

where here Γ are all straight line paths and e_i, e_j are two orthogonal unit vectors. To evaluate $W(\square)$ as $a, b \to 0$ it is convenient to use operators $\hat{\mathrm{x}}^\nu, \hat\partial_\mu$ with the commutation relations

$$\big[\hat{\mathrm{x}}^\mu, \hat{\mathrm{x}}^\nu\big] = 0\,, \qquad \big[\hat\partial_\mu, \hat\partial_\nu\big] = 0\,, \qquad \big[\hat\partial_\mu, \hat{\mathrm{x}}^\nu\big] = \delta_\mu{}^\nu\,, \tag{8.111}$$

which have a representation, acting on vectors $|x\rangle$, $x \in \mathbb{R}^4$, where

$$\hat{\mathrm{x}}^\mu|x\rangle\,, \qquad \hat\partial_\mu|x\rangle = -\partial_\mu|x\rangle\,. \tag{8.112}$$

[57] Kenneth Geddes Wilson, 1936-, American. Nobel prize 1982.

In terms of these operators, since $\hat{\text{x}}^\nu \, e^{-te^\mu \hat{D}_\mu} = e^{-te^\mu \hat{D}_\mu}(\hat{\text{x}}^\nu + te^\nu)$,

$$e^{-te^\mu \hat{D}_\mu}|x\rangle = |x(t)\rangle \, P(\Gamma_{x(t),x}) \, , \qquad \hat{D}_\mu = \hat{\partial}_\mu + A_\mu(\hat{\text{x}}) \, , \quad x^\nu(t) = x^\nu + te^\nu \, , \tag{8.113}$$

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 (8.103) we note that

$$\frac{\partial}{\partial t} \, e^{-te^\mu \hat{D}_\mu}|x\rangle = -e^\mu \hat{D}_\mu \, e^{-te^\mu \hat{D}_\mu}|x\rangle = \bigg(\frac{\partial}{\partial t} \, |x(t)\rangle - |x(t)\rangle \, e^\mu A_\mu(x(t)) \bigg) P(\Gamma_{x(t),x}) \, , \tag{8.114}$$

using (8.112) as well as (8.113). It is then evident that (8.114) reduces to

$$\frac{\partial}{\partial t} \, P(\Gamma_{x(t),x}) = -e^\mu A_\mu(x(t)) \, P(\Gamma_{x(t),x}) \, , \tag{8.115}$$

which is identical to (8.93). For the rectangular closed path in (8.110)

$$\begin{aligned}
&|x\rangle \, P(\Gamma_{x,x+be_j}) \, P(\Gamma_{x+be_j,x+ae_i+be_j}) \, P(\Gamma_{x+ae_i+be_j,x+ae_i}) \, P(\Gamma_{x+ae_i,x}) \\
&= e^{b\hat{D}_j} \, e^{\alpha \hat{D}_i} \, e^{-b\hat{D}_j} \, e^{-\alpha \hat{D}_i}|x\rangle \\
&= e^{ab\,[\hat{D}_j,\hat{D}_i] - \frac{1}{2}a^2 b\,[[\hat{D}_j,\hat{D}_i],\hat{D}_i]] + \frac{1}{2}ab^2\,[\hat{D}_j,[\hat{D}_j,\hat{D}_i]] + \dots} \, |x\rangle \\
&= |x\rangle \, e^{-ab\,F_{ij}(x) - \frac{1}{2}a^2 b\, D_i F_{ij}(x) - \frac{1}{2}ab^2 D_j \, F_{ij}(x) + \dots} \, ,
\end{aligned} \tag{8.116}$$

using the Baker Cambell Hausdorff formula described in 5.4.2 and $[\hat{D}_i, \hat{D}_j] = F_{ij}(\hat{\text{x}})$. Hence, for a N-dimensional representation with $\text{tr}(t_a) = 0$, the leading approximation to (8.110) is just

$$\begin{aligned}
W(\Box) = N &+ \tfrac{1}{2}a^2 b^2 \big(1 + \tfrac{1}{2}a\partial_i + \tfrac{1}{2}b\partial_j + \tfrac{1}{6}a^2\partial_i^2 + \tfrac{1}{6}b^2\partial_j^2 + \tfrac{1}{4}ab\partial_i\partial_j\big) \, \text{tr}\big(F_{ij}F_{ij}\big) \\
&- \tfrac{1}{24}a^4 b^2 \, \text{tr}\big(D_i F_{ij} D_i F_{ij}\big) - \tfrac{1}{24}a^2 b^4 \, \text{tr}\big(D_j F_{ij} D_j F_{ij}\big) \\
&- \tfrac{1}{6}a^3 b^3 \, \text{tr}\big(F_{ij}F_{ij}F_{ij}\big) + \dots \, , \quad \text{no sums on} \;\; i,j \, .
\end{aligned} \tag{8.117}$$

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

$$D_t = \mathbb{1} \, \frac{\partial}{\partial t} + A_t(t) \, , \quad D_s = \mathbb{1} \, \frac{\partial}{\partial s} + A_s(t) \, , \quad A_t(t) = \frac{\partial x^\mu}{\partial t} A_\mu(x) \, , \quad A_s(t) = \frac{\partial x^\mu}{\partial s} A_\mu(x) \, , \tag{8.118}$$

leaving the dependence on s implicit, then

$$[D_t, D_s] = F(t) = \frac{\partial x^\mu}{\partial t} \frac{\partial x^\nu}{\partial s} \, F_{\mu\nu}(x) \, . \tag{8.119}$$

With the definitions in (8.118), (8.99) becomes $D_t P(t) = 0$. Acting with D_s gives

$$D_t D_s P(t,t') = F(t) P(t,t') \, , \qquad D_s P(t,t) = A_s(t) \, , \tag{8.120}$$

which has a straightforward solution giving

$$\frac{\mathrm{d}}{\mathrm{d}s}P(1,0) + A_s(1)P(1,0) - P(1,0)A_s(0) = \int_0^1 \mathrm{d}t\, P(1,t)F(t)P(t,0)\,. \tag{8.121}$$

The result (8.121) may be recast as

$$\begin{aligned}\delta_\Gamma P(\Gamma_{x,y}) + \delta x^\nu A_\nu(x)\, P(\Gamma_{x,y}) - P(\Gamma_{x,y})\, \delta y^\nu A_\nu(y) \\ = \int_{\Gamma_{x,y}} \mathrm{d}z^\mu\, P(\Gamma_{x,z})\, F_{\mu\nu}(z)\delta x^\nu(z)\, P(\Gamma_{z,y})\,,\end{aligned} \tag{8.122}$$

where

$$\Gamma_{x,y} = \Gamma_{x,z} \circ \Gamma_{z,y} \qquad \text{for} \quad z \in \Gamma_{x,y}\,. \tag{8.123}$$

For a Wilson loop

$$\delta_\Gamma W(\Gamma) = \oint_\Gamma \mathrm{d}x^\mu\, \mathrm{tr}\big(F_{\mu\nu}(x)\delta x^\nu(x)\, P(\Gamma_{x,x})\big)\,. \tag{8.124}$$

For a pure Chern-Simons theory then, as a consequence of the dynamical equation (8.66), 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.

9 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 (8.4) with the definitions (8.1) and (8.2). 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

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \, f(x) \, , \tag{9.1}$$

for classes of functions f which are invariant under group transformations belonging to a group G,

$$f(x) = f(x^g) \, , \qquad \text{for} \quad x \underset{g}{\to} x^g \ \text{ for all } \ g \in G \, . \tag{9.2}$$

Necessarily we require

$$(x^{g_1})^{g_2} = x^{g_1 g_2} \, , \qquad (x^g)^{g^{-1}} = x \, , \tag{9.3}$$

and also we assume, under the change of variable $x \to x^g$,

$$\mathrm{d}^n x = \mathrm{d}^n x^g \, . \tag{9.4}$$

The condition (9.4) is an essential condition on the integration measure in (9.1), 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,

$$\mathrm{Orb}(x) = \{x^g\} \, , \qquad H_x = \{h : x^h = x\} \, . \tag{9.5}$$

Clearly two points on the same orbit have isomorphic stability groups since

$$H_{x^g} = g^{-1} H_x g \simeq H_x \subset G \, . \tag{9.6}$$

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 (9.1), 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,

$$\begin{aligned} &\text{for all } x \in \mathbb{R}^n \text{ then } P(x^g) = 0 \ \text{ for some } g \in G \, , \\ &\text{if } P(x_0) = 0 \text{ then } P({x_0}^g) = 0 \ \Rightarrow \ g = h \in H \, , \ {x_0}^h = x_0 \, . \end{aligned} \tag{9.7}$$

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

$$P\big(x_0(\theta)\big) = 0 \quad \Rightarrow \quad \theta^r \ \text{ coordinates on } \ \mathcal{M}\,, \ \ \dim \mathcal{M} = n - \hat{n}\,. \tag{9.8}$$

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

$$\int_G \mathrm{d}\rho(g)\ \delta^{\hat{n}}\big(P(x^g)\big)\ \Delta(x) = 1\,. \tag{9.9}$$

Since by construction $\mathrm{d}\rho(g) = \mathrm{d}\rho(g'g)$ then it is easy to see that

$$\Delta(x^g) = \Delta(x) \quad \text{for all} \quad g \in G\,. \tag{9.10}$$

Using (9.9) in (9.1), and interchanging orders of integration, gives

$$\begin{aligned}
\int_{\mathbb{R}^n} \mathrm{d}^n x\ f(x) &= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x\ \delta^{\hat{n}}\big(P(x^g)\big)\,\Delta(x)\ f(x) \\
&= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x^g\ \delta^{\hat{n}}\big(P(x^g)\big)\,\Delta(x^g)\ f(x^g) \\
&= \int_G \mathrm{d}\rho(g) \int_{\mathbb{R}^n} \mathrm{d}^n x\ \delta^{\hat{n}}\big(P(x)\big)\,\Delta(x)\ f(x)\,.
\end{aligned} \tag{9.11}$$

using the invariance conditions (9.2), (9.4) and (9.10), 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

$$\mathrm{d}\mu(\theta) = \mathrm{d}^n x\ \delta^{\hat{n}}\big(P(x)\big)\,\Delta(x)\,. \tag{9.12}$$

To determine $\Delta(x)$ in (9.9) then, assuming (9.7), if

$$g(\alpha, h) = \exp(\alpha)\,h\,, \qquad \alpha \in \mathfrak{g}/\mathfrak{h}\,, \tag{9.13}$$

we define a linear operator D, which may depend on x_0, such that

$${x_0}^{g(\alpha,h)} = x_0 + D(x_0)\alpha\,, \quad \text{for} \ \ \alpha \approx 0\,, \qquad D(x_0) : \mathfrak{g}/\mathfrak{h} \to \mathbb{R}^n\,. \tag{9.14}$$

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

$$\alpha = \alpha^{\hat{a}} T_{\hat{a}}\,, \tag{9.15}$$

and, with the decomposition in (9.13),

$$\mathrm{d}\rho(g) \approx \prod_{\hat{a}=1}^{\hat{n}} \mathrm{d}\alpha^{\hat{a}}\,\mathrm{d}\rho_H(h) \qquad \text{for} \qquad \alpha^{\hat{a}} \approx 0\,, \tag{9.16}$$

for $\mathrm{d}\rho_H(h)$ the invariant integration measure on H. For x near x_0 we define the linear operator P' by

$$P(x_0 + y) = P'(x_0)y \quad \text{for} \quad y \approx 0\,, \qquad P'(x_0) : \mathbb{R}^n \to \mathbb{R}^{\hat{n}}\,. \tag{9.17}$$

Then in (9.9), with (9.16),

$$\int_G \mathrm{d}\rho(g)\, \delta^{\hat{n}}\big(P(x^g)\big) = \int_G \mathrm{d}\rho(g)\, \delta^{\hat{n}}\big(P({x_0}^g)\big) = V_H \int \mathrm{d}^{\hat{n}}\alpha\, \delta^{\hat{n}}\big(P'(x_0)D(x_0)\alpha\big)$$
$$= V_H\, \frac{1}{\big|\det P'(x_0)D(x_0)\big|}\,, \qquad V_H = \int_H \mathrm{d}\rho_H(h)\,. \tag{9.18}$$

Hence in (9.9)

$$\Delta(x) = \frac{1}{V_H}\big|\det P'(x_0)D(x_0)\big| \qquad \text{for} \qquad x = {x_0}^g\,. \tag{9.19}$$

In a quantum gauge field theory context $\det P'(x_0)D(x_0)$ is the *Faddeev-Popov*[58] *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

$$P(x) = 0 \qquad \Rightarrow \qquad x = x(\theta,\alpha) = x_0(\theta)^{g(\alpha,h)}\,, \tag{9.20}$$

from (9.12) becomes, with a change of variables $x \to \theta, \alpha$,

$$\mathrm{d}\mu(\theta) = \frac{1}{V_H}\, \mathrm{d}^n x\, \delta^{\hat{n}}\big(P(x)\big)\big|\det M(\theta)\big|\,, \qquad M(\theta) = P'(x_0(\theta))D'(x_0(\theta))\,. \tag{9.21}$$

Note that

$$\delta^{\hat{n}}\big(P(x(\theta,\alpha))\big)\big|\det M(\theta)\big| = \delta^{\hat{n}}(\alpha)\,, \tag{9.22}$$

and therefore the measure over $\mathcal{M}$ may also be expressed in terms of the Jacobian from θ, α to x since

$$\mathrm{d}\mu(\theta) = \mathrm{d}^{n-\hat{n}}\theta\, \left|\det\left[\frac{\partial x}{\partial\theta}, \frac{\partial x}{\partial\alpha}\right]\right|_{\alpha=0}\,. \tag{9.23}$$

With these results, for G compact, (9.11) gives

$$\int_{\mathbb{R}^n} \mathrm{d}^n x\, f(x) = V_G \int_{\mathbb{R}^n} \mathrm{d}^n x\, \delta^{\hat{n}}\big(P(x)\big)\,\Delta(x)\, f(x) = V_G \int_{\mathcal{M}} \mathrm{d}\mu(\theta)\, f\big(x_0(\theta)\big)\,. \tag{9.24}$$

As an extension we consider the situation when there is a discrete group W, formed by transformations $\theta \to \theta^{g_i}$, such that

$$W = \big\{g_i : x_0(\theta^{g_i}) = x_0(\theta)^{g(g_i)},\, g(g_i) \in G\big\}\,. \tag{9.25}$$

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 (9.21) by $|W|$ to remove multiple counting so that

$$\mathrm{d}\mu(\theta) = \frac{1}{|W|\,V_H}\, \mathrm{d}^n x\, \delta^{\hat{n}}\big(P(x)\big)\big|\det M(\theta)\big|\,, \tag{9.26}$$

[58] Ludvig Dmitrievich Faddeev, 1934-2017, Russian. Viktor Nikolaevich Popov, 1937-1994, Russian.

9.1 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 (9.1), where $x = (x^1, \dots, x^n)$, we then consider

$$f(x) = F(x^2)\,, \tag{9.27}$$

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

$$x_0 = r(1, 0, \dots, 0, 0)\,, \qquad r > 0\,. \tag{9.28}$$

In this case the stability group, for all $r > 0$, $H \simeq SO(n-1)$ since matrices leaving x_0 in (9.28) invariant have the form

$$R(\hat{R}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{R} \end{pmatrix}, \qquad \hat{R} \in SO(n-1)\,. \tag{9.29}$$

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 (9.28) the corresponding gauge fixing condition, corresponding to $\delta^{\hat{n}}(P(x))$, is

$$\mathcal{F}(x) = \theta(x^1) \prod_{i=2}^{n} \delta(x^i)\,. \tag{9.30}$$

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 (6.31) we have

$$S_{s1} x_0 = r(0, \dots, \underbrace{1}_{s'\text{th place}}, \dots, 0)\,, \qquad s = 2, \dots, n\,, \tag{9.31}$$

so that in (9.13) we may take

$$\alpha = \sum_{s=2}^{n} \alpha_s S_{s1}\,, \tag{9.32}$$

so that

$$\exp(\alpha) x_0 = r(1, \alpha_2, \dots, \alpha_n) \qquad \text{for} \quad \alpha \approx 0\,. \tag{9.33}$$

For the measure we assume a normalisation such that

$$\mathrm{d}\rho_{SO(n)}(R) \approx \mathrm{d}^{n-1}\alpha \, \mathrm{d}\rho_{SO(n-1)}(\hat{R}) \quad \text{for} \quad R = \exp(\alpha) R(\hat{R})\,, \quad \alpha \approx 0\,, \tag{9.34}$$

where $R(\hat{R})$ is given in (9.29). With the gauge fixing function in (9.30)

$$\int_{SO(n)} \mathrm{d}\rho_{SO(n)}(R)\, \mathcal{F}(Rx) = V_{SO(n-1)} \int \mathrm{d}^{n-1}\alpha \prod_{s=2}^{n} \delta\big(\alpha_s |x^1|\big) = V_{SO(n-1)} \frac{1}{|x^1|^{n-1}} . \tag{9.35}$$

Hence

$$\Delta(x) = \frac{1}{V_{SO(n-1)}}\, r^{n-1} , \qquad x^2 = r^2 , \quad r > 0 . \tag{9.36}$$

With this (9.24) becomes

$$\int_{\mathbb{R}^n} \mathrm{d}^n x\, F(x^2) = V_{SO(n)} \int_{\mathbb{R}^n} \mathrm{d}^n x\, \mathcal{F}(x)\, \Delta(x)\, F(x^2) = \frac{V_{SO(n)}}{V_{SO(n-1)}} \int_0^\infty \mathrm{d}r\, r^{n-1} F(r^2) . \tag{9.37}$$

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

$$\int_{\mathbb{R}^n} \mathrm{d}^n x\, e^{-x^2} = \pi^{\frac{1}{2}n} = \frac{V_{SO(n)}}{V_{SO(n-1)}} \int_0^\infty \mathrm{d}r\, r^{n-1} e^{-r^2} = \frac{V_{SO(n)}}{V_{SO(n-1)}}\, \tfrac{1}{2}\, \Gamma(\tfrac{1}{2}n) , \tag{9.38}$$

giving

$$\frac{V_{SO(n)}}{V_{SO(n-1)}} = S_n = \frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)} , \tag{9.39}$$

where S_n is the volume of S^{n-1}. Since $V_{SO(2)} = 2\pi$, or $V_{SO(1)} = 1$, in general

$$V_{SO(n)} = 2^{n-1} \frac{\pi^{\frac{1}{4}n(n+1)}}{\prod_{i=1}^{n} \Gamma\big(\frac{1}{2} i\big)} . \tag{9.40}$$

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

$$\int_{\mathbb{C}^n} \mathrm{d}^{2n} Z\, F(\bar{Z}Z) , \qquad \bar{Z}Z = \sum_{i=1}^{n} |z_i|^2 , \tag{9.41}$$

and where

$$\mathrm{d}^{2n} Z = \prod_{i=1}^{n} \mathrm{d}^2 z_i , \qquad \mathrm{d}^2 z = \mathrm{d}x \mathrm{d}y \quad \text{for} \quad z = x + iy . \tag{9.42}$$

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

$$Z_0 = r(1, 0, \dots, 0, 0) , \qquad r > 0 . \tag{9.43}$$

The stability group $H \simeq U(n-1)$ corresponding to matrices

$$U(\hat{U}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{U} \end{pmatrix} , \qquad \hat{U} \in U(n-1) . \tag{9.44}$$

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 (9.43) the gauge fixing condition becomes

$$\mathcal{F}(Z) = \theta(\operatorname{Re} z_1)\delta(\operatorname{Im} z_1)\prod_{i=2}^{n}\delta^2(z_i)\,, \qquad \delta^2(z) = \delta(x)\delta(y)\,, \quad z = x+iy\,. \tag{9.45}$$

In terms of the generators defined in (6.2) we let

$$\alpha = i\alpha_1 R^1{}_1 + \sum_{s=2}^{n}\left(\alpha_s R^s{}_1 - \alpha_s{}^* R^1{}_s\right), \qquad \alpha_1 \in \mathbb{R}\,, \quad \alpha_s \in \mathbb{C}\,,\ s \geq 2\,. \tag{9.46}$$

Hence

$$\exp(\alpha)Z_0 = r(1+i\alpha_1, \alpha_2, \dots, \alpha_n)\,, \qquad \alpha \approx 0\,, \tag{9.47}$$

and we take

$$\mathrm{d}\rho_{U(n)}(U) \approx \mathrm{d}\alpha_1 \textstyle\prod_{s=2}^{n} \mathrm{d}^2\alpha_s\ \mathrm{d}\rho_{U(n-1)}(\hat{U}) \quad \text{for} \quad U = \exp(\alpha)U(\hat{U})\,, \quad \alpha \approx 0\,. \tag{9.48}$$

With these results

$$\int \mathrm{d}\rho_{U(n)}(U)\ \mathcal{F}(UZ) = V_{U(n-1)}\,\frac{1}{|z_1|^{2n-1}}\,, \tag{9.49}$$

which implies

$$\Delta(Z) = \frac{1}{V_{U(n-1)}}\,r^{2n-1}\,, \qquad \bar{Z}Z = r^2\,, \quad r > 0\,. \tag{9.50}$$

Finally

$$\int_{\mathbb{C}^n}\mathrm{d}^{2n}Z\ F(\bar{Z}Z) = V_{U(n)}\int_{\mathbb{C}^n}\mathrm{d}^{2n}Z\ \mathcal{F}(Z)\,\Delta(Z)\,F(\bar{Z}Z) = \frac{V_{U(n)}}{V_{U(n-1)}}\int_0^\infty \mathrm{d}r\ r^{2n-1}F(r^2)\,. \tag{9.51}$$

Corresponding to (9.39), (9.51) requires

$$\frac{V_{U(n)}}{V_{U(n-1)}} = S_{2n}\,. \tag{9.52}$$

Taking $V_{U(1)} = 2\pi$ we have, with our normalisation,

$$V_{U(n)} = 2^n\,\frac{\pi^{\frac{1}{2}n(n+1)}}{\prod_{i=1}^{n}\Gamma(i)}\,. \tag{9.53}$$

Since $U(n) \simeq SU(n)\times U(1)/\mathbb{Z}_n$

$$V_{U(n)} = \frac{2\pi}{n}\,V_{SU(n)}\,. \tag{9.54}$$

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

$$\int_{\mathbb{H}^n}\mathrm{d}^{4n}Q\ F(\bar{Q}Q)\,, \qquad \bar{Q}Q = \sum_{i=1}^{n}|q_i|^2\,, \tag{9.55}$$

and where

$$\mathrm{d}^{4n}Q = \prod_{i=1}^{n} \mathrm{d}^4 q_i \,, \qquad \mathrm{d}^4 q = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v \quad \text{for} \quad z = x + iy + ju + kv \,. \tag{9.56}$$

$\bar{Q}Q$ 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 (1.121). As before we choose

$$Q_0 = r(1,0,\dots,0,0) \,, \qquad r > 0 \,. \tag{9.57}$$

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 (9.44). 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

$$\mathcal{F}(Q) = \theta(\mathrm{Re}\, q_1)\delta^3(\mathrm{Im}\, q_1) \prod_{i=2}^{n} \delta^4(q_i) \,, \qquad \delta^4(q) = \delta(x)\delta(y)\delta(u)\delta(v) \,, \quad q = x + iy + iu + iv \,. \tag{9.58}$$

In terms of the generators defined in (6.2) we let

$$\alpha = \alpha_1 R^1{}_1 + \sum_{s=2}^{n} \left(\alpha_s R^s{}_1 - \overline{\alpha_s} R^1{}_s \right) , \qquad \alpha_s \in \mathbb{H} \,, \quad \mathrm{Re}\,\alpha_1 = 0 \,. \tag{9.59}$$

and

$$\mathrm{d}\rho_{Sp(n)}(M) \approx \mathrm{d}^3\alpha_1 \textstyle\prod_{s=2}^{n} \mathrm{d}^4\alpha_s \, \mathrm{d}\rho_{Sp(n-1)}(\hat{M}) \qquad \alpha \approx 0 \,. \tag{9.60}$$

Hence we find

$$\Delta(Q) = \frac{1}{V_{Sp(n-1)}} \, r^{4n-1} \,, \qquad \bar{Q}Q = r^2 1 \,, \quad r > 0 \,. \tag{9.61}$$

The integral in (9.55) becomes

$$\int_{\mathbb{H}^n} \mathrm{d}^{4n}Q \, F(\bar{Q}Q) = V_{Sp(n)} \int_{\mathbb{H}^n} \mathrm{d}^{4n}Q \, \mathcal{F}(Q)\, \Delta(Q)\, F(\bar{Q}Q) = \frac{V_{Sp(n)}}{V_{Sp(n-1)}} \int_0^\infty \mathrm{d}r \, r^{4n-1} F(r^2) \,, \tag{9.62}$$

and corresponding to (9.39), (9.62) requires

$$\frac{V_{Sp(n)}}{V_{Sp(n-1)}} = S_{4n} \,. \tag{9.63}$$

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

$$V_{Sp(1)} = \int \mathrm{d}^4 q \; \delta(|q| - 1) = S_4 = 2\pi^2 \,, \tag{9.64}$$

just as in (5.155). Hence

$$V_{Sp(n)} = 2^n \, \frac{\pi^{n(n+1)}}{\prod_{i=1}^{n} \Gamma(2i)} \,. \tag{9.65}$$

The results for the group volumes in (9.40), (9.53) and (9.65) depend on the conventions adopted in the normalisation of the group invariant integration measure which are here determined by (9.34), (9.48) and (9.60) in conjunction with (9.32), (9.46) and (9.59) respectively.

9.2 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

$$\int \mathrm{d}^{\frac{1}{2}n(n+1)}X\, f(X)\,, \quad X = X^T\,, \quad d^{\frac{1}{2}n(n+1)}X = \prod_{i=1}^{n} \mathrm{d}X_{ii} \prod_{1 \le i < j \le n} \mathrm{d}X_{ij}\,, \tag{9.66}$$

and we assume the invariance

$$f(X) = f(RXR^{-1})\,, \qquad R \in SO(n)\,. \tag{9.67}$$

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

$$RXR^{-1} = \Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}, \tag{9.68}$$

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

$$\dim\{X : X = X^T\} - \dim SO(n) = \tfrac{1}{2}n(n+1) - \tfrac{1}{2}n(n-1) = n\,, \tag{9.69}$$

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 (9.67) we may write

$$f(X) = \hat{f}(\lambda)\,, \qquad \lambda = (\lambda_1, \dots \lambda_n)\,. \tag{9.70}$$

However there is a discrete stability group for Λ. The diagonal matrices corresponding to reflections in the i-direction

(9.71)

$$R_i = i \begin{array}{c} \qquad\qquad i \\ \begin{pmatrix} \begin{smallmatrix}1&0\\0&1\end{smallmatrix} & \cdots\cdots\cdots & & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \begin{smallmatrix}1&&\\&-1&\\&&1\end{smallmatrix} & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots\cdots\cdots & & \begin{smallmatrix}1&0\\0&1\end{smallmatrix} \end{pmatrix} \end{array} \in O(n)\,, \quad i = 1, \dots, n\,, \qquad R_i^2 = I_n\,. \tag{9.72}$$

generate the discrete group

$$\{R_{a_1 \dots a_n} = R_1{}^{a_1} \dots R_{n-1}{}^{a_n} : a_i = 0, 1\} \simeq \mathbb{Z}_2{}^{\times n}\,, \tag{9.73}$$

such that for any element

$$R_{a_1 \dots a_n} \Lambda R_{a_1 \dots a_n}{}^{-1} = \Lambda\,. \tag{9.74}$$

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

$$
R_\sigma \Lambda R_\sigma{}^{-1} = \Lambda_\sigma = \begin{pmatrix} \lambda_{\sigma(1)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_{\sigma(n)} \end{pmatrix} . \tag{9.75}
$$

For the permutation $(i\,i{+}1)$ the associated matrix is

$$
\begin{array}{rl} & \qquad\qquad\qquad i \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (9.76) \\ R_{(i\,i+1)} = i & \begin{pmatrix} {\scriptstyle\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}} & \cdots\cdots\cdots & & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & {\scriptstyle\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}} & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots\cdots\cdots & & {\scriptstyle\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}} \end{pmatrix} , \qquad i = 1, \dots, n-1 , \end{array} \tag{9.77}
$$

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 $\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n}$. However $\det R_\sigma = P_\sigma = \pm 1$, with P_σ defined in (7.85), 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 (9.25),

$$
W = \big(\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n}\big)/\mathbb{Z}_2 , \qquad |W| = 2^{n-1} n! . \tag{9.78}
$$

It is possible to restrict W to $\mathbb{Z}_2{}^{\times n-1}$, formed by $R_{a_1 \dots a_n}$ with $\sum_i a_i$ even, by requiring that the eigenvalues in (9.68) are ordered so that $\lambda_i \le \lambda_{i+1}$. However the choice of W in (9.78) is generally more convenient.

Taking $X_0 = \Lambda$, as in (9.68), the corresponding gauge fixing condition is

$$
\mathcal{F}(X) = \prod_{1 \le i < j \le n} \delta(X_{ij}) . \tag{9.79}
$$

For a rotation

$$
R(\alpha) = \exp(\alpha) , \qquad \alpha = -\alpha^T , \tag{9.80}
$$

and the group invariant integration is then assumed to be normalised such that, for R as in (9.80),

$$
\mathrm{d}\rho_{SO(n)}\big(R(\alpha)\big) \approx \prod_{1 \le i < j \le n} \mathrm{d}\alpha_{ij} , \qquad \alpha \approx 0 . \tag{9.81}
$$

With these assumptions, applying (9.9),

$$
\frac{1}{\Delta(X)} = \int_{SO(n)} \mathrm{d}\rho_{SO(n)}(R)\, \mathcal{F}\big(RXR^{-1}\big) = \prod_{1 \le i < j \le n} \int \mathrm{d}\alpha_{ij}\, \delta\big(\alpha_{ij}(\lambda_j - \lambda_i)\big) , \tag{9.82}
$$

so that

$$
\Delta(X) = |\hat{\Delta}(\lambda)| \qquad \text{for} \qquad \hat{\Delta}(\lambda) = \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j) . \tag{9.83}
$$

The resulting $SO(n)$ invariant integration over symmetric matrices becomes

$$\begin{aligned}\int \mathrm{d}^{\frac{1}{2}n(n+1)}X\, f(X) &= \frac{V_{SO(n)}}{|W|}\int \mathrm{d}^{\frac{1}{2}n(n+1)}X\, \mathcal{F}(X)\Delta(X)\, f(X)\\ &= \frac{V_{SO(n)}}{2^{n-1}n!}\int \mathrm{d}^n\lambda\, |\hat{\Delta}(\lambda)|\, \hat{f}(\lambda)\,. \end{aligned}\tag{9.84}$$

Since the normalisations chosen in (9.80) and (9.81) are compatible with those assumed previously we may use (9.40) for $V_{SO(n)}$.

For the particular example

$$f(X) = e^{-\frac{1}{2}\kappa\,\mathrm{tr}(X^2)}\,, \qquad \mathrm{tr}(X^2) = \sum_{i=1}^{n} X_{ii}{}^2 + 2\sum_{1\le i<j\le n} X_{ij}{}^2 = \sum_{i=1}^{n}\lambda_i{}^2\,, \tag{9.85}$$

then

$$\int \mathrm{d}^{\frac{1}{2}n(n+1)}X\, e^{-\frac{1}{2}\kappa\,\mathrm{tr}(X^2)} = 2^{\frac{1}{2}n}\Big(\frac{\pi}{\kappa}\Big)^{\frac{1}{4}n(n+1)}\,. \tag{9.86}$$

Using (9.40) this defines a normalised probability measure for the eigenvalues for a Gaussian ensemble of symmetric real matrices

$$\mathrm{d}\mu(\lambda)_{\text{symmetric matrices}} = \frac{\kappa^{\frac{1}{4}n(n+1)}}{2^{\frac{3}{2}n}\prod_{i=1}^{n}\Gamma\big(1+\frac{1}{2}i\big)}\prod_{i=1}^{n}\mathrm{d}\lambda_i\, |\hat{\Delta}(\lambda)|\, e^{-\frac{1}{2}\kappa\sum_i \lambda_i{}^2}\,. \tag{9.87}$$

There is a corresponding discussion for complex hermitian $n\times n$ matrices when the integrals are of the form

$$\int \mathrm{d}^{n^2}X\, f(X)\,, \quad X = X^\dagger\,, \quad d^{n^2}X = \prod_{i=1}^{n}\mathrm{d}X_{ii}\prod_{1\le i<j\le n}\mathrm{d}^2X_{ij}\,, \tag{9.88}$$

where f satisfies

$$f(X) = f(UXU^{-1})\,, \qquad U\in U(n)\,. \tag{9.89}$$

Just as before hermitian matrices may be diagonalised

$$UXU^{-1} = \Lambda\,, \tag{9.90}$$

where the diagonal elements of Λ are the eigenvalues of X as in (9.68). In this case there is a non trivial continuous subgroup of $U(n)$ leaving Λ invariant formed by the diagonal matrices

$$U_0(\beta) = \begin{pmatrix} e^{i\beta_1} & 0 & \dots & 0\\ 0 & e^{i\beta_2} & & \vdots\\ \vdots & & \ddots & \vdots\\ 0 & \dots\dots\dots & & e^{i\beta_n}\end{pmatrix}, \tag{9.91}$$

and hence we may identify

$$H \simeq U(1)^{\times n}\,. \tag{9.92}$$

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

$$\mathcal{F}(X) = \prod_{1\le i<j\le n} \delta^2(X_{ij})\,. \tag{9.93}$$

In this case we may write for arbitrary $U \in U(n)$,

$$U(\alpha,\beta) = \exp(\alpha)U_0(\beta)\,, \qquad \alpha = -\alpha^\dagger\,, \quad \alpha_{ii} = 0 \text{ all } i\,, \tag{9.94}$$

and the group invariant integration is then assumed to be normalised such that, for U as in (9.94),

$$\mathrm{d}\rho_{U(n)}\big(U(\alpha,\beta)\big) \approx \prod_{1\le i<j\le n} \mathrm{d}^2\alpha_{ij} \prod_{i=1}^n \mathrm{d}\beta_i\,, \qquad \alpha\approx 0\,,\ 0\le\beta_i<2\pi\,. \tag{9.95}$$

With these assumptions

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(U)\, \mathcal{F}\big(UXU^{-1}\big) = (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^2\alpha_{ij}\, \delta^2\big(\alpha_{ij}(\lambda_j-\lambda_i)\big)\,. \tag{9.96}$$

Since

$$\delta^2(\lambda z) = \frac{1}{|\lambda|^2}\,\delta^2(z)\,, \tag{9.97}$$

this gives

$$\Delta(X) = \frac{1}{(2\pi)^n} \prod_{1\le i<j\le n} (\lambda_i-\lambda_j)^2 = \frac{1}{(2\pi)^n}\,\hat\Delta(\lambda)^2\,. \tag{9.98}$$

The result for $U(n)$ invariant integration over hermitian matrices becomes

$$\begin{aligned}\int \mathrm{d}^{n^2}X\, f(X) &= \frac{V_{U(n)}}{n!}\int \mathrm{d}^{n^2}X\,\mathcal{F}(X)\Delta(X)\,f(X)\\ &= \frac{V_{U(n)}}{n!\,(2\pi)^n}\int \mathrm{d}^n\lambda\,\hat\Delta(\lambda)^2\,\hat f(\lambda)\,,\end{aligned} \tag{9.99}$$

where we may use (9.53) for $V_{U(n)}$.

For a Gaussian function

$$\int \mathrm{d}^{n^2}X\, e^{-\frac12\kappa\,\mathrm{tr}(X^2)} = 2^{\frac12 n}\Big(\frac{\pi}{\kappa}\Big)^{\frac12 n^2}\,, \qquad \mathrm{tr}(X^2) = \sum_{i=1}^n X_{ii}{}^2 + 2\sum_{1\le i<j\le n}|X_{ij}|^2 = \sum_{i=1}^n \lambda_i^2\,. \tag{9.100}$$

Using (9.53) this defines a normalised probability measure for the eigenvalues for a Gaussian ensemble of hermitian matrices

$$\mathrm{d}\mu(\lambda)_{\text{hermitian matrices}} = \frac{\kappa^{\frac12 n^2}}{(2\pi)^{\frac12 n}\,\prod_{i=1}^n i!}\prod_{i=1}^n \mathrm{d}\lambda_i\,\hat\Delta(\lambda)^2\,e^{-\frac12\kappa\sum_i\lambda_i^2}\,. \tag{9.101}$$

Extending this to quaternionic hermitian $n\times n$ matrices the relevant integrals are

$$\int \mathrm{d}^{n(2n-1)}X\, f(X)\,, \quad X=\bar X\,, \quad d^{n(2n-1)}X = \prod_{i=1}^n \mathrm{d}X_{ii}\prod_{1\le i<j\le n}\mathrm{d}^4X_{ij}\,, \tag{9.102}$$

where $\bar{X}$ is defined by (1.114) and integration over quaternions is given by (9.56). f is now assumed to satisfy

$$f(X) = f(MXM^{-1})\,, \qquad M \in U(n,\mathbb{H})\,. \tag{9.103}$$

Such quaternionic matrices may be diagonalised so that, for a suitable $M \in U(n,\mathbb{H})$,

$$MXM^{-1} = \Lambda\,, \qquad \lambda_i = \bar{\lambda}_i\,. \tag{9.104}$$

Using the correspondence with $2n \times 2n$ complex matrices provided by (1.115) and (1.79), when $M \to \mathcal{M} \in USp(2n,\mathbb{C})$ and $X \to \mathcal{X}$ where

$$\mathcal{X} = \mathcal{X}^\dagger\,, \qquad \mathcal{X} = -J\mathcal{X}^T J\,. \tag{9.105}$$

The eigenvalues of $\mathcal{X}$ must then be $\pm\lambda_i$, $i = 1,\dots n$, and (9.104) 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 $\{\lambda_i\}$,

$$\mathcal{M}\mathcal{X}J\mathcal{M}^T = \begin{pmatrix} \begin{smallmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{smallmatrix} & 0 & & \\ 0 & \begin{smallmatrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{smallmatrix} & & \\ & & \ddots & \\ & & & \begin{smallmatrix} 0 & \lambda_n \\ -\lambda_n & 0 \end{smallmatrix} \end{pmatrix} \quad \text{for} \;\; \mathcal{M} \in U(2n)\,. \tag{9.106}$$

In (9.104) the subgroup of $U(n,\mathbb{H})$ leaving Λ invariant is formed by quaternionic matrices

$$M_0(q) = \begin{pmatrix} q_1 & 0 & \dots & 0 \\ 0 & q_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & q_n \end{pmatrix}, \qquad |q_i| = 1\,, \tag{9.107}$$

giving

$$H \simeq U(1,\mathbb{H})^{\times n}\,. \tag{9.108}$$

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

$$\mathcal{F}(X) = \prod_{1\le i<j\le n} \delta^4(X_{ij})\,. \tag{9.109}$$

In this case we may write for arbitrary $M \in U(n,\mathbb{H})$,

$$M(\alpha,q) = \exp(\alpha)M_0(q)\,, \qquad \alpha = -\bar{\alpha}\,, \quad \alpha_{ii} = 0 \text{ all } i\,, \tag{9.110}$$

and the group invariant integration is then assumed to be normalised such that, for M as in (9.110),

$$\mathrm{d}\rho_{U(n,\mathbb{H})}\big(M(\alpha,q)\big) \approx \prod_{1\le i<j\le n} \mathrm{d}^4\alpha_{ij} \prod_{i=1}^{n} \mathrm{d}^4 q_i\, \delta(|q_i|-1)\,, \qquad \alpha \approx 0\,. \tag{9.111}$$

With these assumptions and using (9.64)

$$\int_{U(n,\mathbb{H})} \mathrm{d}\rho_{U(n,\mathbb{H})}(M)\, \mathcal{F}\big(MXM^{-1}\big) = (2\pi^2)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4\alpha_{ij}\, \delta^4\big(\alpha_{ij}(\lambda_j-\lambda_i)\big)\,. \tag{9.112}$$

In this case

$$\Delta(X) = \frac{1}{(2\pi^2)^n} \prod_{1\le i<j\le n} (\lambda_i - \lambda_j)^4 = \frac{1}{(2\pi^2)^n}\, \hat{\Delta}(\lambda)^4 \,. \tag{9.113}$$

The result for $U(n,\mathbb{H})$ invariant integration over quaternion hermitian matrices becomes

$$\begin{aligned} \int \mathrm{d}^{n(2n-1)}X\, f(X) &= \frac{V_{Sp(n)}}{n!} \int \mathrm{d}^{n^2}X\, \mathcal{F}(X)\Delta(X)\, f(X) \\ &= \frac{V_{Sp(n)}}{n!\,(2\pi^2)^n} \int \mathrm{d}^n\lambda\, \hat{\Delta}(\lambda)^4\, \hat{f}(\lambda)\,, \end{aligned} \tag{9.114}$$

where we may use (9.65) for $V_{Sp(n)}$.

For the Gaussian integral

$$\int \mathrm{d}^{n(2n-1)}X\, e^{-\frac{1}{2}\kappa\, \mathrm{tr}(X^2)} = 2^{\frac{1}{2}n}\Big(\frac{\pi}{\kappa}\Big)^{\frac{1}{2}n(2n-1)}\,, \qquad \mathrm{tr}(X^2) = \sum_{i=1}^n {X_{ii}}^2 + 2\sum_{1\le i<j\le n} |X_{ij}|^2 = \sum_{i=1}^n {\lambda_i}^2\,. \tag{9.115}$$

Using (9.65) we therefore obtain a normalised probability measure for the eigenvalues for a Gaussian ensemble of hermitian quaternionic matrices

$$\mathrm{d}\mu(\lambda)_{\text{hermitian quaternionic matrices}} = \Big(\frac{2}{\pi}\Big)^{\frac{1}{2}n} \frac{\kappa^{\frac{1}{2}n(2n-1)}}{\prod_{i=1}^n (2i)!} \prod_{i=1}^n \mathrm{d}\lambda_i\, \hat{\Delta}(\lambda)^4\, e^{-\frac{1}{2}\kappa\sum_i {\lambda_i}^2}\,. \tag{9.116}$$

9.2.1 Large n Limits

The results for the eigenvalue measure $\mathrm{d}\mu(\lambda)$, given by (9.87), (9.101) and (9.116) 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

$$\mathrm{d}\mu(\lambda) = N_n\, \mathrm{d}^n\lambda\, e^{-W(\lambda)}\,, \qquad W(\lambda) = \frac{1}{2}\kappa\sum_i {\lambda_i}^2 - \tfrac{1}{2}\beta \sum_{i,j,i\ne j} \ln|\lambda_i - \lambda_j|\,, \tag{9.117}$$

where $\beta = 1,2,4$ and we may order the the eigenvalues so that

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n\,. \tag{9.118}$$

For a minimum $W(\lambda)$ is stationary when

$$\kappa\,\lambda_i = \beta\sum_{j\ne i} \frac{1}{\lambda_i - \lambda_j}\,. \tag{9.119}$$

In the large n limit we may approximate λ_i by a smooth function,

$$\lambda_i \to \lambda(x)\,, \quad x = \frac{i}{n}\,, \qquad \sum_{i=1}^n = n\int_0^1 \mathrm{d}x = n\int \mathrm{d}\lambda\, \rho(\lambda)\,, \quad \rho(\lambda) = \frac{\mathrm{d}x}{\mathrm{d}\lambda} > 0\,, \tag{9.120}$$

where $\rho(\lambda)$ determines the eigenvalue distribution and is normalised since

$$\int \mathrm{d}\lambda\, \rho(\lambda) = \int_0^1 \mathrm{d}x = 1\,. \tag{9.121}$$

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 (9.119) or, taking the large n limit,

$$\frac{\kappa}{n\beta}\,\lambda = P\int \mathrm{d}\mu\, \rho(\mu)\,\frac{1}{\lambda-\mu}\,, \tag{9.122}$$

where P denotes that the principal part prescription is used for the singularity in the integral at $\mu = \lambda$.

(9.122) is an integral equation for ρ. To solve this we define the function

$$F(z) = \int_{-R}^{R} \mathrm{d}\mu\, \rho(\mu)\,\frac{1}{z-\mu} \sim \frac{1}{z} \quad \text{as} \quad z \to \infty\,, \tag{9.123}$$

using (9.121) and assuming

$$\rho(\mu) > 0\,, \quad |\mu| < R\,, \qquad \rho(\mu) = 0\,, \quad |\mu| > R\,. \tag{9.124}$$

$F(z)$ is analytic in z save for a cut along the real axis from $-R$ to R. The integral equation requires

$$F(\mu \pm i\epsilon) = \frac{\kappa}{n\beta}\,\mu \mp i\pi\,\rho(\mu)\,, \qquad |\mu| < R\,. \tag{9.125}$$

Requiring $F(z) = \mathrm{O}(z^{-1})$ for large z this has the unique solution

$$F(z) = \frac{\kappa}{n\beta}\big(z - \sqrt{z^2 - R^2}\big)\,. \tag{9.126}$$

The large z condition in (9.123) requires

$$R^2 = \frac{2n\beta}{\kappa}\,. \tag{9.127}$$

This then gives

$$\rho(\lambda) = \frac{2}{\pi R^2}\sqrt{R^2 - \lambda^2}\,. \tag{9.128}$$

This is Wigner's semi-circle distribution and is relevant for nuclear energy levels.

9.3 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

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(U)\, f(U)\,, \tag{9.129}$$

where

$$f(U) = f(VUV^{-1}) \qquad \text{for all} \qquad V \in U(n)\,. \tag{9.130}$$

Just as for hermitian matrices U can be diagonalised so that

$$VUV^{-1} = U_0(\theta)\,, \qquad \theta = (\theta_1, \dots, \theta_n)\,, \tag{9.131}$$

where U_0 is defined in (9.91). 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 $H = U(1)^{\times n}$. The remaining discrete symmetry group in this case is then

$$W_{U(n)} \simeq \mathcal{S}_n\,, \tag{9.132}$$

since, for any permutation $\sigma \in \mathcal{S}_n$, there is a $R_\sigma \in O(n)$ such that

$$R_\sigma U_0(\theta) R_\sigma{}^{-1} = U_0(\theta_\sigma)\,, \qquad \theta_\sigma = (\theta_{\sigma(1)}, \dots, \theta_{\sigma(n)})\,. \tag{9.133}$$

Thus we use the gauge fixing condition

$$\mathcal{F}(U) = \prod_{1 \le i < j \le n} \delta^2(U_{ij})\,. \tag{9.134}$$

Using the same results as given in (9.94) and (9.95) we then get

$$\int_{U(n)} \mathrm{d}\rho_{U(n)}(V)\, \mathcal{F}\big(VUV^{-1}\big) = (2\pi)^n \prod_{1 \le i < j \le n} \int \mathrm{d}^2\alpha_{ij}\, \delta^2\big(\alpha_{ij}(e^{i\theta_j} - e^{i\theta_i})\big)\,, \tag{9.135}$$

so that, using (9.97),

$$\begin{aligned} \Delta(U) = \frac{1}{(2\pi)^n} \prod_{1 \le i < j \le n} |e^{i\theta_j} - e^{i\theta_i}|^2 &= \frac{1}{(2\pi)^n} \prod_{1 \le i < j \le n} \big(2 \sin \tfrac{1}{2}(\theta_i - \theta_j)\big)^2 \\ &= \frac{1}{(2\pi)^n}\, \hat{\Delta}(e^{i\theta}) \hat{\Delta}(e^{-i\theta})\,, \end{aligned} \tag{9.136}$$

with the definition (9.83). The basic formula (9.26) then gives an integration measure over the θ_i's

$$\mathrm{d}\mu_{U(n)}(\theta) = \frac{1}{n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \prod_{1 \le i < j \le n} \big(2 \sin \tfrac{1}{2}(\theta_i - \theta_j)\big)^2\,, \qquad 0 \le \theta_i \le 2\pi\,. \tag{9.137}$$

By restricting $f(U) = 1$ in (9.129) 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 $U(n) \simeq SU(n) \times U(1)/\mathbb{Z}_n$, 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 (9.132), we still have

$$W_{SU(n)} \simeq \mathcal{S}_n \,. \tag{9.138}$$

Restricting (9.137) to $SU(n)$ we then obtain

$$\mathrm{d}\mu_{SU(n)}(\theta) = \frac{1}{n!\,(2\pi)^{n-1}} \prod_{i=1}^{n-1} \mathrm{d}\theta_i \prod_{1\le i<j\le n} \big(2\sin\tfrac{1}{2}(\theta_i - \theta_j)\big)^2 \,, \qquad \theta_n = -\sum_{i=1}^{n-1} \theta_i \,. \tag{9.139}$$

For real orthogonal matrices in a similar fashion

$$\int_{SO(n)} \mathrm{d}\rho_{SO(n)}(R)\, f(R) \,, \qquad f(R) = f(SRS^{-1}) \quad \text{for all} \quad S \in SO(n) \,. \tag{9.140}$$

In this case it is necessary to distinguish between even and odd n. For any $R \in SO(2n)$ it can be transformed to

$$SRS^{-1} = R_0(\theta) = \begin{pmatrix} r(\theta_1) & 0 & \dots & 0 \\ 0 & r(\theta_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots & & r(\theta_n) \end{pmatrix} , \qquad S \in SO(2n) \,, \tag{9.141}$$

where $R_0(\theta)$ is written as a $n \times n$ matrix of 2×2 blocks with

$$r(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} . \tag{9.142}$$

In (9.141) $S \sim SR_0(\beta)$, for arbitrary $\beta = (\beta_1, \dots, \beta_n)$, so that the stability group for $R_0(\theta)$ is then $SO(2)^{\times n}$. The discrete group defined by $2n \times 2n$ matrices $\{S\} \in O(2n)$ such that $SR_0(\theta)S^{-1} = R_0(\theta')$ is $\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n}$, with the permutation group $\mathcal{S}_n$ formed by $\{R_\sigma \times \mathbb{1}_2\}$ and $\mathbb{Z}_2{}^{\times n}$ generated by

$$R_i = \begin{array}{cc} & \begin{matrix} \phantom{\mathbb{1}_2} & & \phantom{\mathbb{1}_2} & i & \phantom{\mathbb{1}_2} & & \phantom{\mathbb{1}_2} \end{matrix} \\ i & \begin{pmatrix} \mathbb{1}_2 & \dots & & & & & 0 \\ \vdots & \ddots & & & & & \vdots \\ & & \mathbb{1}_2 & & & & \\ & & & \sigma_3 & & & \\ & & & & \mathbb{1}_2 & & \\ & & & & & \ddots & \\ 0 & \dots & & & & & \mathbb{1}_2 \end{pmatrix} \end{array} \in O(2n) \,, \quad i = 1, \dots, n \,, \qquad R_i^2 = \mathbb{1}_{2n} \,, \tag{9.143}$$

since $\sigma_3\, r(\theta)\, \sigma_3 = r(2\pi - \theta)$, for $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Restricting to the subgroup formed my matrices with determinant one

$$W_{SO(2n)} \simeq \big(\mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n}\big)/\mathbb{Z}_2 \,. \tag{9.144}$$

Writing $R \in SO(2n)$ in terms of 2×2 blocks R_{ij}, $i, j = 1, \dots n$, the gauge fixing condition is then taken as

$$\mathcal{F}(R) = \prod_{1\le i<j\le n} \delta^4(R_{ij}) \,, \tag{9.145}$$

with the definitions

$$\delta^4(A) = \delta(a)\delta(b)\delta(c)\delta(d)\,, \quad \mathrm{d}^4 A = \mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c\,\mathrm{d}d \qquad \text{for} \qquad A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right). \tag{9.146}$$

For a general rotation $S \in SO(2n)$ we may write

$$S = e^A S_0(\beta)\,, \qquad A^T = -A\,, \qquad A_{ii} = 0 \ \text{ all } i\,, \tag{9.147}$$

and then

$$\mathrm{d}\rho_{SO(2n)}(S) \approx \prod_{1\le i<j\le n} \mathrm{d}^4 A_{ij} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad A \approx 0\,. \tag{9.148}$$

Using (9.148) is then sufficient to obtain

$$\int_{SO(2n)} \mathrm{d}\rho_{SO(2n)}(S)\; \mathcal{F}\big(S R_0(\theta) S^{-1}\big) = (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4 A_{ij}\; \delta^4\big(A_{ij} r(\theta_j) - r(\theta_i) A_{ij}\big)\,. \tag{9.149}$$

With

$$\delta^4\big(A r(\theta) - r(\theta') A\big) = \frac{1}{4(\cos\theta - \cos\theta')^2}\, \delta^4(A)\,, \tag{9.150}$$

we then get for $SO(2n)$

$$\Delta(R) = \frac{1}{(2\pi)^n} \prod_{1\le i<j\le n} \big(2(\cos\theta_i - \cos\theta_j)\big)^2 = \frac{1}{(2\pi)^n} \big(\hat{\Delta}(2\cos\theta)\big)^2\,, \tag{9.151}$$

where $\hat{\Delta}$ is defined by (9.83).

Combining the ingredients the measure for integration reduces in the $SO(2n)$ case to an integral over the n θ_i's given by

$$\mathrm{d}\mu_{S0(2n)}(\theta) = \frac{1}{2^{n-1} n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \,\big(\hat{\Delta}(2\cos\theta)\big)^2\,. \tag{9.152}$$

For $SO(2n+1)$ (9.141) may be modified, by introducing one additional row and column, to

$$S R S^{-1} = R_0(\theta) = \begin{pmatrix} r(\theta_1) & 0 & \dots & 0 & 0 \\ 0 & r(\theta_2) & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & & & r(\theta_n) & 0 \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}, \qquad S \in SO(2n+1)\,, \tag{9.153}$$

with $r(\theta)$ just as in (9.142). Instead of (9.143) we may now take

$$R_i = \begin{array}{c} \\ i \end{array} \begin{array}{c} \begin{array}{ccccccc} & & & i & & & \end{array} \\ \begin{pmatrix} \mathbb{1}_2 & \dots & \dots & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & & & & \vdots \\ \vdots & & \mathbb{1}_2 & & & & \vdots \\ \vdots & & & \sigma_3 & & & \vdots \\ \vdots & & & & \mathbb{1}_2 & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & & & & & \mathbb{1}_2 & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & -1 \end{pmatrix} \end{array} \in SO(2n+1)\,, \quad i = 1, \dots, n\,, \tag{9.154}$$

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

$$W_{SO(2n+1)} \simeq \mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n} \,. \tag{9.155}$$

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

$$\mathcal{F}(R) = \prod_{1\le i<j\le n} \delta^4(R_{ij}) \prod_{i=1}^{n} \delta^2(R_{i\,n+1}) \,, \tag{9.156}$$

with $\delta^2\big(\begin{smallmatrix} a \\ b \end{smallmatrix}\big) = \delta(a)\,\delta(b)$, similarly to (9.146). Expressing $S \in SO(2n+1)$ in the same form as (9.147) we now have

$$\mathrm{d}\rho_{SO(2n+1)}(S) \approx \prod_{1\le i<j\le n} \mathrm{d}^4 A_{ij} \prod_{i=1}^{n} \mathrm{d}^2 A_{i\,n+1} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad A \approx 0 \,, \tag{9.157}$$

so that

$$\begin{aligned} &\int_{SO(2n+1)} \mathrm{d}\rho_{SO(2n+1)}(S)\, \mathcal{F}\big(S R_0(\theta) S^{-1}\big) \\ &= (2\pi)^n \prod_{1\le i<j\le n} \int \mathrm{d}^4 A_{ij}\, \delta^4\big(A_{ij} r(\theta_j) - r(\theta_i) A_{ij}\big) \prod_{i=1}^{n} \int \mathrm{d}^2 A_{i\,n+1}\, \delta^2\big((I_2 - r(\theta_i)) A_{i\,n+1}\big) \,. \end{aligned} \tag{9.158}$$

In the $SO(2n+1)$ case this implies

$$\Delta(R) = \frac{1}{(2\pi)^n} \big(\hat{\Delta}(2\cos\theta)\big)^2 \prod_{i=1}^{n} \big(2\sin\tfrac{1}{2}\theta_i\big)^2 \,, \tag{9.159}$$

and in consequence

$$\mathrm{d}\mu_{S0(2n+1)}(\theta) = \frac{1}{2^n n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \big(2\sin\tfrac{1}{2}\theta_i\big)^2 \big(\hat{\Delta}(2\cos\theta)\big)^2 \,. \tag{9.160}$$

The remaining case to consider is for integrals over $M \in Sp(n) \simeq U(n,\mathbb{H})$ of the form

$$\int_{Sp(n)} \mathrm{d}\rho_{Sp(n)}(M)\, f(M) \,, \qquad f(M) = f(NMN^{-1}) \quad \text{for all} \quad N \in Sp(n) \,. \tag{9.161}$$

By a suitable transformation the quaternion matrix M can be reduced to the diagonal form

$$NMN^{-1} = M_0(\theta) = \begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & e^{i\theta_n} \end{pmatrix} , \qquad N \in Sp(n) \,, \tag{9.162}$$

As before $N \sim NM_0(\beta)$ so the stability group is $U(1)^{\times n}$. 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

$$N_i = \begin{array}{cc} & \begin{array}{ccccc} & & i & & \end{array} \\ i & \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & 1 & & & \vdots \\ \vdots & & j & & \vdots \\ \vdots & & & 1 & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix} \end{array} \in Sp(n) \, , \quad i = 1, \dots, n \, . \tag{9.163}$$

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

$$W_{Sp(n)} \simeq \mathcal{S}_n \ltimes \mathbb{Z}_2{}^{\times n} \, . \tag{9.164}$$

For the $Sp(n)$ case we take

$$\mathcal{F}(M) = \prod_{1 \le i < j \le n} \delta^4(M_{ij}) \prod_{i=1}^{n} \delta^2(M_{ii}) \, , \tag{9.165}$$

where, for any quaternion q, $\delta^4(q)$ is defined as in (9.58) and also here

$$\delta^2(q) = \delta(u)\delta(v) \qquad \text{for} \qquad q = x + iy + ju + kv \, . \tag{9.166}$$

Writing then, for any $N \in Sp(n)$,

$$N = e^{\alpha} M_0(\beta) \, , \qquad \alpha_{ij} = -\bar{\alpha}_{ji} \in \mathbb{H} \, , \; i \neq j \, , \quad \alpha_{ii} = ju_i + kv_i \, , \tag{9.167}$$

we have

$$\mathrm{d}\rho_{Sp(n)}(N) \approx \prod_{1 \le i < j \le n} \mathrm{d}^4 \alpha_{ij} \prod_{i=1}^{n} \mathrm{d}^2 \alpha_{ii} \prod_{i=1}^{n} \mathrm{d}\beta_i \qquad \text{for} \qquad \alpha \approx 0 \, , \tag{9.168}$$

so that

$$\begin{aligned} &\int_{Sp(n)} \mathrm{d}\rho_{Sp(n)}(N) \; \mathcal{F}\big(N M_0(\theta) N^{-1}\big) \\ &= (2\pi)^n \prod_{1 \le i < j \le n} \int \mathrm{d}^4 \alpha_{ij} \; \delta^4\big(\alpha_{ij} \, e^{i\theta_j} - e^{i\theta_i} \, \alpha_{ij}\big) \prod_{i=1}^{n} \int \mathrm{d}^2 \alpha_{ii} \; \delta^2\big(\alpha_{ii} \, e^{i\theta_i} - e^{i\theta_i} \, \alpha_{ii}\big) \, . \end{aligned} \tag{9.169}$$

For this case we may use

$$\begin{aligned} \delta^4\big(\alpha \, e^{i\theta} - e^{i\theta'} \, \alpha\big) &= \frac{1}{4(\cos\theta - \cos\theta')^2} \, \delta^4\big(\alpha\big) \, , \\ \delta^2\big(\alpha \, e^{i\theta} - e^{i\theta} \, \alpha\big) &= \frac{1}{4\sin^2\theta} \, \delta^2\big(\alpha\big) \quad \text{for} \quad \alpha = ju + kv \, , \end{aligned} \tag{9.170}$$

to obtain

$$\Delta(M) = \frac{1}{(2\pi)^n} \, \big(\hat{\Delta}(2\cos\theta)\big)^2 \prod_{i=1}^{n} \big(2\sin\theta_i\big)^2 \, . \tag{9.171}$$

Hence

$$\mathrm{d}\mu_{Sp(n)}(\theta) = \frac{1}{2^n n!\,(2\pi)^n} \prod_{i=1}^{n} \mathrm{d}\theta_i \left(2\sin\theta_i\right)^2 \left(\hat{\Delta}(2\cos\theta)\right)^2 . \tag{9.172}$$

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 $SO(4) \simeq (Sp(1)\times Sp(1))/\mathbb{Z}_2$, $\mathrm{d}\mu_{SO(4)}(\theta_1-\theta_2,\theta_1+\theta_2) = 2\,\mathrm{d}\mu_{Sp(1)}(\theta_1)\,\mathrm{d}\mu_{Sp(1)}(\theta_2)$ 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 $SO(6) \simeq SU(4)/\mathbb{Z}_2$, $\mathrm{d}\mu_{SO(6)}(\theta_2+\theta_3,\theta_3+\theta_1,\theta_1+\theta_2) = 2\,\mathrm{d}\mu_{SU(4)}(\theta_1,\theta_2,\theta_3)$.

9.4 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 (8.26),

$$A(t) \underset{g}{\longrightarrow} A(t)^{g(t)} = g(t)A(t)g(t)^{-1} - \partial_t g(t)\,g(t)^{-1} , \tag{9.173}$$

where here we take

$$A(t) = -A(t)^\dagger \in \mathfrak{u}(n) , \qquad g(t) \in U(n) . \tag{9.174}$$

The essential functional integral has the form

$$\int \mathrm{d}[A]\, f(A) , \qquad f(A^g) = f(A) , \tag{9.175}$$

where we restrict to $t \in S^1$ by requiring the fields to satisfy the periodicity conditions

$$A(t) = A(t+\beta) , \qquad g(t) = g(t+\beta) . \tag{9.176}$$

In one dimension there are no local gauge the discussion in 8.3 and the periodicity requirement (9.176), the gauge invariant function f in (9.175) should have the form

$$f(A) = \hat{f}(U) \qquad \text{where} \qquad \hat{f}(U) = \hat{f}(gUg^{-1}) \quad \text{for all} \quad g \in U(n) . \tag{9.177}$$

In particular

$$P_\beta(U) = \mathrm{tr}(U) , \tag{9.178}$$

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*[59] *loop*.

The general discussion for finite group invariant integrals can be directly applied to the functional integral (9.175). It is necessary to choose a convenient gauge fixing condition. For any $A(t)$ there is a gauge transformation $g(t)$ such that

$$A(t)^{g(t)} = iX , \qquad X^\dagger = X . \tag{9.179}$$

In consequence we may choose a gauge condition $\partial_t A(t) = 0$ or equivalently take

$$\mathcal{F}[A] = \delta'[A] , \tag{9.180}$$

[59] Alexander Markovich Polyakov, 1945-, Russian.

where $\delta'[A]$ is a functional δ-function, δ' denoting the exclusion of constant modes. For a general Fourier expansion on S^1

$$A(t) = iX + \sum_{n\neq 0} A_n\, e^{2\pi i n t/\beta}\,, \qquad X^\dagger = X\,, \quad {A_n}^\dagger = -A_{-n}\,, \tag{9.181}$$

where X is a hermitian and A_n are complex $n\times n$ matrices, then

$$\delta'[A] = \prod_{n>0} \mathcal{N}_n\, \delta^{2n^2}\big(A_n\big)\,. \tag{9.182}$$

$\mathcal{N}_n$ is a normalisation factor which is chosen later. With the expansion (9.181) the functional integral can also be defined by taking

$$\mathrm{d}[A] = \mathrm{d}^{n^2} X \prod_{n>0} \frac{1}{\mathcal{N}_n}\, \mathrm{d}^{2n^2} A_n\,. \tag{9.183}$$

The integral (9.9) defining the Faddeev Popov determinant then becomes

$$\int_{\mathcal{G}} \mathrm{d}\mu(g)\; \delta'[A^g] \quad \text{where} \quad A(t) = (iX)^{g(t)} \quad \text{for some} \quad g(t)\,, \tag{9.184}$$

and where $\mathrm{d}\mu(g)$ is the invariant measure for the gauge group $\mathcal{G}$. From (9.173) for an infinitesimal gauge transformation

$$(iX)^{g(t)} = iX + i[\lambda(t), X] - \partial_t \lambda(t) \qquad \text{for} \quad g(t) \approx \mathbb{1} + \lambda(t)\,, \quad \lambda(t)^\dagger = -\lambda(t)\,. \tag{9.185}$$

If

$$g(t) = g_0\big(\mathbb{1} + \lambda(t)\big) \quad \text{for} \quad \lambda(t)\approx 0\,, \quad \lambda(t) = \sum_{n\neq 0} \lambda_n\, e^{2\pi i n t/\beta}\,, \quad {\lambda_n}^\dagger = -\lambda_{-n}\,, \tag{9.186}$$

then we may take

$$\mathrm{d}\mu(g) \approx \mathrm{d}\rho_{U(n)}(g_0)\; \mathrm{d}[\lambda]\,, \qquad \mathrm{d}[\lambda] = \prod_{n>0} \mathrm{d}^{2n^2}\lambda_n\,. \tag{9.187}$$

Hence from (9.184) we define

$$\int_{\mathcal{G}} \mathrm{d}\mu(g)\; \delta'\big[(iX)^g\big] = \frac{V_{U(n)}}{\Delta(X)}\,, \tag{9.188}$$

where

$$\begin{aligned} \frac{1}{\Delta(X)} &= \int \mathrm{d}[\lambda]\; \delta'\big[i[\lambda, X] - \partial_t\lambda\big] = \prod_{n>0} \mathcal{N}_n \int \mathrm{d}^{2n^2}\lambda_n\; \delta^{2n^2}\bigg(\frac{2\pi n}{i\beta}\lambda_n - i[X,\lambda_n]\bigg) \\ &= \prod_{n>0} \int \mathrm{d}^{2n^2}\lambda_n\; \delta^{2n^2}\bigg(\lambda_n + \frac{\beta}{2\pi n}[X,\lambda_n]\bigg) \quad \text{for} \quad \mathcal{N}_n = \bigg(\frac{\beta}{2\pi n}\bigg)^{2n^2}\,, \end{aligned} \tag{9.189}$$

which gives

$$\Delta(X) = \prod_{n>0} \bigg(\det\Big(\mathbb{1}_{n^2} + \frac{\beta}{2\pi n}\, X^{\mathrm{ad}}\Big)\bigg)^2\,. \tag{9.190}$$

The essential functional integral in (9.175) then reduces to just an integral over hermitian matrices X,

$$\int \mathrm{d}[A]\, f(A) = \frac{1}{V_{U(n)}} \int \mathrm{d}^{n^2}X\, \Delta(X)\, f(iX)\,. \tag{9.191}$$

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 (9.68). In terms of these

$$\text{eigenvalues}\{X^{\mathrm{ad}}\} = \lambda_i - \lambda_j\,, \quad i,j = 1, \dots, n\,. \tag{9.192}$$

Hence

$$\det\Big(\mathbb{1}_{n^2} + \frac{\beta}{2\pi n} X^{\mathrm{ad}}\Big) = \prod_{1 \le i < j \le n} \Big(1 - \frac{(\lambda_i - \lambda_j)^2 \beta^2}{4\pi^2 n^2}\Big). \tag{9.193}$$

Using

$$\prod_{n>0} \Big(1 - \frac{\theta^2}{\pi^2 n^2}\Big) = \frac{\sin\theta}{\theta}\,, \tag{9.194}$$

we get

$$\Delta(X) = \prod_{1 \le i < j \le n} \left(\frac{\sin \frac{1}{2}(\lambda_i - \lambda_j)\beta}{\frac{1}{2}(\lambda_i - \lambda_j)\beta}\right)^2. \tag{9.195}$$

As a consequence of (9.99) we further express (9.191) in terms of an integral over the eigenvalues $\{\lambda_i\}$ using

$$\frac{1}{V_{U(n)}} \int \mathrm{d}^{n^2}X \to \frac{1}{n!\,(2\pi)^n} \int \mathrm{d}^n\lambda \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j)^2\,. \tag{9.196}$$

Using this in conjunction (9.195) in (9.191) gives finally

$$\int \mathrm{d}[A]\, f(A) = \frac{1}{\beta^{n^2}} \int \mathrm{d}\mu_{U(n)}(\beta\lambda)\, f(i\Lambda)\,, \tag{9.197}$$

with the measure for integration over $U(n)$ determined by (9.137).

Although the freedom of constant gauge transformations has been used in transforming $X \to \Lambda$ there is also a residual gauge freedom given by

$$g(t) = e^{2\pi i r t/\beta}\mathbb{1}\,, \quad r = 0, \pm 1, \pm 2, \dots \qquad \Rightarrow \qquad \Lambda^{g(t)} = \Lambda - \frac{2\pi r}{\beta}\mathbb{1}\,. \tag{9.198}$$

For this to be a symmetry for $f(iX) = f(i\Lambda)$ we must have

$$f(iX) = \hat{f}(e^{-i\beta X})\,, \tag{9.199}$$

where $\hat{f}$ is defined in terms of the line integral over t in (9.177). The final result (9.197) shows that the functional integral over $A(t)$ reduces after gauge fixing just to invariant integration over the unitary matrix U.