Group Theory

Prof. J. J. Binney

Oxford University

Books

Most of the material of this course is covered in *Symmetry in Physics* by J.P. Elliott & P.G. Dawber (MacMillan). Unfortunately, the material is spread over both volumes of this two-volume work, although the majority will be found in Vol. 1.

There are useful sections on group theory in several compendia on mathematical physics, for example *The Mathematics of Physics and Chemistry* by H. Margenau & G.M. Murphy (Van Nostrand Reinhold).

1 Basic Concepts

1.1 Group axioms

A group is a set $\mathcal{G}$ of elements $a, b, \ldots$ and a binary operation 'multiplication' such that

1. ab is in $\mathcal{G}$ $\forall$ a, b;
2. $\exists$ an identity element e such that $ae = e$ $\forall$ a;
3. multiplication is associative: $(ab)c = a(bc)$ $\forall$ a, b, c;
4. each element a as an inverse a^{-1} such that $aa^{-1} = e$

The number of elements g is the **order** of $\mathcal{G}$. g may not be finite.

Note:

We also have $a^{-1}a = e$: Let x be the inverse of a^{-1}, i.e. $a^{-1}x = e$. Then $a^{-1}a = (a^{-1}a)e = a^{-1}(aa^{-1})x = a^{-1}x = e$. Similarly associativity ensures that $ae = e \quad \Leftrightarrow \quad ea = e$.

Example 1.1

The integers $\ldots, -2, -1, 0, 1, \ldots$ form a group of infinite order under addition with 0 the identity.

Example 1.2

Under addition the integers $0, \ldots, N-1$ modulo $N > 0$ for a group. It is called $\mathcal{C}_N$, the **cyclic group** of order N.

Example 1.3

The complex numbers $\mathrm{e}^{2\pi \mathrm{i} r/N}$ with $0 \leq r < N$ form the group $\mathcal{C}_N$ under multiplication.

Example 1.4

The rotations which bring a regular N-gon into coincidence with itself form the group $\mathcal{C}_N$. In the case $N = 3$ the group has 3 members, say r_1, r_2 and e, being rotations by $2\pi/3$, $4\pi/3$ and 0 radians, respectively. The multiplication table is then

	e	r_1	r_2
e	e	r_1	r_2
r_1	r_1	r_2	e
r_2	r_2	e	r_1

Notes:

(i) In any multiplication table, every row and column must contain every element exactly once.

(ii) The table above is symmetric because $\mathcal{C}_N$ is an **Abelian group**, i.e., $r_i r_j = r_j r_i \quad \forall \quad i, j$.

Example 1.5

A **permutation** of $N > 0$ objects $x_1, \ldots, x_N$ is any 1–1 mapping of the set $\{x_i\}$ onto itself. A convenient notation for a permutation is

$$p = \begin{pmatrix} 1, & 2, & \ldots, & N \\ x_1, & x_2, & \ldots, & x_N \end{pmatrix}.$$

The **symmetric group** $\mathcal{S}_N$ is the group formed by all such permutations.

Exercise (1):

Show that[1]

$$\begin{pmatrix} 1,2,3 \\ 2,1,3 \end{pmatrix} \begin{pmatrix} 1,2,3 \\ 3,2,1 \end{pmatrix} = \begin{pmatrix} 1,2,3 \\ 3,1,2 \end{pmatrix}.$$

[1] Here the permutation executed first is that on the right. Some authors use the opposite convention.

Example 1.6

The rotations in and out of the plane which bring a regular N-gon into coincidence with itself form the **dihedral group** $\mathcal{D}_N$. Label the vertices $1 - N$ as you go anti-clockwise around the N-gon. Then vertex 1 can be moved to the site of any vertex, and the remaining vertices can be laid down either anti-clockwise or clockwise. So the order of $\mathcal{D}_N$ is $g = 2N$. $\mathcal{C}_N$ is clearly embedded within $\mathcal{D}_N$ as a **subgroup**, i.e., a subset of elements that forms a group by itself.

Exercise (2):

Construct the multiplication table of $\mathcal{D}_3$ and hence show that this group is not Abelian.

$\mathcal{C}_N$ and $\mathcal{D}_N$ are both clearly subgroups of $\mathcal{S}_N$. This is a specific example of a wide-ranging theorem:[2]

Theorem 1 (Cayley)

Every group is (isomorphic to) a permutation group.

Proof: Associate with each element b of $\mathcal{G}$ the permutation

$$p_b \equiv \begin{pmatrix} a_1, & a_2, & \ldots, & a_N \\ ba_1, & ba_2, & \ldots, & ba_N \end{pmatrix}.$$

Then the permutation $p_b p_c$ maps $a_j \to ca_j \to bca_j$. It is therefore equal to p_{bc}. □

1.2 Generators

Most groups may be **generated** by a small subset of elements. For example, $\{-1, 1\}$ generates the integers, while $\{-2, 2\}$ generates the subset of the even integers. $\mathrm{e}^{2\pi \mathrm{i}/3}$ generates $\mathcal{C}_3$.

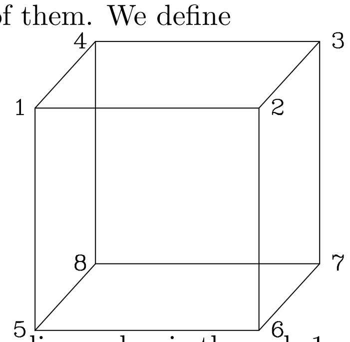
Exercise (3):

What is the smallest set that generates $\mathcal{D}_3$?

Example 1.7

We can construct the symmetry group $\mathcal{O}_h$ of the cube (the **full octahedral group**) by labeling the 8 corners 1–8 and then considering permutations of them. We define

$$a \equiv \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 2, & 3, & 4, & 1, & 6, & 7, & 8, & 5 \end{pmatrix}$$
$$b \equiv \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 1, & 4, & 8, & 5, & 2, & 3, & 7, & 6 \end{pmatrix}$$
$$c \equiv \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 5, & 6, & 7, & 8, & 1, & 2, & 3, & 4 \end{pmatrix}$$



a rotates by $\pi/2$ about a vertical axis, b rotates by $2\pi/3$ about the diagonal axis through 1, and c is a reflection about the horizontal mid-plane. We shall see that $\{a, b, c\}$ generates $\mathcal{O}_h$

1.3 Cosets & classes

Let $\mathcal{H}$ be a subgroup of $\mathcal{G}$. Then for each $a \in \mathcal{G}$ the set of elements $\mathcal{H}a$ that can be obtained by multiplying a member of $\mathcal{H}$ by a is called a **left coset** of $\mathcal{H}$ in $\mathcal{G}$. The **right coset** $a\mathcal{H}$ is similarly defined. We have

[2] It is conventional to allow isomorphic groups to be distinct. But if we are to distinguish between representations and underlying abstract groups, it seems natural to regard isomorphic groups as two representations of a single group.

Theorem 2

Two left cosets of $\mathcal{H}$ in $\mathcal{G}$ are either identical or disjoint. Similarly, two right cosets of $\mathcal{H}$ in $\mathcal{G}$ are either identical or disjoint.

Proof: Suppose $h_i a = h_j b$. Then $a = h_i^{-1} h_j b$, so for any $h_k \in \mathcal{H}$ we have $h_k a = h_k h_i^{-1} h_j b = h_l b$, from which it follows that $\mathcal{H}a \subseteq \mathcal{H}b$. Similarly, $\mathcal{H}b \subseteq \mathcal{H}a$, so $\mathcal{H}a = \mathcal{H}b$. □

Exercise (4):

Show that for each left coset $\mathcal{H}a$ there is a right coset $b\mathcal{H}$ comprising the inverses of the elements of $\mathcal{H}a$. This result establishes a 1-1 relationship between left and right cosets.

The number of elements in a coset is clearly the order of $\mathcal{H}$. So we have

Theorem 3 (Lagrange)

The order of a subgroup $\mathcal{H}$ of $\mathcal{G}$ is a factor of the order of $\mathcal{G}$.

The following shows that most groups have proper subgroups $\mathcal{H} \subset \mathcal{G}$:

Theorem 4

Let $\mathcal{G}$ be a group of order greater than one. Then if $\mathcal{G}$ has no subgroup other than itself and e, then $\mathcal{G}$ is a finite cyclic group of prime order.

Proof: Consider the element $b \neq e$. Then the set $\{e, b, b^2, b^3, \ldots\}$ forms a cyclic group and this must coincide with $\mathcal{G}$. If this group were of infinite order, the set $\{e, b^2, b^4, \ldots\}$ would form a proper subgroup, contrary to hypothesis. If the finite order of $\mathcal{G}$ had a factor r, then the set $\{b^r, b^{2r}, \ldots\}$ would form a proper subgroup, contrary to hypothesis. □

Example 1.8

The elements a, b of the full octahedral group defined in Example 7 generate the **octahedral group** $\mathcal{O}$, a subgroup of $\mathcal{O}_h$ that takes any vertex into any other:

$$1 \xrightarrow{a} 2 \xrightarrow{a} 3 \xrightarrow{a} 4 \xrightarrow{b} 5 \xrightarrow{a} 6 \xrightarrow{a} 7 \xrightarrow{a} 8 \xrightarrow{a} 5 \xrightarrow{b} 2 \xrightarrow{a^3} 1.$$

The elements e, b, b^2 that leave 1 in place form a subgroup $\mathcal{H}_1$ of $\mathcal{O}$. We can decompose $\mathcal{O}$ into cosets by multiplying $\mathcal{H}_1$ by elements that move 1 successively to positions 2,3,...:

$$\mathcal{O} = \mathcal{H}_1 + a\mathcal{H}_1 + a^2\mathcal{H}_1 + a^3\mathcal{H}_1 + ba^3\mathcal{H}_1 + aba^3\mathcal{H}_1 + a^2ba^3\mathcal{H}_1 + a^3ba^3\mathcal{H}_1$$

Thus the order of $\mathcal{O}$ is $3 \times 8 = 24$. The reflection c cannot be expressed as a product of a and b. But $c^2 = 1$, $ac = ca$ and $bc = ca^2ba^2$, so the group generated by $\{a, b, c\}$ can be decomposed into $\mathcal{O} + c\mathcal{O}$. This group is actually the whole of $\mathcal{O}_h$,

$$\mathcal{O}_h = \mathcal{O} + c\mathcal{O},$$

so $\mathcal{O}_h$ is of order 48.

If $\exists\ x \in \mathcal{G}$ such that $a = x^{-1}bx$, then a and b are said to be **conjugate** elements. This is obviously a symmetric relation and it is reflexive: $a = e^{-1}ae$. It is, moreover, also transitive: if b and c are conjugate, then $\exists\ y$ s.t. $b = y^{-1}cy$. So $a = x^{-1}y^{-1}cyx = (yx)^{-1}c(yx)$. The **class** $S(a)$ is the set of all mutually conjugate elements that contains a. We shall see that in practical applications it is often enough to know what class an element belongs to.

Notes:

(i) e is always in a class by itself. This class is denoted E.

(ii) In an Abelian group, every element constitutes a class.

Example 1.9

The dihedral group $\mathcal{D}_3$ has three classes: (i) $\{e\}$, (ii) the set of rotations in the triangle's plane, (iii) rotations whose axis lies in the plane. One denotes these classes by E, $2C_3$ and $3C_2$.

Exercise (5):

Prove the above statements about $\mathcal{D}_3$ by showing, for example, that $r_2 = p_1^{-1} r_1 p_1$, where r_n represents rotation by $2n\pi/3$ and p_n is the rotation by π that leaves vertex n undisplaced.

Example 1.10

The classes of the octahedral group $\mathcal{O}$ (see Example 8) are denoted E, $8C_3$, $3C_2$, $6C_2$, $6C_4$. That all three-fold rotations about a diagonal form a class, follows because, in the notation of Example 8, (i) any rotation about a dagonal b' satisfies either $b' = x^{-1}bx$ or $b' = x^{-1}b^2x$, where x is a rotation that brings to 1 the vertex about which b' rotates and (ii) $b = y^{-1}b^2y$, where y is reflection in the $1-3-7-5$ plane. There are four diagonals and two three-fold rotations around each, so this class contains 8 elements. Similar arguments show why two-fold rotations about each of three axes through the middles of faces form a class, as do the two-fold rotations about each of six axes through the mid-points of edges etc.

Given two groups $\mathcal{G}_1$ and $\mathcal{G}_2$ we can form a third group $\mathcal{G}_3 \equiv \mathcal{G}_1 \times \mathcal{G}_2$: the members of $\mathcal{G}_3$ are of the pairs $A \equiv (a_1, a_2)$, where $a_i \in \mathcal{G}_i$, $i = 1, 2$ and the product of two elements of $\mathcal{G}_3$ is defined by $AB = (a_1b_1, a_2b_2)$. $\mathcal{G}_3$ is the **direct product** of $\mathcal{G}_1$ and $\mathcal{G}_2$.

Example 1.11

The full octahedral group $\mathcal{O}_h$ is the direct product of the octahedral group $\mathcal{O}$ and the group $\mathcal{S}_2$ formed by e and the element c of Example 7.

Exercise (6):

Show that the classes of $\mathcal{G}_1 \times \mathcal{G}_2$ are simply pairs (p_1, p_2) of classes p_i of $\mathcal{G}_i$.

2 Representations

A representation of a group $\mathcal{G}$ is a mapping of $\mathcal{G}$ into the set of linear transformations T of a vector space V, that preserves the group's product structure. Thus a representation assigns to each $a \in \mathcal{G}$ a linear transformation $T(a)$ such that $T(a)T(b) = T(ab) \quad \forall \quad a, b \in \mathcal{G}$.

Once we have chosen a coordinate system for V, each point in V effectively becomes a position vector $\mathbf{v}$, and each linear transformation of V becomes a matrix $\mathbf{T}$:

$$\mathbf{v} \to \mathbf{v}' = \mathbf{T} \cdot \mathbf{v} \quad \text{or in component form} \quad v_i' = \sum_j T_{ij} v_j. \tag{2.1}$$

Preservation of the group product implies that

$$\sum_j T_{ij}(a) T_{jk}(b) = T_{ik}(ab) \tag{2.2}$$

Example 2.1

The mapping of $\mathcal{G}$ need not be 1–1; the **trivial representation** simply maps every element of $\mathcal{G}$ into the unit matrix $\mathbf{I}$. A 1–1 representation is called a **faithful representation**.

Example 2.2

The matrix that rotates the (x, y) plane by $\pi/2$ is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so this is a 2-dimensional representation of $\mathcal{C}_4$:

$$\begin{gathered} e \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ r_1 \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad r_2 = r_1^2 \to \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad r_3 = r_1^3 \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \end{gathered} \tag{2.3}$$

2.1 Equivalent representations

Given one representation $\mathbf{T}(a)$, we can generate many others: let S be any invertible linear transformation. Then $a \to \mathbf{T}'(a) \equiv \mathbf{S} \cdot \mathbf{T}(a) \cdot \mathbf{S}^{-1}$ provides a representation:

$$\mathbf{T}'(a) \cdot \mathbf{T}'(b) = \mathbf{S} \cdot \mathbf{T}(a) \cdot \mathbf{S}^{-1} \cdot \mathbf{S} \cdot \mathbf{T}(b) \cdot \mathbf{S}^{-1} = \mathbf{S} \cdot \mathbf{T}(a) \cdot \mathbf{T}(b) \cdot \mathbf{S}^{-1} = \mathbf{T}'(ab).$$

Representations related as $\mathbf{T}(a)$ and $\mathbf{T}'(a)$ are, are called **equivalent representations**. Equivalence is a reflexive and transitive relation and so groups representations into classes. We concentrate on **unitary representations**, i.e., ones that map $\mathcal{G}$ into unitary matrices ($\mathbf{T}^\dagger = \mathbf{T}^{-1}$) because we have

Theorem 5 (Maschke)

Every class of equivalent representations of a finite group $\mathcal{G}$ contains unitary representations.

Proof: Let $\mathbf{S}^2 \equiv \sum_a \mathbf{T}^\dagger(a) \cdot \mathbf{T}(a)$; $\mathbf{S}^2$ is Hermitian, so it can be diagonalized, and its square root $\mathbf{S}$ defined as the (Hermitian) matrix with the same eigen-vectors and eigen-values that are the square roots of those of $\mathbf{S}^2$. Since $\mathbf{v}^\dagger \cdot \mathbf{S}^2 \cdot \mathbf{v} = \sum_a |\mathbf{T}(a) \cdot \mathbf{v}|^2 > 0 \;\forall\, \mathbf{v}$, no eigen-value of $\mathbf{S}^2$ can vanish and $\mathbf{S}$ has an inverse. Now

$$\begin{aligned} \mathbf{T}^\dagger(b) \cdot \mathbf{S}^2 \cdot \mathbf{T}(b) &= \sum_a \mathbf{T}^\dagger(b) \cdot \mathbf{T}^\dagger(a) \cdot \mathbf{T}(a) \cdot \mathbf{T}(b) \\ &= \mathbf{S}^2, \end{aligned} \tag{2.4}$$

since as a runs through $\mathcal{G}$, $\mathbf{T}(a) \cdot \mathbf{T}(b) = \mathbf{T}(ab)$ runs through all the matrices in the representation as surely as $\mathbf{T}(a)$ does. Pre- and post-multiplying (2.4) by $\mathbf{S}^{-1}$, we obtain

$$\left(\mathbf{S}^{-1} \cdot \mathbf{T}^\dagger(b) \cdot \mathbf{S}\right) \cdot \left(\mathbf{S} \cdot \mathbf{T}(b) \cdot \mathbf{S}^{-1}\right) = \mathbf{I} \tag{2.5}$$

which proves that $\mathbf{T}'(b) \equiv \mathbf{S} \cdot \mathbf{T}(b) \cdot \mathbf{S}^{-1}$ is unitary. □

2.2 Generation of inequivalent representations

Now consider a scheme for turning a representation into an inequivalent one. An obvious generalization of (2.3) yields a 2-dimensional representation of $\mathcal{C}_N$ (or, indeed, of $\mathcal{D}_N$):

$$r_k \to \mathbf{r}(k) \equiv \begin{pmatrix} \cos(2k\pi/N) & \sin(2k\pi/N) \\ -\sin(2k\pi/N) & \cos(2k\pi/N) \end{pmatrix}. \tag{2.6}$$

We can use these matrices to make coordinate changes rather than point transformations:

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix} \equiv \mathbf{r}(k) \begin{pmatrix} x \\ y \end{pmatrix}. \tag{2.7}$$

Now consider the effect of this coordinate transformation on a quadratic form

$$Q(x, y) \equiv a_{11}x^2 + 2a_{12}xy + a_{22}y^2. \tag{2.8}$$

Substituting primed for unprimed coordinates, we get

$$\begin{aligned} Q(x', y') &= a'_{11}x'^2 + 2a'_{12}x'y' + a'_{22}y'^2 \\ &= a_{11}(r_{11}x' + r_{12}y')^2 + 2a_{12}(r_{11}x' + r_{12}y')(r_{21}x' + r_{22}y') + a_{22}(r_{21}x' + r_{22}y')^2 \\ &= \left(a_{11}r_{11}^2 + 2a_{12}r_{11}r_{21} + a_{22}r_{21}^2\right)x'^2 \\ &\quad + 2\left(a_{11}r_{11}r_{12} + a_{12}(r_{11}r_{22} + r_{12}r_{21}) + a_{22}r_{21}r_{22}\right)x'y' \\ &\quad + \left(a_{11}r_{12}^2 + 2a_{12}r_{12}r_{22} + a_{22}r_{22}^2\right)y'^2, \end{aligned} \tag{2.9}$$

where r_{ij} is shorthand for $r_{ij}(-k)$. Comparing coefficients in the first and last lines, we see that our original matrix induces a linear transformation of the coefficients a_{ij}:

$$\mathbf{a} \to \mathbf{a}' \equiv \begin{pmatrix} a'_{11} \\ a'_{12} \\ a'_{22} \end{pmatrix} = \begin{pmatrix} r_{11}^2 & r_{11}r_{21} & r_{21}^2 \\ r_{11}r_{12} & r_{11}r_{22} + r_{12}r_{21} & r_{21}r_{22} \\ r_{12}^2 & r_{12}r_{22} & r_{22}^2 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{22} \end{pmatrix}. \tag{2.10}$$

To see that the transformation $\mathbf{a} \to \mathbf{a}'$ constitutes a representation of $\mathcal{G}$, we need a more abstract notation. We can write the quadratic form $Q(x, y)$ as $\mathbf{x}^\dagger \cdot \mathbf{Q} \cdot \mathbf{x}$, where

$$\mathbf{x} \equiv \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{Q} \equiv \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}. \tag{2.11}$$

Now

$$\begin{aligned} \mathbf{x}'^\dagger \cdot \mathbf{Q}' \cdot \mathbf{x}' &= \mathbf{x}^\dagger \cdot \mathbf{Q} \cdot \mathbf{x} \\ &= \mathbf{x}'^\dagger \cdot \mathbf{r}^\dagger(a^{-1}) \cdot \mathbf{Q} \cdot \mathbf{r}(a^{-1}) \cdot \mathbf{x}'. \end{aligned} \tag{2.12}$$

So we define an operator O on the matrices $\mathbf{Q}$ of quadratic forms like this:

$$\begin{aligned} \mathbf{Q}' = O(a)\mathbf{Q} &\equiv \mathbf{r}^\dagger(a^{-1}) \cdot \mathbf{Q} \cdot \mathbf{r}(a^{-1}) \\ &= \mathbf{r}(a) \cdot \mathbf{Q} \cdot \mathbf{r}^\dagger(a), \end{aligned} \tag{2.13}$$

where use has been made of the unitarity of $\mathbf{r}(a)$. Then on successively applying the linear transformations representing two group elements a, b we have

$$\begin{aligned} O(b)O(a)\mathbf{Q} &= \mathbf{r}(b) \cdot \mathbf{r}(a) \cdot \mathbf{Q} \cdot \mathbf{r}^\dagger(a) \cdot \mathbf{r}^\dagger(b) \\ &= \mathbf{r}(ba) \cdot \mathbf{Q} \cdot \mathbf{r}^\dagger(ba) \\ &= O(ba)\mathbf{Q}. \end{aligned} \tag{2.14}$$

which demonstrates that the action of O on the matrices $\mathbf{Q}$ preserves group multiplication and therefore provides a representation.

Exercise (7):

Prove that the representation (2.13) is unitary in the sense that $a_{11}^2 + 2a_{12}^2 + a_{22}^2$ is invariant.

It is easy to see that the scheme just described generalizes trivially to quadratic forms in any number of variables: starting from an n-dimensional representation, we can immediately derive a $\frac{1}{2}n(n+1)$-dimensional representation by considering transformations of quadratic forms of n variables.

Here's another scheme for generating new representations from old: suppose the $m \times m$ matrices $\mathbf{S}(a)$ provide one representation of $\mathcal{G}$, and the $n \times n$ matrices $\mathbf{T}(a)$ provide another representation. Then a third, mn-dimensional, representation is the **direct product representation** $U = S \times T$:

$$\mathbf{U}(a) \cdot \mathbf{x} \otimes \mathbf{y} = (\mathbf{S}(a) \cdot \mathbf{x}) \otimes (\mathbf{T}(a) \cdot \mathbf{y}). \tag{2.15}$$

Here $\mathbf{x}$ is an m-component vector and $\mathbf{y}$ is an n-component one and $\otimes$ stands for the 'tensor product' of the vectors belonging to the two spaces. (The tensor product of two vectors $\mathbf{u}$, $\mathbf{v}$ is the pair $(\mathbf{u}, \mathbf{v})$. If $\mathbf{u} = \sum_{i=1}^m u_i \hat{\mathbf{e}}_i^{(1)}$ lies in an m-dimensional space and $\mathbf{v} = \sum_{i=1}^n v_i \hat{\mathbf{e}}_i^{(2)}$ lies in an n-dimensional space, then $\mathbf{u} \otimes \mathbf{v}$ lies in the mn-dimensional space spanned by the vectors $(\hat{\mathbf{e}}_i^{(1)}, \hat{\mathbf{e}}_j^{(2)})$, with respect to which its components are the numbers $u_i v_j$.) This *is* a valid representation because we have

$$U_{(i,j)(k,l)}(a) = S_{ik}(a) T_{jl}(a) \tag{2.16}$$

and

$$
\begin{aligned}
\big(\mathbf{U}(a)\cdot\mathbf{U}(b)\big)_{(i,j)(p,q)} &= \sum_{kl} U_{(i,j)(k,l)}(a)U_{(k,l)(p,q)}(b) \\
&= \sum_{kl} S_{ik}(a)S_{kp}(b)T_{jl}(a)T_{lq}(b) \\
&= S_{ip}(ab)T_{jq}(ab) \\
&= U_{(i,j)(p,q)}(ab).
\end{aligned}
\tag{2.17}
$$

Example 2.3

Consider a representation of $\mathcal{G}$ by some 2×2 complex matrices $\mathbf{M}(a)$. These act on complex 2-vectors $\boldsymbol{\eta}$:

$$\boldsymbol{\eta}\to\boldsymbol{\eta}'\equiv\mathbf{M}(a)\cdot\boldsymbol{\eta} \tag{2.18a}$$

The complex conjugate of $\mathbf{M}$ yields a related representation

$$\boldsymbol{\eta}\to\boldsymbol{\eta}'=\mathbf{M}^*(a)\cdot\boldsymbol{\eta}. \tag{2.18b}$$

Consider the direct product representation $M\times M^*$. This acts on four-component vectors $\boldsymbol{\eta}\otimes\boldsymbol{\rho}$, which can be conveniently written in matrix form

$$\boldsymbol{\eta}\otimes\boldsymbol{\rho}=\begin{pmatrix}\eta_1\rho_1 & \eta_1\rho_2\\ \eta_2\rho_1 & \eta_2\rho_2\end{pmatrix}. \tag{2.19}$$

Applying a typical element of the $M\times M^*$ representation, we find

$$
\begin{aligned}
\boldsymbol{\eta}\otimes\boldsymbol{\rho}\to(\boldsymbol{\eta}\otimes\boldsymbol{\rho})' &= \begin{pmatrix}\eta_1'\rho_1' & \eta_1'\rho_2'\\ \eta_2'\rho_1' & \eta_2'\rho_2'\end{pmatrix} \\
\Rightarrow\quad (\boldsymbol{\eta}\otimes\boldsymbol{\rho})'_{ij} &= M_{ip}\eta_p M^*_{jq}\rho_q \\
&= \big(\mathbf{M}\cdot(\boldsymbol{\eta}\otimes\boldsymbol{\rho})\cdot\mathbf{M}^\dagger\big)_{ij}.
\end{aligned}
\tag{2.20}
$$

(Notice the similarity with equation (2.13).) The physical motivation of this example is this. Let

$$\mathbf{X}=\begin{pmatrix}ct+z & x-\mathrm{i}y\\ x+\mathrm{i}y & ct-z\end{pmatrix}\quad\Rightarrow\quad|\mathbf{X}|=c^2t^2-x^2-y^2-z^2. \tag{2.21}$$

Then $s^2\equiv c^2t^2-x^2-y^2-z^2$ is invariant under transformations of the form $\mathbf{X}\to\mathbf{X}'=\mathbf{M}\cdot\mathbf{X}\cdot\mathbf{M}^\dagger$ with $|\mathbf{M}|=1$. Moreover, we shall see that every Lorentz transformation can be associated with a different matrix $\mathbf{M}$, so this is a faithful representation of the Lorentz group. Thus there is a close connection between the Lorentz group and the group $SL(2,\mathcal{C})$ formed by the 2×2 matrices $\mathbf{M}$ with unit determinant.

Exercise (8):

Show that the subgroup $SU(2)$ of $SL(2,\mathcal{C})$ that comprises unitary matrices generates spatial rotations by the mechanism of Example 3. [Hint: consider $\mathrm{Tr}(\mathbf{X})$.]

2.3 Irreducible representations

Pick one vector $\mathbf{v}$ from the space U acted on by the representation $\mathbf{T}(a)$ and from it generate the set $\{\mathbf{T}(a)\cdot\mathbf{v}\}$ $a\in\mathcal{G}$. The vectors of this set span a vector space V. For most choices of $\mathbf{v}$, $V=U$. Sometimes, however, V will be of lower dimension than U. That is, V will be a subspace of U in the same way that the (x,y) plane is a 2-dimensional subspace of 3-dimensional (x,y,z) space.

For example, when the matrices of (2.10) act on the quadratic form $Q_0\equiv x^2+y^2$, the resulting polynomial Q_0' is always the same; that is, if $a_{11}=a_{22}=1$, $a_{12}=0$, then $a_{11}'=a_{22}'=1$, $a_{12}'=0$ no matter what matrix of the representation (2.10) we employ. We say that Q_0 forms an **invariant subspace** under the action of the group. In this case V is the 1-dimensional space spanned by Q_0.

Once an invariant subspace V of dimension $m < n$, say, has been detected, it is natural to use basis vectors such that the first m of these span V, and to complement these by $n-m$ basis vectors that are each perpendicular to every vector in V. These last vectors span what is called the **orthogonal complement** of V in U.

The matrices $\mathbf{T}(a)$ representing $\mathcal{G}$ are block diagonal in this basis: The first column of a matrix gives the image of the first basis vector $\hat{\mathbf{e}}_1 = (1, 0, \ldots)$, the second column gives the image of $\hat{\mathbf{e}}_2 = (0, 1, 0 \ldots)$ etc. Since $\hat{\mathbf{e}}_1$ through $\hat{\mathbf{e}}_m$ lie in the invariant subspace V, their images do too, so these must have vanishing components parallel to $\hat{\mathbf{e}}_{n-m+1}$ through $\hat{\mathbf{e}}_n$, that is, their last $n-m$ components vanish, and thus the bottom left-hand corner of every $\mathbf{T}$ is filled with zeros. Similarly, for $k > m$, $\hat{\mathbf{e}}_k$ lies in the orthogonal complement of V. Now let $j \leq m$ and $k > m$ and consider

$$\begin{aligned}\alpha \equiv \hat{\mathbf{e}}_j \cdot \mathbf{T}(a) \cdot \hat{\mathbf{e}}_k &= (\mathbf{T}^\dagger(a) \cdot \hat{\mathbf{e}}_j) \cdot \hat{\mathbf{e}}_k \\ &= (\mathbf{T}(a^{-1}) \cdot \hat{\mathbf{e}}_j) \cdot \hat{\mathbf{e}}_k,\end{aligned} \tag{2.22}$$

where we have exploited the unitarity of the representation. From the first equality it follows that α is the component along $\hat{\mathbf{e}}_j$ of the image under $\mathbf{T}(a)$ of $\hat{\mathbf{e}}_k$. The last equality shows that this vanishes. Hence the top right-hand corner of every matrix $\mathbf{T}$ is filled with zeros.

The block diagonal structure of $\mathbf{T}$ means that the n-dimensional representation provided by $\mathbf{T}(a)$ decomposes into two completely independent representations of dimension m and $n-m$. We write

$$T = T^{(1)} \dotplus T^{(2)} \tag{2.23}$$

to indicate this fact, where $T^{(1)}$ symbolizes the representation provided by the matrices that form the upper left blocks of each $\mathbf{T}$ (which transform V into itself), and $T^{(2)}$ symbolizes the representation provided by the matrices at lower right of each $\mathbf{T}$.

Sometimes the invariant subspace V can itself be broken down into a smaller invariant subspace and its orthogonal complement. Eventually, though, the original representation space U will have been broken into a series of **irreducible**, mutually orthogonal subspaces $V^{(1)}, \ldots, V^{(\alpha)}$, such that $V^{(i)}$ is spanned by $\{\mathbf{v}, \mathbf{T}(a_1) \cdot \mathbf{v}, \mathbf{T}(a_2) \cdot \mathbf{v}, \ldots\}$, where $\mathbf{v} \in V^{(i)}$ is *any* vector, and $a_1, \ldots$ is a sufficient set of elements of $\mathcal{G}$. The representation of $\mathcal{G}$ formed by the i^{th} blocks of the $\mathbf{T}(a)$ (which transform $V^{(i)}$ into itself) form an **irreducible representation** of $\mathcal{G}$, or an **irrep** for short.

Example 2.4

Obviously the 1-dimensional representation spanned by $Q_0(x, y) = x^2 + y^2$ forms an irrep of $\mathcal{C}_N$. It isn't a very interesting representation, being the trivial representation. The orthogonal complement of Q_0 consists of quadratic forms $Q(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ such that $a_{11} + a_{22} = 0$, i.e., ones of the form $Q = a_- x^2 + 2a_+ xy - a_- y^2$. Convenient basis vectors for this subspace are $Q_- \equiv x^2 - y^2$ and $Q_+ \equiv 2xy$. $Q_\pm$ forms an invariant subspace under the representation of $\mathcal{C}_4$ provided by the matrices of (2.10). But under the representation of $\mathcal{C}_3$ obtained on putting $N = 3$ into (2.6), they transform into each other:

$$\begin{aligned}\mathbf{r}_1 &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \mathbf{r}_2 = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \\ O_1 &= \begin{pmatrix} \frac{1}{4} & -\frac{\sqrt{3}}{4} & \frac{3}{4} \\ \frac{\sqrt{3}}{4} & -\frac{1}{2} & -\frac{\sqrt{3}}{4} \\ \frac{3}{4} & \frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix}, \quad O_2 = \begin{pmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{3}{4} \\ -\frac{\sqrt{3}}{4} & -\frac{1}{2} & \frac{\sqrt{3}}{4} \\ \frac{3}{4} & -\frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix}.\end{aligned} \tag{2.24}$$

Hence

$$\begin{aligned}O_1 Q_- &= -\tfrac{1}{2}x^2 + \sqrt{3}xy + \tfrac{1}{2}y^2 = -\tfrac{1}{2}Q_- + \tfrac{\sqrt{3}}{2}Q_+ \\ O_1 Q_+ &= -\tfrac{\sqrt{3}}{4}x^2 - xy + \tfrac{\sqrt{3}}{4}y^2 = -\tfrac{\sqrt{3}}{4}Q_- - \tfrac{1}{2}Q_+\end{aligned} \tag{2.25}$$

But the space spanned by $Q_\pm$ isn't irreducible: the polynomials

$$Q^\pm \equiv (x \pm \mathrm{i}y)^2 = Q_- \pm \mathrm{i}Q_+ \tag{2.26}$$

are simply multiplied by a phase factor $\mathrm{e}^{4\pi\mathrm{i}k/3}$ when (x, y) are transformed with $\mathbf{r}_1$ or $\mathbf{r}_2$. So the 3-dimensional representation T of $\mathcal{C}_3$ obtained from (2.10) decomposes into the sum of three irreps

$$T = T^{(0)} \dot{+} T^{(+)} \dot{+} T^{(-)}, \tag{2.27}$$

where $T^{(0)}$ is the trivial representation provided by Q_0, and $T^{(\pm)}$ are the faithful representations provided by $Q^\pm$.

Example 2.5

Let's enlarge the group of the last example from $\mathcal{C}_3$ to $\mathcal{D}_3$. The extension can be accomplished by adding to $\{r_1\}$ the generator

$$p_1 \equiv \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.28}$$

The matrix O_p generated by p_1 through (2.10) turns Q^+ into Q^-:

$$O_p Q^+ = (-x + \mathrm{i}y)^2 = Q^-. \tag{2.29}$$

Hence Q^+ alone doesn't form a representation of $\mathcal{D}_3$. In fact the 2-dimensional space spanned by $Q^\pm$ forms a 2-dimensional irrep of $\mathcal{D}_3$.

Exercise (9):

Write out the matrices of O_1 and O_p in the basis (Q_0, Q^+, Q^-) and thus examine the block-diagonal structure of a representation when it has been reduced to the sum of irreps.

3 Schur's lemmas and orthogonality relations

Schur's little theorems provide the key to reducing a given representation to its constituent irreps. They concern a linear map A between two irreducible representation spaces L_1 and L_2 (which may coincide) of dimension s_1 and s_2, respectively. When we choose bases for L_1 and L_2, A will be represented by a $s_2 \times s_1$ matrix $\mathbf{A}$.

Theorem 6 (Schur I)

Let T be an irrep of $\mathcal{G}$ on the space L and let A be a linear map of L into itself such that $\mathbf{T}(a) \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{T}(a)$ $\forall a \in \mathcal{G}$. Then either (i) $\mathbf{A} = \lambda\mathbf{I}$, or (ii) $\mathbf{A} = 0$.

In other words, if you get the same result going two ways around this diagram,

$$\begin{array}{ccc} L & \xrightarrow{A} & L \\ \downarrow T & & \downarrow T \\ L & \xrightarrow{A} & L \end{array} \tag{3.1}$$

then $\mathbf{A}$ is either a multiple of the identity, or zero.

Proof: Let $\mathbf{v}$ be an eigen-vector of $\mathbf{A}$: $\mathbf{A} \cdot \mathbf{v} = \lambda\mathbf{v}$. Then for any $a \in \mathcal{G}$

$$\mathbf{A} \cdot \mathbf{T}(a) \cdot \mathbf{v} = \mathbf{T}(a) \cdot \mathbf{A} \cdot \mathbf{v} = \lambda\mathbf{T}(a) \cdot \mathbf{v}. \tag{3.2}$$

Thus all vectors $\mathbf{u}$ of the form $\mathbf{u} = \mathbf{T}(a) \cdot \mathbf{v}$ are eigenvectors of $\mathbf{A}$ with eigen-value λ. If $\lambda \neq 0$, these vectors span a subspace of L that is invariant under $\mathcal{G}$. Since L is irreducible, this subspace must coincide with L. □

Theorem 7 (Schur II)

Let $T^{(i)}$ (i=1,2) be two inequivalent irreps of $\mathcal{G}$ on the spaces L_i and let A be a linear map $L_1 \to L_2$ such that $\mathbf{T}^{(2)}(a) \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{T}^{(1)}(a)$ $\forall a \in \mathcal{G}$. Then $\mathbf{A} = 0$.

Proof: The vectors $\mathbf{u} \in L_2$ of the form

$$\mathbf{u} = \mathbf{T}^{(2)}(a) \cdot \mathbf{A} \cdot \mathbf{v} = \mathbf{A} \cdot \mathbf{T}^{(1)}(a) \cdot \mathbf{v} \tag{3.3}$$

for $a \in \mathcal{G}$ and any fixed $\mathbf{v} \in L_1$ form a subspace $V \subset L_2$ which is invariant under $\mathcal{G}$. Since L_2 is irreducible, V must either be null ($\Rightarrow \mathbf{A} = 0$) or the whole of L_2. If $V = L_2$, we must have $s_1 \geq s_2$. If $s_1 = s_2$, $\mathbf{A}$ would have an inverse and then $\mathbf{T}^{(2)}(a) = \mathbf{A} \cdot \mathbf{T}^{(1)}(a) \cdot \mathbf{A}^{-1}$, and the $T^{(i)}$ would be equivalent, which they are not. Hence either $\mathbf{A} = 0$ or $s_1 > s_2$. But in the latter case $\mathbf{A}$ must map certain vectors to zero and by (3.3) the space spanned by such vectors would be a subspace of L_1 invariant under $\mathcal{G}$, contrary to assumption. Hence $\mathbf{A} = 0$.

Theorem 8

Abelian groups only have 1-dimensional irreps

Proof: When $\mathcal{G}$ is Abelian, all representing matrices commute: $\mathbf{T}(a) \cdot \mathbf{T}(b) = \mathbf{T}(ab) = \mathbf{T}(ba) = \mathbf{T}(b) \cdot \mathbf{T}(a)$ $\forall$ $b \in \mathcal{G}$. Hence $\mathbf{T}(a)$ satisfies the requirements Theorem 6 imposes on $\mathbf{A}$, so it must be a multiple of the identity $\mathbf{I}$. But if every $\mathbf{T}(a)$ were a multi-dimensional identity, the representation would be reducible. Since it is not reducible, the representation must be one-dimensional. □

3.1 Orthogonality relations

Let $\mathbf{X}$ be any matrix that maps vectors from L_α to L_β, where both L's are irreducible spaces. Then the matrix

$$\mathbf{A} \equiv \sum_{b \in \mathcal{G}} \mathbf{T}^{(\beta)}(b) \cdot \mathbf{X} \cdot \mathbf{T}^{(\alpha)}(b^{-1}) \tag{3.4}$$

satisfies the hypotheses of theorems 6 and 7:

$$\begin{aligned}
\mathbf{T}^{(\beta)}(a) \cdot \mathbf{A} &= \sum_{b \in \mathcal{G}} \mathbf{T}^{(\beta)}(a) \cdot \mathbf{T}^{(\beta)}(b) \cdot \mathbf{X} \cdot \mathbf{T}^{(\alpha)}(b^{-1}) \\
&= \sum_{b \in \mathcal{G}} \mathbf{T}^{(\beta)}(ab) \cdot \mathbf{X} \cdot \mathbf{T}^{(\alpha)}\big((ab)^{-1}\big) \cdot \mathbf{T}^{(\alpha)}(a) \\
&= \mathbf{A} \cdot \mathbf{T}^{(\alpha)}(a),
\end{aligned} \tag{3.5}$$

where we have exploited the fact that, for fixed a, ab runs over $\mathcal{G}$ as b does. Hence by theorems 6 and 7

$$A_{ij} = \sum_{b \in \mathcal{G}} \sum_{p,q} T^{(\beta)}_{ip}(b) X_{pq} T^{(\alpha)}_{qj}(b^{-1}) = \lambda \delta_{\alpha\beta} \delta_{ij}. \tag{3.6}$$

Let's choose $\mathbf{X}$ so that it has only one non-zero entry, namely that in the $l^{\rm th}$ row and $m^{\rm th}$ column: $X_{pq} = \delta_{pl}\delta_{qm}$. Doing the sums over p and q we now find

$$\sum_{b \in \mathcal{G}} T^{(\beta)}_{il}(b) T^{(\alpha)}_{mj}(b^{-1}) = \lambda \delta_{\alpha\beta} \delta_{ij}. \tag{3.7}$$

Notice that none of the subscripts on the left are summed. To find the value of λ, we set $\alpha = \beta$, $i = j$ and sum over i:

$$\begin{aligned}
\lambda s_\alpha &= \sum_{b \in \mathcal{G}} \sum_i T^{(\alpha)}_{mi}(b^{-1}) T^{(\alpha)}_{il}(b) \\
&= \sum_{b \in \mathcal{G}} \delta_{ml} \\
&= g \delta_{ml}.
\end{aligned} \tag{3.8}$$

Substituting this back into (3.7), we have

$$\sum_{b\in\mathcal{G}} T^{(\beta)}_{il}(b)T^{(\alpha)}_{mj}(b^{-1}) = g/s_\alpha \delta_{\alpha\beta}\delta_{ij}\delta_{lm}. \tag{3.9}$$

When $\mathbf{T}$ is unitary a further simplification is possible since $\mathbf{T}(b^{-1}) = \mathbf{T}^{-1}(b) = \mathbf{T}^\dagger(b)$. Using this in (3.9) yields

$$\sum_{b\in\mathcal{G}} T^{(\beta)}_{il}(b)T^{(\alpha)*}_{jm}(b) = g/s_\alpha \delta_{\alpha\beta}\delta_{ij}\delta_{lm}. \tag{3.10}$$

This is the fundamental orthogonality relation for unitary representations of groups.

Example 3.1

$\mathcal{C}_3$ has the unitary irrep $e = 1$, $r_1 = \mathrm{e}^{2\pi\mathrm{i}/3}$, $r_2 = \mathrm{e}^{4\pi\mathrm{i}/3}$. Set $\alpha = \beta$, $i = j = 1$ and $l = m = 1$ in (3.10). Then the left side gives $1 + \mathrm{e}^{2\pi\mathrm{i}/3}\mathrm{e}^{-2\pi\mathrm{i}/3} + \mathrm{e}^{4\pi\mathrm{i}/3}\mathrm{e}^{-4\pi\mathrm{i}/3} = 3$, as does the right side.

Exercise (10):

Verify the orthogonality relation (3.10) for two different irreps of $\mathcal{D}_3$.

3.2 Characters of representations

The trace of $\mathbf{T}(a)$, $\mathrm{Tr}(\mathbf{T})$ is invariant under similarity transformations $\mathbf{T} \to \mathbf{T}' \equiv \mathbf{S}^{-1}\cdot\mathbf{T}\cdot\mathbf{S}$. Hence the trace is the same for all equivalent representations (§2.1) and is the same for every $\mathbf{T}$ that represents the elements of any given class (§1.3). These properties make $\mathrm{Tr}(\mathbf{T})$ an important number and it has its own name: the **character** $\chi^{(\alpha)}(a)$ of $a \in \mathcal{G}$ in the representation α. Setting $l = i$ and $j = m$ in (3.10) and summing over i and j, we obtain an orthogonality relation for characters

$$\sum_{p\in\text{classes of }\mathcal{G}} c_p \chi^{(\beta)}(p)\chi^{(\alpha)*}(p) = g\delta_{\alpha\beta}. \tag{3.11}$$

3.3 Reduction of representations

The orthogonality relation (3.11) enables us to determine if an irrep occurs in a given representation, and, if so, the number m_α of times it occurs. Since the trace is invariant under the similarity transformation that brings a given representation $\mathbf{T}$ to block diagonal form, by taking the trace of the block-diagonal form it follows that

$$\chi(\mathbf{T}(a)) = \sum_\beta m_\beta \chi^{(\beta)}(a); \quad \text{i.e.,} \quad \chi(p) = \sum_\beta m_\beta \chi^{(\beta)}(p). \tag{3.12}$$

We multiply through by $c_p\chi^{(\alpha)*}(p)$ and sum over classes p

$$\begin{aligned}\sum_{p\in\text{classes of }\mathcal{G}} c_p \chi^{(\alpha)*}(p)\chi(p) &= \sum_\beta m_\beta \sum_p c_p \chi^{(\alpha)*}(p)\chi^{(\beta)}(p) \\ &= m_\alpha g.\end{aligned} \tag{3.13}$$

Thus as soon as we know the characters $\chi_\alpha(p)$ of every class in every irrep, we can determine m_α, the number of times the α^{th} representation occurs in $\mathbf{T}(a)$.

3.4 The regular representation

The key to finding all irreps is a dumb representation. The **regular representation** is the $g \times g$-dimensional representation in which rows and columns are labelled by the elements of $\mathcal{G}$ and each row or column of $\mathbf{T}(a)$ has exactly one non-zero entry:

$$T^{(R)}_{a,b}(c) = \delta_{a,bc} \quad \Rightarrow \quad \chi^{(R)}(c) = \sum_a \delta_{a,ac} = g\delta_{c,e}. \tag{3.14}$$

i.e. in the regular representation, the character of every element other than e is zero, while the character of e is g.

The decomposition of $T^{(R)}$ into its constituent irreps is now straightforward. We multiply

$$\chi^{(R)}(a) = \sum_\beta m_\beta \chi^{(\beta)}(a) \tag{3.15}$$

by $\chi^{(\alpha)*}(a)$ and sum over $a \in \mathcal{G}$:

$$\begin{aligned}\sum_{a\in\mathcal{G}} \chi^{(\alpha)*}(a)\chi^{(R)}(a) &= s_\alpha g \\ &= m_\alpha g.\end{aligned} \tag{3.16}$$

Thus $m_\alpha = s_\alpha$ and each representation occurs the same number of times as its dimension. Moreover, from (3.15) with $a = e$ we have $g = \sum_\alpha m_\alpha s_\alpha$, so

$$g = \sum_\alpha s_\alpha^2. \tag{3.17}$$

In particular, the number of irreps of any finite group is finite.

Example 3.2

$g(\mathcal{D}_3) = 6$ and we know of irreps of dimensionality $s = 1$ and $s = 2$. Since $1^2 + 2^2 = 5$, there must be a second 1-dimensional irrep.

Theorem 9

The number n_r of irreps of $\mathcal{G}$ is equal to the number n_c of classes in $\mathcal{G}$

Proof: Equation (3.10) can be interpreted as expressing the mutual orthogonality of g vectors

$$\mathbf{t}^{(\alpha ij)} \equiv \sum_a T^{(\alpha)}_{ij}(a)\,\hat{\mathbf{e}}_a, \tag{3.18a}$$

where we have associated with each $a \in \mathcal{G}$ an abstract unit vector $\hat{\mathbf{e}}_a$:

$$\hat{\mathbf{e}}_a \cdot \hat{\mathbf{e}}_b = \delta_{ab}. \tag{3.18b}$$

The set of characters in the irrep α is associated with a particular linear combination of the $\mathbf{t}$'s, the **character vector**

$$\boldsymbol{\chi}^{(\alpha)} \equiv \sum_{i=1}^{s_\alpha} \mathbf{t}^{(\alpha ii)} = \sum_a \chi^{(\alpha)}(a)\hat{\mathbf{e}}_a \quad \Rightarrow \quad \chi^{(\alpha)}(a) = \hat{\mathbf{e}}_a \cdot \boldsymbol{\chi}^{(\alpha)}. \tag{3.19}$$

Since all elements in the same class have the same character, we have $\hat{\mathbf{e}}_a \cdot \boldsymbol{\chi}^{(\alpha)} = \hat{\mathbf{e}}_c \cdot \boldsymbol{\chi}^{(\alpha)}$ whenever a and c are in the same class. So any linear combination of the $\boldsymbol{\chi}^{(\alpha)}$ also has equal components along all $\hat{\mathbf{e}}$'s of a given class. Consequently, the $\boldsymbol{\chi}^{(\alpha)}$ for $\alpha = 1, \ldots, n_r$ cannot span the whole g-dimensional space of the $\hat{\mathbf{e}}_a$.

Now the character vectors $\boldsymbol{\chi}^{(\alpha)}$ (i) lie in the n_c-dimensional subspace V_c of all vectors $\mathbf{v}$ that have equal components along every $\hat{\mathbf{e}}_a$ of a given class, and (ii) are linearly independent by the orthogonality relation (3.11). These facts together imply $n_r \leq n_c$. What we now show is that the character vectors form a basis for V_c, and thus that $n_r = n_c$.

Any $\mathbf{v} \in V_c$ is such that

$$\hat{\mathbf{e}}_a \cdot \mathbf{v} = \hat{\mathbf{e}}_c \cdot \mathbf{v}, \quad \text{where} \quad a = bcb^{-1}, \tag{3.20}$$

so summing over $b \in \mathcal{G}$ we get

$$\begin{aligned}
\hat{\mathbf{e}}_a \cdot \mathbf{v} &= \frac{1}{g} \sum_{b \in \mathcal{G}} \hat{\mathbf{e}}_c \cdot \mathbf{v} \qquad (a \text{ constant, } c(b) \text{ varying}) \\
&= \frac{1}{g} \sum_{b} \hat{\mathbf{e}}_c \cdot \sum_{\alpha ij} v_{\alpha ij} \mathbf{t}^{(\alpha ij)} \\
&= \frac{1}{g} \sum_{\alpha ij} v_{\alpha ij} \sum_{b} \hat{\mathbf{e}}_c \cdot \mathbf{t}^{(\alpha ij)} \\
&= \frac{1}{g} \sum_{\alpha ij} v_{\alpha ij} \sum_{b} T_{ij}^{(\alpha)}(c) \\
&= \frac{1}{g} \sum_{\alpha ij} v_{\alpha ij} \sum_{b} T_{ij}^{(\alpha)}(b^{-1}ab) \\
&= \frac{1}{g} \sum_{\alpha ij} v_{\alpha ij} \sum_{pq} \sum_{b} T_{ip}^{(\alpha)}(b^{-1}) T_{pq}^{(\alpha)}(a) T_{qj}^{(\alpha)}(b) \\
&= \frac{1}{g} \sum_{\alpha i} v_{\alpha ii} \sum_{q} (g/s_\alpha) T_{qq}^{(\alpha)}(a),
\end{aligned} \tag{3.21}$$

where in obtaining the last line use has been made of the orthogonality relation (3.9). Multiplying (3.21) through by $\hat{\mathbf{e}}_a$, summing over a and expressing the result in terms of the character vectors, we find

$$\mathbf{v} = \sum_{\alpha} \frac{1}{s_\alpha} \Big(\sum_{i} v_{\alpha ii} \Big) \boldsymbol{\chi}^{(\alpha)}. \tag{3.22}$$

This demonstrates that any vector in our n_c-dimensional subspace can be written as a linear combination of the n_r character vectors, and thus that $n_r = n_c$. □

Since $\mathcal{G}$ has as many irreps as classes, the **character table**, which gives the character of elements in each class in each irrep, is square. Usually each row gives data for a given irrep, and each column corresponds to a given class:

	E	$2C_3$	$3C_2$
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

Character table of $\mathcal{D}_3$

One- two- three- and four-dimensional irreps are denoted by A, E, T and U, respectively.

3.5 Reduction of a direct product representation

Equation (2.15) defines the direct product $S \times T \equiv U$ of two reps S and T. We suppose these are the irreps $T^{(\alpha)}$ and $T^{(\beta)}$, respectively, and ask what irreps $S \times T$ contains. From equation (2.16) we see that in $S \times T$, $a \in \mathcal{G}$ has character $\chi^{(\alpha\times\beta)} = \chi^{(\alpha)}(a)\chi^{(\beta)}(a)$. So taking the trace of

$$T^{(\alpha)} \times T^{(\beta)} = \dot{\sum_{\gamma}} m_\gamma T^{(\gamma)}, \tag{3.23}$$

we have

$$\chi^{(\alpha\times\beta)}(a) = \sum_\gamma m_\gamma \chi^{(\gamma)}(a). \tag{3.24}$$

With (3.11) it follows that

$$m_\gamma = \frac{1}{g} \sum_{p\in\text{classes of } \mathcal{G}} c_p \chi_p^{(\gamma)*} \chi_p^{(\alpha)} \chi_p^{(\beta)}. \tag{3.25}$$

For example, the 4-dimensional rep $E \times E$ obtained by taking the direct product of the 2-dimensional irrep E of $\mathcal{D}_3$ with itself decomposes into two one-dimensional and a two-dimensional irreps as follows:

$$\begin{aligned} m_{A_1}(E\times E) &= \tfrac{1}{6}(1\times 4 + 2\times 1 + 3\times 0) = 1,\\ m_{A_2}(E\times E) &= \tfrac{1}{6}(1\times 4 + 2\times 1 + 3\times 0) = 1,\\ m_E(E\times E) &= \tfrac{1}{6}(1\times 8 + 2\times(-1) + 3\times 0) = 1. \end{aligned} \tag{3.26}$$

4 Applications to quantum mechanics

In quantum mechanics the state of a system is described by the ket $|\psi\rangle$, from which any question about the system can be answered as well as is humanly possible. For any system we can define an operator $R(\hat{\mathbf{n}}, \theta)$ associated with rotating the system by angle θ about the unit vector $\hat{\mathbf{n}}$; that is, $R|\psi\rangle$ describes a state of the system in every respect the same as that described by $|\psi\rangle$, except that the first state is rotated with respect to the second.

The operators $R(\hat{\mathbf{n}}, \theta)$ form a representation of the group $\mathcal{R}(3)$ of 3-dimensional rotations; the representation space V is the Hilbert space of all physically admissible kets $|\psi\rangle$.

The **parity operator** P is defined be such that the state $P|\psi\rangle$ describes the state we would see if we observed our laboratory in a mirror. (Mathematically, $P|\psi\rangle$ is obtained from $|\psi\rangle$ by inverting all spatial coordinates and taking account of the intrinsic parities of every particle in the system.) P provides a representation of $\mathcal{S}_2$. The operators $R(\hat{\mathbf{n}}, \theta)$ and P etc form a representation of the group $O(3)$ of orthogonal 3-dimensional transformations.

More generally, if we apply to the ket $|\psi\rangle$ the operator Λ associated with a Lorentz transformation, we obtain the ket that describes a system that is boosted and rotated with respect to the original system. The operators corresponding to all possible Lorentz transformations provide a representation of the group $\mathcal{L}$ of Lorentz transformations.

Because the space of kets is infinite-dimensional, quantum mechanics provides physical contexts for a much wider range of representations than does classical physics. This wide range makes a study of possible representations worthwhile.

4.1 Function spaces

Let's look at the way in which high-dimensional representations can arise in the simplest quantum case, that of a single spin-zero particle. Then the ket $|\psi\rangle$ can be represented by a wavefunction $\psi(\mathbf{x}) \equiv \langle \mathbf{x}|\psi\rangle$. Now suppose we have a representation T of some group $\mathcal{G}$ in ordinary three-dimensional space $\mathcal{R}^3$. That is, to every $a \in \mathcal{G}$ we have a mapping $\mathcal{R}^3 \to \mathcal{R}^3$

$$\mathbf{x} \to \mathbf{x}' \equiv \mathbf{T}(a)\cdot\mathbf{x}. \tag{4.1}$$

To each of these mappings there corresponds a mapping on the space of wavefunctions:

$$\psi \to \psi' \equiv \widetilde{T}(a)\psi \quad \text{where} \quad \psi'(\mathbf{x}') \equiv \psi(\mathbf{x}) = \psi\big(\mathbf{T}(a^{-1})\cdot\mathbf{x}'\big). \tag{4.2}$$

The operators $\widetilde{T}$ form a representation of $\mathcal{G}$:

$$\begin{aligned}\big(\widetilde{T}(a)\widetilde{T}(b)\psi\big)(\mathbf{x}) &= \big(\widetilde{T}(b)\psi\big)(\mathbf{T}(a^{-1})\cdot\mathbf{x}) \\ &= \psi\big(\mathbf{T}(b^{-1})\cdot\mathbf{T}(a^{-1})\cdot\mathbf{x}\big) \\ &= \psi\big(\mathbf{T}(b^{-1}a^{-1})\cdot\mathbf{x}\big) \\ &= \psi\big(\mathbf{T}((ab)^{-1})\cdot\mathbf{x}\big) \\ &= \big(\widetilde{T}(ab)\psi\big)(\mathbf{x}).\end{aligned} \tag{4.3}$$

Notice the way in which the apparently gratuitous introduction of the inverse in (4.2) ensures that the operators $\widetilde{T}$ provide a representation.

Example 4.1

In §2 we studied the representation of $\mathcal{C}_N$ induced in the space of functions on the plane that are quadratic in the coordinates.

Example 4.2

Suppose we have an insulating sphere and that its surface bears electric charge of density $\sigma(\theta,\phi)$. Then after we have rotated the sphere through angle θ_0 about the axis $\hat{\mathbf{n}}$, the charge density on the sphere has a new functional form $\sigma'(\theta,\phi)$. Specifically

$$\sigma'(\theta,\phi) = \sigma\big(R(\hat{\mathbf{n}},-\theta_0)(\theta,\phi)\big). \tag{4.4}$$

Since $R(\hat{\mathbf{n}},-\theta_0) = R^{-1}(\hat{\mathbf{n}},\theta_0)$, the inverse arises naturally in this physical illustration.

Starting from any charge density σ, we can generate a representation of $\mathcal{R}(3)$ by rotating our sphere every which way. The representation generated will be finite-dimensional only if the initial charge density is special in that any distribution we obtain by rotating the sphere can be expressed as a linear combination of just a few basic distributions. (Clearly, a uniform distribution trivially satisfies this condition.)

4.2 Commuting observables

Now consider some observable Q which is necessarily the same for a system described by $|\psi\rangle$ and for the 'rotated' system described by $T|\psi\rangle$, where $T = T(a)$ is a typical member of some representation of a group $\mathcal{G}$. Let the kets $|n\rangle$ form a complete set of eigenkets of Q. Then under our hypothesis regarding Q

$$Q|n\rangle = q_n|n\rangle \quad \Rightarrow \quad Q(T|n\rangle) = q_n(T|n\rangle) = TQ|n\rangle. \tag{4.5}$$

Since the set $\{|n\rangle\}$ is complete, it follows that Q and T commute: $[Q,T]=0$. From (4.5) it also follows that $[Q,T]=0 \;\Rightarrow |n\rangle$ and $T|n\rangle$ both yield value q_n on a measurement of Q. In other words, invariance of an observable Q under $\mathcal{G}$ is equivalent to $[Q,T]=0$ for any rep T of $\mathcal{G}$.

We have seen the importance of sub-spaces that are invariant under the operations of a group. The essence of Schur's lemmas is the idea that when any operator Q commutes with all the operators $T(a)$ of a representation, the sub-space $V(q)$ associated with each eigen-value q of Q must be invariant under the group and therefore provide a representation of $\mathcal{G}$:

$$Q(T|q\rangle) = T(Q|q\rangle) = q(T|q\rangle). \tag{4.6}$$

If the dimensionality $\dim\big(V(q)\big)$ is greater than unity, q is clearly a degenerate eigen-value. Often $V(q)$ provides an irrep of $\mathcal{G}$, so $\dim\big(V(q)\big) = s_\alpha$ for some α. Hence a knowledge of the irreps of $\mathcal{G}$ leads to a knowledge of the possible degeneracies of Q's eigen-values. When $\dim\big(V(q)\big) \neq s_\alpha$ for some α, one says that the degeneracy is 'accidental'. In reality this is often the signal that Q commutes with operators representing a bigger group $\mathcal{G}'$ of which $\mathcal{G}$ is merely a sub-group, and that $\dim\big(V(q)\big)$ equals the dimensionality of one of the irreps of $\mathcal{G}'$.

For example, the total angular momentum operator $J^2 = J_x^2 + J_y^2 + J_z^2$ commutes with every $R(\hat{\mathbf{n}}, \theta)$, since two systems which are identical in all respects except their orientation clearly have the same total angular momentum. We shall see that $\mathcal{R}(3)$ has irreps of dimension $2j+1$ for $j = 0, \frac{1}{2}, 1, \ldots$, and this makes it natural that the eigen-values $j(j+1)\hbar^2$ of J^2 are $(2j+1)$-fold degenerate.

Example 4.3

Let's go back to our charged sphere and suppose that our initial charge density σ can be expanded as a linear superposition of spherical harmonics Y_l^m of the same order l. Then from Legendre's equation we know that

$$L^2\sigma \equiv -\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)\sigma = l(l+1)\sigma. \tag{4.7}$$

But by grinding away at coordinate changes $(\theta, \phi) \to (\theta', \phi') = R(\theta, \phi)$ we can show that the operator L^2 looks the same in any system of spherical polars. So $L^2\sigma' = L^2\sigma = l(l+1)\sigma$ and it follows that σ' can also be expanded in terms of the eigenfunctions of L^2 with eigenvalue $l(l+1)$. That is, it follows that σ' is also a linear superposition of the Y_l^m for fixed l. This demonstrates that in this case the representation of $\mathcal{R}(3)$ generated by our charged sphere is $(2l+1)$-dimensional. We shall see that it is actually an irrep.

Exercise (11):

The irrep obtained when $l = 1$ is equivalent to the representation provided by ordinary 3-dimensional vectors. Show this as follows: write $\mathrm{Y}_{11} = (\alpha/\sqrt{2})\sin\theta\mathrm{e}^{\mathrm{i}\phi}$, $\mathrm{Y}_{10} = \alpha\cos\theta$, $\mathrm{Y}_{1-1} = -(\alpha/\sqrt{2})\sin\theta\mathrm{e}^{-\mathrm{i}\phi}$, where $\alpha \equiv \sqrt{3/4\pi}$. Show that the Cartesian coordinates (x, y, z) of the point $\mathbf{x}$ on the unit sphere that represents (θ, ϕ) are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha^{-1}\mathbf{U}\cdot\begin{pmatrix} \mathrm{Y}_{11} \\ \mathrm{Y}_{10} \\ \mathrm{Y}_{1-1} \end{pmatrix} \qquad \text{where} \quad \mathbf{U} \equiv \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ \frac{-\mathrm{i}}{\sqrt{2}} & 0 & \frac{-\mathrm{i}}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}. \tag{4.8}$$

Verify that $\mathbf{U}$ is unitary. Let (x', y', z') be the coordinates of the point (θ, ϕ) with respect to rotated axes, and (θ', ϕ') be the corresponding polar angles. Then $\mathbf{x}' = \mathbf{R}\cdot\mathbf{x}$, where $\mathbf{R}$ is a rotation matrix. Show that

$$\begin{pmatrix} \mathrm{Y}'_{11} \\ \mathrm{Y}'_{10} \\ \mathrm{Y}'_{1-1} \end{pmatrix} = \mathbf{U}^\dagger\cdot\mathbf{R}\cdot\mathbf{U}\cdot\begin{pmatrix} \mathrm{Y}_{11} \\ \mathrm{Y}_{10} \\ \mathrm{Y}_{1-1} \end{pmatrix}, \tag{4.9}$$

where $\mathrm{Y}'_{11} \equiv \mathrm{Y}_{11}(\theta', \phi')$ etc. Hence show that if we expand the charge density $\sigma(\theta, \phi)$ as

$$\sigma(\theta, \phi) = \sum_{m=-1}^{1} \sigma_m \mathrm{Y}_{1m}(\theta, \phi), \tag{4.10a}$$

we have that

$$\sigma'(\theta', \phi') = \sum_{m=-1}^{1} \sigma'_m \mathrm{Y}_{1m}(\theta', \phi'), \tag{4.10b}$$

where

$$\begin{aligned} \sigma'_n &\equiv \sum_m \sigma_m (\mathbf{U}^\dagger\cdot\mathbf{R}^T\cdot\mathbf{U})_{mn} = \sum_m (\mathbf{U}^T\cdot\mathbf{R}\cdot\mathbf{U}^*)_{nm}\sigma_m \\ &= \sum_m (\mathbf{V}^\dagger\cdot\mathbf{R}\cdot\mathbf{V})_{nm}\sigma_m, \end{aligned} \tag{4.10c}$$

where $V \equiv U^*$ is unitary.

4.3 Dynamical symmetries

The dynamics of a quantum-mechanical system is encoded in its Hamiltonian H. In particular, dynamical symmetries of the system are encoded in H. For example empty space is believed to be isotropic, so an isolated atom can be rotated through any angle without changing its energy. Consequently, R commutes with the Hamiltonian H of any isolated system; $[H, R] = 0$.

So long as weak interactions are not in play, the mirror state $P|E\rangle$ of any eigen-state $|E\rangle$ of H will also be an eigen-ket of H, so $[H, P] = 0$.

By the arguments of the last subsection, the eigen-values of H are liable to be degenerate with degeneracy equal to the dimensionality of the irreps of $\mathcal{R}(3)$ and/or $\mathcal{S}_2$. For example, the energy levels of isolated atoms are well known to be $(2j+1)$-fold degenerate.[3] By contrast, the irreps of the Abelian group $\mathcal{S}_2$ are all 1-dimensional, and this is reflected in the energy levels of particles that move in even one-dimensional potentials being non-degenerate (the ground state is generally of even-parity). Here's a less trivial application

Theorem 10 (Bloch)

There is a complete set of wavefunction $\psi(\mathbf{x})$ of a particle of well-defined energy in a periodic potential $V(\mathbf{x})$ which all satisfy

$$\psi(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{a}}\psi(\mathbf{x}-\mathbf{a}), \tag{4.11}$$

where $\mathbf{a}$ is any lattice vector and $\mathbf{k}$ is some vector.

Proof: The particle's Hamiltonian is invariant under the Abelian group $\mathcal{G}$ of displacements by lattice vectors. Hence the eigen-functions of a given energy will form a rep of an Abelian group. Since the irreps of $\mathcal{G}$ are 1-dimensional, we can choose a complete set of energy eigen-functions that each forms the basis for a 1-dimensional irrep of $\mathcal{G}$. Let $T(\mathbf{a})$ be the unitary operator associated with displacement by a lattice vector $\mathbf{a}$. Then by the 1-dimensionality of the irrep we have

$$\begin{aligned}\psi(\mathbf{x}-\mathbf{a}) &= \big(T(\mathbf{a})\psi\big)(\mathbf{x})\\ &= f(\mathbf{a})\psi(\mathbf{x}),\end{aligned} \tag{4.12}$$

where f is some number. Moreover, by the unitarity of T, f must be of the form $f = \mathrm{e}^{\mathrm{i}\phi(\mathbf{a})}$. By compounding displacements by $\mathbf{a}$ and $\mathbf{b}$ we see that

$$\phi(\mathbf{a}+\mathbf{b}) = \phi(\mathbf{a}) + \phi(\mathbf{b}). \tag{4.13}$$

So

$$\phi(\mathbf{a}) = -(k_x a_x + k_y a_y + k_z a_z) \quad \text{where} \quad \mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z \tag{4.14}$$

expresses $\mathbf{a}$ as a linear combination of the basic lattice displacements $\mathbf{e}_i$. Writing $\mathbf{k} = \sum_i k_i\widetilde{\mathbf{e}}_i$ where $\widetilde{\mathbf{e}}_x \equiv \mathbf{e}_y \times \mathbf{e}_z/(\mathbf{e}_x \cdot \mathbf{e}_y \times \mathbf{e}_z)$ etc., completes the proof. □

Bloch's theorem assures us that we can label each energy eigenfunction of a periodic potential by the vector $\mathbf{k}$ that characterizes its irrep of the potential's translation group. It has a useful corollary

Corollary

The wavefunction $\psi(\mathbf{x})$ of a particle of well-defined energy in a periodic potential may be written

$$\psi(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}u(\mathbf{x}) \quad \textit{where} \quad u(\mathbf{x}) = u(\mathbf{x}-\mathbf{a}) \tag{4.15}$$

with $\mathbf{a}$ any lattice vector.

Proof: Define $u(\mathbf{x}) \equiv \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\psi(\mathbf{x})$. Then

$$u(\mathbf{x}-\mathbf{a}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{a}-\mathbf{x})}\psi(\mathbf{x}-\mathbf{a}).$$

With (4.11) the r.h.s. immediately reduces to $u(\mathbf{x})$. □

[3] Hydrogen's energy levels are 'accidentally' degenerate because hydrogen's H is actually invariant under $O(4)$ rather than just $\mathcal{R}(3)$.

Example 4.4

The stationary states of an isolated atom can be taken to be states of well defined total angular momentum j, and the states of a given energy E usually constitute basis vectors for an irrep $D^{(j)}$ of $\mathcal{R}(3)$.

An atom in a crystal experiences an electrostatic field that lacks the full $\mathcal{R}(3)$ symmetry of empty space; the field's rotational symmetry group is just the point group $\mathcal{G}$ of the lattice, which is a finite-dimensional sub-group of $\mathcal{R}(3)$. The handful of matrices in $D^{(j)}$ that correspond to elements of $\mathcal{R}(3)$ that are also in $\mathcal{G}$ clearly provide a representation of $\mathcal{G}$. This $D^{(j)}$ representation of $\mathcal{G}$ is unlikely to be an irrep because there are so many fewer elements in $\mathcal{G}$ than in $\mathcal{R}(3)$ from which to derive the operators we need if we are to be able to turn any given vector into any other, as required by the definition of an irrep.

If the shifts in the atom's energy that are produced by the introduction of the crystal field are small compared with the differences in energy between states of the unperturbed atom of different j, these shifts may be calculated by first-order, degenerate perturbation theory using only the states of some given j. That is, to a reasonable approximation, there are $2j+1$ eigen-states $|C,k\rangle$ of the Hamiltonian that includes the crystal field which are linear combinations of the $2j+1$ states $|E_j,m\rangle$ of unperturbed energy E_j:

$$|C,k\rangle = \sum_{m=-j}^{j} a_{km}|E_j,m\rangle. \tag{4.16}$$

Now we know that if any perturbed energy E'_k is degenerate, the eigen-kets (4.16) associated with it will support a representation (and probably an irrep) of the point group $\mathcal{G}$. So the question is, what irreps of $\mathcal{G}$ can we construct by taking linear combinations of the basis vectors of the $D^{(j)}$ representation of $\mathcal{G}$?

But this is precisely the question we addressed in §3: what irreps does a given representation contain? The answer lies in the character table of $\mathcal{G}$. Let's assume that the crystal is a cubic one. Then $\mathcal{G}$ will be the octahedral group of example 8, whose character table is this:

	E	$8C_3$	$3C_2$	$6C_2$	$6C_4$
A_1	1	1	1	1	1
A_2	1	1	1	−1	−1
E	2	−1	2	0	0
T_1	3	0	−1	−1	1
T_2	3	0	−1	1	−1

(4.17)

Character table of $\mathcal{O}$

$\mathcal{O}$ has one-, two- and three-dimensional irreps. Writing

$$D^{(j)} = \dot{\sum_\alpha} m_\alpha^{(j)} T^{(\alpha)}, \tag{4.18}$$

and using the standard orthogonality relation (3.11), we find

$$m_\beta^{(j)} = \frac{1}{g} \sum_{p\in\text{classes of }\mathcal{O}} c_p \chi^{(\beta)*}(p)\chi^{(j)}(p). \tag{4.19}$$

To complete the calculation we need to know the characters of elements of $\mathcal{O}$ in the spin-j representation of $\mathcal{R}(3)$. In §8 we shall see that all rotations through angle ϕ have character $\sin(j+\frac{1}{2})\phi/\sin\frac{1}{2}\phi$. This enables us to show (see Exercise 12), for example, that the spin-3 representation decomposes under $\mathcal{O}$ into $A_2 \dot{+} T_1 \dot{+} T_2$, and the 7-fold degeneracy of an isolated spin-3 atom (F-level in idiotic spectroscopic notation) should be split by a cubic crystal field into a singlet level and two triplets.

Actually the last statement is too sweeping to be true since it assumes that the splittings produced by the crystal field are small compared with the energy differences between states of different total angular momentum j. This condition is satisfied in the case of the rare-earths, which have incomplete $4f$ shells buried inside complete $5s$ and $5p$ shells, but in most cases it does not hold.

In most cases the best scheme is that in which one assumes that electrostatic interactions between electrons lock their orbital angular momenta together so that the system is characterized by a well defined total orbital angular momentum L. The crystal field then divides the $(2L+1)$-fold degenerate energy level associated with L into multiplets, and finally spin-orbit coupling splits up these multiplets into individual states.

Exercise (12):
Show that the characters of the classes of $\mathcal{O}$ in the $D^{(3)}$ irrep of $\mathcal{R}(3)$ are $(7,1,-1,-1,-1)$ in the order given in the character table above. Hence verify the decomposition of $D^{(3)}$ under $\mathcal{O}$ given in Example 4.

4.4 Two-particle states & Clebsch–Gordon coefficients

A trick often employed in physics is to break a system C down into two simpler sub-systems, A and B, whose dynamics can be exactly solved, and then to treat the original system as one in which A and B interact weakly. When we adopt this tack we represent states of C as compounds of states of A and B. Let the kets $|\alpha,i\rangle$, $|\beta,j\rangle$ and represent states i and j of A and B. Then from the product rule for combining probabilities it follows that the kets

$$|i,j\rangle \equiv |\alpha,i\rangle|\beta,j\rangle \tag{4.20}$$

represent states of the entire system C. For example, A and B might both be electrons moving in a central Coulomb potential, so that C is a model of a helium atom.

Now suppose the set $\{|\alpha,i\rangle\, i=1,\dots\}$ provides the α irrep of $\mathcal{G}$, while the set $\{|\beta,j\rangle\, j=1,\dots\}$ provides the β irrep of $\mathcal{G}$. Then it is clear that under $\mathcal{G}$ the kets $|i,j\rangle$ transform amongst themselves according to the representation $T^{(\alpha\times\beta)}$:

$$\begin{aligned} T(a)|i,j\rangle &= T^{(\alpha)}_{ik}(a)T^{(\beta)}_{jl}(a)|\alpha,k\rangle|\beta,l\rangle \\ &= T^{(\alpha\times\beta)}_{(ij)(kl)}(a)|\alpha,k\rangle|\beta,l\rangle. \end{aligned} \tag{4.21}$$

From the theory of §3.5 we know that the direct product representation can be decomposed into irreps of $\mathcal{G}$. That is, the matrices $\mathbf{T}^{(\alpha\times\beta)}$ can be simultaneously reduced to block diagonal form by a suitable unitary transformation $\mathbf{T}^{(\alpha\times\beta)} \to \mathbf{C}^\dagger\cdot\mathbf{T}^{(\alpha\times\beta)}\cdot\mathbf{C}$. This transformation is accomplished by using as basis vectors in the Hilbert space of C's kets linear combinations of the kets $|i,j\rangle$ defined by (4.20). The new basis vectors are

$$|\gamma,t,k\rangle = \sum_{ij} C(\gamma,t,k;\alpha,\beta,i,j)|\alpha,i\rangle|\beta,j\rangle. \tag{4.22}$$

The ket $|\gamma,t,k\rangle$ is the $k^{\rm th}$ basis vector of the $t^{\rm th}$ copy of the γ irrep of $\mathcal{G}$ in the decomposition of $T^{(\alpha\times\beta)}$. In most practical applications, the decomposition does not contain more than one copy of any representation, and the t label is formally redundant. However, even then it serves to indicate that its ket is one that describes a state of the entire system C, rather than of the component systems A and B.

The numbers $C(\gamma,t,k;\alpha,\beta,i,j)$ that form the matrix $\mathbf{C}$ can be deduced from a knowledge of the irreps of $\mathcal{G}$ alone, that is, without reference to the specific systems A, B and C (see §8 for this in the case $\mathcal{G}=\mathcal{R}(3)$). They are called Clebsch–Gordon coefficients.

Note:

The notation used here is non-standard. The conventional notation for $C(\gamma, t, k; \alpha, \beta, i, j)$ is $C(\alpha\beta\gamma t, ijk)$ but this seems to obscure the natural grouping of the indices.

Example 4.5

A two-electron ion has orbital angular momentum $l = 1$ and is polarized so that all this angular momentum is parallel to the z axis ($m = 1$). On energetic grounds, it is thought that one of the ion's electrons is in a state with $l = 1$ and the other has $l = 2$. One wishes to know the probability that the $l = 2$ electron has all its angular momentum parallel to the z axis (i.e. have $m = 2$). Wavefunctions of orbital angular momentum l provide the $D^{(l)}$ representation of $\mathcal{R}(3)$. The member of the $l = 1$ irrep provided by the ion's electronic wavefunction can be expressed as a linear combination of single-electron wavefunctions according to (4.22):

$$|1, t, 1\rangle = \sum_{i=-1}^{1} \sum_{j=-2}^{2} C(1, t, 1; 1, 2, i, j)|1, i\rangle|2, j\rangle, \tag{4.23}$$

where the C's are those appropriate to $\mathcal{R}(3)$. The probability we require is the mod-square of the coefficient of $|2, 2\rangle$.

The unitarity of the matrix C ensures that the kets $|\gamma, t, k\rangle$ for system C that are defined by (4.22) are orthogonal for two different values of γ. But at this stage it is not self-evident that kets for system A or B that belong to two different irreps of $\mathcal{G}$ are orthogonal. We shall see that when $\mathcal{G}$ is a compact Lie group, this orthogonality is guaranteed by the existence of a Casimir operator. For a finite group the demonstration of orthogonality goes like this. Let $\langle\alpha, i, /| \equiv T(a)\langle\alpha, i|$, then by the unitarity of the irreps, for any $a \in \mathcal{G}$ we can write

$$\langle\alpha, i|\beta, k\rangle = \langle\alpha, i, /|\beta, k, /\rangle = \sum_{jl} T_{ij}^{(\alpha)*}(a) T_{kl}^{(\beta)}(a)\langle\alpha, j|\beta, l\rangle. \tag{4.24}$$

Now we average both sides of the equation over all elements of $\mathcal{G}$ and use our basic orthogonality relation (3.10):

$$\begin{aligned}\langle\alpha, i|\beta, k\rangle &= \sum_{jl} \frac{1}{g} \sum_{a} T_{ij}^{(\alpha)*}(a) T_{kl}^{(\beta)}(a)\langle\alpha, j|\beta, l\rangle \\ &= \delta_{\alpha\beta}\delta_{ik}\frac{1}{s_\alpha}\sum_{j}\langle\alpha, j|\beta, j\rangle.\end{aligned} \tag{4.25}$$

This demonstrates (i) the required orthogonality of kets between irreps, and (ii) that $\langle\alpha, i|\alpha, i\rangle$ does not depend on i. (With properly normalized kets the last is anyway obvious, but the demonstration above assumes nothing about normalization, and we shall presently find this important.)

4.5 Wigner–Eckart theorem

Till now we have been concentrating on transformations between kets that represent different (if physically related) systems. In this subsection we consider representations that arise from a change of coordinate system. For example, when we rotate our Cartesian axes, it may be convenient to use new basis kets, for example those which correspond to states of well defined angular momentum around the new z-axis. These new basis kets will be linear combinations of the old, and the matrices defining these linear combinations for all possible reorientations of our axes will provide representations of $\mathcal{R}(3)$.

Similarly, once we've rotated axes, we'll want to work with new momentum operators p_i' that are linear combinations of the old ones:

$$p_i' = \sum_j R_{ij} p_j. \tag{4.26}$$

Thus in addition to the quintessential quantum-mechanical representations furnished by the Hilbert space of kets, we also have representations by sets of operators. These latter representations are the only ones known to classical physics.

Abstracting from the example above, let us suppose that we have a set $\{|\beta, i\rangle\}$ of kets that transform according to the β irrep of $\mathcal{G}$, and a set $\{Q_i\}$ of operators that transform according to the α irrep of $\mathcal{G}$:

$$T(a)Q_i = \sum_j T_{ij}^{(\alpha)}(a)Q_j. \tag{4.27}$$

(Such an operator set is called irreducible.) Then the kets obtained by applying Q_i to $|\beta, j\rangle$ transform according to the $T^{(\alpha)} \times T^{(\beta)}$ representation:

$$\begin{aligned} T(a)\big(Q_i|\beta, j\rangle\big) &= \sum_{kl} T_{ik}^{(\alpha)}(a)T_{jl}^{(\beta)}(a)Q_k|\beta, l\rangle \\ &= \sum_{kl} T_{(ij)(kl)}^{(\alpha\times\beta)}(a)Q_k|\beta, l\rangle. \end{aligned} \tag{4.28}$$

The matrix $C(\gamma, t, k; \alpha, \beta, i, j)$ introduced in equation (4.22) block-diagonalizes the matrices $\mathbf{T}^{(\alpha\times\beta)}(a)$, so it gives the linear combinations of the objects $Q_i|\beta, j\rangle$ that transform according to each irrep δ of $\mathcal{G}$. Let's denote these objects $|\delta, t, j; Q\rangle$. Then

$$Q_k|\beta, l\rangle = \sum_{\delta,t,j} C^*(\delta, t, j; \alpha, \beta, k, l)|\delta, t, j; Q\rangle. \tag{4.29}$$

In a number of practical applications one wishes to calculate the matrix element $\langle\gamma, i|Q_j|\beta, k\rangle$. Dotting (4.29) through by $\langle\gamma, i|$ we have

$$\langle\gamma, i|Q_k|\beta, l\rangle = \sum_{\delta,t,j} C^*(\delta, t, j; \alpha, \beta, k, l)\langle\gamma, i|\delta, t, j; Q\rangle. \tag{4.30}$$

Now the same line of reasoning that was used to derive (4.25) can be deployed to show that the matrix element on the rhs of (4.30) vanishes for $\gamma \neq \delta$ and/or $i \neq j$. In fact, since $|\delta, t, j; Q\rangle$ transforms under $\mathcal{G}$ according to the δ irrep, we have

$$\langle\gamma, i|\delta, t, j; Q\rangle = \delta_{\gamma\delta}\delta_{ij}\frac{1}{s_\gamma}\sum_m\langle\gamma, m|\delta, t, m; Q\rangle. \tag{4.31}$$

This shows not only that the left side vanishes unless $\gamma = \delta$ and $i = j$, but also that it is independent of i – in other words its only non-vanishing component is a function of γ alone. One usually writes

$$\langle\gamma, i|\delta, t, j; Q\rangle = \delta_{\gamma\delta}\delta_{ij}\langle\gamma||Q||\gamma\rangle \quad \text{no sum on } \gamma. \tag{4.32}$$

This useful result is called the **Wigner-Eckart theorem**. $\langle\gamma||Q||\gamma\rangle$ is called a **reduced matrix element** of Q. Plugging (4.32) back into (4.30) we have

$$\langle\gamma, i|Q_k|\beta, l\rangle = \sum_t C(\gamma, t, i; \alpha, \beta, k, l)\langle\gamma||Q||\gamma\rangle. \tag{4.33}$$

In most applications only one value of t arises. In any event, the physically interesting matrix element on the left is determined by the Clebsch-Gordon coefficients up to one normalizing number per irrep γ.

5 Important Lie Groups

A Lie group is a set that, in addition to satisfying the group axioms of §1, has all the structure of a differentiable manifold:

Definition:

A Lie group is a set $\mathcal{G}$ such that: (i) $\mathcal{G}$ is a group; (ii) $\mathcal{G}$ is an analytic manifold; (iii) the mapping $(a, b) \to ab$ of $\mathcal{G} \otimes \mathcal{G}$ into $\mathcal{G}$ is analytic.

Clearly $\mathcal{G}$ is an infinite group.[4] The interplay between its two very different structures – the algebraic structure of the group and the analytic structure of the manifold – has far-reaching consequences.

Example 5.1

The (Abelian) group of translations in $\mathcal{R}^3$ is a Lie group; any translation can be labelled by the vector $\mathbf{a}$ which we add to the position-vector of any point in order to move it, and the components a_x etc of $\mathbf{a}$ constitute a (globally-valid) chart for the group. Thus, *qua* manifold, this Lie group is identical with $\mathcal{R}^3$.

Example 5.2

The group $\mathcal{R}(3)$ of three-dimensional rotations is a Lie group. The elements of $\mathcal{R}(3)$ are the rotations $R(\hat{\mathbf{n}}, \theta)$ by angle θ about the axis $\hat{\mathbf{n}}$. It is easy to see that

$$R(\hat{\mathbf{n}}, \theta)\mathbf{r} = (\hat{\mathbf{n}} \cdot \mathbf{r})\hat{\mathbf{n}} + \cos\theta(\mathbf{r} - (\hat{\mathbf{n}} \cdot \mathbf{r})\hat{\mathbf{n}}) + \sin\theta(\hat{\mathbf{n}} \times \mathbf{r}). \tag{5.1}$$

It is often convenient to write

$$R(\mathbf{a}) \equiv R(\hat{\mathbf{n}}, \theta) \quad \text{where} \quad \mathbf{a} \equiv \theta\hat{\mathbf{n}}. \tag{5.2}$$

The three (independent) components of $\mathbf{a}$ constitute a chart for $\mathcal{R}(3)$. So locally $\mathcal{R}(3)$ is homeomorphic to $\mathcal{R}^3$. Globally it is quite unlike $\mathcal{R}^3$ because $R(\hat{\mathbf{n}}, 2m\pi) = e \ \forall \ \hat{\mathbf{n}}, m$.

Example 5.3

The **group of Mobius transformations** comprises the transformations of the complex plane

$$z \to z' = \frac{az + b}{cz + d} \quad \text{where} \quad a, b, c, d \in \mathcal{C} \quad \text{and} \quad ad - bc = 1. \tag{5.3}$$

This Lie group has 6 real dimensions because its four complex parameters are constrained by 1 equation (which has both a real and an imaginary part).

Exercise (13):

Prove that the set of Mobius transformations forms a group.

Example 5.4

$GL(n, \mathcal{R})$ is the set of all $n \times n$ real matrices $\mathbf{M}$ with non-zero determinant. It forms an n^2-dimensional Lie group under multiplication and is called the **real general linear group**. Analogously $GL(n, \mathcal{C})$ is the **complex general linear group** and has real-dimensionality $2n^2$. $GL(n, \mathcal{R})$ is divided by the condition $|\mathbf{M}| \neq 0$ into two disconnected components, that in which $|\mathbf{M}| > 0$, and that in which $|\mathbf{M}| < 0$. $GL(n, \mathcal{C})$ is not so divided because a complex-valued $|\mathbf{M}|$ can pass smoothly from real positive values to real negative values.

Example 5.5

The **special linear group** $SL(n, \mathcal{R})$ is the subset of $GL(n, \mathcal{R})$ in which $|\mathbf{M}| = 1$. It has dimension $n^2 - 1$. $SL(n, \mathcal{C})$ is similarly defined and has real-dimension $2n^2 - 2$. Example 3 showed how the elements of $SL(2, \mathcal{C})$ generate Lorentz transformations.

[4] Unless, a pedant would point out, it is of dimension zero.

Example 5.6

The real orthogonal group $O(n)$ is the subset of $GL(n,\mathcal{R})$ in which $\mathbf{M}^\dagger = \mathbf{M}^{-1}$. In component form the equation $\mathbf{M}^\dagger \cdot \mathbf{M} = \mathbf{I}$ reads

$$B_{ik} \equiv \sum_j M_{ji} M_{jk} = \delta_{ik}. \tag{5.4}$$

This imposes $\frac{1}{2}n(n+1)$ independent conditions on the components of $\mathbf{M}$ ($\mathbf{B}$ is necessarily symmetric). Consequently, the group has dimension $n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$. In particular for $n = 3$ its dimension is 3 in agreement with Example 2, which discusses the component in which $|\mathbf{M}| = +1$. Taking the trace of (5.4) we have $\sum_{ij} M_{ij}^2 = n$, which demonstrates that the elements of $O(n)$ are confined to the (n^2-1)-sphere $S^{(n^2-1)}$ in n^2-dimensional space. Moreover, it is the inverse image of the closed[5] set $\{\mathbf{I}\}$ under the continuous map of $GL(n,\mathcal{R})$ to itself that is obtained by compounding this sequence of maps: $\mathbf{M} \to (\mathbf{M},\mathbf{M}) \to (\mathbf{M},\mathbf{M}^\dagger) \to \mathbf{M}\cdot\mathbf{M}^\dagger$. Since the inverse-image of a closed set under a continuous map is closed, this implies that $O(n)$ is a closed subset of $S^{(n^2-1)}$. These two facts imply that $\mathcal{O}(n)$ forms a compact manifold.[6] Its compactness has far-reaching consequences.

Example 5.7

The **special orthogonal group** $SO(n)$ is the subgroup of $O(n)$ formed by matrices $\mathbf{M}$ with $|\mathbf{M}| = +1$. For $n = 3$ it coincides with $\mathcal{R}(3)$.

Example 5.8

The **unitary group** $U(n)$ is the subset of $GL(n,\mathcal{C})$ in which $\mathbf{M}^\dagger = \mathbf{M}^{-1}$. The off-diagonal components of the constraint equation $\mathbf{B} \equiv \mathbf{M}\cdot\mathbf{M}^\dagger = \mathbf{I}$ imposes $n(n-1)$ real conditions on the components of $\mathbf{M}$ ($\mathbf{B}$ is necessarily Hermitian), while the diagonal components impose a further n constraints (the diagonal components of $\mathbf{B}$ are necessarily real). Hence $U(n)$ has real-dimensionality $2n^2 - n(n-1) - n = n^2$.

Example 5.9

The **special unitary group** $SU(n)$ is the subset of $U(n)$ in which $|\mathbf{M}| = 1$. This is the most important Lie group for physicists. It has dimension $n^2 - 1$, so $SU(2)$, like $O(3)$ has dimension 3, and we shall find that these two groups are closely related. The Lorentz group will prove to be a direct product of $SU(2)$ with itself. The standard electroweak model is built around a direct product of $U(1)$ and $SU(2)$. Chromodynamics, the theory of how hadrons are built up of quarks, is built around $SU(3)$. Not so long ago it seemed that $SU(5)$ might provide a unified theory of strong, weak and electromagnetic interactions.

6 General theory of Lie Groups

The literature on Lie groups suffers from a kind of schizophrenia in that there is very little overlap between pure-mathematically oriented books and those used by physicists. The former contain general theory and few useful results, while the latter are full of 'it may be shown' statements and numerous useful results.

Elliott & Dawber is a good example of the latter genre. I find it rather unsatisfying in that (i) the discussion is almost entirely in terms of representations rather than the underlying

[5] Closed is here used in the sense of general topology, which recognizes open sets, closed sets and sets that are both open and closed. An open set is one such that every point in the set sits in a neighbourhood ('nhd') of points of the set. A closed set is the complement of an open set. Thus the subset of the circle $\{\theta;\ 2 > \theta > 3\}$ is open because no matter how close θ is to either 2 or 3, there are always points closer still, so we may find a small interval (nhd) in which θ sits. The complement set $\{\theta; \theta \leq 2 \text{ or } \theta \geq 3\}$ is closed.

[6] A **compact manifold** is one such that every open covering contains a finite sub-cover. In effect, a compact manifold doesn't have space for an infinite number of reasonably disjoint open sets.

Box 1: Basics of Differential Geometry

Let M be a manifold. A **chart** x for an open set $W \subset M$ is a 1-1, mapping $x : W \to \mathcal{R}^n$ of W into Euclidean n-space. The coordinates $x^i(p)$ of the image point are "the coordinates of p in the chart x".

The analytic, real-valued functions on M form an ∞-dimensional vector space $A(M)$. The **tangent vectors** at p are defined to be the linear operators $v : A \to \mathcal{R}$ that satisfy

$$\begin{aligned} v(\alpha f + \beta g) &= \alpha v(f) + \beta v(g), \\ v(fg) &= v(f)g(p) + f(p)v(g); \end{aligned} \qquad \alpha, \beta \in \mathcal{R},\ f, g \in A(M). \tag{B1.1}$$

The first condition states linearity, the second that tangent vectors work like differential operators. In fact, in coordinate form v can be expanded

$$v = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x^i} \quad \Rightarrow \quad v(f) = \sum_{i=1}^{n} v_i \frac{\partial \overline{f}}{\partial x^i}, \tag{B1.2}$$

where $\overline{f}$ is the coordinate form of f in the chart x. That is, $f(p) = \overline{f}\big(\mathbf{x}(p)\big)$.

A **vector field** V (a.k.a. 'infinitesimal transformation') is a mapping $V : A(M) \to A(M)$ that satisfies conditions analogous to (B1.1)

$$\begin{aligned} V(\alpha f + \beta g) &= \alpha V(f) + \beta V(g), \\ V(fg) &= V(f)g + fV(g); \end{aligned} \qquad \alpha, \beta \in \mathcal{R},\ f, g \in A(M). \tag{B1.3}$$

If U is a second vector field, the compound (UV) is also a linear map $UV : A(M) \to A(M)$, but it is not a vector field because it does not satisfy the second of conditions (B1.3). The commutator $[U, V] \equiv UV - VU$ is a vector field, however:

$$\begin{aligned} [U,V](fg) &= UV(fg) - VU(fg) = U\big(V(f)g + fV(g)\big) - V\big(U(f)g + fU(g)\big) \\ &= U\big(V(f)\big)g + V(f)U(g) + U(f)V(g) + fU\big(V(g)\big) \\ &\quad - V\big(U(f)\big)g - U(f)V(g) - V(f)U(g) - fV\big(U(g)\big) \\ &= [U,V](f)g - f[U,V](g). \end{aligned} \tag{B1.4}$$

$[U, V]$ is called the **Lie product** of U and V.

Let N be a second manifold and h a mapping $h : M \to N$. Then h induces a mapping $h* : T_p(M) \to T_{h(p)}(N)$ between the tangent spaces of mapped points. Indeed, let $v \in T_p(M)$. Then $(h * v) \in T_{h(p)}(N)$ is defined by its action on functions $f \in A(N)$:

$$(h * v)(f) = v(f \circ h) \equiv v\big(f(h())\big). \tag{B1.5}$$

$h * v$ measures the gradient in the direction of v of the composite function $f \circ h : M \to \mathcal{R}$

group, and (ii) it is irritating to perpetually read 'it may be shown that' without having the least clue *how* the demonstration might run.

This section represents an attempt to bridge this unfortunate gulf between the intellectual worlds of mathematicians and physicists. I am by no means satisfied with either the strength or the accessibility of my bridge. But the gulf is both wide and deep, and my powers puny.

If you are not stoutly shod with a robust knowledge of coordinate-free differential geometry, I recommend that rather than struggling to clamber over my bridge, you allow yourself to be magically wafted over the gulf, and skip to §7.

6.1 Summary of the theory

The key feature of a Lie group is that its structure is largely determined by its structure in a nhd of e. Consequently, we shall be able to characterize Lie groups by taking a close look at the structure of the tangent space at e. This dependence of the entire structure on a nhd of e arises because axiom (iii) above ensures that the map

$$l_a : \mathcal{G} \to \mathcal{G};\ l_a b = ab \tag{6.1}$$

constitutes a homeomorphism between a nhd of b and a nhd of ab. Thus taking b to lie in a nhd of e, we obtain a homeomorphism between a nhd of e and a nhd of the arbitrary point a. On account of this homeomorphism, the validity of many results can be immediately extended from nhds of e to the group as a whole.

In outline the argument runs as follows – Box 1 defines much of the notation and summarizes the underlying differential geometry. In addition we need

Definition:

A Lie algebra is a vector space Λ on which a binary product $[,]$ is defined that (i) is antisymmetric ($[X, Y] = -[Y, X]$), (ii) is linear in each slot ($[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$), and (iii) satisfies the **Jacobi identity**:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \tag{6.2}$$

- We use the homeomorphism (6.1) to associate a vector field L^v with every $v \in T_e(\mathcal{G})$. The set of all such vector fields we call $\Lambda(\mathcal{G})$. It is a vector space.
- We show that the Lie product $[L^u, L^v]$ of any two of these vector fields is itself the vector field $L^{[uv]}$ associated with a vector $[uv] \in T_e(\mathcal{G})$. This establishes that $\Lambda(\mathcal{G})$ is a Lie algebra. The map $(u, v) \to [uv]$ makes $T_e(\mathcal{G})$ a Lie algebra also. It is isomorphic to $\Lambda(\mathcal{G})$, and it's often convenient to regarded it, rather than $\Lambda(\mathcal{G})$, as $\mathcal{G}$'s Lie algebra.
- Let $e_1, \ldots, e_n$ form a basis of $T_e(\mathcal{G})$. Then $[e_i e_j]$, being an element of $T_e(\mathcal{G})$, can be expressed as a linear combination of the e_i:

$$[e_i e_j] = \sum_k c_{ij}^k e_k. \tag{6.3}$$

The coefficients c_{ij}^k are called $\mathcal{G}$'s **structure constants**.

- We show that for every vector field $V \in \Lambda(\mathcal{G})$ we can find a **complete integral curve** $e^v(t)$ through e ($t \in \mathcal{R}$). That is, we can find a path from e through $\mathcal{G}$ whose tangent at any point $e^v(t)$ is the local value $V_{e^v(t)}$ of V – in terms of the map $e^v : \mathcal{R} \to \mathcal{G}$ we have

$$e^v * \frac{\mathrm{d}}{\mathrm{d}t} = V_{e^v(t)}. \tag{6.4}$$

This path is complete in the sense that $e^v(t)$ is defined for all t.

- The dependence of $e^v(t)$ on t and v is such that $e^v(t) = e^{tv}$. In honour of this result we write $\exp(tv) \equiv e^v(t)$. We have

$$\begin{aligned} \exp(0v) = e\ \forall\ v \in T_e(\mathcal{G}) \quad \text{and} \\ \exp\big((t_1 + t_2)v\big) = \exp(t_1 v)\exp(t_2 v)\ \forall\ t_1 t_2 \in \mathcal{R}, \end{aligned} \tag{6.5}$$

as the exponential notation suggests. The map e^v is called the **exponential map**.

- It follows that to each $v \in T_e(\mathcal{G})$ there corresponds a 1-dimensional subgroup $\mathcal{H}^v \subset \mathcal{G}$. This is the Abelian group formed by the elements $a \in \mathcal{G}$ that lie on the path e^v through e:

$$\mathcal{H}^v = \{e^v(t);\ t \in \mathcal{R}\}. \tag{6.6}$$

- Every 1-dimensional subgroup of $\mathcal{G}$ proves to be a set of the form

$$\{a = \exp(tv),\ v \text{ fixed}\}. \tag{6.7}$$

When $\mathcal{G}$ is compact and connected, every element $a \in \mathcal{G}$ is of the form $a = \exp(v)$ for some $v \in T_e(\mathcal{G})$.

- The map $\exp : T_e(\mathcal{G}) \to W \subset \mathcal{G};\ v \to \exp(v)$, where W is a (probably large) nhd of e, proves to be a diffeomorphism of the linear vector space $T_e(\mathcal{G})$ and $\mathcal{G}$.
- The map $l_a \circ e^v : \mathcal{R} \to \mathcal{G};\ t \to ae^v(t)$ provides an integral curve of $V \equiv L^v$ through any point $a \in \mathcal{G}$:

$$(l_a \circ e^v) * \frac{\mathrm{d}}{\mathrm{d}t} = V_{ae^v(t)}. \tag{6.8}$$

These results show that $\mathcal{G}$ is to a great extent characterized by its Lie algebra $\Lambda(\mathcal{G})$. This reduces a problem in differential geometry to one in algebra.

6.2 Proof of key results

By (B1.5) the mapping l_a defined by (6.1) induces a mapping

$$l_a * : T_b(\mathcal{G}) \to T_{l_a(b)}(\mathcal{G}) = T_{ab}(\mathcal{G}). \tag{6.9}$$

So putting $b = e$ we see that any $v \in T_e(\mathcal{G})$ can be extended to a vector field L^v valid throughout $\mathcal{G}$:

$$L^v_a \equiv l_a * v \quad \forall\, v \in \mathcal{G}. \tag{6.10}$$

Definition:

A **left invariant** vector field X is one such that

$$l_a * X = X \quad \forall\, a \in \mathcal{G}. \tag{6.11}$$

Theorem 11

L^v is left invariant. Moreover, every left-invariant field is of the form L^v for some $v \in T_e(\mathcal{G})$.

Proof: Since for any $f \in A(\mathcal{G})$, $\big(l_c * (l_a * v)\big)(f) = (l_a * v)(f \circ l_c) = v(f \circ l_c \circ l_a) = v(f \circ l_{ca}) = l_{ca} * v(f)$, we have that L^v is left invariant:

$$(l_c * L^v)_{ca} = l_c * L^v_a = l_c * (l_a * v) = l_{ca} * v. \tag{6.12}$$

If X is left-invariant, then $X_a = l_a * X_e = L^x_a$, where $x \equiv X_e$. □

This theorem shows that the space $\Lambda(\mathcal{G})$ of left-invariant fields is isomorphic with $T_e(\mathcal{G})$. In particular, these two spaces have the same finite dimension.

Theorem 12

The Lie product of any two left-invariant fields is itself left-invariant

Proof:

$$\begin{aligned}(l_a * [X,Y])(f) &= \big(l_a * (XY)\big)(f) - \big(l_a * (YX)\big)(f)\\ &= X\big(Y(f \circ l_a)\big) - Y\big(X(f \circ l_a)\big)\\ &= X\big(l_a * Y(f)\big) - Y\big(l_a * X(f)\big)\\ &= (XY - YX)(f) = [X,Y](f). \qquad \Box\end{aligned} \tag{6.13}$$

Since every left-invariant field is of the form L^v for some $v \in T_e(\mathcal{G})$, for each $u, v \in T_e(\mathcal{G})$ there is a unique vector $[uv] \in T_e(\mathcal{G})$ such that $L^{[uv]} = [L^u, L^v]$. This association makes $T_e(\mathcal{G})$ into a Lie algebra.

Since $\Lambda(\mathcal{G})$ is a finite-dimensional vector field, we can choose a basis of left-invariant fields $E_1, \ldots, E_n$. The Lie product $[E_i, E_j]$, being left-invariant, can be expressed as a linear combination of the E_k:

$$[E_i, E_j] = \sum_k c_{ij}^k E_k, \tag{6.14}$$

where the c_{ij}^k are the structure constants introduced in (6.3). Equation (6.14) explains why they are called 'constants': the coefficients here could in principle vary with position in $\mathcal{G}$, but they do not.

Exercise (14):

Prove that

$$c_{jk}^i c_{im}^l + c_{km}^i c_{ij}^l + c_{mj}^i c_{ik}^l = 0. \tag{6.15}$$

The proof that the exponential map $\exp(tv)$ can be evaluated for arbitrarily large t is tedious, but the basic idea is simple: By expressing the equation

$$e^v * \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_t = L^v_{e^v(t)} \tag{6.16}$$

in coordinate form,[7] we may show that it can be solved in a sufficiently small nhd W of e. Hence we may assume that $e^v(t)$ is defined for $|t| < \epsilon > 0$. Then we show that for $|t_1|, |t_2| < \epsilon/2$, $e^v(t_1)e^v(t_2) = e^v(t_1 + t_2)$ and go on to argue that for $\epsilon < |t| < n\epsilon$ we may define $e^v(t)$ to be

$$e^v(t) = \Big(e^v(t/n)\Big)^n, \tag{6.17}$$

that is, as the n^{th} power of $e^v(t/n)$.

6.3 Concrete calculations

After this survey of the general theory of Lie groups, the logical next step is to obtain coordinate representations of the quantities we have encountered along the way – the left-invariant fields L^v, structure constants c_{ij}^k, exponential map etc – for the most important of the groups listed in §5. Unfortunately, at this stage we come up against a block: for nearly all these groups it is fiendishly difficult to express the result of the group product $c = ab$ in terms of the group's natural chart. For example, in the case of $\mathcal{R}(3)$ the natural chart is provided by the three components of the vector $\mathbf{a}$ defined in (5.2), and to proceed we need to obtain the functional dependence $\mathbf{a}''(\mathbf{a}', \mathbf{a})$ where the $\mathbf{a}$'s are defined by $R(\mathbf{a}'') = R(\mathbf{a}')R(\mathbf{a})$. This problem of determining the **composition function** for the group in the natural chart is a well-defined but usually excessively nasty problem.

Now all the groups of §5 are submanifolds (and subgroups) of either $GL(n, \mathcal{R})$ or $GL(n, \mathcal{C})$. The GL groups are not mere manifolds; each $a \in GL(n, \mathcal{R})$ is effectively a matrix $\mathbf{A}$, and matrices can be multiplied by numbers and added. So the GL groups are linear vector spaces. Consequently, in their case the machinery of differential geometry is not necessary: every tangent space is isomorphic to the entire manifold $\mathcal{G}$, and since we can now difference elements $a \in \mathcal{G}$, we can associate each $v \in T_e(\mathcal{G})$ with the slope $\mathrm{d}a/\mathrm{d}\lambda$ of a path $a(\lambda)$ from e through $\mathcal{G}$.

Moreover, for the GL groups the problem of determining the coordinate form of the composition function is straightforward. The natural chart is provided by the n^2 components of

[7] This is $(\mathrm{d}/\mathrm{d}t)x^i(e^v(t)) = L^v_{e^v(t)}x^i = v_j\psi^i_j(e^v(t))$, where $\psi^i_j(a) \equiv \frac{\partial x^i(ab)}{\partial x^j(b)}|_{b=e}$

the matrix $\mathbf{A}$ that represents the element a: $x^{(i,j)}(a) = A_{ij}$. Hence the composition function is trivial: if $c = ab$, then

$$C_{ij} = \sum_k A_{ik} B_{kj}. \tag{6.18}$$

If a, b, c happen all to belong to a subgroup $\mathcal{H} \subset GL(n, \mathcal{R})$, for example to $\mathcal{R}(3)$, then (6.18) still provides a valid relationship between the components of the matrices that represent these elements. It is just that the matrix elements aren't all independent, and therefore don't provide a valid chart for the subgroup. The way we get around this problem is to calculate quantities for elements of $\mathcal{H} \subset GL(n, \mathcal{R})$ in their capacity as elements of $GL(n, \mathcal{R})$.

Theorem 13

Let $\mathcal{G} \subset GL(n, \mathcal{R})$ be a Lie group, and $u, v \in T_e(\mathcal{G})$ have components u_{ij}, v_{ij} in the natural chart of $GL(n, \mathcal{R})$. Then the components of $[uv] \in T_e(\mathcal{G})$ are the elements of the matrix $\mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u}$.

In other words, the components of the vector $[uv] \in T_e(\mathcal{G})$ that generates the left-invariant field $[L^u, L^v]$ form the commutator of the matrices formed by the components of u and v.

Proof: Since (i) $T_e(\mathcal{G}) \subset T_e(GL(n, \mathcal{R}))$, (ii) all elements of $\mathcal{G}$ are also elements of $GL(n, \mathcal{R})$, and (iii) multiplication in $\mathcal{G}$ coincides with multiplication in $GL(n, \mathcal{R})$, we can proceed under the assumption that everything we require pertains to $GL(n, \mathcal{R})$. We first find the coordinate representation of L^v for $v \in T_e\big(GL(n, \mathcal{R})\big)$. v can be written as a linear combination of derivatives $\partial/\partial M_{ij}$ with respect to individual matrix elements M_{ij}:

$$v = \sum_{kl} v_{kl} \frac{\partial}{\partial M_{kl}}\bigg|_{\mathbf{M}=\mathbf{I}}. \tag{6.19}$$

Now let f be an arbitrary function on $GL(n, \mathcal{R})$. Then by (6.10)

$$L^v_a f = \sum_{ijkl} v_{kl} \frac{\partial \overline{f}(\mathbf{B})}{\partial B_{ij}} \frac{\partial B_{ij}}{\partial M_{kl}}\bigg|_{\mathbf{M}=\mathbf{I}}, \quad \text{where} \quad \mathbf{B} \equiv \mathbf{A} \cdot \mathbf{M}. \tag{6.20}$$

Evaluating the derivative of $\mathbf{B}$ at $\mathbf{M} = \mathbf{I}$, we easily find

$$L^v_a f = \sum_{ikl} v_{kl} A_{ik} \frac{\partial \overline{f}(\mathbf{A})}{\partial A_{il}}. \tag{6.21}$$

Now that we have the coordinate representation of L^v we can evaluate its Lie product with L^u:

$$\begin{aligned}
[L^u, L^v]_e(f) &= L^u_e\big(L^v(f)\big) - L^v_e\big(L^u(f)\big) \\
&= u_{kl} A_{ik} \frac{\partial}{\partial A_{il}} \left(v_{rs} A_{jr} \frac{\partial \overline{f}}{\partial A_{js}} \right)\bigg|_{\mathbf{A}=\mathbf{I}} - (v \leftrightarrow u) \\
&= \big(u_{kl} v_{rs} - v_{kl} u_{rs}\big) \delta_{ik} \delta_{ij} \delta_{lr} \frac{\partial \overline{f}}{\partial A_{js}}\bigg|_{\mathbf{A}=\mathbf{I}} \\
&= (u_{kr} v_{rs} - v_{kr} u_{rs}) \frac{\partial \overline{f}}{\partial A_{ks}}\bigg|_{\mathbf{A}=\mathbf{I}} \\
&= (\mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u})_{ks} \frac{\partial \overline{f}}{\partial A_{ks}}\bigg|_{\mathbf{A}=\mathbf{I}}. \qquad \square
\end{aligned} \tag{6.22}$$

Theorem 14

In the natural chart $\mathbf{x}$ of $GL(n, \mathcal{R})$, the exponential map is for $v \in \mathcal{G} \subset GL(n, \mathcal{R})$ given by

$$\mathbf{x}\big(\exp(tv)\big) = \mathbf{I} + t\mathbf{v} + \tfrac{1}{2!}(t\mathbf{v})^2 + \cdots. \tag{6.23}$$

Proof: Let $\mathbf{S}(t)$ be the given power series. Then $\mathrm{d}\mathbf{S}/\mathrm{d}t = \mathbf{S}(t)\cdot\mathbf{v}$ and $\mathbf{S}(0) = \mathbf{I}$. On the other hand, the coordinate representation of $L^v_{s(t)}$ is

$$v_{ij}\frac{\partial}{\partial M_{ij}}\big(S_{kl}(t)M_{lm}\big) = v_{ij}S_{kl}(t)\delta_{il}\delta_{jm} \\ = \big(\mathbf{S}(t)\cdot\mathbf{v}\big)_{km}. \tag{6.24}$$

Thus $\mathbf{S}(t) = \mathbf{x}(e^v_t)$ satisfies the coordinate form of the equation $e^v_t * (\mathrm{d}/\mathrm{d}t) = L^v_{e^v(t)}$. □

7 Infinitesimal generators

We now look at the Lie algebras Λ of the most important Lie groups $\mathcal{G}$ of §5. In §6 we saw that it is most convenient to treat the elements of all these groups as elements of the embedding $GL(n,)$ group, which enables us to associate each $a \in \mathcal{G}$ and each $v \in \Lambda(\mathcal{G})$ with an $n \times n$ matrix $\mathbf{A}$ or $\mathbf{v}$.

By theorem 14, every such $\mathbf{v}$ is associated with the 1-dimensional subgroup of $GL(n,\mathcal{R})$ whose elements are $\mathbf{A}(t) = \exp(t\mathbf{v})$. So if we happen to know of such a subgroup, we can recover its **infinitesimal generator v** through

$$\mathbf{v} = \left.\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\right|_{t=0}. \tag{7.1}$$

7.1 The Lie algebra $\Lambda(\mathcal{R}(3))$

Consider the subgroup of rotations through angle t around the z-axis. We have

$$\mathbf{A}(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad\Rightarrow\quad \mathbf{v}\equiv\mathbf{v}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7.2}$$

Similarly, differentiating the matrices for rotation by angles t about the x and y-axes, we find from (7.1) their generators to be

$$\mathbf{v}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad \mathbf{v}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{7.3}$$

From §§5 and 6 we know that $\Lambda(\mathcal{R}(3))$ is 3-dimensional. Hence these three generators provide a basis for $\Lambda(\mathcal{R}(3))$. From theorem 13 we know that the Lie product of $u, v \in \Lambda(\mathcal{G})$ is just the commutator of the matrices $\mathbf{u}$, $\mathbf{v}$. We easily find

$$[\mathbf{v}_i, \mathbf{v}_j] = \epsilon_{ijk}\mathbf{v}_k \qquad (i, j, k \in \{1, 2, 3\}). \tag{7.4}$$

Exercise (15):

Verify by explicit calculation that the matrix $\exp(t\mathbf{v}_3)$ coincides with $\mathbf{A}$ in (7.2).

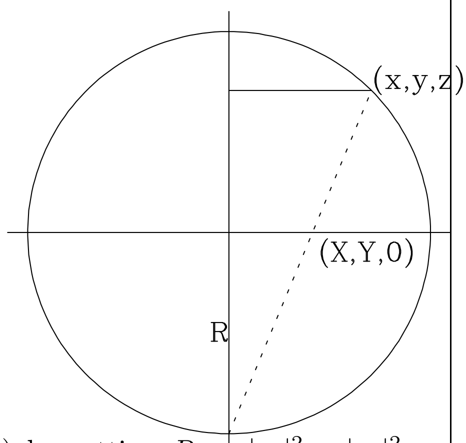
7.2 The Lie algebra $\Lambda(SU(2))$

For very small t, $\mathbf{A}(t) = \exp(t\mathbf{v}) \simeq \mathbf{I} + t\mathbf{v}$. If $\mathbf{A}$ is to belong to $SU(2)$, we must have

$$\mathbf{A}\cdot\mathbf{A}^\dagger = \mathbf{I} \quad\Rightarrow\quad \mathbf{I} = (\mathbf{I} + t\mathbf{v} + \cdots)\cdot(\mathbf{I} + t\mathbf{v}^\dagger + \cdots) \quad\Rightarrow\quad \mathbf{v} + \mathbf{v}^\dagger = 0 \tag{7.5}$$

Box 2: The Relation between $SU(2)$ and $\mathcal{R}(3)$

Spatial rotations move points around spheres. Consider the sphere of radius R. Positions on this sphere are most simply described by stereographically projecting points (x, y, z) on the sphere to points $(X, Y, 0)$ in the plane $z = 0$. Positions in this plane are given by $\zeta \equiv (X + \mathrm{i}Y)/R$:



$$x + \mathrm{i}y = 2R\frac{\zeta}{1+\zeta\zeta^*} \quad ; \quad z = R\frac{1-\zeta\zeta^*}{1+\zeta\zeta^*},$$

Writing $\zeta = \eta_2/\eta_1$, we find

$$x + \mathrm{i}y = 2R\frac{\eta_1^*\eta_2}{|\eta_1|^2+|\eta_2|^2} \quad ; \quad z = R\frac{|\eta_1|^2-|\eta_2|^2}{|\eta_1|^2+|\eta_2|^2}.$$

We fix the length of the complex 2-vector $\boldsymbol{\eta} \equiv (\eta_1, \eta_2)$ by setting $R = |\eta_1|^2 + |\eta_2|^2$. This normalization is clearly invariant under unitary transformations $\boldsymbol{\eta} \to \boldsymbol{\eta}' \equiv \mathbf{M}\cdot\boldsymbol{\eta}$. Moreover, one may easily show that the squared distance $(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2$ between any two points $\mathbf{r}_1$ and $\mathbf{r}_2$ on the sphere is unchanged when $\mathbf{r}_1$ is mapped to $\mathbf{r}_1'$ and $\mathbf{r}_2$ to $\mathbf{r}_2'$ by the operation $\boldsymbol{\eta} \to \boldsymbol{\eta}'$. So unitary transformations of the two-dimensional complex space of the $\boldsymbol{\eta}$'s generate rotations of the sphere. A complex 2-vector $\boldsymbol{\eta}$ is called a **Weyl spinor**.

In words, the infinitesimal generator $\mathbf{v}$ must be skew-Hermitian. We require, moreover, that

$$|\mathbf{A}| = 1 \quad \Rightarrow \quad \begin{vmatrix} 1+tv_{11}+\cdots & tv_{12}+\cdots \\ -(tv_{12}+\cdots)^* & 1+tv_{22}+\cdots \end{vmatrix} \quad \Rightarrow \quad v_{11}+v_{22}=0. \tag{7.6}$$

Thus $\mathbf{v}$ is of the form

$$\mathbf{v} = \mathrm{i}\begin{pmatrix} \alpha & \beta^* \\ \beta & -\alpha \end{pmatrix} = \mathrm{i}\begin{pmatrix} 0 & \beta^* \\ \beta & 0 \end{pmatrix} + \mathrm{i}n_z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{7.7}$$

where β is complex and n_z real. Now writing $\beta = n_x + \mathrm{i}n_y$ $(n_i \in \mathcal{R})$, we may rewrite (7.7) in the form

$$\mathbf{v} = \mathrm{i}\big(n_x\boldsymbol{\sigma}_x + n_y\boldsymbol{\sigma}_y + n_z\boldsymbol{\sigma}_z\big), \tag{7.8}$$

where the $\boldsymbol{\sigma}_i$ are the usual Pauli matrices. Thus the Pauli matrices form a basis for the Lie algebra of $SU(2)$.

Theorem 15

$\Lambda(SU(2))$ is isomorphic to $\Lambda(\mathcal{R}(3))$

Proof: Define $\mathbf{v}_i \equiv -\frac{\mathrm{i}}{2}\boldsymbol{\sigma}_i$. Then by (7.8) the $\mathbf{v}_i$ form a basis for $\Lambda(SU(2))$. Moreover from the usual commutation relation $[\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j] = 2\mathrm{i}\epsilon_{ijk}\boldsymbol{\sigma}_k$ of the Pauli matrices, we have $[\mathbf{v}_i, \mathbf{v}_j] = \epsilon_{ijk}\mathbf{v}_k$, which is precisely the commutation relationship (7.4) of $\Lambda(R(3))$. □

Since the structure of a Lie group is very nearly determined by its Lie algebra, this isomorphism of $\Lambda(SU(2))$ and $\Lambda(\mathcal{R}(3))$ implies that $\mathcal{R}(3)$ is closely related to $SU(2)$. Box 2 relates the groups by an explicit construction.

7.3 The Lie algebra $\Lambda(SU(3))$

An argument closely analogous to that deployed in the case of $SU(2)$ shows that a generator of $SU(3)$ must be of the form

$$\mathbf{v} = \mathrm{i}\begin{pmatrix} \alpha & \beta^* & \gamma^* \\ \beta & \delta & \epsilon^* \\ \gamma & \epsilon & 1-\alpha-\delta \end{pmatrix}, \tag{7.9}$$

where $\alpha, \delta \in \mathcal{R}$ and all the other numbers are complex. Any matrix of the form (7.9) can be written as a linear superposition of the matrices

$$\mathbf{v}_1 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{v}_2 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{v}_3 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7.10a}$$

$$\begin{array}{l} \mathbf{v}_4 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{v}_5 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix} \\ \mathbf{v}_6 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{v}_7 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \end{array}, \quad \mathbf{v}_8 = -\tfrac{\mathrm{i}}{2}\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{7.10b}$$

Since the $\mathbf{v}_i$ for $i \leq 3$ are just the Pauli matrices with an extra row and column of zeros tacked on, it is clear that these generators form a subalgebra of $\Lambda(SU(3))$ which is isomorphic to $\Lambda(SU(2))$. This subalgebra is associated with a subgroup of $SU(3)$ that's isomorphic to $SU(2)$. Similarly, $\{\mathbf{v}_6, \mathbf{v}_7, (\mathbf{v}_8 - \frac{1}{2}\mathbf{v}_3)\}$ and $\{\mathbf{v}_4, \mathbf{v}_5, (\mathbf{v}_8 + \frac{1}{2}\mathbf{v}_3)\}$ generate subalgebras isomorphic to $\Lambda(SU(2))$ that are associated with other $SU(2)$ subgroups of $SU(3)$.

8 Representations of Lie groups

Theorem 16

Every representation of a Lie group $\mathcal{G} \subset GL(n,)$ in terms of linear operators $T(a) : U \to U$ on a vector space U provides a representation of the group's Lie algebra $\Lambda(\mathcal{G})$.

Proof: The elements $v \in T_e(\mathcal{G})$ of $\Lambda(\mathcal{G})$ are effectively the same kind of $n \times n$ matrices as are the elements $a \in \mathcal{G}$. Hence with each $v \in T_e(\mathcal{G})$ we can associate an operator $T(\mathbf{v}) : U \to U$ by precisely the rule by which $a \in \mathcal{G}$ is represented as an operator. Moreover, by the requirement that composition of the operators T be compatible with group multiplication, we have

$$T([uv]) = T(\mathbf{u}\cdot\mathbf{v} - \mathbf{v}\cdot\mathbf{u}) = T(\mathbf{u})T(\mathbf{v}) - T(\mathbf{v})T(\mathbf{u}). \quad \square \tag{8.1}$$

The converse of the last theorem – that every representation of $\Lambda(\mathcal{G})$ gives rise to a representation of $\mathcal{G}$ – is almost but not quite true.[8] The operators $\mathbf{A} = \exp\big(T(\mathbf{v})\big)$ with $v \in \Lambda(\mathcal{G})$ form a group $\mathcal{G}'$ under multiplication, and in the nhd of e, $\mathcal{G}'$ is isomorphic with $\mathcal{G}$. The global structure of $\mathcal{G}$ and $\mathcal{G}'$ may differ, however.

8.1 Representations of $\Lambda(SU(2)) = \Lambda(\mathcal{R}(3))$

We now construct all finite-dimensional representations of $\Lambda(SU(2))$, which by theorem 15 is isomorphic to $\Lambda(\mathcal{R}(3))$. The main lines of the argument will be familiar from elementary quantum mechanics courses. We suppose that the representing operators $T(\mathbf{v})$ act on a space U whose elements it is convenient to take to be kets $|\psi\rangle$.

[8] Elliott & Dawber are hopelessly confused about this. On p. 128 they cite 3 'theorems'. The first is meaningless. The second is true but weak, the third is false.

We start by defining the Hermitian operators

$$J_i \equiv \mathrm{i}T(\mathbf{v}_i) \quad , \quad J^2 \equiv \sum_{i=1}^{3} J_i^2 \quad , \quad J_\pm = J_1 \pm \mathrm{i}J_2, \tag{8.2}$$

where $i = 1, 2, 3$ and the $\mathbf{v}_i$ are the infinitesimal generators of theorem 15. By the same theorem, the new operators satisfy the commutation relations

$$[J_i, J_k] = \mathrm{i}\epsilon_{ijk}J_k \quad , \quad [J_i, J^2] = 0 \quad , \quad [J_\pm, J^2] = 0 \quad , \quad [J_3, J_\pm] = \pm J_\pm. \tag{8.3}$$

It is also straightforward to show that

$$\begin{aligned} J^2 &= J_+J_- + J_3^2 - J_3 \\ &= J_-J_+ + J_3^2 + J_3. \end{aligned} \tag{8.4}$$

Since J_3 and J^2 commute, we may choose a basis for U consisting of kets $|\beta, m\rangle$ which are mutual eigenkets of J_3 and J^2:

$$J_3|\beta, m\rangle = m|\beta, m\rangle \quad , \quad J^2|\beta, m\rangle = \beta|\beta, m\rangle. \tag{8.5}$$

Since $[J_-, J^2] = 0$, $J_-|\beta, m\rangle$ is also an eigenket of J^2 with eigenvalue β. It is an eigenket of J_3, though with eigenvalue $m - 1$ rather than m:

$$\begin{aligned} J_3(J_-|\beta, m\rangle) &= J_-(J_3|\beta, m\rangle) + [J_3, J_-]|\beta, m\rangle \\ &= (m-1)(J_-|\beta, m\rangle). \end{aligned} \tag{8.6}$$

Hence

$$J_-|\beta, m\rangle = \alpha_-(\beta, m)|\beta, m-1\rangle, \tag{8.7}$$

where $\alpha_-(\beta, m)$ is a normalization constant chosen so that every $|\beta, m\rangle$ has unit norm. Multiplying both sides of (8.7) by is adjoint, we find with (8.4)

$$\begin{aligned} \alpha_-^2(\beta, m) &= \langle\beta, m|J_+J_-|\beta, m\rangle \\ &= \beta - m(m-1). \end{aligned} \tag{8.8}$$

Successively multiplying $|\beta, m\rangle$ by J_- and renormalizing, we can clearly generate a number of basis vectors. To get some others, we multiply by J_+. We first prove that this reverses the action of multiplying by J_-:

$$\begin{aligned} J_+J_-|\beta, m\rangle &= \alpha_-(\beta, m)J_+|\beta, m-1\rangle \\ &= (J^2 - J_3^2 + J_3)|\beta, m\rangle \\ &= (\beta - m(m-1))|\beta, m\rangle. \end{aligned} \tag{8.9}$$

Then it is straightforward to show that applying J_+ to $|\beta, m\rangle$ we get a multiple of $|\beta, m+1\rangle$:

$$\begin{aligned} J_3(J_+|\beta, m\rangle) &= J_+J_3|\beta, m\rangle + [J_3, J_+]|\beta, m\rangle \\ &= (m+1)(J_+|\beta, m\rangle). \end{aligned} \tag{8.10}$$

Hence

$$J_+|\beta, m\rangle = \alpha_+|\beta, m+1\rangle, \quad \text{where} \quad \alpha_+^2 = \beta - m(m+1). \tag{8.11}$$

We have established that starting from a single ket $|\beta, m\rangle$ we can generate a ladder of kets that differ in their eigenvalues with respect to J_3. If $\dim(U)$ is to be finite, the ladder must have a top and a bottom, which it will only if α_+ vanishes for m sufficiently large and α_- vanishes for m sufficiently small:

$$\begin{aligned} \alpha_+(\beta, \mathrm{m}_{\max}) = 0 \quad &\Rightarrow \quad \beta - m_{\max}(m_{\max}+1) = 0 \\ \alpha_-(\beta, \mathrm{m}_{\min}) = 0 \quad &\Rightarrow \quad \beta - m_{\min}(m_{\min}-1) = 0. \end{aligned} \tag{8.12}$$

Subtracting the right-hand eqns of (8.12) we easily find that $m_{\max} = -m_{\min} \equiv j$. But $m_{\max} - m_{\min}$ is an integer, so j is an integral number of half-integers. $\dim(U) = 2j+1$ is then an integer.

Now we are in a position to construct representations of $\Lambda(SU(2))$ of any dimension $2j+1$. We adopt the eigenkets $|j,m\rangle \equiv |\beta,m\rangle$ of J_3 as our basis kets. Then the matrix of J_3 is

$$J_3 = \mathrm{diagonal}(j, j-1, \ldots, -j). \tag{8.13}$$

The entries in the matrix of $J_\pm$ are the numbers

$$\begin{aligned} (J_\pm)_{mn} &= \langle \beta, m | J_\pm | \beta, n\rangle \\ &= \alpha_\pm(\beta, n)\delta_{m,n\pm1}. \end{aligned} \tag{8.14}$$

The representing matrix $T(\mathbf{v})$ of any $v \in \Lambda(SU(2))$ can now be constructed by taking the appropriate linear combination of $T(\mathbf{v}_3)$, $T(\mathbf{v}_+)$ and $T(\mathbf{v}_-)$.

The construction just described does not generate *every* representation of $\Lambda(SU(2))$. Indeed, we could have assigned some eigenvalue m_0 of J_3 two eigenkets, say $|j, m_0, +\rangle$ and $|j, m_0, -\rangle$, in hopes of constructing a representation of dimensionality greater than $2j+1$. But then from each of $|j, m_0, \pm\rangle$, J_+ and J_- would have generated an entire ladder of kets $|j, m, \pm\rangle$ and at the end of the day we could have ordered the eigenvectors so that the matrix $(J_3)_{mn}$ was block-diagonal, each of the two blocks being a copy of our former matrix $(J_3)_{mn}$, and similarly for the matrices $(J_\pm)_{mn}$.

This argument makes it plausible that our construction yields all the representations of $\Lambda(SU(2))$ associated with irreps of $SU(2)$, and this is in fact the case. Moreover, we can go on to construct the irreps themselves by exponentiating every real linear combination of the J_i $(i = 1,2,3)$. The $(2j+1)^{\mathrm{th}}$ representation $D^{(j)}$ obtained in this way is called the **spin-j** representation of $SU(2)$. Books such as Elliott & Dawber (§20.5) give the matrices that result from the exponentiation.

When j is an integer, $D^{(j)}$ is not a faithful representation of $SU(2)$. In particular, both e and the element associated with the matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ are mapped into the $(2j+1)$-dimensional identity matrix. More generally, with j an integer $D^{(j)}$ maps both $\exp(t\mathbf{v})$ and $\exp\big((t+\pi)\mathbf{v}\big)$ into the same matrix, where $\mathbf{v} = \sum_i n_i \mathbf{v}_i$ with $|\mathbf{n}| = 1$ is a linear combination of the generators of theorem 15.

However, with $j > 0$ an integer, $D^{(j)}$ does provide a faithful representation of $\mathcal{R}(3)$. In fact $D^{(1)}$ is equivalent to the representation provided by ordinary 3-vectors, being the representation provided by the Y_1^m studied in exercise 11. For non-integral j, $D^{(j)}$ does not provide a representation of $\mathcal{R}(3)$. To understand the problem, consider a typical 1-parameter subgroup of $\mathcal{R}(3)$, say $\{\exp(tv)\}$ with $v \in T_e\big(\mathcal{R}(3)\big)$. The elements of this subgroup represent rotations about some axis, through an angle that is proportional to t. As t is increased, we move first away from e, and then return to it when $t = t_0$, say. With the correct choice of generator $\mathbf{v}$, $\exp(tv)$ is represented in $D^{(j)}$ by the matrix $\exp(t\mathbf{v})$. For integer j we have $\exp(t_0\mathbf{v}) = \mathbf{I}$, but for non-integer j, $\exp(t_0\mathbf{v}) = -\mathbf{I}$; in this case e is being mapped into a matrix other than $\mathbf{I}$, and this violates the representation requirement that $\mathbf{T}(e)\cdot\mathbf{T}(a) = \mathbf{T}(ea) = \mathbf{T}(a)$.

It is often said that for non-integer j, $D^{(j)}$ provides a 'double-valued representation' of $\mathcal{R}(3)$, but to my mind this is an unnecessary and confusing circumlocution. $\mathcal{R}(3)$ and $SU(2)$ are different groups even though they share the same Lie algebra. $D^{(j)}$ always provides an irrep of $SU(2)$, though the irrep is faithful only for non-integer j. The unfaithful irreps of $SU(2)$ are faithful irreps of $\mathcal{R}(3)$ because $\mathcal{R}(3)$ is equivalent to $SU(2)$ with points identified in pairs.

Box 3: The Topology of $SU(2)$ and $\mathcal{R}(3)$

$SU(2)$ is diffeomorphic to the 3-sphere S^3. One can conceptualize S^3 as a solid sphere of radius 2 with its surface $r=2$ identified as a single point. If $e \leftrightarrow \mathbf{I}$ lies at $r=0$, then $r=2$ is the element $\overline{e} \leftrightarrow -\mathbf{I}$. $\mathcal{R}(3)$ is associated with the half of $SU(2)$ at $r \leq 1$ with diametrically opposed points on the 'equator' $r=1$ identified; points on $r=1$ correspond to rotations through π and the identification arises because a rotation by $-\pi$ is equivalent to a rotation by π about the same axis.

The sphere of $SU(2)$ is simply connected but $\mathcal{R}(3)$ is not: A path which passes from $r=0$ to the equator at some point p, and then emerges from the diametrically opposed point to return to e, cannot be continuously contracted to a point.

If by an extension of the identification of points on the equator $r=1$, we identify every point at $r>1$ with the diametrically opposed point at $r<1$, then $SU(2)$ provides a 'double covering' of $\mathcal{R}(3)$.

8.2 Characters of the $D^{(j)}$ irreps

Theorem 17

The classes of $\mathcal{R}(3)$ comprise rotations about a given angle ϕ.

Proof: Let $a, c \in \mathcal{R}(3)$ be given. Then the rotation axes $\mathbf{e}_a$, $\mathbf{e}_c$ of these rotations are uniquely defined by $a\mathbf{e}_a = \mathbf{e}_a$, $c\mathbf{e}_c = \mathbf{e}_c$. Let b be a rotation that transforms $\mathbf{e}_c$ into $\mathbf{e}_a$: $b\mathbf{e}_c = \mathbf{e}_a$. Then $(b^{-1}ab)\mathbf{e}_c = \mathbf{e}_c$, so $(b^{-1}ab)$ is a rotation about $\mathbf{e}_c$. By choosing coordinates such that $\hat{\mathbf{z}}$ lies along $\mathbf{e}_c$, and representing $(b^{-1}ab)$ and c in matrix form $\mathbf{B}^{-1}\cdot\mathbf{A}\cdot\mathbf{B}$ and $\mathbf{C}$, we see that these rotations are identical iff their rotation angles ϕ are equal. Moreover, $\mathrm{Tr}(\mathbf{A}) = 1 + 2\cos\phi_a = \mathrm{Tr}(\mathbf{B}^{-1}\cdot\mathbf{A}\cdot\mathbf{B})$. Hence their rotation angles are equal iff $\phi_a = \phi_c$. Thus $c = b^{-1}ab$ for some b, and thus a and c are in the same class, iff a and c are rotations through the same angle. □

Theorem 18

The character $\chi^{(j)}$ of a rotation through angle ϕ in the $D^{(j)}$ irrep is

$$\chi^{(j)}(\phi) = \frac{\sin(j+\frac{1}{2})\phi}{\sin\frac{1}{2}\phi}. \tag{8.15}$$

Proof: By the last theorem it suffices to consider rotations about the z-axis. The matrices of such rotations are of the form

$$\exp(-\phi \mathrm{i}\mathbf{J}_3) = \mathrm{diagonal}\big(\mathrm{e}^{-\mathrm{i}j\phi}, \mathrm{e}^{-\mathrm{i}(j-1)\phi}, \ldots, \mathrm{e}^{\mathrm{i}j\phi}\big). \tag{8.16}$$

The stated result now follows from the usual formula for the sum of a geometric progression.

8.3 Integration over a Lie group

Several of the proofs of §§1–4 involve taking averages of functions f on $\mathcal{G}$:

$$\langle f\rangle \equiv \frac{1}{g}\sum_{b\in\mathcal{G}} f(b). \tag{8.17}$$

When $\mathcal{G}$ has uncountably many elements, the meaning of such sums is not immediately evident. However, inspection of their occurrences will show that what is required is (i) that $\langle f\rangle = 1$ when $f \equiv 1$ and (ii) that for any function f on $\mathcal{G}$

$$\langle f\rangle = \frac{1}{g}\sum_{b\in\mathcal{G}} f(ab) \quad \forall\, a\in\mathcal{G}. \tag{8.18}$$

Hence when $\mathcal{G}$ is a Lie group we replace (8.17) with

$$\langle f\rangle = \int \mathrm{d}^d\mathbf{x}(b)\,\rho(b)f(b), \tag{8.19}$$

where $\mathbf{x}$ is a chart for $\mathcal{G}$ and $\rho(b)$ is 'volume element' that satisfies these conditions:

$$\begin{gathered}\int \mathrm{d}^d\mathbf{x}(b)\,\rho(b)=1\\ \int \mathrm{d}^d\mathbf{x}(b)\,\rho(b)f(ab)=\int \mathrm{d}^d\mathbf{x}(b)\,\rho(b)f(b)\quad \forall\ f \text{ on } \mathcal{G} \text{ and } a\in\mathcal{G}.\end{gathered} \tag{8.20}$$

The first condition can be satisfied only if the group forms a compact manifold.

We trivially have

$$\int \mathrm{d}^d\mathbf{x}(ab)\,\rho(ab)f(ab)=\int \mathrm{d}^d\mathbf{x}(b)\,\rho(b)f(b), \tag{8.21}$$

and by standard calculus

$$\mathrm{d}^d\mathbf{x}(ab)=\mathrm{d}^d\mathbf{x}(b)\frac{\partial\big(\mathbf{x}(ab)\big)}{\partial\big(\mathbf{x}(b)\big)}, \tag{8.22}$$

where the coordinates $\mathbf{x}(b)$ of b are being used as alternative coordinates for the point ab and the fraction containing ∂'s is the corresponding Jacobian. With (8.21), the second equation of (8.20) is thus equivalent to the requirement that

$$\frac{\partial\big(\mathbf{x}(ab)\big)}{\partial\big(\mathbf{x}(b)\big)}\rho(ab)=\rho(b). \tag{8.23}$$

Setting $b=e$ we obtain an expression for ρ evaluated at a general point in terms of its value at e:

$$\rho(a)=\rho(e)\Big/\frac{\partial\big(\mathbf{x}(ab)\big)}{\partial\big(\mathbf{x}(b)\big)}\bigg|_{b=e}. \tag{8.24}$$

The number $\rho(e)$ is obviously determined by the first of conditions (8.20).

Let's see how this works out in the specific case of $\mathcal{R}(3)$. Equation (5.2) defines a suitable chart. If we use polar coordinates to describe the vector $\mathbf{a}$, then it is reasonably obvious that $\mathrm{d}^3\mathbf{x}(a)\,\rho(a)=\rho(|\mathbf{a}|)\mathrm{d}|\mathbf{a}|\mathrm{d}\Omega$, where $\mathrm{d}\Omega$ is an element of solid angle; that is, ρ depends only on the angle of the rotation a, not on the direction of its axis. This observation enables us to take $\mathbf{a}$ to be a rotation about the z-axis. Then we have to calculate the vector $\mathbf{c}$ which characterizes the product of

$$\mathbf{a}=\begin{pmatrix}\cos a & -\sin a & 0\\ \sin a & \cos a & 0\\ 0 & 0 & 1\end{pmatrix} \tag{8.25}$$

with

$$\mathbf{b}=\mathbf{I}+\mathbf{v}_i\delta b_i, \tag{8.26}$$

where $i=1,2,3$ and the $\mathbf{v}_i$ are the infinitesimal generators of $\mathcal{R}(3)$. Finally we subtract off $\mathbf{a}$, divide through by each b_i in turn and calculate the determinant of the resulting matrix – see §A4.3 of Elliott & Dawber for details. The result is

$$\rho(a)=\frac{\sin^2\frac12\phi}{2\pi^2}. \tag{8.27}$$

We are mostly interested in functions $f(a)$ which are also only functions of $|\mathbf{a}|$ because they are functions only of class, and in this case $\mathrm{d}\Omega$ can be immediately integrated out.

Exercise (16):

Verify that

$$\frac{1}{\pi}\int_0^\pi \mathrm{d}\phi\,\sin^2(\tfrac12\phi)\chi^{(j_1)}(\phi)\chi^{(j_2)}(\phi)=\delta_{j_1j_2}. \tag{8.28}$$

Explain why the integral's upper limit is not 2π.

8.4 Casimir operators

In §8.1 the operator J^2 played an important rôle although it is not a generator of $SU(2)$. Its job is simply to tell us which irrep of $SU(2)$ a given ket belongs to: $J^2|\psi\rangle = j(j+1)|\psi\rangle$ if $|\psi\rangle$ is a linear combination of basis kets of the $D^{(j)}$ irrep of $SU(2)$. Operators with this property of identifying the irrep any ket belongs to, are called **Casimir operators**. $SU(2)$ has only one Casimir operator, but more complex groups have more than one; for example $SU(3)$ has two.

Exercise (17):

The key property of a Casimir operator is that it commutes with every representing operator $T(a)$. Prove that the operator

$$C \equiv \sum_{ij} g^{ij} T(\mathbf{v}_i) T(\mathbf{v}_j) \tag{8.29}$$

has this property, where the $\mathbf{v}_i$ are $\mathcal{G}$'s infinitesimal generators, $[\mathbf{v}_i, \mathbf{v}_j] = \sum_k c_{ij}^k \mathbf{v}_k$, and g^{ij} is the inverse of the 'metric' tensor $g_{ij} \equiv \sum_{kl} c_{il}^k c_{jk}^l$. [Hint: first prove, in an obvious notation, that $[C, T_p] = g^{ij} c_{jp}^k (T_i T_k + T_k T_i)$, then use (6.15) to show that $c_{ijk} \equiv g_{ip} c_{jk}^p$ is totally antisymmetric.]

Exercise (18):

Show that for $SU(2)$ the operator C of the previous exercise reduces to J^2.

9 Representations of $SU(3)$

Equations (7.10) give infinitesimal generators $\mathbf{v}_i$, $i = 1 - 8$, of $SU(3)$. Let T map elements of $SU(3)$ into operators on a space U of kets $|\psi\rangle$. Then T maps the $\mathbf{v}_i$ into operators on U that have the same commutation relations as have the $\mathbf{v}_i$. We define

$$\begin{aligned} &Y \equiv \mathrm{i}T(\tfrac{4}{3}\mathbf{v}_8) \\ I_3 &\equiv \mathrm{i}T(\mathbf{v}_3) \quad , \quad I_\pm \equiv \mathrm{i}T(\mathbf{v}_1 \pm \mathrm{i}\mathbf{v}_2) \\ U_3 &\equiv \tfrac{3}{4}Y - \tfrac{1}{2}I_3 \quad , \quad U_\pm \equiv \mathrm{i}T(\mathbf{v}_6 \pm \mathrm{i}\mathbf{v}_7) \\ V_3 &\equiv -\tfrac{3}{4}Y - \tfrac{1}{2}I_3 \quad , \quad V_\pm \equiv \mathrm{i}T(\mathbf{v}_4 \pm \mathrm{i}\mathbf{v}_5)\ . \end{aligned} \tag{9.1}$$

It is straightforward to check that the I, U, and V sets of operators all enjoy the commutation relations associated with $SU(2)$ generators:

$$[I_3, I_\pm] = \pm I_\pm \quad , \quad [U_3, U_\pm] = \pm U_\pm \quad , \quad [V_3, V_\pm] = \pm V_\pm\,. \tag{9.2}$$

Also, Y commutes with all three I_i operators, and, since all the other subscript-3 operators are linear combinations of Y and I_3, it follows that all the operators in the set $\{Y, I_3, U_3, V_3\}$ commute with one another. For the rest, we have

$$\begin{aligned} [I_3, U_\pm] &= \mp\tfrac{1}{2}U_\pm \quad , \quad [Y, U_\pm] = \pm U_\pm \\ [I_3, V_\pm] &= \mp\tfrac{1}{2}V_\pm \quad , \quad [Y, V_\pm] = \mp V_\pm \end{aligned} \tag{9.3}$$

From the $SU(2)$-like commutation relations of the I, U and V families it follows that the possible eigenvalues of I_3, U_3 and V_3 are $\ldots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \ldots$.

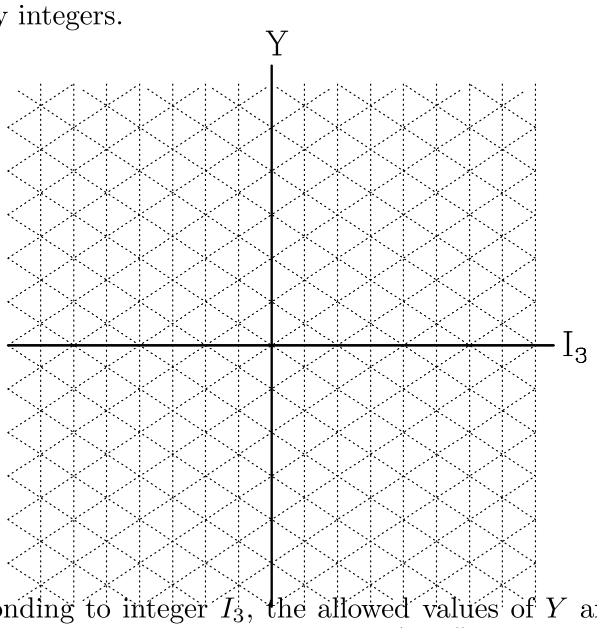
Exercise (19):

Show that $[I_+, V_-] = 0$.

We can assume that the basis kets of our irrep are mutual eigenkets of all the operators in the set $\{Y, I_3, U_3, V_3\}$. We take the independent operators of the set to be I_3 and Y. Our basis kets can then be arranged as points in a plane with (x, y) coordinates given by their eigenvalues w.r.t. (I_3, Y). In this plane kets lie on the intersections of the three sets of lines

$$I_3 = \frac{l}{2}, \quad \begin{array}{l} U_3 = \frac{3}{4}Y - \frac{1}{2}I_3 = \frac{m}{2}, \\ V_3 = -\frac{3}{4}Y - \frac{1}{2}I_3 = \frac{n}{2}, \end{array} \tag{9.4}$$

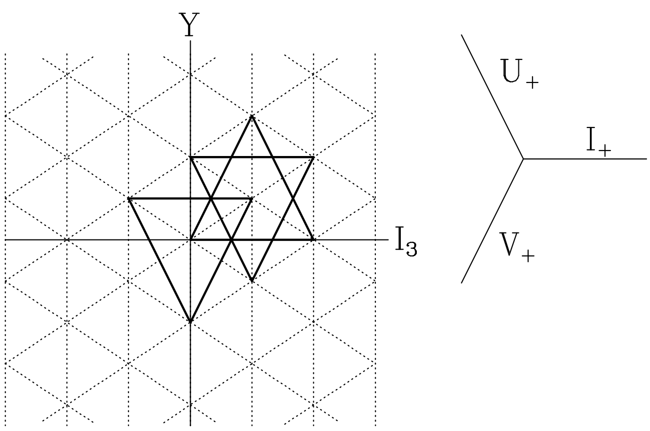
where l, m, n are arbitrary integers.



On vertical lines corresponding to integer I_3, the allowed values of Y are $\ldots, -\frac{2}{3}, 0, \frac{2}{3}, \frac{4}{3}, \ldots$. When I_3 is half-integer, the allowed values of Y are $\ldots, \frac{1}{3}, 1, \frac{5}{3}, \ldots$.

An irrep is generated by starting at any allowed point and moving from there by applying the operators of the representation of the Lie Algebra. Since any such operator may be expressed as a linear combination of the eight operators Y, I_3, $I_\pm$, $U_\pm$, $V_\pm$, we have only to consider how we move under the actions of the ladder operators of this set. I_+ moves one horizontally to the right, U_+ moves one up and to the left, and V_+ moves one down and to the left. Basic triangles are generated by successively applying I_+, $U_\pm$ and $V_\pm$, either in the order $I_+U_+V_+$ or in the order $I_+V_+U_+$. The points associated with a given irrep all lie on the vertices of such triangles when they are used to tessellate the plane. Three tessellations are possible, corresponding to the 3 mutually intersecting triangles that have vertices at

$$\begin{array}{rl} \text{either} & (0,0),\ (\frac{1}{2},1),\ (1,0) \\ \text{or} & (0,\frac{2}{3}),\ (1,\frac{2}{3}),\ (\frac{1}{2},-\frac{1}{3}) \\ \text{or} & (-\frac{1}{2},\frac{1}{3}),\ (\frac{1}{2},\frac{1}{3}),\ (0,-\frac{2}{3})\ . \end{array}$$



At some point as one moves away from the origin by applying the ladder operators, one's progress will be halted by a ladder operator annihilating the current state, e.g., $I_+|\psi\rangle = 0$; one has reached the boundary of the irrep. It is conventional to classify an irrep $D^{(\lambda\mu)}$ by the values $\lambda \equiv 2I_3$, $\mu \equiv 2U_3$ associated with the top-right corner of the irrep; the top-right corner being defined to be the ket $|\mathrm{gw}\rangle$ with the largest value of Y, and for that Y-value, the largest value of I_3. $|\mathrm{gw}\rangle$ is called the **ket of greatest weight**; by definition one has $I_+|\mathrm{gw}\rangle = \mathrm{V}_-|\mathrm{gw}\rangle = \mathrm{U}_+|\mathrm{gw}\rangle = 0$.

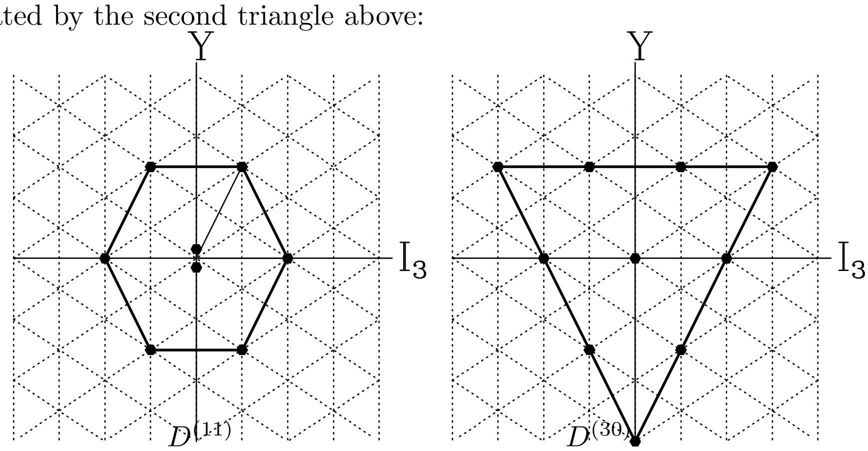
If $\mu = 0$, $|\mathrm{gw}\rangle$ has no U-spin, so you can't move NW–SE from $|\mathrm{gw}\rangle$ by applying $U_\pm$. You can only move W or SW by applying either I_- or V_+ and the irrep forms a triangle. If $\mu > 0$, you can move either SE, W or SW as you please, and the irrep forms a hexagon.

If $\mu > 0$, you can reach the point that lies SW of $|\mathrm{gw}\rangle$ either directly by applying V_+, or you can reach it by applying the product I_-U_-. Two different kets are generated by the two different routes.

Exercise (20):

Use the result of Exercise 19 to show that $|\phi\rangle \equiv V_-|\mathrm{gw}\rangle$ is such that $I_+|\phi\rangle = 0$. Hence deduce that $|\phi\rangle$ and $I_-U_-|\mathrm{gw}\rangle$ differ in their eigen-values w.r.t. $I^2 \equiv I_1^2 + I_2^2 + I_3^2$. Show further that if $V_+^2|\mathrm{gw}\rangle$ and $U_-^2|\mathrm{gw}\rangle$ are non-zero, *three* kets correspond to the location on the grid of $V_+^2|\mathrm{gw}\rangle$. Generalize this result.

The fundamental, 3-dimensional, irrep is $D^{(1,0)}$. The two other physically important irreps are both generated by the second triangle above:



The hexagonal irrep $D^{(11)}$ has two basis kets at the origin and is thus 8-dimensional, while the triangular irrep $D^{(30)}$ has one basis ket at the origin, and is thus 10-dimensional.

In general the irreps have dimension $1, 3, 6, 8, 10, 15, 24, 27, \ldots$ and it is an important problem to explain why only the $1, 8$ and 10-dimensional irreps feature in Nature.

As in the case of $SU(2)$, the complex-conjugates of matrices of an irrep of $SU(3)$ provide a different irrep of $SU(3)$. Now the 3-dimensional matrices of the fundamental irrep are of the form $\mathbf{A} = \exp(t\mathbf{v})$, where $\mathbf{v}$ is a real linear combination of the $\mathbf{v}_i$ of equations (7.10). So complex-conjugating the $\mathbf{A}$ is equivalent to complex-conjugating the $\mathbf{v}_i$. From eqs (7.10) we see that this changes the sign of $\mathbf{v}_3$ and $\mathbf{v}_8$, and thus, by the linearity of the representing function $T()$, it changes the sign of I_3 and Y. Consequently, the figures in the (I_3, Y) plane formed by the basis kets of complex-conjugate irreps, can be obtained from the figures of the ordinary irreps by inverting everything through the origin. Since all figures are symmetric on reflection in the Y-axis, reflection through the origin is equivalent to turning the figure up-side-down.

9.1 Quarks and $SU(3)$

A quark has three 'flavour' states: it can be an up quark, a down quark or a strange quark. Let $|u\rangle$, $|d\rangle$ and $|s\rangle$ be the kets describing these three states. A general quark state is a linear superposition of these flavour eigenstates:

$$|\psi\rangle = \sum_{k=u,d,s} \psi_k |k\rangle. \tag{9.5}$$

ψ_u gives the amplitude for the quark to be an up-quark, ψ_d the amplitude for it to be a down-quark, and ψ_s the amplitude for it to be a strange-quark.

Elements $a \in SU(3)$ are supposed to transform flavours into one another:

$$|u'\rangle \equiv T(a)|u\rangle = \sum_{k=u,d,s} T_{ku}(a)|k\rangle. \tag{9.6}$$

Consequently we have

$$\begin{pmatrix} \psi'_u \\ \psi'_d \\ \psi'_s \end{pmatrix} = \mathbf{T}(a) \cdot \begin{pmatrix} \psi_u \\ \psi_d \\ \psi_s \end{pmatrix}. \tag{9.7}$$

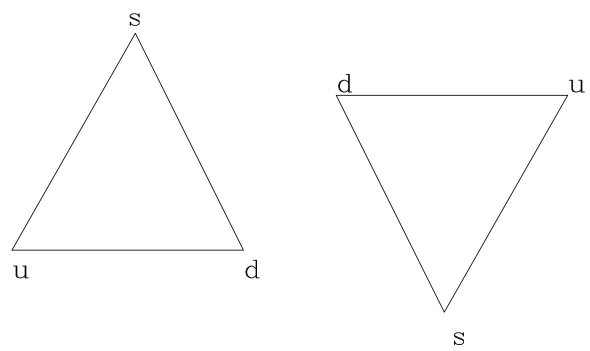
The quark's Hamiltonian is not actually invariant under such transformations, but it is sufficiently nearly so for these transformations to be of physical interest.

Clearly quarks provide a 3-dimensional rep of $SU(3)$, and it is natural to assume that this is an irrep. We know of two 3-dimensional irreps of $SU(3)$: $D^{(10)}$ and its conjugate $D^{(01)}$. We assume that the quarks furnish the $D^{(10)}$ irrep and the anti-quarks the $D^{(01)}$ irrep. In terms of components, we are postulating that (9.6) should be paired with

$$T(a)|\overline{u}\rangle = \sum_{k=u,d,s} T^*_{ku}|\overline{k}\rangle, \tag{9.8}$$

where a bar indicates an anti-quark.

In the (Y, I_3) plane we have



We identify I_3 with isospin and Y with hypercharge, so that in its u-state a quark has quantum numbers $Y = \frac{1}{3}, I_3 = \frac{1}{2}, I = \frac{1}{2}$, while in its s-state it has $Y = -\frac{2}{3}, I_3 = I = 0$. Whatever its state it has $B = 3$ and its charge is $Q = I_3 + \frac{1}{2}Y$. It is straightforward to see that

$$\begin{array}{lll} I_-|u\rangle = |d\rangle, & I_+|d\rangle = |u\rangle, & I_\pm|s\rangle = 0, \\ V_+|u\rangle = |s\rangle, & V_-|s\rangle = |u\rangle, & V_\pm|d\rangle = 0, \\ U_-|d\rangle = |s\rangle, & U_+|s\rangle = |d\rangle, & U_\pm|u\rangle = 0, \end{array}$$

and

$$\begin{array}{lll} I_+|\overline{u}\rangle = |\overline{d}\rangle, & I_-|\overline{d}\rangle = |\overline{u}\rangle, & I_\pm|\overline{s}\rangle = 0, \\ V_-|\overline{u}\rangle = |\overline{s}\rangle, & V_+|\overline{s}\rangle = |\overline{u}\rangle, & V_\pm|\overline{d}\rangle = 0, \\ U_+|\overline{d}\rangle = |\overline{s}\rangle, & U_-|\overline{s}\rangle = |\overline{d}\rangle, & U_\pm|\overline{u}\rangle = 0. \end{array}$$

Mesons are supposed to be bound states of a quark and an anti-quark. Consequently the general meson wavefunction is has 9 components, and under $SU(3)$ these transform according to the direct-product representation (cf. (9.7))

$$\pi'_{ij} = \sum_{k,l=u,d,s} T_{ik} T^*_{jl} \pi_{kl}. \tag{9.9}$$

Since these transformations change the likely nature of the constituent quarks, they correspond to mapping from states in which the meson is likely to be, say, a K^+ particle, to one in which it is, say, an η.

Experiences teaches that kets corresponding to definite elementary particles should provide the basis kets of irreps. That is, if we start with a ket associated with the meson definitely being a K^+ and transform this state with every available $SU(3)$ operator, we expect to explore an *irreducible* invariant subspace. This implies that it is possible to group mesons into multiplets such that the mesons in each multiplet constitute the basis kets of an irrep of $SU(3)$.

These multiplets should place each elementary particle at a well defined point in the (I_3, Y) plane in which we display irreps of $SU(3)$. Since the particles in a given irrep should all have the same baryon number B, and strangeness is given by $S = Y - B$, objects containing strange quarks ($S = -1$) will lie below objects with zero strangeness. Since charge $Q = I_3 + \frac{1}{2}Y$, charged objects will lie upwards and to the right.

We know that direct product representations are rarely irreps. So we expect the mesons to constitute a smaller multiplet than a 9-tuple. To discover what we do expect, we have to decompose the direct product rep into irreps.

The same problem of the decomposition of a direct product rep arises in connection with the baryons, which are supposed to be bound states of 3 quarks; consequently, the baryonic wavefunction transforms under $SU(3)$ as a third-rank tensor:

$$\psi'_{ijk} = \sum_{l,m,n=u,d,s} T_{il} T_{jm} T_{kn} \psi_{lmn}. \tag{9.10}$$

9.2 Young's tableaux

The key to decomposing into irreps the direct product reps provided by meson and baryon fields is symmetry under exchange of indices. The situation is closely analogous to the familiar case of ordinary tensors; the 9-dimensional rep of $\mathcal{R}(3)$ provided by the second-rank tensors M_{ij} decomposes into the 5-dimensional $l = 2$ irrep provided by symmetric, traceless tensors, the 3-dimensional irrep of antisymmetric tensors (axial vectors) and the scalar irrep associated with the tensor's trace. Thus each symmetry class of tensors provides an irrep of $\mathcal{R}(3)$. The same applies in the case of $SU(3)$. Four exercises lay the foundation of the theory:

Exercise (21):
Show that if ψ_{ijk} is symmetric (antisymmetric) under the exchange $i \leftrightarrow j$, then so is the object ψ' defined by (9.10).

Exercise (22):
Show that the Levi-Civita symbol ϵ_{ijk} is invariant under $SU(3)$. That is, show that if ϵ_{ijk} is defined in the usual way, then the quantity ϵ' derived from it according to (9.10) has the same pattern of zeros and ones. [Hint: for any 3×3 matrix $\mathbf{M}$, $|\mathbf{M}|\epsilon_{pqr} = \epsilon_{ijk} M_{pi} M_{qj} M_{rk}$.]

Exercise (23):
Given that the antisymmetric matrix x_{jk} transforms according to $D^{(10)} \times D^{(10)}$, show that $x_i \equiv \epsilon_{ijk} x_{jk}$ transforms according to the conjugate irrep $D^{(10)*}$.

Exercise (24):
Given that the symmetric matrix x_{jk} transforms according to $D^{(10)} \times D^{(10)}$, show that $\mathrm{Tr}(x_{ij})$ is not invariant under $SU(3)$. Show that $\mathrm{Tr}(x_{ij})$ would be invariant if x_{jk} transformed according to $D^{(10)} \times D^{(10)*}$.

Exercise 21 demonstrates that irreps must be confined to a given symmetry class. In fact they coincide with the symmetry classes as we can show by comparing the latter's dimensions with those of the $SU(3)$ irreps. Exercise 23 hints how we should handle reps that involve the conjugate irreps: when dealing with symmetry classes it is helpful for all indices to transform in the same way. So we replace an index which transforms according to $D^{(10)*}$ by an antisymmetrized pair that transform according to $D^{(10)}$.

Tensors that transform according to $D^{(10)} \times D^{(10)}$ break into two classes, the symmetric ones and the antisymmetric ones. The latter provide a 3-dimensional irrep ($D^{(10)*}$), and the former provide a 6-dimensional irrep, $D^{(20)}$. This state of affairs is represented by one of Young's tableaux:

$$\Box \times \Box = \Box\Box + \begin{array}{c}\Box\\ \Box\end{array} \tag{9.11}$$

Each block stands for an index. Indices in horizontally adjacent blocks are symmetrized, those in vertically stacked blocks are antisymmetrized.

Now consider tensors that transform according to $D^{(10)*} \times D^{(10)}$. We now start with three indices, two of which are already antisymmetrized:

$$\begin{array}{c}\Box\\ \Box\end{array} \times \Box = \begin{array}{c}\Box\\ \Box\\ \Box\end{array} + \begin{array}{l}\Box\Box\\ \Box\end{array} \tag{9.12}$$

This tableau says that the extra index can either be antisymmetrized with the other two (the column) or be symmetrized with one of the other two (the ups-side-down L). The column provides a 1-dimensional irrep (there is only one free component of a totally antisymmetric 3rd rank tensor) and the L provides the 8-dimensional irrep $D^{(11)}$.

Exercise (25):

Justify the last claim as follows. Let ψ_{ijk} be anti-symmetric in the first pair of indices and symmetric in the second pair. Show that it is symmetric in the third pair. Show that all 27 components can be written down once the values of 9 have been determined. Finally show that these 9 values satisfy a linear constraint. [Hint: the latter involves ϵ_{ijk}.]

Finally we consider tensors that transform as $D^{(10)} \times D^{(10)} \times D^{(10)}$.

$$\square \times \square \times \square = \begin{array}{c}\square\\ \square\\ \square\end{array} + \begin{array}{ll}\square & \square\\ \square & \end{array} + \square\square\square \tag{9.13}$$

The first two irreps on the right we recognize, and the third is the 10-dimensional irrep $D^{(30)}$ provided by totally symmetric 3rd rank tensors.

Note:

There is an independent component of such a tensor for every number of occurrences of 1,2,3 in sets of three digits. We can choose 1 0,1,2 or 3 times. If we choose 1 0 times, we can choose 2 0,1,2 or 3 times and in each case the number of 3s we need to make up our set of three digits is determined. Similarly, if 1 is chosen once, we can choose 2 0,1 or 2 times. In this way we arrive at a total of ten options.

The spin-$\frac{1}{2}$ baryons fill out the 8-dimensional irrep $D^{(11)}$ and the spin-$\frac{3}{2}$ baryons fill out the 10-dimensional irrep $D^{(30)}$.

9.3 Colour

Why is there no 6-tuple of mesons corresponding to $D^{(20)}$ when the more complex $D^{(11)}$ irrep is realized in Nature? QCD answers this question by postulating that a second $SU(3)$ group acts on quark fields, and that the action $\int \mathrm{d}^4\mathbf{x}\,\mathcal{L}$ of these fields is such that only the trivial 1-dimensional irrep of this group can be observed. This postulate restricts the combinations of quarks inside hadrons to those which give rise to direct product reps of $SU(3)_{\rm colour}$ which contain the trivial irrep represented by

$$\begin{array}{c}\square\\ \square\\ \square\end{array}.$$

Since the combination rules of $SU(3)_{\rm flavour}$ are the same as those for $SU(3)_{\rm colour}$, it follows that the only multiplets expected are those supporting irreps of $SU(3)_{\rm flavour}$ that occur in conjunction with the trivial irrep. Since the reduction of $D^{(10)} \times D^{(10)}$ does not contain the trivial irrep, qq mesons do not occur, and therefore the $D^{(20)}$ irrep of $SU(3)_{\rm flavour}$, which they would support, does not occur either.

The three basic colours are red, green and blue and the $SU(3)_{\rm colour}$ singlet ket is totally anti-symmetric under an exchange of colours between the quarks, being

$$|C\rangle \equiv \frac{1}{\sqrt{6}} \sum_{ijk=rgb} \epsilon_{ijk}|ijk\rangle, \tag{9.14}$$

where $|ijk\rangle \equiv |i\rangle|j\rangle|k\rangle$ is the ket describing the state in which the first quark has colour i, the second colour j and the third colour k. Since quarks are fermions, it follows that whatever factors multiply the colour component of the wavefunction must together be invariant on exchange of quark labels.

Three factors are expected: the flavour ket, the spin ket and the spatial wavefunction. For the low-lying states of interest, the quarks have zero orbital angular momentum, so the last of these is totally symmetric. Hence the product of the flavour and spin kets should be totally symmetric.

A spin-$\frac{3}{2}$ object can be constructed from 3 spin-$\frac{1}{2}$ objects in only one way, that is by forming the totally symmetric ket

$$|\tfrac{3}{2}\tfrac{3}{2}\rangle = |\uparrow\rangle|\uparrow\rangle|\uparrow\rangle \tag{9.15}$$

and then using the usual ladder operators to construct $|\frac{3}{2}\frac{1}{2}\rangle$ etc. It is easy to show that all these kets are totally symmetric. Hence the flavour kets of particles of the 10-tuple of spin-$\frac{3}{2}$ baryons must be totally symmetric too. In fact, the ket of the object at the 10-tuple's top right-hand corner is

$$|\Delta^{++}\rangle = |uuu\rangle. \tag{9.16}$$

Applying I_- we get another totally symmetric ket:

$$|\Delta^{+}\rangle = \frac{1}{\sqrt{3}}I_-|\Delta^{++}\rangle = \frac{1}{\sqrt{3}}\big(|duu\rangle + |udu\rangle + |uud\rangle\big). \tag{9.17}$$

The flavour kets of the entire 10-tuple can be obtained by further applications of I_-, U_-, etc, and they are all totally symmetric.

Now what about the proton's flavour ket? The proton lies at the top-right corner of an octet, and we know that such irreps are provided by objects that are antisymmetric in one pair of indices and symmetric in the remaining two. So we look for numbers α, β, γ such that

$$|\psi\rangle \equiv \alpha|duu\rangle + \beta|udu\rangle + \gamma|uud\rangle \tag{9.18}$$

is antisymmetric under exchange of the first two flavours and symmetric under the exchange of the second two. That is,

$$\begin{aligned} -|\psi\rangle &= \alpha|udu\rangle + \beta|duu\rangle + \gamma|uud\rangle \\ +|\psi\rangle &= \alpha|duu\rangle + \beta|uud\rangle + \gamma|udu\rangle. \end{aligned} \tag{9.19}$$

The first equation implies $\alpha = -\beta$ and $\gamma = 0$, while the second implies $\beta = \gamma$. So the two together require $\alpha = \beta = \gamma = 0$. In other words, we cannot construct a ket with the required symmetry from the given pieces; we must consider linear combinations which include strange quarks.

These combinations are not physically interesting, however, because the combination of the spin and orbital kets with which they have to be combined to achieve overall symmetry of the non-colour part of the ket, are associated with large energies. The reason is that the product $|\text{spin}\rangle|\text{orbit}\rangle$ would need to be antisymmetric under exchange of one pair of quarks and symmetric under exchange of the other two pairs. Now it is easy to check that no linear combination of the two $S = \frac{1}{2}\hbar$, $m = \frac{1}{2}\hbar$ kets[9]

$$\begin{aligned} |\tfrac{1}{2},\tfrac{1}{2};0\rangle &\equiv \tfrac{1}{\sqrt{2}}\big(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle\big)|\uparrow\rangle \\ |\tfrac{1}{2},\tfrac{1}{2};1\rangle &\equiv \tfrac{1}{\sqrt{6}}\big(2|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle\big) \end{aligned} \tag{9.20}$$

has the required symmetry. So if we are to multiply kets from an $SU(3)$ octet by a combination $|\text{spin}\rangle|\text{orb}\rangle$ that ensures overall symmetry, we are going to have to employ an orbital ket with non-trivial symmetry. Such orbital kets correspond to states in which the quarks have non-zero angular momentum about the overall centre of mass, and thus to states in which our composite particle has appreciable internal kinetic energy. These states are presumably unstable.

An $S = \frac{1}{2}\hbar$ ket with zero orbital angular momentum that is totally symmetric under the exchange of quarks, *can* be constructed when we mix spin and flavour in a non-trivial way; then our pieces are

$$\left.\begin{matrix} |duu\rangle \\ |udu\rangle \\ |uud\rangle \end{matrix}\right\} \times \left\{\begin{matrix} |\downarrow\uparrow\uparrow\rangle \\ |\uparrow\downarrow\uparrow\rangle \\ |\uparrow\uparrow\downarrow\rangle \end{matrix}\right. . \tag{9.21}$$

[9] The first of the kets (9.20) is obtained by adding a further particle to the spin-zero combination of the other two, while the second state is obtained by adding a further particle to the spin-one combinbation of the other two in such a way as to achieve overall spin-half.

We first construct kets $|u\uparrow u\uparrow d\downarrow\rangle$ and $|u\uparrow u\downarrow d\uparrow\rangle$ which are (a) totally symmetric in the quarks and (b) such that the quark in the d-state is always spin-down, and always spin-up, respectively. It is easy to see that

$$\begin{aligned}|u\uparrow u\uparrow d\downarrow\rangle &= \sqrt{\tfrac{1}{3}}\Big[\big(|uud\rangle|\uparrow\uparrow\downarrow\rangle + |udu\rangle|\uparrow\downarrow\uparrow\rangle + |duu\rangle|\downarrow\uparrow\uparrow\rangle\big)\Big],\\ |u\uparrow u\downarrow d\uparrow\rangle &= \sqrt{\tfrac{1}{6}}\Big[\big(|udu\rangle + |duu\rangle\big)|\uparrow\uparrow\downarrow\rangle + \big(|udu\rangle + |uud\rangle\big)|\downarrow\uparrow\uparrow\rangle\\ &\quad + \big(|uud\rangle + |duu\rangle\big)|\uparrow\downarrow\uparrow\rangle\Big].\end{aligned} \tag{9.22}$$

Now we argue that $|\psi_p\rangle = \cos\theta|u\uparrow u\uparrow d\downarrow\rangle + \sin\theta|u\uparrow u\downarrow d\uparrow\rangle$ for suitable θ and determine θ from the condition that $I_3|\psi_p\rangle = 0$. This gives

$$|\psi_p\rangle = \sqrt{\tfrac{2}{3}}|u\uparrow u\uparrow d\downarrow\rangle - \sqrt{\tfrac{1}{3}}|u\uparrow u\downarrow d\uparrow\rangle. \tag{9.23}$$

Notice that this vanishes if we set all the spin-kets to unity.

Exercise (26):
Determine the ket $|\psi_n\rangle$ of the neutron.

10 The Lorentz Groups

Lorentz transformations are the linear transformations $\mathbf{x} \to \mathbf{x}' = \boldsymbol{\Lambda}\cdot\mathbf{x}$ which preserve the inner product

$$\begin{aligned}\mathbf{x}\cdot\mathbf{y} &\equiv x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3\\ &= \mathbf{x}\cdot\boldsymbol{\eta}\cdot\mathbf{y} \equiv x^\mu\eta_{\mu\nu}y^\nu\end{aligned} \tag{10.1}$$

in the sense that $\mathbf{x}'\cdot\mathbf{y}' = \mathbf{x}\cdot\mathbf{y}\ \forall\mathbf{x},\mathbf{y}$. These transformations clearly form a group. It has four distinct components:

(i) The **proper orthochronous transformations** are those for which $\det(\boldsymbol{\Lambda}) = 1$ and ${\Lambda^0}_0 \geq 1$. These transformations form the component $\mathcal{L}$ within which the identity e sits.

(ii) The transformations with $\det(\boldsymbol{\Lambda}) = -1$ and ${\Lambda^0}_0 \geq 1$ all involve space inversion and form a second connected component. Because this component does not include e it does not by itself form a group, but can be added to $\mathcal{L}$ to form a group $\mathcal{L}_s$.

(iii) The transformations with $\det(\boldsymbol{\Lambda}) = -1$ and ${\Lambda^0}_0 \leq 1$ all involve time inversion and form a third connected component. This component does not include e either, but can be added to $\mathcal{L}$ to form a group $\mathcal{L}_t$.

(iv) When elements from the last two components are multiplied together, a fourth component is generated that comprises transformations with $\det(\boldsymbol{\Lambda}) = 1$ and ${\Lambda^0}_0 \leq 1$. All four components taken together form the **full Lorentz group** $\mathcal{L}_{st}$.

Exercise (27):
Show that if $|\mathbf{x}'| = |\mathbf{x}|$ for every 4-vector $\mathbf{x}$, where $\mathbf{x}' \equiv \boldsymbol{\Lambda}\cdot\mathbf{x}$ etc., then $\mathbf{x}'\cdot\mathbf{y}' = \mathbf{x}\cdot\mathbf{y}$ for every $\mathbf{x}$, $\mathbf{y}$,

The rotation group $\mathcal{R}(3)$ is clearly a subgroup of all four Lorentz groups. Each rotation is characterized by the 3-vector $\mathbf{a}$ that specifies its axis and angle. A boost is likewise characterized by a 3-vector $\mathbf{b}$ that specifies its direction $\hat{\mathbf{b}}$ and velocity $|\mathbf{b}|$. The matrix of a boost parallel to the x^1-axis is

$$\mathbf{Q}(b,0,0) = \begin{pmatrix}\cosh b & -\sinh b & 0 & 0\\ -\sinh b & \cosh b & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \tag{10.2}$$

The matrix of a general boost can be deduced from

$$\begin{aligned}\mathbf{Q}(\mathbf{b})\cdot\mathbf{x} = \mathbf{x} - (x^0, x^b\hat{\mathbf{b}}) + \big(x^0\cosh b - x^b\sinh b, (-x^0\sinh b + x^b\cosh b)\hat{\mathbf{b}}\big),\\ \text{where}\quad x^b \equiv \mathbf{x}\cdot\hat{\mathbf{b}}.\end{aligned} \tag{10.3}$$

Any $\boldsymbol{\Lambda}$ may be written as a product $\mathbf{Q}(\mathbf{b})\cdot\mathbf{R}(\mathbf{a})$ of a boost and a rotation: Define $\hat{\mathbf{e}}_0 \equiv (1,0,0,0)$ and let $\hat{\mathbf{e}} \equiv \boldsymbol{\Lambda}\cdot\hat{\mathbf{e}}_0$. Then (10.3) yields $\mathbf{Q}(\mathbf{b})\cdot\hat{\mathbf{e}}_0 = (\cosh b, -\sinh b\,\hat{\mathbf{b}})$, so $\exists$ unique $\mathbf{Q}(\mathbf{b})$ such that $\hat{\mathbf{e}} = \mathbf{Q}(\mathbf{b})\cdot\hat{\mathbf{e}}_0$. Now $\hat{\mathbf{e}}_0 = \mathbf{Q}^{-1}\cdot\boldsymbol{\Lambda}\cdot\hat{\mathbf{e}}_0$, so $\mathbf{Q}^{-1}\cdot\boldsymbol{\Lambda}$ must be a pure rotation $\mathbf{R}(\mathbf{a})$. Hence

$$\boldsymbol{\Lambda} = \mathbf{Q}(\mathbf{b})\cdot\mathbf{R}(\mathbf{a}) \tag{10.4}$$

as was claimed. It follows that the components of $\mathbf{a}$ and $\mathbf{b}$ constitute a chart for $\mathcal{L}$, and that the latter is a 6-dimensional Lie group. Since $|\mathbf{b}|$ is unbounded, the group is not compact. Consequently, Maschke's theorem does not apply and we cannot insist on representations being unitary.

10.1 The Lie Algebra $\Lambda(\mathcal{L})$

$\Lambda(\mathcal{L})$ is generated by 6 infinitesimal generators. Three of these we have already in the generators $\mathbf{v}_i$ of $\Lambda(\mathcal{R}(3))$ (see §7.1). A further 3 generators are

$$\mathbf{w}_i \equiv \left.\frac{\mathrm{d}\mathbf{Q}(b\hat{\mathbf{e}}_i)}{\mathrm{d}b}\right|_{b=0} \qquad i=1,2,3. \tag{10.5}$$

From (10.3) one easily finds

$$\mathbf{w}_1 = \begin{pmatrix}0&-1&0&0\\-1&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix},\ \mathbf{w}_2 = \begin{pmatrix}0&0&-1&0\\0&0&0&0\\-1&0&0&0\\0&0&0&0\end{pmatrix},\ \mathbf{w}_3 = \begin{pmatrix}0&0&0&-1\\0&0&0&0\\0&0&0&0\\-1&0&0&0\end{pmatrix}. \tag{10.6}$$

These boost generators, unlike the rotation ones, are not skew-symmetric.

The commutation relations are

$$[\mathbf{v}_i,\mathbf{v}_j] = \epsilon_{ijk}\mathbf{v}_k,\quad [\mathbf{w}_i,\mathbf{w}_j] = -\epsilon_{ijk}\mathbf{v}_k,\quad [\mathbf{v}_i,\mathbf{w}_j] = \epsilon_{ijk}\mathbf{w}_k.$$

The linear combinations

$$\mathbf{A}_i \equiv \tfrac{1}{2}(\mathbf{v}_i + \mathrm{i}\mathbf{w}_i) \tag{10.7}$$

and their complex-conjugates $\mathbf{A}_i^*$ have the same commutation relations as two independent sets of generators of $\Lambda(SU(2))$:

$$[\mathbf{A}_i,\mathbf{A}_j] = \epsilon_{ijk}\mathbf{A}_k,\quad [\mathbf{A}_i,\mathbf{A}_j^*] = 0,\quad [\mathbf{A}_i^*,\mathbf{A}_j^*] = \epsilon_{ijk}\mathbf{A}_k^*. \tag{10.8}$$

As with any other Lie group, any representation T of $\mathcal{L}$ automatically furnishes a representation of the Lie algebra $\Lambda(\mathcal{L})$ and we can use the images $T(\mathbf{A}_i)$ and $T(\mathbf{A}_i^*)$ of the $\mathbf{A}$'s under the representing map to identify which representation we have. The procedure is this. We define operators on our representing space U (whose elements we denote $|\psi\rangle$) by

$$\begin{aligned} J_i &\equiv \mathrm{i}T(\mathbf{A}_i), & K_i &\equiv \mathrm{i}T(\mathbf{A}_i^*),\\ J_\pm &\equiv J_1 \pm \mathrm{i}J_2, & K_\pm &\equiv K_1 \pm \mathrm{i}K_2.\end{aligned} \tag{10.9}$$

Since J_3 commutes with K_3, we can pick out a complete set of mutual eigenkets $|m_j, m_k\rangle$, where the labels are the relevant eigenvalues:

$$J_3|m_j,m_k\rangle = m_j|m_j,m_k\rangle,\quad K_3|m_j,m_k\rangle = m_k|m_j,m_k\rangle. \tag{10.10}$$

Arguing as in §8.1, we demonstrate that on applying $J_\pm$ to $|m_j,m_k\rangle$, we generate a new eigenket $|m_j\pm 1, m_k\rangle$, and similarly for $K_\pm$. If $\dim(U)$ is to be finite, m_j and m_k must have minimum and maximum values, $j \equiv m_j(\max) = -m_j(\min)$ and $k \equiv m_k(\max) = -m_k(\min)$, with j and k whole half-integers. The dimension of our representation is then $\dim(U) = (2j+1)(2k+1)$ and we can label the representation by the pair (j,k).

Exercise (28):

Show that the 4×4 representation of $\mathcal{L}$ given by the usual Λ matrices is the $(\frac{1}{2}, \frac{1}{2})$ representation by calculating the matrices of the Casimir operators J^2 and K^2 in this representation.

The best route to the irreps of $\mathcal{L}$ starts from Example 3. This shows that every element of $SL(2,\mathcal{C})$ corresponds to an element of $\mathcal{L}$. We already know from Box 2 that any rotation can be generated by a unitary member of $SL(2,\mathcal{C})$ so to demonstrate that any element of $\mathcal{L}$ can be generated it suffices to show that any boost along, say, the z-axis can be accomplished.

Exercise (29):

Complete this demonstration by showing that

$$\mathbf{M} = \begin{pmatrix} e^{-b/2} & 0 \\ 0 & e^{b/2} \end{pmatrix}$$

accomplishes a boost to speed $v = c \tanh b$ parallel to the z-axis.

Exercise (30):

A general element of $\mathcal{L}$ can be obtained by exponentiating a suitable linear combination of $\mathcal{L}$'s six generators $\mathbf{v}_i$, $\mathbf{w}_k$. Show that the corresponding element of $SL(2,\mathcal{C})$ is given by

$$\mathbf{M} = \exp\left[\tfrac{i}{2}(\mathbf{a} + i\mathbf{b}) \cdot \boldsymbol{\sigma}\right]. \tag{10.11}$$

Check that with $\mathbf{b} = 0$ $\mathbf{M}$ corresponds to a rotation through angle $|\mathbf{a}|$.

Exercise (31):

Calculate the Casimir operators J^2 and K^2 in the representation provided by $SL(2,\mathcal{C})$ and thus show that this is the $(\frac{1}{2}, 0)$ irrep. The complex-conjugate representation M^* is the $(0, \frac{1}{2})$ irrep.

In fact $SL(2,\mathcal{C})$ stands in the same relation to $\mathcal{L}$ that $SU(2)$ does to $\mathcal{R}(3)$: it is the universal covering group of $\mathcal{L}$ and each element of $\mathcal{L}$ corresponds to *two* elements of $SL(2,\mathcal{C})$: $\mathbf{M}$ and $-\mathbf{M}$ always generate the same element of $\mathcal{L}$.

We may construct higher-dimensional irreps by considering objects with more than two spinor indices. For example, let $\alpha_{A\dot{B}\dot{C}}$ have the transformation law

$$\boldsymbol{\alpha} \to \boldsymbol{\alpha}' = \sum_{D,\dot{E},\dot{F}=1,2} M_{AD} M^*_{\dot{B}\dot{E}} M^*_{\dot{C}\dot{F}} \alpha_{D\dot{E}\dot{F}} \,. \tag{10.12}$$

As here, spinor indices are often written in capitals to distinguish them from regular tensor indices, and indices that should be transformed with $\mathbf{M}^*$ rather than with $\mathbf{M}$ are dotted. Objects that have the transformation law of (10.12) provide a representation of $\mathcal{L}$ that contains the $(\frac{1}{2}, 1)$ representation, and similarly for other values of (j, k).

Note:

In the notation of (10.12) each 4-vector v^μ corresponds to the two-index spinor $v_{A\dot{B}}$, formed by the matrix

$$\begin{pmatrix} v^0 + v^3 & v^1 - iv^2 \\ v^1 + iv^2 & v^0 - v^3 \end{pmatrix}.$$

10.2 The Lorentz group with inversions

Let $i \in \mathcal{L}_s$ correspond to spatial inversions. In the usual tensor notation, this has the matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{10.13}$$

Physically it is clear that $iQ(\mathbf{b})i = Q(-\mathbf{b})$, that is, that doing an inversion, then a boost along $\mathbf{b}$ and then another inversion, has the same effect as doing a boost along $-\mathbf{b}$. From (10.13) this may be readily checked algebraically. So i does not commute with boosts. On the other hand, it does commute with rotations: $iR(\mathbf{a})i = R(\mathbf{a})$.

Writing $Q(\mathbf{b}) = e + (\mathbf{b}\cdot\mathbf{w}) + \cdots$ and $R(\mathbf{a}) = e + (\mathbf{a}\cdot\mathbf{v}) + \cdots$ we deduce the implications of all this for the relations of i and the generators $\mathbf{w}$ and $\mathbf{v}$:[10]

$$\begin{matrix} \mathbf{w}i = -i\mathbf{w} \\ \mathbf{v}i = i\mathbf{v} \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} \mathbf{A}i = i\mathbf{A}^* \\ \mathbf{A}^*i = i\mathbf{A} \end{matrix} \tag{10.14}$$

By changing the sign of $\mathbf{w}$, which is the imaginary component of the $SU(2)$ generator $\mathbf{A}$, i changes $\mathbf{A}$ into the other $SU(2)$ generator, $\mathbf{A}^*$.

Now let $|m_j m_k\rangle$ be a ket in the (j,k) representation of $\mathcal{L}$. Then the operator $T(i)$ takes this into a ket of the (k,j) representation. To show this, we calculate the action of J_3 on $T(i)|m_j m_k\rangle$:

$$\begin{aligned} J_3 T(i)|m_j m_k\rangle &= \mathrm{i}T(A_3 i)|m_j m_k\rangle \\ &= \mathrm{i}T(iA_3^*)|m_j m_k\rangle \\ &= T(i)K_3|m_j m_k\rangle \\ &= m_k T(i)|m_j m_k\rangle. \end{aligned} \tag{10.15}$$

The claimed result now follows because we know that there are $2k+1$ values of m_k.

If $j \neq k$, it is clear that $T(i)|m_j m_k\rangle$ is a brand new ket and not linearly dependent on the $|m_j m_k\rangle$, which span the representation space of $\mathcal{L}$. We conclude that to each irrep of $\mathcal{L}$ with $j \neq k$ there corresponds an irrep of $\mathcal{L}_s$ that has dimension $2(2j+1)(2k+1)$. For example the 2-dimensional irrep $(\frac{1}{2},0)$ of $\mathcal{L}$ provided by $SL(2,\mathcal{C})$ yields a 4-dimensional irrep of $\mathcal{L}_s$ that is also denoted $(\frac{1}{2},0)$. The kets of the latter irrep are called **Dirac spinors**.

When $j = k$, $T(i)|m_j m_k\rangle$ *is* linearly dependent on the $|m_j m_k\rangle$, so adding i does not double the dimension of the irrep. To prove this we consider the $(2j+1)^2$-dimensional space U spanned by the kets

$$|m_j m_k; s\rangle \equiv |m_j m_k\rangle + T(i)|m_k m_j\rangle. \tag{10.16}$$

(Notice that for all the $|m_j m_k; s\rangle$ to be well defined we require $j = k$.) U is invariant under $T(i)$. Moreover,

$$\begin{aligned} J_3|m_j m_k; s\rangle &= m_j\big(|m_j m_k\rangle + T(i)|m_k m_j\rangle\big) \\ J_+|m_j m_k; s\rangle &= \sqrt{j(j+1) - m_j(m_j+1)}\big(|m_j+1\, m_k\rangle + T(i)|m_k m_j+1\rangle\big) \\ J_-|m_j m_k; s\rangle &= \sqrt{j(j+1) - m_j(m_j-1)}\big(|m_j-1\, m_k\rangle + T(i)|m_k m_j-1\rangle\big), \end{aligned} \tag{10.17}$$

so U is invariant under all $a \in \mathcal{L}$. Hence U supports a $(2j+1)^2$-dimensional irrep of $\mathcal{L}_s$ and must coincide with our original representation space for $\mathcal{L}$. The fact that the usual 4-vectors provide the $(\frac{1}{2},\frac{1}{2})$ irreps of both $\mathcal{L}$ and $\mathcal{L}_s$ is a particular case of this general result.

[10] Here the boldness of $\mathbf{w}$ does not indicate that it is a matrix, but that it is a vector (with matrices for components).

10.3 The Poincaré group

In addition to rotations and boosts, space-time is invariant under translations. The Poincaré group[11] $\mathcal{P}$ is the group formed by adding translations to $\mathcal{L}_{st}$. The action of a translation is simply

$$\mathbf{x} \to \mathbf{x}' = \epsilon\mathbf{x} = \mathbf{x} + \boldsymbol{\epsilon}, \tag{10.18}$$

where $\boldsymbol{\epsilon}$ is a 4-vector. $\mathcal{P}$ has 10 parameters – the 6 parameters of $\mathcal{L}$ and the 4 components of $\boldsymbol{\epsilon}$.

The irreps of $\mathcal{P}$ are of interest because the Lagrangian density of any free field is invariant under any $p \in \mathcal{P}$ with the consequence that any field obtained by transforming a given solution of the field equations with an element of $\mathcal{P}$, is also a solution of the field equations. Hence it is natural to seek solutions of the field eqns that together form the basis for an irrep of $\mathcal{P}$. Such sets of fields are fundamental to perturbative field theory, and are closely identified with the concept of a 'particle'.

The translations form an Abelian subgroup $\mathcal{T} \subset \mathcal{P}$. The rule for interchanging $\Lambda \in \mathcal{L}$ and $\epsilon \in \mathcal{T}$ is deduced as follows:

$$\begin{aligned} \Lambda\epsilon\mathbf{x} &= \boldsymbol{\Lambda} \cdot (\mathbf{x} + \boldsymbol{\epsilon}) \\ &= \boldsymbol{\Lambda} \cdot \mathbf{x} + \boldsymbol{\Lambda} \cdot \boldsymbol{\epsilon} \\ &= (\Lambda \cdot \boldsymbol{\epsilon})\Lambda\mathbf{x}. \end{aligned} \tag{10.19}$$

That is, translating by $\boldsymbol{\epsilon}$ and then transforming with $\boldsymbol{\Lambda}$ has the same effect as transforming with $\boldsymbol{\Lambda}$ and then translating by $\boldsymbol{\Lambda} \cdot \boldsymbol{\epsilon}$.

From (10.19) we can deduce the commutation relations between the infinitesimal generators of the translations P_μ and those Q_α, say, of $\mathcal{L}$ (here $\alpha = 1, \ldots, 6$ and $Q_1 \equiv T(v_1)$, say). We have

$$T(\epsilon) = I + \epsilon^\mu P_\mu + \cdots \quad , \quad T(\Lambda(\mathbf{a})) = I + a^\alpha Q_\alpha + \cdots \tag{10.20}$$

Notice that P_μ and Q_α are operators on the representation space U, and that $\mathbf{a}$ and $\boldsymbol{\epsilon}$ are small.

$$\begin{aligned} 0 &= T(\Lambda)T(\epsilon) - T(\Lambda \cdot \boldsymbol{\epsilon})T(\Lambda) \\ &= (I + a^\alpha Q_\alpha)(I + \epsilon^\mu P_\mu) - (I + \Lambda^\mu{}_\nu(\mathbf{a})\epsilon^\nu P_\mu)(I + a^\alpha Q_\alpha) \end{aligned} \tag{10.21}$$

Writing $\Lambda^\mu{}_\nu(\mathbf{a}) \simeq \delta^\mu_\nu + a^\beta v^\mu_{\beta\nu}$, where the 4×4 matrix $\mathbf{v}_1$ is the x-generator of $\mathcal{L}$ and $\mathbf{v}_4 = \mathbf{w}_1$ is the generator of x-boosts etc, (10.21) reduces to

$$0 = a^\alpha \epsilon^\mu \big(Q_\alpha P_\mu - P_\mu Q_\alpha\big) - a^\beta v^\mu_{\beta\nu} \epsilon^\nu P_\mu. \tag{10.22}$$

Deleting redundant infinitesimals we have finally

$$[Q_\alpha, P_\mu] = v^\nu_{\alpha\mu} P_\nu. \tag{10.23}$$

This expresses the commutator of two generators of $\mathcal{P}$ as a linear combination of the generators, so the numbers $v^\mu_{\beta\nu}$ are some of $\mathcal{P}$'s structure constants.

Exercise (32):

Does any P_μ commute with any of the Q_α?

Crucially, the new generators do not commute with the Casimir operators J^2, K^2 of $\mathcal{L}$. This tells us that the P_μ map kets that belong to different irreps of $\mathcal{L}$ into one another, with the result that the irreps of $\mathcal{P}$ are bigger than those of $\mathcal{L}$.

Let U be the space on which the operators $T(p)$ of an irrep of $\mathcal{P}$ act. The translations form an Abelian subgroup $\mathcal{T} \subset \mathcal{P}$ and we know that all irreps of Abelian groups are 1-dimensional. Let $|\mathbf{k}\rangle \in U$ provide an irrep of $\mathcal{T}$. Then by the arguments deployed in the proof of Bloch's theorem (eq. (4.11)), we have for $\epsilon \in \mathcal{T}$

$$T(\epsilon)|\mathbf{k}\rangle = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\epsilon}}|\mathbf{k}\rangle, \tag{10.24}$$

[11] A.k.a. the **inhomogeneous Lorentz group**.

where we have assumed that our ket has been labelled by the 4-vector $\mathbf{k}$ that determines its response to each $T(\epsilon)$.

Since U is irreducible, any ket in it can be generated by applying an appropriate $T(p)$ to some chosen starting ket $|\mathbf{k}_0\rangle$. Now

$$\mathbf{k} = \boldsymbol{\Lambda} \cdot \mathbf{k}_0 \quad \Rightarrow \quad |\mathbf{k}\rangle = T(\Lambda)|\mathbf{k}_0\rangle. \tag{10.25}$$

To see that this is so, we apply $T(\epsilon)$ to $T(\Lambda)|\mathbf{k}_0\rangle$:

$$T(\epsilon)T(\Lambda)|\mathbf{k}_0\rangle = T(\Lambda)T(\Lambda^{-1} \cdot \boldsymbol{\epsilon})|\mathbf{k}_0\rangle, \tag{10.26}$$

where (10.19) has been used. Evaluating the r.h.s. with (10.24) we find that $T(\Lambda)|\mathbf{k}_0\rangle$ satisfies the defining relation of $|\mathbf{k}\rangle$:

$$\begin{aligned} T(\epsilon)\big(T(\Lambda)|\mathbf{k}_0\rangle\big) &= \mathrm{e}^{\mathrm{i}\mathbf{k}_0 \cdot \boldsymbol{\Lambda}^{-1} \cdot \boldsymbol{\epsilon}}\big(T(\Lambda)|\mathbf{k}_0\rangle\big) \\ &= \mathrm{e}^{\mathrm{i}(\boldsymbol{\Lambda} \cdot \mathbf{k}_0) \cdot \boldsymbol{\epsilon}}\big(T(\Lambda)|\mathbf{k}_0\rangle\big) \\ &= \mathrm{e}^{\mathrm{i}\mathbf{k} \cdot \boldsymbol{\epsilon}}\big(T(\Lambda)|\mathbf{k}_0\rangle\big). \end{aligned} \tag{10.27}$$

Equation (10.25) demonstrates that for $\mathbf{k} \neq 0$ the irreps of $\mathcal{P}$ are ∞-dimensional; all the $|\mathbf{k}\rangle$ obtained from a given $|\mathbf{k}_0\rangle$ must be linearly independent of each other because they each have a different eigenvalue w.r.t. some $T(\epsilon)$.

We have seen that the rep of a subgroup $\mathcal{H} \subset \mathcal{G}$ furnished by an irrep of $\mathcal{G}$ is usually reducible. It's interesting to ask how small $\mathcal{H} \subset \mathcal{P}$ must be if an irrep of $\mathcal{P}$ is to reduce to finite-dimensional irreps. Eqn. (10.25) shows that $\mathcal{L} \subset \mathcal{P}$ is still too large. The irrep *does* reduce to an infinite number of finite-dimensional irreps when we consider the **little group** $\mathcal{L}_\mathbf{k}$ of $\mathbf{k}$. This is the subgroup of $\mathcal{L}$ under which $\mathbf{k}$ is invariant. If $\mathbf{k}$ is time-like, $\mathcal{L}_\mathbf{k}$ is just the group $\mathcal{R}(3)$ of all rotations in the frame in which $\mathbf{k} = (|k|, 0, 0, 0)$, i.e., $\mathbf{k}$'s rest-frame. We know that the irreps of this group are the $(2j+1)$-dimensional ones associated with states of definite angular momentum. If $\mathbf{k}$ is null, the little group is similar to the two-dimensional Euclidean group and the irreps are one-dimensional.

Two Casimir operators can be constructed for $\mathcal{P}$. One is $P^\mu P_\mu$ and detects the value of $|k|$ associated with the irrep. The other is more complex. It reports the nature of $\mathbf{k}$'s little group by returning either the total spin S^2 that characterizes rotations in $\mathbf{k}$'s rest frame, or if $\mathbf{k}$ is null and there is no rest frame, the helicity.

Exercise (33):

Find the action of $P^\mu P_\mu$ on $|\mathbf{k}\rangle$.