MA PH 464 - Group Theory in Physics

Lecture Notes

MA PH 464 - Group Theory in Physics

Lecture Notes

Vincent Bouchard
University of Alberta

March 30, 2020

These are lecture notes for MA PH 464 offered at the University of Alberta.

Motivation

Symmetry is the language of modern physics. When confronted with a physical system, the theoretical physicist's main tool to formulate a mathematical description of the system is to study its symmetries. Those correspond to transformations of the physical system that leaves the physics invariant.

The mathematical language for studying transformations and symmetries is group theory. This is why group theory is omnipresent in modern theoretical physics. We study what kind of transformations leave a physical system invariant; those transformations generally form an abstract group (the symmetry group of the system), and then we use the language of group theory to describe the physical system.

Groups are defined abstractly. But what they consist in is a set of symmetries (of an object, a theory, etc.) that is closed under composition and contains inverse operations. Most symmetries of nature are of that sort, which is why group theory appears universally.

But in fact in physics what we are interested in is how groups act on something: an object, a theory. This is what representation theory is about: it represents the elements of a group as acting on something. While in geometry, one may ask what the symmetries of a particular object are, in representation theory, the question is turned around: given a group, what kind of objects does it act on? This is the central theme of representation theory.

In particular, in physics groups of symmetries act on the space of solutions of a physical system. This is given by a representation of the symmetry group. Understanding the details of this representation tells us a lot about the properties of the possible solutions (or particles, fields, etc.)

Therefore, by studying group theory and representation theory, we will learn a lot about physical entities. We will see how the spin of a particle arises naturally in terms of representations, how particles in the Standard Model transform according to representations of the gauge group $SU(3) \times SU(2) \times U(1)$, how these representations naturally give rise to the idea of Grand Unified Theories, etc. Cool stuff!

References

These lecture notes are not meant to be original. They are heavily influenced by already existing lecture notes and textbooks on group theory in physics. The main references that I have been using are:

- Anthony Zee, Group Theory in a Nutshell for Physicists, Princeton University Press, 2016.
- Dimitri Vedensky, lecture notes on Group Theory.
- Jock McOrist, lecture notes on Representation Theory.
- Eugene A. Lim, lecture notes on Symmetry in Physics.

Contents

Chapter 1

Basic concepts of group theory

1.1 Groups

Objectives

You should be able to:

- Recall the definition of a group.
- Recognize various groups that occur frequently in physics and mathematics.
- Show in particular examples that the group axioms are satisfied.
- Distinguish between discrete, finite and continuous groups.
- Determine whether a group is abelian.

1.1.1 Definition

In physics, we think of symmetries in terms of transformations of a system (think of rotations), that we can compose with each other. What we want to do here is capture the abstract properties that are shared by all such symmetry transformations. This is what the abstract concept of a group does. Basically, a group is a set with a binary operation (composition) that satisfies a bunch of axioms.

Definition 1.1.1 Group. A **group** is a set G with a binary operation $\cdot$, which we call composition or multiplication, that satisfies the following axioms:

- *Closure*: For all $a, b \in G$, $a \cdot b \in G$.
- *Associativity*: For all $a, b, c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- *Identity*: There exists a group element $e \in G$, known as the identity, such that $e \cdot a = a \cdot e = a$ for all $a \in G$.
- *Inverses*: For all $a \in G$, there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

◊

It is easy to show a few elementary properties of groups:

1. The identity element of a group is always unique.
2. For each element $a \in G$, there is a unique inverse $a^{-1} \in G$.
3. In a group, $a \cdot b = a \cdot c$ implies $b = c$, and $b \cdot a = c \cdot a$ implies $b = c$, as expected.

Checkpoint 1.1.2 Prove these three properties.

We can further categorize groups according to some of their properties:

Definition 1.1.3 Abelian groups. If $a \cdot b = b \cdot a$ for all $a, b \in G$, i.e. composition is commutative, then we say that G is **abelian**. We say that it is **non-abelian** otherwise. ♢

Definition 1.1.4 Finite groups. If G has a finite number of elements, we say that the group is **finite**. We call $|G|$ the **order** of G. ♢

Definition 1.1.5 Continuous and discrete groups. If we write $G = \{g_\alpha\}$, then the index α can be either discrete or continuous. Note that a discrete group can well be infinite. ♢

1.1.2 Examples

You already know many examples of groups. Let us look at a few of them:

Example 1.1.6 Rotations. Rotations in two-dimensional space form a group, denoted by $SO(2)$, which is an abelian continuous group. Rotations in three-dimensional space also form a group, which is denoted by $SO(3)$, which happens to be non-abelian. Similarly, rotations in n dimensions form the non-abelian group $SO(n)$. □

Example 1.1.7 $\mathbb{R}$ and $\mathbb{Z}$ under addition. The set $\mathbb{R}$ with the operation of addition forms an abelian continuous group. If we restrict to the subset of integers $\mathbb{Z}$, we still get an infinite abelian group, but it is now discrete. □

Example 1.1.8 The trivial group. The group with only one element $\{1\}$ forms a group under multiplication, but it is a rather boring one. It is called the **trivial group**. □

Example 1.1.9 The cyclic group $\mathbb{Z}_N$. The two square roots of 1, $\{1, -1\}$, form the group $\mathbb{Z}_2$ under multiplication, which is an abelian discrete group of order 2. More generally, the N'th roots of unity, $\{e^{2\pi ik/N}\}_{k=0,\ldots,N-1}$ form the group $\mathbb{Z}_N$ under multiplication, which is an abelian discrete group of order N. □

Example 1.1.10 The unitary group $U(1)$. Complex numbers of norm 1, namely $e^{i\theta}$ with $\theta \in [0, 2\pi)$, form an abelian continuous group under multiplication called $U(1)$. □

Example 1.1.11 Addition of integers mod N. Addition of integers mod N generates an abelian discrete group of order N. The elements of the set are $\{0, 1, \ldots, N-1\}$, with the operation of addition mod N. This has in fact the same abstract group structure as $\mathbb{Z}_N$ defined above. □

Example 1.1.12 The symmetric group S_n. The set of invertible mappings $f : S \to S$ of a finite set S with n elements forms a group, which is denoted S_n and called **symmetric group** (or permutation group). The elements of the group S_n are permutations of n objects. There are $n!$ such distinct permutations, so S_n has order $n!$, and is non-abelian for $n \geq 3$. □

Example 1.1.13 The general linear groups $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$. The set of invertible linear maps $f : V \to V$ of a vector space V forms a continuous group, known as the **general linear group of** V and denoted by $GL(V)$. If $V = \mathbb{R}^n$, we write $GL(n, \mathbb{R})$; if $V = \mathbb{C}^n$, we write $GL(n, \mathbb{C})$. We can also think of those as the sets of invertible $n \times n$ matrices with the operation of matrix multiplication (with either real or complex entries). □

Example 1.1.14 The special linear groups $SL(n, \mathbb{R})$ and $SL(n, \mathbb{C})$. The set of $n \times n$ matrices with unit determinant under matrix multiplication also forms a continuous group, since $\det(AB) = \det(A)\det(B)$. This is called the **special linear group** $SL(n, \mathbb{R})$ for matrices with real entries, and $SL(n, \mathbb{C})$ for matrices with complex entries. □

Example 1.1.15 The Lorentz and Poincare groups. Lorentz transformations in Minkowski space form a continuous group. If we also include translations, we obtain the Poincare group. □

1.2 Subgroups

Objectives

You should be able to:

- Determine whether a subset of a group is a subgroup.
- Construct the orthogonal and unitary groups as subgroups of the general linear groups.
- Determine the centre of a group.
- Determine the cyclic subgroup generated by an element of a group.

1.2.1 Definition

The notion of a subgroup is rather obvious. But these are important in physics, for instance in the context of symmetry breaking. Suppose that a physical system is symmetric under a certain group of transformation. Then one could think of adding a perturbation to the system, for instance an extra term in the Lagrangian or Hamiltonian, that lowers the symmetry group to a subgroup of symmetries. This is captured by the concept of a subgroup.

The idea of subgroups is just like subspaces for vector spaces, which consist of subspaces that are vector spaces in their own right.

Definition 1.2.1 Subgroups. A subset H of a group G is a **subgroup** if it is a group in its own right. We write $H \subset G$. ◊

Since H is a group in its own right, it follows that any subgroup of G contains the identity element $e \in G$. According to the definition, the group G is a subgroup of itself. Also, for any G, the subset $\{e\}$ containing only the identity element is a subgroup, known as the *trivial subgroup*. Those two subgroups are rather boring; we call them *improper*. We are interested in studying *proper* subgroups.

1.2.2 Examples

Some examples of subgroups:

Example 1.2.2 Rotations in 3D around a fixed axis. Consider only rotations in three dimensions around a fixed axis. This gives a subgroup $SO(2) \subset SO(3)$. □

Example 1.2.3 Permutations that leave some objects fixed. Consider only permutations of n objects that leave $n - m$ objects fixed. This gives a subgroup $S_m \subset S_n$, with $m < n$. □

Example 1.2.4 The alternating group A_n. We have not defined the parity of a permutation yet, but if we consider the subgroup of permutations that are even, we get the so-called **alternating subgroup** $A_n \subset S_n$. □

Example 1.2.5 Cyclic subgroup in $\mathbb{Z}_n$. Consider the cyclic group $\mathbb{Z}_n$, represented as the n'th roots of unity under multiplication. Let m be an integer that divides n. Then there are a subset of n'th roots of unity that are also m'th roots of unity; those form the subgroup $\mathbb{Z}_m \subset \mathbb{Z}_n$. For instance, if we write $\mathbb{Z}_4 = \{1, -i, -1, i\}$, then we have a subgroup $\mathbb{Z}_2 = \{1, -1\} \subset \mathbb{Z}_4$. □

Example 1.2.6 Integers and reals. The real numbers $\mathbb{R}$ form a group under addition. If we restrict to the integers $\mathbb{Z}$, then we get a subgroup $\mathbb{Z} \subset \mathbb{R}$ under addition. □

Example 1.2.7 The subgroup $m\mathbb{Z}$. The set of all multiples of a positive integer m, denoted by $m\mathbb{Z}$, is a subgroup of $\mathbb{Z}$ with respect to addition. In fact, all subgroups of $\mathbb{Z}$ are of this form. □

We can in fact construct many subgroups of $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$. Recall that we can think of $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$ either as groups of real and complex invertible matrices (respectively), or as groups of invertible linear transformations on $\mathbb{R}^n$ and $\mathbb{C}^n$. We can construct subgroups by restricting to matrices of special types, or by looking at linear transformations that preserve special structures on $\mathbb{R}^n$ or $\mathbb{C}^n$.

Example 1.2.8 The special linear groups. Since $SL(n, \mathbb{R})$ and $SL(n, \mathbb{C})$ consist of $n \times n$ matrices with determinant 1, they form subgroups of $GL(, n\mathbb{R})$ and $GL(n, \mathbb{C})$ respectively. □

Example 1.2.9 The orthogonal group $O(n)$. We can start with $GL(n, \mathbb{R})$, and consider the subset of orthogonal matrices, which we denote by $O(n)$. Recall that a matrix $M \in GL(n, \mathbb{R})$ is orthogonal if $M^T M = I$.

The subset $O(n)$ is a subgroup of $GL(n, \mathbb{R})$, which we now check. First, the identity $I \in O(n)$. Second, if $A, B \in O(n)$, then $AB \in O(n)$, since $(AB)^T AB = B^T A^T AB = B^T B = I$. Therefore the subset $O(n)$ is closed under matrix multiplication. Finally, if $A \in O(n)$, then its inverse A^{-1} is also orthogonal, and hence in $O(n)$. Indeed, since $A^{-1} = A^T$, we have $(A^{-1})^T A^{-1} = (A^T)^T A^T = AA^T = I$.

Orthogonal matrices also have a special interpretation in terms of linear transformations of $\mathbb{R}^n$. We want to look at linear transformations that preserve special structures on $\mathbb{R}^n$. In this case, we are interested in preserving the Euclidean scalar product on $\mathbb{R}^n$. Recall that given two column vectors $u, v \in \mathbb{R}^n$, the Euclidean scalar product is given by

$$u^T v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

If we apply a linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$, to get new vectors $u' = Mu$ and $v' = Mv$, then

$$(u')^T v' = (Mu)^T (Mv) = u^T M^T M v.$$

Thus $(u')^T v' = u^T v$ for all $u, v \in \mathbb{R}^n$ if and only if $M^T M = I$. In other words, M is orthogonal, i.e. in $O(n)$. Therefore, $O(n)$ can be thought of as the subgroup of invertible linear transformations of $\mathbb{R}^n$ that preserve the Euclidean scalar product.

These transformations preserve length and angle. Indeed, the length square of a vector u is $|u|^2 = u^T u$, thus if $u' = Mu$ with M orthogonal, then $|u'|^2 = |u|^2$, and the length is preserved. But then, since $u^T v = |u||v| \cos\theta$ with θ the angle between the vectors u and v, it follows from $(u')^T v' = u^T v$ and the fact that $|u'| = |u|$ and $|v'| = |v|$ that $\theta' = \theta$, i.e. the angle between the vectors is also preserved.

But what transformations preserve length and angle? Rotations and reflections! So $O(n)$ is the group generated by rotations and reflections in n dimensions. □

Example 1.2.10 The special orthogonal group $SO(n)$. Consider $n \times n$ orthogonal matrices M. Then $M^T M = I$, and $\det(M^T M) = \det(M)^2 = \det(I) = 1$, and hence $\det(M) = \pm 1$. It is easy to convince yourself that rotations correspond to the subgroup of transformations with $\det(M) = 1$ (which is the component of $O(n)$ that is connected to the identity). We call this subgroup the **special orthogonal group** $SO(n) \subset GL(n, \mathbb{R})$. This is the group of rotations in n dimensions. □

Checkpoint 1.2.11 Check that $SO(n)$ is a subgroup of $O(n)$.

Example 1.2.12 The unitary group $U(n)$ and special unitary group $SU(n)$. We can do a similar construction for $GL(n, \mathbb{C})$. Thinking of it as the group of $n \times n$ invertible matrices with complex entries, we can first restrict to the subset of unitary matrices, namely matrices $M \in GL(n, \mathbb{C})$ such that $M^\dagger M = I$. Here we defined the **adjoint** (or Hermitian conjugate) of a matrix M as $M^\dagger = (M^*)^T$, where M^* denotes the matrix M with the entries complex conjugated. Following the exact same steps as in the orthgonal case for real matrices, one can show that the subset of unitary matrices is a subgroup, which we call the **unitary group** $U(n) \subset GL(n, \mathbb{C})$. As for orthogonal matrices, $(\det M)^2 = 1$, and if we restrict to matrices with $\det M = 1$, we get the **special unitary group** $SU(n) \subset GL(n, \mathbb{C})$.

We can also think of unitary matrices as linear transformations of $\mathbb{C}^n$. What kind of transformations are those? They are the transformations that preserve the standard inner product on $\mathbb{C}^n$. Recall that given two vectors $u, v \in \mathbb{C}^n$, we define the inner product as

$$(u^*)^T v = u_1^* v_1 + u_2^* v_2 + \ldots + u_n^* v_n.$$

Thus, just as for orthogonal transformations, it is easy to see that unitary transformations preserve the inner product on $\mathbb{C}^n$. So the unitary subgroup $U(n) \subset GL(n, \mathbb{C})$ can be understood as the subgroup of invertible linear transformations on $\mathbb{C}^n$ that preserve the standard inner product. □

Example 1.2.13 The Lorentz group and its generalizations. Recall that the orthogonal group $O(n)$ can be thought of as the subgroup of invertible linear transformations of $\mathbb{R}^n$ that preserve the Euclidean scalar product. But instead of looking at the Euclidean notion of distance on $\mathbb{R}^n$, we could have started with a different scalar product. For instance, we could take $\mathbb{R}^4$ to be Minkowski space, that is, equipped with the Minkowski scalar product:

$$u \cdot v = -u_0 v_0 + u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Note the different sign for the first term, which corresponds to the time direction in Minkowski space. We can look at the subset of linear transformations of $\mathbb{R}^4$ that leave this notion of scalar product invariant. It forms a subgroup, which is denoted by $O(1,3) \subset GL(4, \mathbb{R})$. This is nothing but the **Lorentz**

group of special relativity! Indeed, one can check that elements of $O(1,3)$ are Lorentz transformations. So one can think of Lorentz transformations as "rotations in Minkowski space".

(More precisely, $O(1,3)$ is obtained by composing Lorentz transformations with parity and time reversal transformations, just as $O(n)$ is obtained by composing reflections with rotations. Lorentz transformations correspond to the component of $O(1,3)$ connected to the identity.)

In general, if we start with $\mathbb{R}^n$ with a scalar product of the form

$$u \cdot v = -u_1 v_1 - \ldots - u_p v_p + u_{p+1} v_{p+1} + \ldots + u_n v_n,$$

the subset of linear transformations that preserve this inner product is denoted by $O(p, n-p)$, and is a subgroup of $GL(n, \mathbb{R})$. □

Example 1.2.14 The symplectic group $Sp(2n, \mathbb{R})$. Another interesting subgroup that plays an important role in physics is the symplectic group. Let $x, y \in \mathbb{R}^{2n}$, and J be the $2n \times 2n$ symplectic matrix $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, where I is the $n \times n$ identity matrix. The subset of $GL(2n, \mathbb{R})$ that leave the "anti-symmetric bilinear form" $x^T J y$ invariant is a subgroup of $GL(2n, \mathbb{R})$, which is called the **sympletric group** and is denoted by $Sp(2n, \mathbb{R})$. It is for instance fundamental in the more geometrical treatment of Hamiltonian mechanics. □

Many of those groups will play a fundamental in physics. In particular, we will see that $SO(3)$ and $SU(2)$ are intimately related, and so are $SO(1,3)$ and $SL(2, \mathbb{C})$.

1.2.3 A few universal subgroups

There are a few particular subgroups that are rather universal. In some way, what we are doing now parallels what we you did in linear algebra for vector spaces. We want to define the analog of the span of a subset, and the commutator of two elements, but in the context of abstract groups.

Definition 1.2.15 Subgroup generated by a subset of a group. Let S be a subset of finite group G. The **subgroup generated by** S, denoted by $\langle S \rangle$, is the union of S and all inverses and products of the elements in S

In particular, if S consists of a single element $S = \{a\}$ with $a \in G$, then we write $\langle a \rangle$ for the subgroup generated by a, which consists in the collection of integer powers of a. Note that because G is finite, we know that at some point $a^k = e$ for some integer k. $\langle a \rangle$ is called the **cyclic subgroup generated by** a. ♢

In fact, based on this, we can define the notion of cyclic groups.

Definition 1.2.16 Cyclic group. A **cyclic group** is a group generated by a single element. In other words, it is equal to one of its cyclic subgroups $G = \langle g \rangle$ for some element $g \in G$, called a generator of G. ♢

Given a particular element of a group, we can also look at all elements that commute with it.

Definition 1.2.17 Centralizer and the centre of a group. Let $a \in G$. The set of elements of G that commuate with a, denoted by $C_G(a)$, is called the **centralizer of a in** G. The set of elements of G that commute with all elements of G is called the **centre of** G, and denoted by $Z(G)$. ♢

It is easy to to show that both $C_G(a)$ and $Z(G)$ are subgroups of G. Moreover, $Z(G)$ is abelian by definition, and, clearly, G is abelian if and only if $Z(G) = G$, since all its elements commute with each other.

Remark 1.2.18 The definitions that we are making look very similar to the analogous definitions for vector places, but with 0 replaced by the identity element e. Indeed, the definitions that we are proposing reduce to the previous propositions for vector spaces when the group operation is considered to be vector space addition. In this case, the identity element for addition in a vector space is the zero vector, so e would be replaced by 0, as needed.

We can also define the "commutator" of two elements of a group G. But we need to keep in mind that we only have one operation to work with, namely group multiplication. The commutator should somehow measure how non-abelian a group G is.

Definition 1.2.19 The commutator subgroup of a group. Let $a, b \in G$. The **commutator** $[a, b]$ of a and b is defined by

$$[a, b] = a^{-1} \cdot b^{-1} \cdot a \cdot b.$$

If a and b commute, then $[a, b] = e$, thus we can think of the commutator as measuring how far from being abelian a group is. The subgroup generated by all commutators of elements G is called the **commutator subgroup** of G. $\diamondsuit$

If G is abelian, then $[a, b] = e$ for all $a, b \in G$. The implication goes in the other direction as well. Thus a group G is abelian if and only if its commutator subgroup is the trivial subgroup $\{e\} \subset G$.

1.3 Direct product

Objectives

You should be able to:

- Calculate the direct product of two groups.
- Determine when a group is the direct product of two of its subgroups.

We can also take products of groups to construct bigger groups. The natural operation is called direct product. This is analogous to the notion of direct sums of vector subspaces, where vectors are written as sums of vectors of different subspaces, such that the only common vector in the distinct subspaces is the zero vector.

Here we will first take the point of view where we study when a group is a direct product of two of its subgroups. We will then see how we can use the Cartesian product of two sets to naturally construct direct products of groups.

Definition 1.3.1 The direct product. A group G is said to be the **direct product** of its subgroups H_1 and H_2, which is denoted by $G = H_1 \times H_2$, if the following conditions are satisfied:

1. All elements of H_1 commute with all elements of H_2;
2. The group identity is the only common element to H_1 and H_2;
3. Every $g \in G$ can be written as $g = h_1 h_2$ for some $h_! \in H_1$ and $h_2 \in H_2$.

$\diamondsuit$

We can use this definition to construct bigger groups by taking direct products of groups:

Proposition 1.3.2 *Let G and H be groups. The Cartesian product $G \times H$ can* *be given a group structure by defining the operation*

$$(g, h) \star (g', h') = (gg', hh')$$

for all $g, g' \in G$ and $h, h' \in H$. The resulting group is the direct product of the subgroups (G, e) and (e, H), regarded as subgroups of $G \times H$.

Checkpoint 1.3.3 Prove this proposition.

Example 1.3.4 The Klein four-group, aka $\mathbb{Z}_2 \times \mathbb{Z}_2$. An example of a direct product is the group $\mathbb{Z}_2 \times \mathbb{Z}_2$ (also called the “Klein four-group”). Using the construction above, one can think of this group as consisting of the four elements $(1, 1), (1, -1), (-1, 1), (-1, -1)$, with the operation being component-wise multiplication. Note that this group is different from the cyclic group of order 4 $\mathbb{Z}_4$; for instance, in $\mathbb{Z}_4 = \{1, -i, -1, i\}$ only two elements (1 and -1) square to the identity, while in $\mathbb{Z}_2 \times \mathbb{Z}_2$ all four elements square to the identity. □

This notion of direct product is what is needed to make sense, for instance, of the gauge group of the Standard Model in particle physics, which is given by $SU(3) \times SU(2) \times U(1)$.

1.4 Multiplication table

Objectives

You should be able to:

- Construct the multiplication table of a finite group.

1.4.1 The idea

So far we studied various examples of groups by giving explicit realizations of them. But in the end groups can be understood entirely abstractly, without referring to a particular realization of it. In this section we will focus on finite groups.

To specify a finite group all that we need to know is the number of elements in the set, and the result of multiplying these elements together. This is neatly encoded in the form of a **multiplication table**. Suppose that G is a finite group of order n. Then its multiplication table will be a square $n \times n$ array with rows and columns labelled by elements of the group, and entries corresponding to the products of these elements:

Table 1.4.1 An example of a multiplication table

	$\cdots$	g_j	$\cdots$
$\vdots$	$\ddots$		
g_i		$g_i g_j$	
$\vdots$			$\ddots$

1.4.2 The rearrangement theorem

Before we construct multiplication tables, an interesting fact to note is that each element of a group can only appear once and only once in each row, and same for each column. This is sometimes formally stated as the “rearrangement theorem”:

Theorem 1.4.2 Rearrangement theorem. *Let $G = \{g_1, \ldots, g_n\}$ be a finite group, and pick an element $g_k \in G$. Then*

$$g_k G = \{g_k g_1, \ldots, g_k g_n\}$$

contains each element of the group once and only once.

Proof. Suppose that two elements of $g_k G$ are equal: $g_k g_i = g_k g_j$ for $i \neq j$. Then multiply by g_k^{-1} on the left to get $g_i = g_j$. This is a contradiction, since $i \neq j$. Thus all elements of $g_k G$ must be distinct. Since $|g_k G| = n = |G|$, it follows that all elements of G must appear once and only once in $g_k G$. Thus the sets G and $g_k G$ are identical up to reordering of elements, hence the name of the theorem. ■

1.4.3 Groups of order 2 and 3

With this under our belt let us start by constructing the multiplication table for a group of order two. Let us write $G = \{e, a\}$, where e is the identity element. We know that $e^2 = e$, $ea = a$ and $ae = a$. Since each element can appear once and only once in each row and column of the multiplication table, it follows that we must have $a^2 = e$, to get:

Table 1.4.3 The multiplication table for the group of order 2

	e	a
e	e	a
a	a	e

Let us now move to groups with three elements $G = \{e, a, b\}$. We know by definition of the identity that

Table 1.4.4 Constructing the multiplication table for the group of order 3

	e	a	b
e	e	a	b
a	a		
b	b		

Further, the product ab cannot be equal to a (resp. b), since otherwise we could multiply by a^{-1} (resp. b^{-1}) on the left to conclude that $b = e$ (resp. $a = e$). Thus we must have $ab = e$. We then complete the table using the requirement that each element appears once and only once in each row and column (this is like sudoku! :-)), to conclude that

Table 1.4.5 Multiplication table for the group of order 3

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Again, this is the unique possibility. Hence there is a unique abstract group with three elements. Further, it is abelian, since the table is symmetric about its diagonal.

Going to higher order however we lose uniqueness. For groups with four elements, there are two distinct multiplication tables (corresponding to the groups $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$, see Example 1.3.4). And for higher order groups there are many possibilities. I will not construct multiplication tables for groups of higher order, but if you like this kind of thing feel free to do it yourself as an exercise. Try for instance the symmetric group S_3, which has order $3! = 6$. Lots of fun!

1.5 Presentation

Objectives

You should be able to:

- Describe a group in terms of its presentation.
- Recall the general properties of the modular group.

If you have constructed the multiplication table for S_3 as an exercise, you have probably come to the realization that writing down multiplication tables becomes quite annoying rather quickly for higher order groups. There has to be a better way of encoding the abstract group structure of finite groups without having to write down the multiplication table each time. One such way is to write down the "presentation" of a group.

The idea is to specify a set of **generators** S of the group, which is a subset of elements of the group from which all other elements can be obtained by group multiplication, and the essential **relations** R that the generators satisfy. We write the **presentation** as $\langle S|R\rangle$. Without getting into too much detail, let me simply give a few examples of presentations.

Example 1.5.1 $\mathbb{Z}_2$, $\mathbb{Z}_n$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$. As a basic example, consider the group $\mathbb{Z}_2$ of two elements. Its presentation would be given as:

$$\langle a|a^2 = e\rangle.$$

Similarly, the cyclic group with n elements $\mathbb{Z}_n$ would be written as:

$$\langle a|a^n = e\rangle.$$

The direct product $\mathbb{Z}_2 \times \mathbb{Z}_2$ would be written as:

$$\langle a, b|a^2 = b^2 = e, ab = ba\rangle.$$

□

Example 1.5.2 The modular group. As a more interesting example, let us look at the so-called modular group, which is important in various areas of physics, such as string theory and conformal field theory. Consider 2×2 matrices $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries and such that $\det M = 1$. This group is denoted by $SL(2, \mathbb{Z})$, the special linear group of 2×2 matrices with integer entries. If we also impose the identification $M = -M$ in $SL(2, \mathbb{Z})$ we end up with the group $PSL(2, \mathbb{Z})$, the projective special linear group of 2×2 matrices with integer entries, which is also known as the **modular group**. This group is important because it is the group of linear fractional transformations of the upper half of the complex plane, where it acts as:

$$z \mapsto \frac{az + b}{cz + d}, \tag{1.5.1}$$

for $z \in \mathbb{C}$ with $\Im(z) > 0$.

This group looks rather complicated, but in fact one can show that all such

2×2 matrices can be obtained by repeatedly multiplying together two matrices:

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

In terms of fractional linear transformations, what that means is that the general transformation (1.5.1) can be obtained by repeatedly composing the two transformations:

$$S : z \mapsto -\frac{1}{z}, \qquad T : z \mapsto z + 1.$$

Moreover, it is clear that $S^2 = I$, where I is the identity matrix (the identity element in the group), and one can check easily that $(ST)^3 = I$ as well. Thus we can write the presentation of $PSL(2,\mathbb{Z})$ as:

$$\langle S, T | S^2 = I, (ST)^3 = I \rangle.$$

□

1.6 Cosets

Objectives

You should be able to:

- Determine the cosets of a subgroup of a finite group.
- State and prove Lagrange's theorem.
- Use Lagrange's theorem to rule out when a subset cannot be a subgroup of a group.

1.6.1 Definition

Now that we understand subgroups, we can talk about cosets. In fact we have already seen the idea of a coset when we stated the rearrangement theorem, but now instead of taking the whole group G, we will consider subgroups $H \subset G$.

Definition 1.6.1 Cosets. Let $H = \{e, h_1, \ldots, h_r\}$ be a subgroup of G, and pick an element $g \in G$. Then the set:

$$Hg = \{eg, h_1 g, \ldots, h_r g\}$$

is a **right coset** of H. Similarly, the set

$$gH = \{ge, gh_1, \ldots, gh_r\}$$

is a **left coset** of H. Cosets are generally not subgroups of G; they will be only if $g \in H \subset G$. ◊

Example 1.6.2 The cosets of $\mathbb{Z}_2 \subset \mathbb{Z}_4$. Consider the group $\mathbb{Z}_4$ represented as $G = \{1, -i, -1, i\}$ under multiplication. Consider the subgroup $H = \{1, -1\}$. Then the left cosets are:

$$1 \cdot H = \{1, -1\}, \qquad i \cdot H = \{i, -i\},$$
$$-1 \cdot H = \{-1, 1\}, \qquad -i \cdot H = \{-i, i\}.$$

So there are really two distinct cosets, given by H itself and the set $\{i, -i\}$.

Since the group is abelian, the left cosets are the same as the right cosets. In general however left cosets and right cosets are not the same. We also notice that all elements of G appear exactly once in the distinct cosets; as we will see, this is a general result. □

Cosets will turn out to be very useful in physics. Often we need to deal with the fact that many different mathematical quantities give the same physical observable. For instance, the electromagnetic potential is not uniquely fixed; add to it the gradient of any function, and you get the same electric and magnetic fields. Similarly, in quantum mechanics, any two states in a Hilbert space that only differ by a phase give the same physical system. The way to deal with this kind of redundancy mathematically is to define equivalence classes of objects that give the same physics. It turns out that cosets are perfectly suited for that; under suitable conditions, we can see the cosets as being equivalence classes, hence from a physics point of view what we are often really interested in are the cosets of a given group, rather than the group itself.

Thus what we would like to see is whether we can upgrade the cosets to groups. It is clear from the definition that cosets of a subgroup $H \subset G$ are generally not subgroups of G themselves. However, what we can do is look at the "set of cosets" all together; these in fact form a group, which will be called the quotient group. This will be the group of equivalence classes formed by the cosets, which is often the relevant group of objects that we are interested in physically. But we are jumping ahead of ourselves; we will come back to quotient groups shortly.

1.6.2 Cosets and Lagrange's theorem

Going back to cosets, the cool thing about them is that, given a subgroup, there is some sort of decomposition of a group into cosets of the subgroup, as you can see in the example above. This is because of the following simple Lemma:

Lemma 1.6.3 *Two cosets of a subgroups are either identical as sets or have no common elements.*

This lemma will be important to make sense of cosets as equivalence classes, since two elements cannot be in two distinct but not identical equivalence classes.

Proof. Let $H = \{e, h_1, \dots, h_r\}$ be a subgroup of G. Clearly, two cosets either have no common elements or have at least one common element. Let us show that if they have a common element, then they are identical.

Suppose that the left cosets generated by g_i and g_j, $i \neq j$, have a common element, say $g_i h_m = g_j h_n$ for some m and n. Then $g_j^{-1} g_i = h_n h_m^{-1}$, therefore $g_j^{-1} g_i \in H$. It thus follows that the left coset $g_j^{-1} g_i H$ is in fact equal to H itself as a set, since multiplying by an element of H on the left is only rearranging terms (this is the statement of the Rearrangement Theorem 1.4.2 for the subgroup H). Therefore $g_j^{-1} g_i H = H$ as sets, and multiplying by g_j on the left, we get $g_i H = g_j H$, that is, the two cosets are identical as sets.

We conclude that left cosets are either identical as sets or contain not common elements. The same proof goes through for right cosets. ■

This lemma allows us to prove an important theorem about subgroups of finite groups, known as Lagrange's theorem:

Theorem 1.6.4 Lagrange's theorem. *The order of a subgroup $H \subset G$ of a finite group G divides the order of G, that is, $|H|$ divides $|G|$.*

Proof. First, since the identity element e is in the subgroup H, and to construct the cosets of H we multiply by all elements of G, we see that all elements of

G must appear in at least one coset of H. But since two cosets are either identical as sets or have no common elements, it follows that all elements of G must appear in exactly one distinct coset.

Furthermore, one can prove that all cosets have the same number of elements, which is equal to the order of the subgroup $|H|$. Consider a coset gH for some $g \in G$. Suppose that two elements of the coset are identical, $gh_i = gh_j$ for some $h_i \neq h_j$. Multiplying by g^{-1} on the left, we get $h_i = h_j$, which is a contradiction. Therefore all elements of gH must be distinct, and hence the number of elements in gH is the same as the number of elements in H.

Putting this together, we conclude that the distinct cosets are partitioning the group G into non-intersecting bins of size equal to $|H|$. It follows that the number of distinct cosets times $|H|$ must be equal to $|G|$, that is, $|H|$ divides $|G|$. ■

Definition 1.6.5 The index of a subgroup. The number of distinct cosets of a subgroup $H \subset G$ is called the **index** of the subgroup. ♢

Lagrange's theorem is in fact quite powerful to rule out when a subset cannot be a subgroup of a finite group. For instance, Lagrange's theorem implies the following corollaries:

Corollary 1.6.6 More than half implies abelian. *If more than half of the elements of a finite group commute with each other, then the group is abelian.*
Proof. We know that the centre of a group is a subgroup. Thus, by Lagrange's theorem, we know that if the order of the centre of the group is larger than half of the order of the group, then the centre must be the whole group, i.e. the group is abelian. ■

Corollary 1.6.7 Groups of prime order. *All finite groups whose order are prime numbers are cyclic, and have no proper subgroups.*
Proof. By Lagrange's theorem, if a group G has order given by a prime number, then its only subgroups must have either order one (the trivial subgroup) or the same order as the group itself (the group itself). Thus it cannot have proper subgroups. Moreover, if you pick any element $g \in G$, then it generates a cyclic subgroup $\langle g \rangle \subseteq G$. But since G has no proper subgroups, then $\langle g \rangle$ is either the trivial subgroup if g is the identity element, or it is the whole group G for any other element $g \in G$. That is, G is a cyclic group, as it is equal to its non-trivial cyclic subgroups. ■

1.7 Conjugacy classes

Objectives

You should be able to:

- Determine the conjugacy classes of a finite group.

1.7.1 Definition

We have already discussed briefly the idea of equivalence classes. Given a group G, the idea is to define an equivalence relation, which we generally denote by $\equiv$, between elements of the group, which is reflexive ($a \equiv a$), symmetric (if $a \equiv b$ then $b \equiv a$) and transitive (if $a \equiv b$ and $b \equiv c$ then $a \equiv c$). Once we have an equivalence relation, we can define equivalence classes, which are subsets of elements of the group that are equivalent according to $\equiv$.

There is a particularly interesting equivalence relation, known as conjugation, that partitions any finite group into equivalence classes, known as conjugacy classes.

Definition 1.7.1 Conjugacy classes. We say that two elements $a, b \in G$ are **conjugate** if there exists an element $g \in G$ such that $a = gbg^{-1}$. This is an example of an equivalence relation. Then we can form equivalence classes, that is, classes of elements of G that are conjugate to each other. The set of all elements conjugate to one another is called a **conjugacy class**. ♢

Conjugacy classes are composed of elements that somehow behave in a similar way. What "similar" means here depends on the context. As we will see, for the symmetric group S_n, conjugacy classes will be comprised of permutations that have the same cycle structure. In the group of three-dimensional rotations $SO(3)$, we can characterize rotations by specifying an axis of rotation and an angle of rotation. Then one can show that all rotations that have the same angle belong to the same conjugacy class.

As we will see, conjugacy classes are very useful in the study of groups, especially in representation theory.

1.7.2 Properties of conjugacy classes

It is clear that conjugacy classes are disjoint, since an element cannot be in two different conjugacy classes. So we can partition a group G into its conjugacy classes.

It is also interesting to note that all elements in a given conjugacy class have the same order.

Definition 1.7.2 Order of an element of a group. We define the **order of an element** $a \in G$ to be the smallest integer n such that $a^n = e$. ♢

With this definition we can prove the following lemma:

Lemma 1.7.3 *All elements of a given conjugacy class have the same order.*

Proof. Suppose that $a \in G$ has order n. If b is conjugate to a, that is $b = gag^{-1}$ for some $g \in G$, then

$$b^n = (gag^{-1})^n = (gag^{-1})(gag^{-1}) \cdots (gag^{-1}) = ga^n g^{-1} = geg^{-1} = e.$$

It follows that $b^n = e$, and hence the order of b is $\leq n$.

To finish the proof, we need to argue that there cannot be an integer $m < n$ such that $b^m = e$, so that the order of b is precisely n. Suppose that there is such a $m < n$. Then

$$b^m = ga^m g^{-1} = e.$$

Multiplying by g^{-1} on the left and g on the right, we get that $a^m = e$. But n must be the smallest integer such that $a^n = e$, therefore this is a contradiction. It follows that the order of b is equal to n. ■

We can also prove the following interesting result:

Lemma 1.7.4 *Each element in the centre of a group G constitutes a conjugacy class by itself. And hence the identity is always a conjugacy class by itself, and for an Abelian group G, each element forms its own distinct class.*

Proof. Let $h \in Z(G)$ be an element of the centre of G. Then for all $g \in G$, we have $hg = gh$, and hence for all $g \in G$, $ghg^{-1} = h$, therefore the only element conjugate to h is h itself.

The two corollaries follow directly, since the identity is always an element of the centre of a group, and if G is abelian, then $Z(G) = G$. ■

1.8 The symmetric group S_n

Objectives

You should be able to:

- Sketch the proof of Cayley's theorem.
- Use standard notations (permutations, cycles) to describe elements of the symmetry group S_n.
- Calculate the parity of a permutation.
- Describe conjugacy classes in S_n in terms of cycle structures and integer partitions.

Many of the concepts that we just studied abstractly become alive when we study the symmetric group S_n, which is the group of permutations of n elements. Everything becomes clear and explicit, since we kind of all known intuitively how permutations work.

1.8.1 Cayley's theorem

But in fact studying the symmetric group S_n is fundamental, because of what is known as **Cayley's theorem**, which basically states that any finite group G of order n is isomorphic (that is, identical as abstract group) to a subgroup of the symmetric group S_n. More precisely, the statement of Cayley's theorem is that G is a subgroup of the group of permutations of the elements of G, which is S_n since G has order n. So from this point of view, studying symmetric groups and their subgroups basically means studying all finite groups.

Cayley's theorem is quite deep, since it says that we can realize any abstract finite group as a subgroup of a permutation group. In other words, it connects abstract group theory to the very concrete world of permutations. Isn't it cool? As we will see, Cayley's theorem gives an example of a nice representation which exists for all finite groups, and is called the **regular representation**.

Why is Cayley's theorem true? Without going through a formal proof, let us see how it goes. Consider the multiplication table of a finite group $G = \{g_1, \ldots, g_n\}$. The i'th row of the table is given by the elements $\{g_i g_1, \ldots, g_i g_n\}$. By the Rearrangement Theorem Theorem 1.4.2, this is the same set of elements as G, but in a different order. Thus we can assign to the i'th row an element of the permutation group of n elements, that is an element of S_n, which acts as $\{g_1, \ldots, g_n\} \mapsto \{g_i g_1, \ldots, g_i g_n\}$. As a result, to every element $g_i \in G$ we can assign a permutation $s_i \in S_n$; we have constructed a map from G to a subgroup of S_n, and one can show that this map is a group isomorphism. This is the essence of Cayley's theorem.

Example 1.8.1 **$\mathbb{Z}_3$ as a subgroup of S_3.** As an example, consider the group of three elements $G = \mathbb{Z}_3$ with multiplication table Table 1.4.5. We see that the map assigns to the element $e \in G$ the identity permutation $e \in S_3$; to $a \in G$ it assigns the cyclic permutation $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$ in S_3; and to $b \in G$ it assigns the cyclic permutation $1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2$. Thus we have identified $G = \mathbb{Z}_3$ with the alternating subgroup $A_3 \subset S_3$ consisting of even permutations. □

1.8.2 Notation

Now let us study the symmetric group S_n in more detail. First we introduce standard notation. An element of S_n corresponds to a permutation of n elements. It is customary to denote it by its action on the n elements $\{1, \dots, n\}$. For instance, the cyclic rotation of the n elements, which takes 1 to 2, 2 to 3, etc., would be denoted by

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}.$$

Generically, a permutation π that takes 1 to $\pi(1)$, 2 to $\pi(2)$, and so on, would be denoted by

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ \pi(1) & \pi(2) & \cdots & \pi(n-1) & \pi(n) \end{pmatrix}.$$

Example 1.8.2 The symmetric group S_3. As an example, consider S_3. It has $3! = 6$ elements; the identity permutation (π_1), three permutations corresponding to a single exchange of two elements (π_2, π_3, π_4), and two permutations that permute the three elements (π_5, π_6). In the notation introduced above, the six elements are

$$e = \pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \qquad \pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \pi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

To construct the multiplication table for S_3, we need to multiply elements in S_3, where by multiplication here we mean composition of permutations. For instance, the composition $\pi_4 \circ \pi_5$ can be constructed as follows. Start with π_5; it takes $1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2$. Then, combining with the action of π_4, we get $1 \mapsto 3 \mapsto 2$, $2 \mapsto 1 \mapsto 1$, and $3 \mapsto 2 \mapsto 3$, which is π_2. Therefore, $\pi_4 \circ \pi_5 = \pi_2$.

In fact, we see explicitly that S_3 is non-Abelian. Consider $\pi_5 \circ \pi_4$. We obtain $1 \mapsto 1 \mapsto 3$, $2 \mapsto 3 \mapsto 2$ and $3 \mapsto 2 \mapsto 1$, which is π_3. Therefore $\pi_5 \circ \pi_4 = \pi_3$, hence $\pi_4 \circ \pi_5 \neq \pi_5 \circ \pi_4$. □

1.8.3 Cycles

Since we are dealing with permutations of set of n elements, it is clear that after applying a given permutation a certain number of times, we will come back to the initial elements. In other words, consider for instance the element 1. A given permutation π will take $1 \to \pi(1)$. Applying it again, we will get $1 \to \pi(1) \to \pi(\pi(1))$. After a number of application, say r, we will necessarily get back $\pi^r(1) = 1$. The numbers $1, \pi(1), \pi^2(1), \dots, \pi^{r-1}(1)$, that are reached from 1 by π form what we call a "cycle". More precisely:

Definition 1.8.3 Cycle. Let $\pi \in S_n$, $i \in \{1, \dots, n\}$ and let r be the smallest positive integer such that $\pi^r(i) = i$. Then the set of r distinct elements $\{\pi^k(i)\}_{k=0}^{r-1}$ is called a **cycle** of π of length r, or an **r-cycle generated by** i. We denote a given cycle by $(i\ \pi(i)\ \dots\ \pi^{r-1}(i))$.

It is clear that a given permutation π breaks up the set of n elements $\{1, \dots, n\}$ into disjoint cycles. We can then denote the permutation π by the cycle decomposition of $\{1, \dots, n\}$ that it implies. ◊

Example 1.8.4 The cycle structure of permutations in S_3. Consider the permutation π_5 for S_3 as in the previous example. Start with 1. Then $\pi_5(1) = 3$, $\pi_5^2(1) = \pi_5(3) = 2$, and $\pi_5^3(1) = \pi_5(2) = 1$. Thus π_5 decomposes $\{1, 2, 3\}$ into a single cycle of length 3. In cycle notation, we write $\pi_5 = (132)$. We could have written $\pi_5 = (321)$ or $\pi_5 = (213)$ as well; those are the same permutations. In this notation the numbers defining the cycles can be moved around cyclically without changing anything. Similarly, π_6, which is the other permutation that exchanges the three elements, has cycle decomposition $\pi_6 = (123)$.

Consider instead the permutation π_2. Start with 1. Then $\pi_2(1) = 2$, and $\pi_2^2(1) = \pi_2(2) = 1$. Thus 1 generates a 2-cycle given by (12). There is another cycle however. Start with 3. Then $\pi_2(3) = 3$, hence 3 generates a 1-cycle given by (3). Therefore, the set of three elements is broken up into two cycles by π_2, that is $\pi_2 = (12)(3)$. Similarly, we get $\pi_3 = (13)(2)$, $\pi_4 = (1)(23)$, and $\pi_1 = (1)(2)(3)$. □

With the notion of cycles we can define permutations that are "cyclic":

Definition 1.8.5 Cyclic permutations. If $\pi \in S_n$ has a cycle of length r and all other cycles of π have only one element, then π is called a **cyclic permutation of length** r. ◊

Example 1.8.6 Cyclic permutations in S_3. In the example of S_3, we then have that π_5 and π_6 are cyclic permutations of length 3, while π_2, π_3, π_4 are cyclic permutations of length 2. Thus in S_3 all permutations are cyclic. But for instance in S_4, we have the permutation $\pi = (12)(34)$, which is not cyclic. □

A particularly useful type of cyclic permutation is of length two:

Definition 1.8.7 Transposition. A cyclic permutation of length 2 (i.e. a permutation that only exchanges two elements) is called a **transposition**. ◊

Example 1.8.8 Transpositions in S_3. To come back to S_3, there are three transpositions, namely π_2, π_3 and π_4. □

Now that we understand cycles, we can easily compute composition of permutations. Given two permutations π_1 and π_2, the composition $\pi_1 \circ \pi_2$ can be computed easily by "multiplying cycles".

This is easier understood in terms of an example. Consider the product of cycles (15)(234)(253) in S_5. To find the corresponding permutation, we consider each element of $\{1, \ldots, 5\}$, and, starting from the far right of the product of cycles, we keep track of what the element becomes under the action of the cycles. In our case, consider first 1. The first two cycles on the right leave it invariant, while the last cycle sends its to 5. Thus $1 \to 5$. Consider 2. The first cycle sends it to 5, the second cycle then leaves 5 invariant, and then the last cycle sends it to 1. Thus $2 \to 1$. Consider 3. Then we get $3 \to 2 \to 3 \to 3$. Consider 4. We get $4 \to 4 \to 2 \to 2$. Finally, for 5 we get $5 \to 3 \to 4 \to 4$. The corresponding permutation is then:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 3 & 2 & 4 \end{pmatrix}.$$

(A consistency check here is that each number 1 to 5 only appear once in the final result of the permutation, which must be the case.) Equivalently, we can write the resulting permutation in terms of its cycle structure, which is $\pi = (1542)(3)$. Hence we get that

$$(15)(234)(253) = (1542)(3).$$

This is how we multiply cycles.

Example 1.8.9 Multiplication of cycles vs composition of permutations. Let us just show in an example that multiplication of cycles and composition of permutations are equivalent. Consider $\pi_1 = (123)(45)$ and $\pi_2 = (1)(2)(345)$ in S_5. Suppose that we want to compute the composition $\pi_1 \circ \pi_2$. The resulting permutation is easily calculated to be

$$\pi_1 \circ \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 4 & 1 \end{pmatrix}.$$

We claim that this is equivalent to the product of cycles $(123)(45)(1)(2)(345)$. When we consider products of cycles, we can remove cycles of length 1, since they do not do anything in the product. So we want to compute $(123)(45)(345)$. Under this product, we have $1 \to 1 \to 1 \to 2$, $2 \to 2 \to 2 \to 3$, $3 \to 4 \to 5 \to 5$, $4 \to 5 \to 4 \to 4$, $5 \to 3 \to 3 \to 1$, which is indeed the permutation above. In fact, a few minutes of thought will make it clear that the way we defined products of cycles is precisely equivalent to composition of the corresponding permutations. □

A very interesting result for the symmetric group is that any permutation can in fact be written as the product of a number of transpositions. This is just saying that every permutation can be done in steps, where in each step we only exchange two objects. Let's see how it goes. We start by studying how a cycle can be written as a product of transpositions:

Proposition 1.8.10 Decomposition of cycles into products of transpositions. *An r-cycle $(i_1 i_2 \dots i_r)$ can be decomposed into the product of $r-1$ transpositions:*

$$(i_1 i_2 \dots i_r) = (i_1 i_r)(i_1 i_{r-1}) \cdots (i_1 i_3)(i_1 i_2).$$

Note however that the decomposition is not unique; in particular, since the square of any transposition is the identity, we could insert the square of any transposition in the product on the RHS without changing the end result.

Proof. The proof simply involves multiplying the transpositions on the RHS. Start with i_1. Under the product of transpositions, we get $i_1 \to i_2 \to \dots \to i_2$. Consider i_2: we get $i_2 \to i_1 \to i_3 \to \dots \to i_3$. Iterating, we get that $i_k \to i_{k+1}$ for all $k = 1, \dots, r-1$, and $i_r \to i_1$. Thus the resulting permutation is the cyclic permutation of length r that has cycle structure $(i_1 i_2 \dots i_r)$. ■

Definition 1.8.11 Parity. We define the **parity** of the cycle as being **even** if the number of transpositions in the decomposition is even, and **odd** if the number if transpositions is odd. Note that even though the decomposition is not unique, the parity of the decomposition is uniquely defined. ◊

We can then apply these results to permutations themselves, not just to individual cycles. We obtain:

Proposition 1.8.12 Decomposition of permutations into products of transpositions. *Any permutation can be decomposed as a product of transpositions (and length 1 cycles). The parity (even or odd, according to whether the decomposition has an even or odd number of transpositions) of the permutation is unique.*

In fact, the parity of a permutation can be determined directly from its cycle structure, without having to work out its decomposition into transpositions, since each r-cycle can always be decomposed into the product of $r-1$ transpositions.

Example 1.8.13 Decomposition of a permutation in S_6. As an example, consider the permutation $\pi = (145)(23)(6)$ in S_6. Looking at the cycle structure, the 3-cycle is even (can be decomposed into two transpositions), the 2-cycle is of course odd (it is a transposition itself), and the 1-cycle does not contribute to the parity. Thus π is odd, since $2+1 = 3$. In fact, a decomposition is easy to find following Proposition 1.8.10. We get:

$$\pi = (15)(14)(23)(6).$$

Of course, this decomposition is not unique (for instance (15)(14)(24)(24)(23)(6) would also work), but its parity is uniquely defined. □

Definition 1.8.14 The alternating group. The set of even permutations of n elements is denoted by A_n and is called the **alternating group**. ♢

Clearly, A_n is a subgroup of S_n, since the product of two even permutations is even, it contains the identity, and the inverse of an even permutation is also even.

1.8.4 Conjugacy classes

Now let us study conjugacy classes of S_n. Recall that conjugacy classes are defined as being subsets of elements that are conjugate to each other, where two elements $x, y \in G$ are conjugate if $y = gxg^{-1}$ for some $g \in G$.

The main result for S_n is the following theorem:

Theorem 1.8.15 Conjugacy classes and cycle structure. *Two permutations in S_n are conjugate if and only if they have the same cycle structure, meaning that they have the same number of cycles of equal length.*

To clarify the notation, if we go back to S_3, then the three permuations (12)(3), (13)(2) and (1)(23) have the same cycle structure, hence are conjugate. The two permutations (123) and (132) are then also conjugate, and the identity permutation (1)(2)(3) is of course only conjugate to itself. Thus there are three conjugacy classes in S_3.

Proof. We will only show that conjugate permutations have the same cycle structure, and leave the other direction as an exercise. Consider two permutations $\pi, \sigma \in S_n$. We want to check whether $\sigma \circ \pi \circ \sigma^{-1}$ has the same cycle structure as π. If π and σ are given by the permutations:

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix},$$

then it follows that

$$\sigma \circ \pi \circ \sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ \sigma(\pi(1)) & \sigma(\pi(2)) & \dots & \sigma(\pi(n)), \end{pmatrix},$$

since σ^{-1} takes the element $\sigma(i)$ and sends it to i. Thus we can see the permutation $\sigma \circ \pi \circ \sigma^{-1}$ as applying σ to the symbols in the cycle decomposition of π. The result will generally be different from π, but its cycle structure will necessarily be the same.

The argument may become clearer with a simple example. Consider $\pi = (123)(4) \in S_4$, and $\sigma = (12)(34) \in S_4$. The corresponding permutations are

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

The conjugate permutation is then

$$\sigma \circ \pi \circ \sigma^{-1} = \begin{pmatrix} 2 & 1 & 4 & 3 \\ 1 & 4 & 2 & 3 \end{pmatrix},$$

which has cycle structure $\sigma \circ \pi \circ \sigma^{-1} = (214)(3)$. It is indeed the same cycle structure as π. In fact, it could have been obtained directly by simply applying σ to the symbols in the cycle decomposition $\pi = (123)(4)$, which gives directly $\sigma \circ \pi \circ \sigma^{-1} = (214)(3)$. ■

It follows from Theorem 1.8.15 that all one has to do to find the conjugacy classes in S_n is to construct the possible cycle structures. This in turn is equivalent to partitioning the set $\{1, \ldots, n\}$ into disjoint subsets of various lengths. What we now show is that this problem can be reformulated into the problem of finding integer partitions of the positive integer n.

Consider a permutation $\pi \in S_n$ and its cycle decomposition. Let ν_k be the number of k-cycles in its decomposition. We introduce a new notation which gives the cycle structure of the permutation (i.e. the number of cycles of each length): we write $(1^{\nu_1}, 2^{\nu_2}, \ldots, n^{\nu_n})$ to denote that the cycle decomposition of π has ν_1 cycles of length 1, ν_2 cycles of length 2, and so on.

The total number of elements in the cycle decomposition must of course be n. This means that the sum $\sum_{k=1}^{n} k\nu_k = n$. We can rewrite this sum as:

$$(\nu_1 + \nu_2 + \ldots + \nu_n) + (\nu_2 + \nu_3 + \ldots + \nu_n) + (\nu_3 + \ldots + \nu_n) + \ldots + (\nu_n) = n.$$

We then introduce the notation $\lambda_j := \sum_{k=j}^{n} \nu_k$, so that the sum above becomes

$$\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_n = n.$$

It is also clear that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$, since $\lambda_1 - \lambda_2 = \nu_1$, $\lambda_2 - \lambda_3 = \nu_2$, and so on, and the ν_i are by definition non-negative integers.

The result of this combinatorial analysis is that the cycle structure of permutations in S_n are in one-to-one correspondence with decreasing sequences of non-negative integers $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ such that $\sum_{k=1}^{n} \lambda_k = n$. Such a decreasing sequence of non-negative integers is called a **partition of** n (we usually omit the λ_k's that are zero).

To be precise, we only showed how the λ_k's can be constructed out of the cycle structure, that is out of the ν_i's. But it works the other way around as well. Given a partition of n, that is a sequence of λ_k, then we define $\nu_k = \lambda_k - \lambda_{k+1}$, with $\nu_n = \lambda_n$. ν_k then gives us the number of cycles of length k in the cycle structure associated to the partition.

Example 1.8.16 Conjugacy classes in S_4 and partitions. This construction will become clearer with an example. Let us construct the conjugacy classes in S_4, corresponding to the different cycle structures in S_4. This corresponds to different partitions of 4. The five distinct partitions of 4 are easily found to be 4, $3+1$, $2+2$, $2+1+1$ and $1+1+1+1$. What this tells us directly is that there are five distinct conjugacy classes in S_4.

For fun, let us find out what kind of cycle structures these partitions correspond to. Consider first the partition 4, which means $\lambda_1 = 4$, $\lambda_2 = \lambda_3 = \lambda_4 = 0$. To get the cycle structure, we compute $\nu_1 = \lambda_1 - \lambda_2 = 4$, $\nu_2 = \nu_3 = \nu_4 = 0$. Thus the corresponding cycle structure consists in 4 cycles of length 1. There is in fact a unique permutation of S_4 with this cycle structure, which is the identity permutation $\pi = (1)(2)(3)(4)$.

Consider the partition $3+1$. We get $\nu_1 = 2$, $\nu_2 = 1$, $\nu_3 = \nu_4 = 0$. So the cycle structure consists in 2 cycles of length 1, and one cycle of length 2. The corresponding permutations are transpositions, and there are 6 of them in S_4.

Consider the partition $2+2$. We get $\nu_1 = 0$, $\nu_2 = 2$, $\nu_3 = \nu_4 = 0$. The cycle structure is two cycles of length 2. There are 3 such permutations in S_4.

Consider $2+1+1$. We get $\nu_1 = 1$, $\nu_2 = 0$, $\nu_3 = 1$ and $\nu_4 = 0$. The cycle structure is one cycle of length 3 and one cycle of length 1, which corresponds to cyclic permutations of length 3. There are 8 such permutations in S_4.

Finally, $1+1+1+1$, which gives $\nu_1 = \nu_2 = \nu_3 = 0$, $\nu_4 = 1$. The cycle structure is a single cycle of length 4, corresponding to a cyclic permutation of length 4. There are 6 such permutations in S_4.

The result is that there are five conjugacy classes in S_4, corresponding to five possible cycle structures, corresponding to the five partitions of 4. Counting the number of elements in each conjugacy classes, we end up with $1+6+3+8+6 = 24 = 4!$, which is indeed the order of S_4. □

For those of you who like combinatorics, it is also fun to count in general the number of permutations of S_n of a given cycle structure. We will leave that as an exercise, but the result is the following. Consider a cycle structure $(1^{\nu_1}, 2^{\nu_2}, \ldots, n^{\nu_n})$ with $\sum_k k\nu_k = n$. The number of permutation in S_n with this cycle structure is precisely:

$$\frac{n!}{\prod_{j=1}^n j^{\nu_j} \nu_j!}.$$

Let us check that this formula is consistent with what we wrote above in S_4. Consider for instance the cycle structure given by one cycle of length 3 and one cycle of length 1. The number of permutations with that cycle structure is $\frac{4!}{3 \cdot 1! \cdot 1 \cdot 1!} = 8$, as written above. Similarly, the number of permutations with one cycle of length 2 and two cycles of length 1 is $\frac{4!}{2 \cdot 1! \cdot 1^2 \cdot 2!} = 6$, and so on and so forth.

Remark 1.8.17 Note that partitions of an integer n can be encoded diagrammatically into **Young diagrams**. A Young diagram is a finite collection of boxes arranged in left-justified rows, with the rows "weakly decreasing" (i.e., such that each row has the same or shorter length than the row right above it). It is then clear that Young diagrams with n boxes contain precisely the same information as integer partitions of n, with the integers in the partition given by the length of the rows of the Young diagram. Young diagrams are useful for studying representations of the symmetric group, in which case they are upgraded to **Young tableaux**.

1.9 Normal subgroups

Objectives

You should be able to:

- Prove whether a given subgroup is normal or not.
- Recall the definition of simple groups.

Let us go back to the idea of cosets. We know that the cosets are not subgroups of G, except for the one that includes the identity element. What we want to do now is define a group structure on the "set of cosets". That is, the elements of the new group will be the cosets themselves.

To get a group structure, given two cosets, we want to be able to "multiply them".

Definition 1.9.1 Product of subsets. Given any two subsets S_1 and S_2 of

G, we define the product of S_1 and S_2 as being the subset

$$S_1S_2 = \{s_1s_2 \mid s_1 \in S_1 \text{ and } s_2 \in S_2\}.$$

That is, we multiply each element of S_1 with each element of S_2. ◊

Now we want to give the set of cosets a group structure under this operation of multiplication of cosets. It turns out that this is only possible if the left cosets of H are equal to its right cosets. Or, equivalently, if all elements conjugate to elements of H are also in H. So we give such subgroups a special name:

Definition 1.9.2 Normal subgroups. A subgroup H of G is called **normal** if $Hg = gH$ for all $g \in G$. Equivalently, it is normal if $gHg^{-1} = H$ (as sets) for all $g \in G$. Such subgroups are also called **self-conjugate**, since given any element $h \in H$, all its conjugate elements are also in H. ◊

In fact, a subgroup is normal if and only if for all $h \in H$, all the elements conjugate to h are also in H. That is, $gHg^{-1} \subseteq H$ for all $g \in G$. This may seem like a slighly weaker condition than $gHg^{-1} = H$ for all $g \in G$, but in fact the latter is implied by the former for any group, as one can prove. Therefore, to prove that a subgroup is normal, it is sufficient to show that all elements conjugate to elements of H are also in H, which is often taken as the definition of a normal subgroup.

Checkpoint 1.9.3 Prove that $gHg^{-1} \subseteq H$ for all $g \in G$ implies that $gHg^{-1} = H$ for all $g \in G$.

Lemma 1.9.4 A normal subgroup is the union of conjugacy classes. *A subgroup $H \subseteq G$ is normal if and only if it is the union of conjugacy classes.*

Proof. This is also rather clear. Suppose H is normal. Pick a $h \in H$. Then, for any $g \in G$, $ghg^{-1} \in H$, and hence the conjugacy class of h is included in H. Since this must true for all $h \in H$, we conclude that H is the union of the conjugacy classes of elements in H.

Conversely, if H is the union of conjugacy classes, then we know that the elements conjugate to any element in H are also in H, and hence H is normal. ■

This gives another way of proving that a subgroup is normal: it will be if and only if it "does not break conjugacy classes", i.e. is the union of conjugacy classes.

Before we study examples of normal subgroups, let us introduce one more piece of jargon:

Definition 1.9.5 Simple groups. A group that does not have proper normal subgroups is called **simple**. ◊

Simple, isn't it? All right, let us now look at some examples!

Example 1.9.6 The alternating group is a normal subgroup of the symmetric group. Let us show explicitly that the alternating group A_3 is a normal subgroup of S_3. Recall from Definition 1.8.14 that A_3 consists of the even permutations $\pi_1 = (1)(2)(3), \pi_5 = (132)$ and $\pi_6 = (123)$. To show that it is normal, one could show that all left and right cosets agree. But let us avoid such lengthy calculations and simply argue that the elements of gHg^{-1} must be in H for all $g \in G$. So what we need to show is that all permutations in gHg^{-1} are even, in which case they are in $H = A_3$. Assume that $h \in H$ is even. If g is even, then ghg^{-1} is certainly even. If g is odd, then g^{-1} is also odd, and hence ghg^{-1} is even. Therefore H is a normal subgroup. In fact, this argument works for all alternating subgroups $A_n \subset S_n$, which are all normal. □

Example 1.9.7 Subgroups of abelian groups are always normal. All subgroups of abelian groups are necessarily normal, since $ghg^{-1} = gg^{-1}h = h$ for all $g \in G$ and $h \in H$. Equivalently, each element in an Abelian group constitutes a conjugacy class by itself, and hence all subgroups are unions of conjugacy classes. Thus all subgroups are normal. As an example, the group $H = \{1, -1\}$ under multiplication is a normal subgroup of $G = \{1, i, -1, -i\}$. □

Example 1.9.8 The centre of a group is normal. The centre $Z(G)$ of a group G is always normal. Indeed, for all $g \in G$ and $h \in Z(G)$, we have $ghg^{-1} = gg^{-1}h = h$, since h must commute with all elements of G. Equivalently, each element in the centre of a group constitutes a conjugacy class by itself, and hence the centre of a group is always a union of conjugacy classes. Therefore it is normal. □

Example 1.9.9 Normal subgroups of a direct product. Given a direct product $G_1 \times G_2$, the subgroups (e, G_2) and (G_1, e) are normal. Indeed, it is easy to check that these subgroups are self-conjugate. □

Example 1.9.10 A normal subgroup of the Rubik's cube group. Another interesting example of a normal subgroup is the subgroup C_0 of the 3×3 Rubik's cube group consisting of all moves that leave the position of every block fixed but can change the orientation of the blocks. Why is it normal? Well, if you first do an arbitrary move, than do a move in C_0 which only changes the orientiation of the blocks, and then undo the first move, you will end up with all the block in their original position, but with their orientation potentially changed. In other words, you will end up with an element of C_0. Therefore we conclude that C_0 is a self-conjugate subgroup. □

1.10 Quotient groups

Objectives

You should be able to:

- Given a normal subgroup H of a group G, construct the quotient group G/H.

By definition, for normal subgroups left and right cosets agree, so we can talk about cosets without specifying left or right. Given a normal subgroup $H \subset G$, we can then define a group structure on the set of cosets. This is called the quotient group:

Theorem 1.10.1 Quotient groups. *If H is a normal subgroup of G, then the collection of all distinct cosets of H, denoted by G/H, is a group (with the operation being product of subsets defined in Definition 1.9.1), called the* ***quotient group of*** G ***by*** H*. The order of the quotient group is the number of distinct cosets, which is given by $|G/H| = |G|/|H|$, also called the* ***index*** *of the subgroup $H \subset G$.*

Proof. Consider the collection of all distinct cosets of H in G, which we denote by G/H. We want to give it a group structure. We define group multiplication as being the multiplication of sets defined in Definition 1.9.1. Let us first show that the multiplication of two cosets yield another coset. Consider the left cosets aH and bH for $a, b \in G$. Elements of $(aH)(bH)$ have the form $ah_1bh_2 = a(h_1b)h_2$ for $h_1, h_2 \in H$. Since H is normal, we know that there exists a $h_3 \in N$ such that $h_1b = bh_3$. Thus we can write $ah_1bh_2 = (ab)h_3h_2$, which is an element of the coset $(ab)H$. Hence multiplication of sets defines a group operation $(aH)(bH) = abH$.

Then we can prove that this operation defines a group structure. The identity element is the coset $eH = H$. For any $g \in G$, the inverse of gH is the coset $g^{-1}H$, since $(gH)(g^{-1}H) = gg^{-1}H = H$. Associativity is also clear:

$$((aH)(bH))(cH) = (abH)(cH) = abcH = (aH)(bcH) = (aH)((bH)(cH)).$$

■

Remark 1.10.2 The reason for this construction to be called a quotient group comes from division of integers. Suppose you consider $12/4 = 3$. You can understand this calculation of the quotient as taking 12 objects, and dividing them into disjoint classes of 4 objects; the result is 3 disjoint classes. Here we are doing the same thing, but instead of only talking about the number of things in a set, we have a group structure, so we start with a group, partition the set into subsets, and equip the resulting partition with the structure of a group.

Example 1.10.3 Trivial examples. Note that the quotient group G/G is clearly isomorphic to the trivial group (since it has only one element, which is the identity); and $G/\langle e\rangle$ is isomorphic to G, since the cosets of the subgroup $\langle e\rangle$ are just the elements of G themselves. □

Example 1.10.4 The parity quotient group for permutations. Consider the normal subgroup $A_3 \subset S_3$. Using the notation of Definition 1.8.14, there are two cosets, consisting of the even permutations $E = \{\pi_1, \pi_5, \pi_6\}$ and the odd permutations $O = \{\pi_2, \pi_3, \pi_4\}$, so the order of the quotient group is 2. We know that there is only one abstract group of order 2, with multiplication table given in Table 1.4.3, and using the group operation on cosets we can check that indeed, $E \cdot E = O \cdot O = E$, $E \cdot O = O \cdot E = O$, since composing two even permutations or two odd permutations gives and even permutation, while composing and odd and an even permutation gives an odd permutation. □

Example 1.10.5 Finite cyclic groups as quotients. A fundamental example of quotient groups consists in the construction of finite cyclic groups as quotients. Start with the integers $\mathbb{Z}$ under addition. Let $H = m\mathbb{Z}$ be the subgroup of multiples of the positive integer m. Since $\mathbb{Z}$ is abelian, $m\mathbb{Z}$ is necessarily normal. We then construct the quotient group $\mathbb{Z}/m\mathbb{Z}$. What do elements of $\mathbb{Z}/m\mathbb{Z}$ look like? Consider any integer $k \in \mathbb{Z}$. Recalling that the group operation here is addition, the coset $k\mathbb{Z}$ generated by the integer k is $k + m\mathbb{Z}$, that is k plus mutiples of m. Since any two k_1 and k_2 that differ by a mutiple of m generate the same coset, we see that there are exactly m disjoint cosets $k + m\mathbb{Z}$, generated by $k \in \{0, 1, \ldots, m-1\}$. It follows that the quotient $\mathbb{Z}/m\mathbb{Z}$ is a finite group of order m. Product of cosets (in the sense of Definition 1.9.1)

$$(k_1 + m\mathbb{Z}) \cdot (k_2 + \mathbb{Z}) = k_1 + k_2 + m\mathbb{Z} = k + m\mathbb{Z},$$

where $k = k_1 + k_2 \pmod{m}$.

The group $\mathbb{Z}/m\mathbb{Z}$ is often denoted by $\mathbb{Z}_m$, and we can write its elements simply as $\{0, 1, \ldots, m-1\}$, each of which representing the corresponding coset, with the understanding that the group operation on $\mathbb{Z}_m$ is addition modulo m. This group if of course finite of order m and cyclic; it is generated by 1 (the other elements are just the "powers of 1", with power here meaning repeated application of the operation of addition, i.e. $2 = 1 + 1$, $3 = 1 + 1 + 1$, etc.). In fact, it can be shown that every cyclic group of order m is isomorphic to $\mathbb{Z}_m$, so it is of crucial importance. This is also why we can use the same notation

$\mathbb{Z}_m$ as we used for the group of roots of unity; those are the same abstract groups. □

Example 1.10.6 $SO(2)$ **as a quotient group.** Another interesting example is constructed similarly. Start with the real numbers $\mathbb{R}$, and consider the normal subgroup consisting of integers $\mathbb{Z}$. The cosets have the form $a + \mathbb{Z}$ for $a \in \mathbb{R}$. It is clear that two a_1 and a_2 that are related by the addition of an integer generate the same cosets, hence the cosets are indexed by $a \in [0, 1)$. The quotient group $\mathbb{R}/\mathbb{Z}$ is the group of these cosets, with operation given by adding the cosets $a_1 + a_2 + \mathbb{Z}$, with $a_1 + a_2$ being understood as addition such that if the result is greater than one, than we substract one so that the result is always between 0 and 1.

Let us now argue that $\mathbb{R}/\mathbb{Z}$ is isomorphic to the group of complex numbers of absolute value 1 under multiplication, also called $U(1)$. Consider the mapping $f : \mathbb{R}/\mathbb{Z} \to U(1)$ given by $f(a+\mathbb{Z}) = \mathrm{e}^{2\pi i a}$. Under this mapping, the group operation on cosets (addition of numbers) is mapped to multiplication of complex numbers. We have not defined group homomorphisms and isomorphisms yet (we will revisit this example in Example 1.11.5), but this mapping is an isomorphism, as it is a bijection, and it preserves the group structure. Thus we can think of $\mathbb{R}/\mathbb{Z}$ and $U(1)$ as being the same abstract group. Further, from the point of view of complex numbers of absolute value 1, we can understand $2\pi a$ as an angle, and we see that $U(1)$ is isomorphic to the group of rotations in the complex plane ($SO(2)$). The group operation on $U(1)$, multiplication of exponentials, is then mapped to composition of rotations in $SO(2)$. All in all, we obtain that the quotient group $\mathbb{R}/\mathbb{Z}$ is the same abstract group as the group of two-dimensional rotations $SO(2)$! □

1.11 Homomorphisms and isomorphisms

Objectives

You should be able to:

- Show when a given mapping between groups is a group homomorphism, isomorphism and automorphism.
- Find the kernel of a group homomorphism.
- Use the first isomorphism theorem to show that a given subgroup is normal.

Let us define what we mean more precisely when we say that two groups are "the same" as abstract groups. We want to define the notion of isomorphisms for groups. But to start we define homomorphisms. In a few words, those are maps between groups that preserve the group structure. Isomorphisms are then defined as homomorphisms that are bijective.

1.11.1 Definitions and examples

Definition 1.11.1 Group homomorphism. Let $(G, \star)$ and $(H, \cdot)$ be groups, with corresponding group operations denoted by $*$ and $\cdot$ respectively. A **group homomorphism** $f : G \to H$ is a map that preserves the group structure, that is

$$f(a \star b) = f(a) \cdot f(b) \qquad \text{for all } a, b \in G.$$

◊

In other words, a group homomorphism is a map between sets that is consistent with their group structures.

Example 1.11.2 Trivial homomorphism. The most simple example of a homomorphism is called the **trivial homomorphism**. Take any group $(G, \star)$, and let H be the group consisting of a single element $\{e\}$. Then the trivial homomorphism is given by the mapping $f(a) = e$ for all $a \in G$, which trivially preserves the group structure. For any G with order > 1, it is of course not an isomorphism, since the mapping is not one-to-one (all elements of G are mapped into $e \in H$). □

While a group homomorphism is a mapping between groups that is consistent with their group structures, in general it may "lose information", if many group elements are mapped to the same group element. This is clear by looking at the trivial group homomorphism, which maps all groups to the trivial group with a single element. So we want to define a group isomorphism as an homomorphism that "preserves all information about the group."

Definition 1.11.3 Group isomorphism and automorphism. An **isomorphism** is a group homomorphism that is bijective. We say that two groups are **isomorphic**, denoted by $G \simeq H$, if there is an isomorphism $f : G \to H$. An isomorphism of a group into itself is called an **automorphism**. ◊

Example 1.11.4 The exponential map. An interesting example of an isomorphism is given by the exponential map, $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$. The notation $(\mathbb{R}, +)$ denotes the group of real numbers under addition, while $(\mathbb{R}_{>0}, \cdot)$ denotes the group of positive real numbers under multiplication. The mapping exp takes any real number and outputs its exponential, which is a positive real number. This mapping preserves the group structure (the operation in the first group is addition, while in the second group it is multiplication). What this means is that the exponential of the sum of two real numbers should be equal to the product of the exponential of these real numbers. Indeed:

$$\exp(a + b) = \exp(a) \cdot \exp(b) \qquad \text{for all } a, b \in \mathbb{R}.$$

Thus exp is a group homomorphism. Moreover, from standard properties of the real exponential it is easy to see that the homomorphism is bijective, thus it is an isomorphism. □

Example 1.11.5 $\mathbb{R}/\mathbb{Z}$ and $U(1)$. A similar example was studied in the previous section (see Example 1.10.6). We considered the quotient group $\mathbb{R}/\mathbb{Z}$, and define a mapping $f : \mathbb{R}/\mathbb{Z} \to U(1)$ by $f(a + \mathbb{Z}) = \mathrm{e}^{2\pi i a}$. As in the previous example, it follows from properties of the exponential that this mapping preserves the group structure (being coset multiplication in $\mathbb{R}/\mathbb{Z}$ and multiplication of exponentials in $U(1)$), and hence it is a group homomorphism. Moreover, one can see that it is bijective, therefore it is a group isomorphism. This is what we meant when we said that the two groups were "the same as abstract groups". □

1.11.2 First isomorphism theorem

Given an homomorphism $f : G \to H$, we can define its kernel as being the subset of elements of G that are mapped to the identity element in H. It is straightforward to show that it is a subgroup of G:

Definition 1.11.6 The kernel of an homomorphism. Let G and H be groups, and $f : G \to H$ a homomorphism. The **kernel** of f is

$$\ker(f) = \{g \in G | f(g) = e \in H\}.$$

One can show that $\ker(f)$ is a subgroup of G. ♢

Checkpoint 1.11.7 Prove that $\ker(f)$ is a subgroup of G.

In fact, one of the important theorems in group theory is known as the **first isomorphism theorem** (also sometimes called the fundamental theorem of homomorphisms). It is quite fundamental, as it relates the structure of the kernel and image of any group homomorphism. The statement of the theorem goes as follows:

Theorem 1.11.8 First isomorphism theorem. *Let $f : G \to H$ be a group homomorphism. Then $\ker(f)$ is a normal subgroup of G, and the image $f(G)$ is isomorphic to the quotient group $G/\ker(f)$.*

Proof. We know that $\ker(f)$ is a subgroup of G. Let us show that it is normal. Let $x \in \ker(f)$, and pick any $g \in G$. We want to show that $gxg^{-1} \in \ker(f)$. Since $f : G \to H$ is a group homomorphism, we have:

$$f(gxg^{-1}) = f(g)f(x)f(g^{-1}) = f(g)e_H f(g^{-1}) = f(g)f(g^{-1}) = f(gg^{-1}) = f(e_G) = e_H,$$

where e_H (resp. e_G) denotes the identity element in H (resp. in G). Thus we conclude that gxg^{-1} is in the kernel of f, and hence $\ker(f)$ is normal.

The isomorphism ϕ between $G/\ker(f)$and $f(G)$ is simply given by sending the coset $g\ker(f)$ generated by $g \in G$ to the image $f(g) \in f(G)$ of g under the group homomorphism. ϕ is of course surjective, and it is a good exercise to show that it is also one-to-one, hence and isomorphism. ■

This theorem is quite deep. In particular, it implies the following result:

Corollary 1.11.9 Normal subgroups are kernels of group homomorphisms. *Any normal subgroup $N \subset G$ can be realized as the kernel of a group homomorphism.*

Proof. Given any group G and normal subgroup $N \subset G$, there is a natural group homomorphism $\pi : G \to G/N$, which sends each element of G to the coset to which it belongs. This homomorphism is surjective, but not one-to-one (unless N is trivial). The kernel of π consists of all elements of G that are mapped to the identity coset of N in G, which is just a copy of N itself. Thus one can identify $\ker(\pi) = N$. ■

The first isomorphism theorem (and the direct corollary above, which is also referred to as first isomorphism theorem) is useful for many reasons. In particular, it can be used to show that a given subgroup is normal, by realizing it as the kernel of a group homomorphism. It can also be used to construct quotient groups as images of group homomorphisms.

Example 1.11.10 The determinant homomorphism. Consider the general linear group $GL(n, \mathbb{C})$ and define the mapping $\det : GL(n, \mathbb{C}) \to \mathbb{C}^*$ by taking the determinant of the matrices (zero is not in the image since $GL(n, \mathbb{C})$ contains invertible matrices). The image $\mathbb{C}^*$ here is the group of non-zero complex numbers under multiplication. First, let us argue that det is a group homomorphism. That is, we must show that it respects the group structure (in the first group the operation is matrix multiplication, in the second group it is product of complex numbers). For any two matrices $A, B \in GL(n, \mathbb{C})$, $\det(AB) = \det(A)\det(B)$, therefore det is a group homomorphism.

Since the identity element in the multiplicative group $\mathbb{C}^*$ is 1, the kernel of det is given by the subgroup of $GL(n, \mathbb{C})$ consisting of matrices with unit determinant, that is $SL(n, \mathbb{C})$. The first isomorphism theorem then implies that $SL(n, \mathbb{C})$ is a normal subgroup of $GL(n, \mathbb{C})$. □

We have seen another example of group homomorphism when we discussed parity of permutations in S_n and the definition of the alternating group A_n. In this case the first isomorphism theorem can be used to show directly that A_n is always a normal subgroup of S_n.

Example 1.11.11 The parity homomorphism for S_n. Consider the group $H = \{1, -1\}$ under multiplication. Consider the mapping $\epsilon : S_n \to H$ that assigns 1 to even permutations and -1 to odd permutations. Let us first show that ϵ is a group homomorphism.

For any permutations $\pi_1, \pi_2 \in S_n$, we need to show that $\epsilon(\pi_1 \circ \pi_2) = \epsilon(\pi_1)\epsilon(\pi_2)$. What this means is that we must show that the composition of two even or odd permutations is even, while the composition of an even and an odd permutation is odd. But this is clear by definition of parity. If r_1 (resp. r_2) is the number of transpositions in the decomposition of π_1 (resp. π_2), then the number of transpositions in the decomposition of $\pi_1 \circ \pi_2$ is $r_1 + r_2$, and thus if r_1 and r_2 are both even or odd, then $r_1 + r_2$ is even, while if r_1 is even and r_2 odd, or vice-versa, then $r_1 + r_2$ is odd.

Then ϵ is a group homomorphism, and its kernel is the alternating group A_n consisting of even permutations. It thus follows from the first isomorphism theorem that A_n is a normal subgroup of S_n, and that the quotient group S_n/A_n is isomorphic to H. □

1.12 Fun stuff

Let me end this section on finite groups with three interesting problems to think about.

1. Can you prove that the binomial coefficient

$$\binom{m+n}{m} = \frac{(m+n)!}{m!n!}$$

is always an integer for any positive integers $m, n \in \mathbb{Z}$ using group theory?

2. The set of sensible orientations of a rectangular mattress on a bed forms a group. What group is it? What strategy can you take to rotate periodically your mattress between all its orientations?

3. For more fun stuff have a look at the Futurama theorem... It's really fun! See for instance https://en.wikipedia.org/wiki/The_Prisoner_of_Benda and https://theinfosphere.org/Futurama_theorem, as well as https://arxiv.org/abs/1608.04809 for an interesting generalization.

Chapter 2

Representation theory

What is representation theory? So far we have studied groups as abstract objects, defined by a set and some operation on it satisfying a bunch of axioms. Abstraction is beautiful, but sometimes it is useful to make things more concrete. From a mathematical point of view, this is what representation theory does for you.

The idea is simple. In fact we have already encountered it in the previous section. A representation of a group is basically a concrete realization of a group as acting on something. In this course, we will mostly deal with matrix representations; that is, we will realize the elements of an abstract group as matrices. More formally, we will define a representation as a group homomorphism from an abstract group to a subgroup of $GL(V)$ for some finite-dimensional (complex) vector space V. After choosing a basis on V, we can represent the group elements by $n \times n$ matrices. We will call the dimension n of the vector space the dimension of the representation.

In other words, what we are doing is shifting viewpoint. One can think of group theory as studying the group of symmetries of a given object, or theory, or what not. Instead, we now start with an abstract group, and ask the question: on what kind of objects can this group act on? Representation theory is concerned with studying all possible ways that a given group can act.

But representation theory is also more than just the study of a question of mathematical interest. In fact, in many ways it is the essence of group theory. Groups usually arise because they act on something. They are symmetries of some object, or some physical theories. Thus the elements of the group are really realized concretely, in a given context, as things that operate on some vector space, and after choosing a basis, they can be represented as matrices. So in many contexts, especially in physics, we encounter groups in terms of their representations.

In fact, many physicists think of some groups as their representations. When physicists think of the group $SO(3)$ of rotations in three dimensions, they will often think of the corresponding 3×3 rotation matrices. But $SO(3)$ itself, as a group, is an abstract entity; a mathematician would think of rotations as abstract elements of $SO(3)$, with the abstract group operation of $SO(3)$. The 3×3 matrices form in fact a three-dimensional representation of $SO(3)$ (but there are in fact many more representations of $SO(3)$!)

In any case, in physics groups often arise as symmetry groups of a particular system. So one has a particular physical system, and a group of symmetry acts on it to leave the physical observables invariant. However, the mathematical objects describing the physical theory (or the space of solutions) may not be invariant; for instance, while the observables in quantum field theory may be invariant, the fields themselves will not be. Thus they must transform in a certain way. This is where representation theory comes in; they will transform in some representation of the group of symmetry. This is why group theory and representation theory is so fundamental in modern physics!

Let me give a slightly more concrete example of this to end this introduction. In quantum mechanics, we may consider a Hamiltonian that is invariant under a group of transformations. A very important question then is to find how the solutions of the Schrodinger equation corresponding to this particular Hamiltonian behave under these symmetry transformations. This is answered by representation theory; we can think of the group of symmetries as linear transformations on the vector space of solutions, which is precisely the notion of a representation of the group. This can then be used to classify the eigenfunctions of the Hamiltonian according to how they transform under the symmetry group. Very powerful! So let us study representation theory!

2.1 Representations

Objectives

You should be able to:

- Recall the definition of a representation.
- Construct explicitly a few representations of elementary finite groups.
- Recognize when a representation is faithful.

2.1.1 Definition

Let us start by defining a group representation, and then study a bunch of examples. We will focus here on representations on finite-dimensional complex vector spaces (but we could easily upgrade vector spaces to Hilbert spaces, and we could also consider vector spaces over other fields than $\mathbb{C}$). After choosing a basis, we can think of those as matrix representations.

Definition 2.1.1 Representation of a group. Let G be a group. A **representation of** G is a group homomorphism $T : G \to GL(V)$ (see Definition 1.11.3 for the definition of a group homomorphism), where V is a finite-dimensional complex vector space of dimension n. We call n the **dimension** of the representation.

After choosing a basis on V, we can think of the representation as being given by $n \times n$ complex invertible matrices $T(g)$ for each group element $g \in G$. ♢

In other words, a representation is a mapping that takes group elements into $n \times n$ complex invertible matrices, and this mapping is such that it preserves the group structure. Just as when we discussed homomorphisms vs isomorphisms, we want to distinguish between representations that keep the whole group structure intact, and those that lose some information:

Definition 2.1.2 Faithful representations. If T is one-to-one (that is, if it is a group isomorphism on its image), then we say that it is a **faithful** representation. We say that it is **unfaithful** otherwise. ♢

Now we can ask many questions about groups. Do every group have a representation (aside from the trivial representation)? If so, how many does it have? What are their dimensions? How do we characterize representations, and how do we distinguish between them? This is the essence of representation theory. In other words, we want to characterize and classify all possible ways that a given group can act. That is, we want to understand its representations.

2.1.2 Examples

But before we study these questions, let us look at a few examples of representations.

Example 2.1.3 Identity representation. We start with the most boring example. For any group G, consider the mapping $G \to GL(V)$, with $V = \mathbb{C}$, given by sending $g \mapsto 1$ for all $g \in G$. This is certainly a group homomorphism, but it is a rather boring one. It exists for any group G. It is called the **identity representation** (or trivial representation). For any non-trivial G, it is a dimension one, unfaithful, representation. □

Example 2.1.4 The permutation representation of the symmetric group S_n. Let us look at a more interesting example. Consider the symmetric group S_n, which can be understood as the group of permutations of n objects. Let us focus on S_3 for simplicity. We can think of the three objects as vectors $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Then a permutation can be represented as a 3×3 matrix acting on the vector space spanned by these vectors. That is, we can define a representation $T : S_3 \to GL(V)$ with V the three-dimensional vector space spanned by the vectors $\{v_1, v_2, v_3\}$. For instance, the cyclic permutation $\pi = (123)$ can be represented by the matrix:

$$T(\pi) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

since by matrix multiplication $T(\pi)v_1 = v_2$, $T(\pi)v_2 = v_3$ and $T(\pi)v_3 = v_1$. We could do the same thing for the other permutations in S_3: this gives a three-dimensional representation of S_3. In this way, for S_n we would construct an n-dimensional representation. This is a faithful representation, and is sometimes called the **permutation representation of S_n**. □

Example 2.1.5 The regular representation of a finite group. Using the previous example we can easily construct a faithful representation for all finite groups G. We know from Cayley's theorem that every finite group G of order n is isomorphic to a subgroup of S_n. Using the construction of the n-dimensional representation of S_n above, this gives us for free an n-dimensional representation of G, by keeping only the subset of matrices corresponding to the subgroup isomorphic to G. This is called the **regular representation** of G; its dimension is the order of G.

For example, one can think of the group of order three $\mathbb{Z}_3$ as a subgroup of S_3. It is easy to see that $\mathbb{Z}_3$ is isomorphic to the subgroup of S_3 given by the cyclic permutations $\{e, (123), (132)\}$. Thus, if we denote the elements of $\mathbb{Z}_3$ by $\{e, \omega, \omega^2\}$, we can write a three-dimensional representation as

$$T(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad T(\omega) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad T(\omega^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

You can check that these matrices satisfy the multiplication rules in $\mathbb{Z}_3$. □

Example 2.1.6 Other representations for S_n. Going back to S_n, we can construct other representations of different dimensions. For instance, there is a one-dimensional representation $T : S_n \to GL(V)$ with $V = \mathbb{C}$ given by assigning 1 if $g \in S_n$ is even, and -1 if it is odd. We have already seen that this is a group homomorphism. This is an unfaithful representation.

As a further example, one can construct a two-dimensional representation for S_3. The idea is to think of symmetries of an equilateral triangle. The group of symmetries is the dihedral group of order 6, that is D_3, which turns out to be isomorphic to S_3, as you can check. One can then write two-dimensional matrices that represent the symmetries of the triangle in the plane:

$$T(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(a) = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad T(b) = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$
$$T(c) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(d) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \qquad T(f) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}.$$

□

Example 2.1.7 Two-dimensional representation of the quaternion group. The quaternion group Q is defined by the set $Q = \{1, -1, i, -i, j, -j, k, -k\}$ with multiplication rules $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$. One can show that the following 2×2 complex-valued matrices (i here is the imaginary number) form a faithful representation of Q:

$$T(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(-1) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad T(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$
$$T(-i) = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \qquad T(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad T(-j) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$
$$T(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad T(-k) = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix},$$

□

Example 2.1.8 One-dimensional representations of $\mathbb{Z}_n$. In Example 1.1.9, we defined the cyclic group $\mathbb{Z}_n$ as the set of n'th roots of unity under multiplication. This is in fact a one-dimensional representation of the abstract group $\mathbb{Z}_n$. Indeed, writing the group elements as $a_k = e^{2\pi i k/n}$ for $k \in \{0, 1, \dots, n-1\}$ is in fact a faithful one-dimensional representation of $\mathbb{Z}_n$. But there are other one-dimensional representations. One could write $a_k = e^{2\pi i k \ell/n}$ for some fixed $\ell \in \{0, 1, \dots, n-1\}$; this gives in total n different one-dimensional representations for $\mathbb{Z}_n$. We recover the original one for $\ell = 1$.

Are these all faithful? Certainly not. Setting $\ell = 0$ gives the trivial representation $a_k = 1$ for all $k \in \{0, 1, \dots, n-1\}$, which is clearly not faithful. In general, the representation will be faithful if ℓ and n are coprime. Can you check that?

For instance, for $\mathbb{Z}_3$, if we denote the third root of unity $\omega = e^{2\pi i/3}$, we get three one-dimensional representations, namely $\{1, 1, 1\}, \{1, \omega, \omega^2\}$ and $\{1, \omega^2, \omega\}$. The last two are faithful.

For $\mathbb{Z}_4$, if we denote the fourth root of unity $\alpha = e^{2\pi i/4}$, we get four one-dimensional representations, namely $\{1, 1, 1, 1\}, \{1, \alpha, \alpha^2, \alpha^3\} = \{1, i, -1, -i\}, \{1, \alpha^2, 1, \alpha^2\} = \{1, -1, 1, -1\}$ and $\{1, \alpha^3, \alpha^2, \alpha\} = \{1, -i, -1, i\}$. The first and third ones are not faithful, while the second and fourth ones are. □

Example 2.1.9 Representations of $(\mathbb{R}, +)$. Consider now the group of real numbers $\mathbb{R}$ under addition. To build a representation for this group, we need to rewrite addition as matrix multiplication. How can we do that?

One simple way is to use the exponential map. For any $u \in \mathbb{R}$, we define the one-dimensional matrix $D(u) = e^u$. Then $D(u+v) = D(u)D(v)$, hence it is a group homomorphism. In fact, it is an isomorphism, so it gives a faithful one-dimensional representation.

We can also build a two-dimensional representation easily. Consider the matrices

$$D(u) = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}.$$

Then $D(u+v) = D(u)D(v)$, so it is a group homomorphism. Again, it is one-to-one, so it gives a two-dimensional faithful representation. □

2.2 Properties of representations

Objectives

You should be able to:

- Recall the definition of equivalent representations.
- Recall the definition of reducible and irreducible representations.
- Construct the restriction of a representation of a group G to a given subgroup $H \subset G$.
- Construct the tensor product of two given matrix representations.

2.2.1 Equivalent representations

We defined representations as group homomorphisms between abstract groups G and operators on a vector space V. After choosing a basis, we can think of those as invertible matrices. So if we think of representations in terms of matrices, we want to make sure that we only study "different" representations. What we mean here is that we do not want to distinguish between matrix representations that come from the same operators, but in different choices of bases.

Recall from linear algebra that if you are given a vector equation $\mathbf{b} = A\mathbf{a}$, and that you want to rewrite this equation in a different basis which is obtained by applying the transformation B, we get:

$$B\mathbf{b} = BA\mathbf{a} = BAB^{-1}B\mathbf{a}.$$

So, in this new basis with $\mathbf{b}' = B\mathbf{b}$ and $\mathbf{a}' = B\mathbf{a}$, we get the equation

$$\mathbf{b}' = A'\mathbf{a}',$$

with $A' = BAB^{-1}$. This is what is called a similarity transformation. So we want to say that two representations are the same if they are related by a similarity transformation.

Definition 2.2.1 Equivalent representations. We say that two representations $T : G \to GL(V)$ and $T' : G \to GL(V)$ are **equivalent** if, given a choice of basis for V, all matrices $T(g)$ and $T'(g)$ are related by a similarity transformation:

$$T'(g) = BT(g)B^{-1},$$

for some invertible matrix B. ♢

In other words, one can think of equivalent representations as mapping the group elements to the same operators on V, but in a different choice of basis.

Note that any two distinct one-dimensional representations cannot be equivalent, since complex numbers commute.

2.2.2 Irreducible and reducible representations

Our goal is to understand and classify representations of groups in general. But not all representations are equally interesting. We want to define a notion of representations that are "minimal", in the sense that they are somehow the building block for constructing all representations of a given group. "Minimal" representations should be representations that have no "sub-structure". Let us first define the concept of subrepresentation.

Definition 2.2.2 Subrepresentation. Let $T : G \to GL(V)$ be a representation of a finite group G. A vector subspace $U \subset V$ is said to be T**-invariant** if $T(g)u \in U$ for all $u \in U$. The restriction of T to the subspace U is called a **subrepresentation**. ♢

Thus, if a representation has a subrepresentation, this means that there exists a subspace $U \subset V$ that is closed under the action of the operators in the image of the representation. So it has some sub-structure. We want our "minimal" representations, our building block, to have no such sub-structure.

Definition 2.2.3 Reducible and irreducible representations. If a representation $T : G \to GL(V)$ has a non-trivial subrepresentation, then it is said to be **reducible**. It is otherwise **irreducible**. ♢

Irreducible representations are the minimal representations that we are looking for. They are the true building blocks of representation theory. In fact, a central goal of representation theory is to establish criteria to determine whether a given representation is irreducible or not, and to classify all possible irreducible representations of a given group.

By the way in the beginning of this section we said that we would focus on finite-dimensional representations, but that representations could also be defined for infinite-dimensional vector spaces. However, there is a theorem that says that:

Theorem 2.2.4 Irreducible representations of a finite group. *All irreducible representations of a finite group are finite-dimensional.*

So as far as irreducible representations of finite groups are concerned, it is sufficient to consider only finite-dimensional representations.

2.2.3 Semisimple representations

Let us now go back to the study of general representations. There is an easy way to construct higher-dimensional representations from lower-dimensional ones. Consider the following example. Suppose that I ask you to give me a 6-dimensional representation of, say, S_3. Well, you could say, sure, I'll start with the trivial representation that assigns 1 to all permutations in S_3, and I'll assign the 6×6 matrix:

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

to all permutations in S_3. That is certainly a 6-dimensional representation, but it's a rather boring one.

You could be a little fancier, and start instead with the 3-dimensional representation of S_3 that we constructed in Example 2.1.4. If we let $T(g)$ be the 3×3 matrices associated to the group elements $g \in S_3$ in this representation, you could construct a 6-dimensional representation by constructing the block diagonal matrices:

$$\begin{pmatrix} T(g) & 0 \\ 0 & T(g) \end{pmatrix}$$

Those are examples of direct sums of representations.

Definition 2.2.5 Direct sum of representations. Suppose that $T : G \to GL(V)$ and $S : G \to GL(W)$ are m- and n-dimensional representations of a group G. We construct the $(m+n)$-dimensional **direct sum** representation $T \oplus S : G \to GL(V \oplus W)$ by forming block diagonal matrices, with the m-dimensional matrices $T(g)$ in the upper left block and the n-dimensional matrices $S(g)$ in the bottom right block:

$$T(g) \oplus S(g) = \begin{pmatrix} T(g) & 0 \\ 0 & S(g) \end{pmatrix}.$$

◊

Representations that are constructed as direct sums are certainly not minimal. More precisely, with our definition of minimality above, we can see that all direct sum representations are reducible. Indeed, consider a direct sum representation $T \oplus S : G \to GL(V \oplus W)$. Then the subspaces $V \subset V \oplus W$ and $W \subset V \oplus W$ are certainly T-invariant, as all matrices $T(g)$ are block diagonal, and hence do not mix elements of V with elements of W. Thus the restrictions of $T \oplus S$ to the subspaces V and W are subrepresentations, and hence $T \oplus S$ is reducible.

So we have seen that all direct sum representations are reducible. But this raises an interesting question: is the converse true? Can all reducible representations be written as direct sums of irreducible representations? In other words, can all reducible representations be written in block diagonal form? We will see shortly that the answer is yes for finite groups, but no in general. Thus it makes sense to distinguish between representations that can or cannot be written as direct sums of irreducible representations.

Definition 2.2.6 Semisimple representations. A representation is **semisimple** (also called completely reducible) if it is a direct sum of irreducible representations. ◊

With this definition, the question becomes: are all representations of a given group semisimple? (i.e. either irreducible or direct sums of irreducible representations.) We will come back to this shortly.

Remark 2.2.7 A consequence of the direct sum construction is that for any group, there exists an infinite number of reducible representations, which we can construct as direct sums. This is why, from a classification viewpoint, it is much more interesting to focus on irreducible representations.

However, in physics one often encounters representations that are reducible (in fact semisimple). So it is also interesting to develop methods to find the decomposition of a semisimple representation as a direct sum of irreducible representations. For instance, in the case of a quantum mechanical Hamiltonian H with a symmetry group G, we know that solutions will transform according to some semisimple representation T of G. In other words, we can decompose the space of solutions as a direct sum of invariant subspaces according to the decomposition of the semisimple representation as a direct sum of irreducible representations. As we will see later on, all states in a given invariant subspace must be eigenstates of the Hamiltonian with the same energy eigenvalue. Thus, if we are interested in the energy levels of our system, what we first want to do is decompose the space of states into invariant subspaces, that is, decompose the representation T into irreducible representations of the symmetry group G.

2.2.4 Restriction to subgroups

Reducible representations also arise naturally when one considers a subgroup H of a group G. A representation of G clearly gives a representation of H if one keeps only the matrices corresponding to the elements of the subgroup $H \subset G$. However, when restricted to a subgroup $H \subset G$, a representation that is irreducible for G may become reducible for H. Indeed, the induced representation for H may have a subrepresentation (suppose for example that all matrices corresponding to the subgroup are block diagonal), while the original representation for G may not (suppose that the matrices corresponding to elements of G not in H are not block diagonal). A central theme in physics is to find how a given irreducible representation of G decomposes into a direct sum of irreducible representations of H upon restriction to a subgroup $H \subset G$.

Example 2.2.8 Restriction to $\mathbb{Z}_2 \subset S_3$. As an example of this, let us look back at the two-dimensional representation $T : S_3 \to GL(V)$ that we constructed in Example 2.1.6. We reproduce it here for convenience:

$$T(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(a) = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad T(b) = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

$$T(c) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(d) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \qquad T(f) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}.$$

One can show that this representation is irreducible. Let us now consider the order two subgroup $H = \{e, c\} \subset S_3$. The restriction of T to the subgroup $H \simeq \mathbb{Z}_2$ gives the two-dimensional representation:

$$T(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(c) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is now reducible (in fact semisimple), as a representation of H, since it is in block diagonal form. It is the direct sum of the trivial representation of $\mathbb{Z}_2$ and its non-trivial one-dimensional representation. □

2.2.5 Tensor product representations

We have already seen one way of constructing higher-dimensional representations from lower-dimensional ones. Given two representations $T : G \to GL(V)$ and $S : G \to GL(V')$ of the same group, we have already seen how to construct a $(m+n)$-dimensional representation on $V \oplus V'$ by adjoining the matrices into $(m+n) \times (m+n)$ block diagonal matrices. We called this construction the direct sum, and denoted it by $T \oplus S$.

There is another way to construct higher-dimensional representations from lower-dimensional ones. We can also take the "product" of representations to construct a new mn-dimensional representation: this is called the tensor product representation, and we denote it by $T \otimes S$. What we are really doing here is define a new representation $T \otimes S$ that is an operator on the tensor product of the underlying vector spaces. If you know what this means, great, otherwise I will simply be pedestrian here and construct the matrix representation $T \otimes S$ explicitly.

Let us start with the simpler case where T and S are two-dimensional representations, with matrices:

$$T(g) = \begin{pmatrix} T_{11}(g) & T_{12}(g) \\ T_{21}(g) & T_{22}(g) \end{pmatrix}, \qquad S(g) = \begin{pmatrix} S_{11}(g) & S_{12}(g) \\ S_{21}(g) & S_{22}(g) \end{pmatrix}.$$

Then the 4×4 matrices $T \otimes S(g)$ are constructed by taking the product of all matrix entries (this is not a matrix product of course):

$$\begin{aligned} T \otimes S(g) &= \begin{pmatrix} T_{11}(g)S(g) & T_{12}(g)S(g) \\ T_{21}(g)S(g) & T_{22}(g)S(g) \end{pmatrix} \\ &= \begin{pmatrix} T_{11}(g)S_{11}(g) & T_{11}(g)S_{12}(g) & T_{12}(g)S_{11}(g) & T_{12}(g)S_{12}(g) \\ T_{11}(g)S_{21}(g) & T_{11}(g)S_{22}(g) & T_{12}(g)S_{21}(g) & T_{12}(g)S_{22}(g) \\ T_{21}(g)S_{11}(g) & T_{21}(g)S_{12}(g) & T_{22}(g)S_{11}(g) & T_{22}(g)S_{12}(g) \\ T_{21}(g)S_{21}(g) & T_{21}(g)S_{22}(g) & T_{22}(g)S_{21}(g) & T_{22}(g)S_{22}(g) \end{pmatrix}. \end{aligned}$$

It is easy to see how this generalizes to arbitrary m and n. The $mn \times mn$ tensor product matrices are constructed as:

$$T \otimes S(g) = \begin{pmatrix} T_{11}(g)S(g) & \dots & T_{1n}(g)S(g) \\ \vdots & \ddots & \vdots \\ T_{n1}(g)S(g) & \dots & T_{nn}(g)S(g) \end{pmatrix},$$

where of course one could expand out by writing down the matrices $S(g)$ explicitly as above.

We call the corresponding representation the **tensor product representation**. Such tensor product representations occur frequently in physics, for instance when taking the product of wave-functions in quantum mechanics.

Checkpoint 2.2.9 Check that the tensor product representation constructed above is indeed a representation, i.e. that $T \otimes S : G \to GL(mn, \mathbb{C})$ is a group homomorphism.

One important point is that even if T and S are irreducible representations of G, the tensor product $T \otimes S$ may be reducible. In fact often in particle physics one needs to determine how the tensor product of two irreducible representations decomposes into a sum of irreducible representations.

2.3 Unitary representations

Objectives

You should be able to:

- Determine whether an explicitly given representation is unitary.
- Recall and state the unitarity theorem for finite groups.

One type of representations is particularly important in physics: unitary representations. Those are widely used in quantum mechanics. So studying unitary representations in general allows physicists to learn lots of interesting things about quantum systems. You can look at these slides by Peter Woit if you want to know more about this, but we will come back to this later on in this course.

Mathematically, unitary representations are also generally easier to handle than general representations: they satisfy nice properties not shared by general representations. It turns out that for finite groups, all representations are equivalent to unitary representations. This is not so simple for infinite groups though.

2.3.1 Review of linear algebra

Let us start by reviewing a bit of linear algebra.

Definition 2.3.1 Adjoint of a matrix. Let $A \in GL(n, \mathbb{C})$ be an $n \times n$ matrix with complex-valued entries. We write A^* for its complex conjugate, obtained by complex conjugating the entries. We write $A^\dagger = (A^*)^T$ for the transpose of the complex conjugated matrix. $A^\dagger$ is called the **adjoint** (or Hermitian conjugate). ♢

Remark 2.3.2 Note that in the literature, particularly in the mathematics literature, the notation A^* is often use to denote the adjoint $A^\dagger$.

Definition 2.3.3 Hermitian matrices. A matrix $A \in GL(n, \mathbb{C})$ is **Hermitian** if $A = A^\dagger$. ♢

Hermitian matrices are also all over the place in quantum mechanics, since observables are generally understood in terms of Hermitian operators on the Hilbert space of states. The eigenvalues of these operators correspond to the possible values of the particular observable variables represented by the operators. This interpretation makes sense because of the following fundamental result in linear algebra:

Theorem 2.3.4 Eigenvalues of Hermitian matrices. *All eigenvalues of a Hermitian matrix $A \in GL(n, \mathbb{C})$ are real, and A has n linearly independent eigenvectors that can all be chosen to be orthogonal.*

Another type of complex-valued matrices is important:

Definition 2.3.5 Unitary matrices. We say that a matrix $A \in GL(n, \mathbb{C})$ is **unitary** if $AA^\dagger = I$, that is, $A^\dagger = A^{-1}$. ♢

Unitary matrices preserve the inner product on a complex vector space, just like orthogonal matrices preserve the scalar product on real vector spaces. Given two vectors $u, v \in \mathbb{C}^n$, the standard inner product is defined as $\langle u, v\rangle = u^\dagger v$. Then a unitary transformation A preserves the inner product, since

$$\langle Au, Av\rangle = u^\dagger A^\dagger A v = u^\dagger v = \langle u, v\rangle.$$

This is why unitary operators are important in quantum mechanics: they can be used to do changes of bases while preserving orthogonality between basis vectors according to the inner product. They behave just like rotations but in a complex vector space.

Another nice property of unitary matrices concerns their eigenvalues. While eigenvalues of Hermitian matrices are real, eigenvalues of unitary matrices are generally complex, but of a very particular form: they must lie on the unit circle. In other words, they must have modulus 1. This means that they can be written as $e^{i\theta}$ for some angle $\theta \in [0, 2\pi)$.

It also turns out that unitary and Hermitian matrices are closely related. Indeed, the theorem about eigenvalues of Hermitian matrices above relies on the fundamental statement that any Hermitian matrix can be diagonalized by an appropriate unitary transformation. That is, for any Hermitian matrix H, there exists a unitary matrix U such that

$$D = U^\dagger H U$$

is a diagonal matrix.

Checkpoint 2.3.6 Prove that any Hermitian matrix can be diagonalized by an appropriate unitary transformation.

2.3.2 The unitarity theorem

Let us now go back to representations.

Definition 2.3.7 Unitary representation. If all matrices $T(g)$ of a representation of a group G are unitary, we say that the representation is **unitary**. $\Diamond$

Unitary representations may seem like very special and constrained. Which raises a question: how common are unitary representations?

Let us consider first one-dimensional representations of a finite group G. To any element $g \in G$, a one-dimensional representation T associates a complex number $T(g) = re^{i\theta}$. (This is the polar form of a non-zero complex number, where $r > 0$ is the modulus, and $\theta \in [0, 2\pi)$.) Since G is finite, we know that g has finite order, and hence there must exist an integer n such that $g^n = e$. That is, $T(g^n) = r^n e^{in\theta} = T(e) = 1$. This is only possible if $r = 1$. In other words, what this means is that $T(g)$ has modulus one, and hence is a one-dimensional unitary matrix, since $T^\dagger(g)T(g) = r^2 = 1$. Therefore, all one-dimensional representations of finite groups are unitary. Note that it was key in the argument that the group is finite.

Does there exist a similar statement for higher-dimensional representations of finite groups? This is the content of the important unitarity theorem, the main result that we prove in this section:

Theorem 2.3.8 Unitarity theorem. *Any finite-dimensional representation of a finite group G is equivalent to a unitary representation. In other words, it can be brought into unitary form by a similarity transformation.*

Proof. Let G be a finite group, and $T : G \to GL(V)$ be an n-dimensional representation. The proof is constructive: for any such T, we will construct an equivalent representation that is explicitly unitary.

To this end, let us first introduce the $n \times n$ matrix:

$$H = \sum_{g \in G} T^\dagger(g) T(g).$$

This matrix is interesting. For instance, it satisfies the following invariance property. For any $g' \in G$,

$$\begin{aligned} T^\dagger(g') H T(g') &= \sum_{g \in G} T^\dagger(g') T^\dagger(g) T(g) T(g') \\ &= \sum_{g \in G} (T(g)T(g'))^\dagger T(g) T(g') \\ &= \sum_{g \in G} T(gg')^\dagger T(gg') \\ &= H. \end{aligned} \tag{2.3.1}$$

The last line follows because of the Rearrangement theorem 1.4.2, since multiplying the group elements $g \in G$ by a fixed element g' only rearranges the terms in the sum. This invariance property will be useful later on.

Now notice that H is Hermitian, since $H^\dagger = H$. Thus we know that it has real eigenvalues, and that is can be diagonalized by a unitary matrix. Hence we can write $D = U^\dagger H U$, where D is a diagonal matrix with real entries and U is unitary. Further, we now show that D has real positive entries. We have:

$$\begin{aligned} D &= U^\dagger H U \\ &= \sum_{g\in G} U^\dagger T^\dagger(g) T(g) U \\ &= \sum_{g\in G} A^\dagger(g) A(g), \end{aligned}$$

where we defined $A(g) = T(g)U$. Now consider the j'th diagonal entry D_{jj}. It is given by summing over $g \in G$ the contributions given by $A_j^\dagger(g)A_j(g)$, where $A_j(g)$ denotes the j'th column vector in $A(g)$. Since for each $g \in G$ and each j, $A_j(g)$ is a non-zero vector, then $A_j^\dagger(g)A_j(g) > 0$, and hence $D_{jj} > 0$.

We then define the diagonal matrix $D^{1/2}$ whose entries are the square roots of the entries of D, and $D^{-1/2}$ as the inverse of $D^{1/2}$. We now form the matrices $B(g) = D^{1/2}U^\dagger T(g) U D^{-1/2}$, and their adjoints $B^\dagger(g) = D^{-1/2}U^\dagger T^\dagger(g) U D^{1/2}$. Why are we constructing these matrices? Note that, since U is unitary, and hence $U^\dagger = U^{-1}$, the transformation $B(g) = D^{1/2}U^\dagger T(g) U D^{-1/2}$ is a similarity transformation. Thus the representations furnished by the matrices $B(g)$ and the $T(g)$ are equivalent. Our goal is to show that the new representation $B(g)$ is explicitly unitary, which would prove the theorem, namely that any representation of a finite group is equivalent to a unitary representation.

So let us show that the matrices $B(g)$ are unitary. We have:

$$\begin{aligned} B^\dagger(g)B(g) =& D^{-1/2}U^\dagger T^\dagger(g) U D^{1/2} D^{1/2} U^\dagger T(g) U D^{-1/2} \\ =& D^{-1/2}U^\dagger T^\dagger(g) U D U^\dagger T(g) U D^{-1/2} \\ =& D^{-1/2}U^\dagger T^\dagger(g) H T(g) U D^{-1/2} \\ =& D^{-1/2}U^\dagger H U D^{-1/2} \\ =& D^{-1/2} D D^{-1/2} \\ =& I, \end{aligned}$$

where we used the invariance property (2.3.1). Thus for any finite-dimensional representation T of a finite group G, we have constructed a new, equivalent unitary representation B, given by the set of unitary matrices $B(g) = D^{1/2}U^\dagger T(g) U D^{-1/2}$. We have thus proved that all finite-dimensional representations of finite groups are equivalent to unitary representations. ■

Remark 2.3.9 Note here that the requirement of having a finite group G was crucial in the proof. Otherwise, the expression

$$H = \sum_{g\in G} T^\dagger(g) T(g)$$

doesn't even make sense, since the sum would be over an infinite-dimensional set (or a continuous space if the group is continuous).

Remark 2.3.10 In view of Remark 2.3.9, we may ask: is the unitarity theorem still true for infinite groups, either discrete or continuous? Consider for example the infinite continuous group $(\mathbb{R}, +)$, and the two-dimensional representation:

$$T(u) = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}, \qquad u \in \mathbb{R}$$

that we encountered in Example 2.1.9. We have:

$$\begin{aligned} T^\dagger(u)T(u) &= \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1+u^2 & u \\ u & 1 \end{pmatrix}, \end{aligned}$$

which is not the identity for non-zero $u \in \mathbb{R}$. So this is not a unitary transformation, and one can show that it cannot be brought into a unitary transformation by a similarity transformation. In fact, the same representation restricted to $u \in \mathbb{Z}$ is also not unitary for the infinite but discrete group $(\mathbb{Z}, +)$.

So for what kind of groups, beyond finite groups, is the unitarity theorem true? The general result is that the theorem holds if the group is **compact**. To define compact groups, we first need to define the concept of topological groups. Topological groups are groups G that are given the extra structure of a topology on G, such that the group's binary operation and function mapping elements to their inverses are continuous with respect to this topology. Then, a compact group is a topological group such that its topology is compact.

In the end, they key point is that because of the existence of this compact topology on G, for compact groups we can "replace" the sum over elements of the group by an appropriate integral over the continuous group with respect to some measure (the Haar measure), and the resulting integral then converges. Using this the proof above goes through with minor modifications, and the unitarity theorem holds for compact groups. Examples of compact groups include $SO(n)$ and $SU(n)$, which appear frequently in physics. Note however that the Lorentz group is not compact.

Note that not only the unitarity theorem holds for compact groups, but Theorem 2.2.4 also applies to compact groups: *all irreducible representations of compact groups are finite-dimensional.*

2.4 Semisimplicity

Objectives

You should be able to:

- Recall that every representation of a finite group is semisimple.

In Definition 2.2.3 and Definition 2.2.6 we introduced reducible and semisimple representations. We then asked whether all representations of a given group are semisimple. That is, whether all representations are either irreducible or can be written as direct sums of irreducible representations. With our detour study of unitary representations, we are now in a position to answer this question for finite groups.

We said earlier that unitary representations are quite nice, and satisfy properties that are not shared by all representations. An example of this is the following important theorem.

Theorem 2.4.1 Unitary representations are semisimple. *All finite-dimensional unitary representations are semisimple. That is, all finite-dimensional* *unitary representations are either irreducible or direct sums of irreducible representations.*

Proof. Let $T : G \to GL(V)$ be a unitary representation. If T is irreducible, then we are done. So let us assume that T is reducible. Then there exists a T-invariant subspace $U \subset V$. Let W be the orthogonal complement of U in V. Then $V = U \oplus W$. We now show that W is also T-invariant, and hence the representation T must leave both U and its orthogonal complement W invariant. This means that it must decompose in block diagonal form, i.e. it is a direct sum of representations.

To prove this, we formulate the T-invariance condition using projection operators. Recall a few facts from linear algebra. A projection P on a vector space V is an operator $P : V \to V$ such that $P^2 = P$. Let $U \subset V$ be the range of P, and $W \subset V$ be its kernel. Then one can think of P as projecting vectors onto the subspace U. Further, it follows that $V = U \oplus W$, and that $Q = I - P$ is also a projection, but now onto the subspace $W \subset V$. If we have a notion of inner product on V, we can define an orthogonal projection P: a projection for which the range U and the kernel W are orthogonal. A projection is orthogonal if and only if it is given by a Hermitian matrix.

Let us then introduce an orthogonal (thus Hermitian) projection matrix onto the subspace $U \subset V$, that is $P : V \to U$, with $P^2 = P$ and $P = P^\dagger$. Then the statement that U is T-invariant can be written as the condition that

$$PT(g)P = T(g)P$$

for all $g \in G$. This is because the projection operator P acts as the identity operator on the subspace $U \subset V$, and $Pv \in U$ for all $v \in V$, so the condition that $T(g)u \in U$ for all $u \in U$ will be satisfied if and only if $PT(g)Pv = T(g)Pv$ for all $v \in V$.

Take the Hermitian conjugate on both sides. We get the condition $PT^\dagger(g)P = PT^\dagger(g)$. We know that the matrices $T(g)$ are unitary, and hence $T^\dagger(g) = T^{-1}(g) = T(g^{-1})$. Thus the condition becomes $PT(g^{-1})P = PT(g^{-1})$, but since this must be true for all $g \in G$, we can write it simply as

$$PT(g)P = PT(g)$$

for all $g \in G$. Reorganizing, we get:

$$PT(g)P - PT(g) - T(g)P + T(g) = -T(g)P + T(g),$$

that is,

$$(1 - P)T(g)(1 - P) = T(g)(1 - P).$$

But the projection matrix $1 - P$ projects on the orthogonal complement W of the subspace $U \subset V$. Since $V = U \oplus W$, this means that T must decompose in block diagonal form, i.e. it is a direct sum of representations.

We continue this process inductively. Since the representation is finite-dimensional, the process must stop at some point, and we end up with T being a direct sum of irreducible representations, that is, a semisimple representation.

■

Now we can use this key result to study whether representations of finite groups can be written as direct sums of irreducible representations. Indeed, in Theorem 2.3.8 we showed that all finite-dimensional representations of finite groups are equivalent to unitary representations. Therefore Theorem 2.4.1 implies the following corollary:

Corollary 2.4.2 Finite-dimensional representations of finite groups are semisimple. *All finite-dimensional representations of finites groups are equivalent to semisimple representations. That is, they are either irreducible or equivalent to direct sums of irreducible representations.*

Thus as far as finite groups are concerned, irreducible representations really are the true building blocks, since all finite-dimensional representations can be constructed as direct sums of irreducible representations.

Remark 2.4.3 The key result here is Theorem 2.4.1, which applies to finite-dimensional unitary representations in general, regardless of whether the group is finite or infinite. Moreover, for the case of compact groups, such as $SU(n)$ and $SO(n)$, the unitarity theorem Theorem 2.3.8 also holds. Thus one can generalize Corollary 2.4.2 to compact groups: *all finite-dimensional representations of compact groups are either irreducible or equivalent to direct sums of irreducible representations.*

However, just like the unitarity theorem, this is not true in general for infinite groups. For instance, consider the two-dimensional representation of $(\mathbb{R}, +)$ given by

$$\begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}, \qquad u \in \mathbb{R}$$

This representation is not a direct sum, hence it is not semisimple, but it is reducible. Indeed, the subspace spanned by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is invariant, hence this two-dimensional representation has a one-dimensional subrepresentation, even if it is not block diagonal and cannot be brought into block diagonal form by a similarity transformation.

The outcome of the last two sections is that, as long as finite (or compact) groups are concerned, we only have to deal with direct sums of irreducible representations (up to equivalence). Phew!

2.5 Schur's lemmas

Objectives

You should be able to:

- Recall, state and prove Schur's first and second lemmas.
- Recall that all irreducible representations of finite abelian groups are one-dimensional.

The two lemmas of Schur are fundamental in representation theory. The ultimate goal of the next few sections is to develop tools to determine whether a given representation is irreducible or not. The fundamental result in this direction is called the "great orthogonality theorem". But to be able to prove this theorem, we first need to introduce two lemmas that are interesting in their own right, known as "Schur's lemmas".

Schur's lemmas are concerned with the type of matrices that commute with all matrices of irreducible representations of a finite group G. This will be important to prove the great orthogonality theorem.

2.5.1 Schur's first lemma

Schur's first lemma gives us a criterion to determine when two representations are irreducible.

Lemma 2.5.1 Schur's first lemma. *Let $T : G \to GL(V)$ and $S : G \to GL(U)$ be two irreducible representations of a finite group G. If a matrix A is such that*

$$AT(g) = S(g)A \qquad \forall g \in G,$$

then either $A = 0$, or A is an invertible square matrix and hence the two representations are equivalent.

Proof. Suppose that T is n-dimensional, and S is m-dimensional. First we note that A is a $m \times n$ matrix. Suppose that there is a subspace $W \subset V$ such that $Aw = 0$ if and only if $w \in W$. Let P be the projection matrix on the subspace $W \subset V$. Then $AP = 0$. It follows that

$$AT(g)P = S(g)AP = 0.$$

But then, this implies that $T(g)P \in W$, or, in other words, $T(g)w \in W$ for any $w \in W$. This means that W is a T-invariant subspace. Since W is irreducible, the only possible invariant subspaces are the trivial ones, namely either $W = V$ or $W = \{0\}$. In the first case, this implies that $A = 0$. In the second case, this means that Av is never zero for all non-zero $v \in V$.

We can redo the same argument by starting with a subspace $W' \in U$ such that $uA = 0$ if and only if $u \in W'$. Following the same steps as above, we end up with two possibilities; either $A = 0$ or uA is never zero for all non-zero $u \in U$.

So we are left with two cases: either $A = 0$, or Av and uA are never zero for non-zero $v \in V$ and $u \in U$. We now show that the second case implies that A is an invertible square matrix. First, if the number of rows is less than the number of columns, than there must exist a $v \in V$ such that $Av = 0$, which is a contradiction. Similarly, if the number of columns is less than the number of rows, than there must exist a $u \in U$ such that $uA = 0$, which is a contradiction. Therefore A must be square, that is, $m = n$. Finally, we know that a square matrix has a non-trivial kernel if and only its determinant is zero. Therefore A must have non-zero determinant, or equivalently, it must be invertible.

Therefore, we conclude that either $A = 0$ or A is an invertible square matrix. In the latter case, if then follows that

$$AT(g)A^{-1} = S(g),$$

that is, T and S are equivalent representations. ■

2.5.2 Schur's second lemma

The second lemma studies what kind of matrices commute with all matrices of a given irreducible representation.

Lemma 2.5.2 Schur's second lemma. *Let $T : G \to GL(V)$ be an irreducible representation of a finite group G. If a matrix A commutes with $T(g)$ for all $g \in G$, that is*

$$AT(g) = T(g)A \qquad \forall g \in G,$$

then $A = \lambda I$ for some $\lambda \in \mathbb{C}$. In other words, A is a constant multiple of the identity matrix.

This is a rather powerful statement. In general, if you were given a bunch of matrices $T_1, \ldots, T_n$, there would potentially be many matrices that commute with all of those. However, if the matrices T_i are not arbitrary but actually form an irreducible representation of a finite group, things are much more

constrained: only multiples of the identity matrix commute with the T_i.
Proof. The second lemma is a direct consequence of the first one. A must have at least one eigenvalue, which we call λ. Then $(A - \lambda I)$ has determinant zero, and hence is not invertible. But since $AT(g) = T(g)A$, it follows that:

$$(A - \lambda I)T(g) = T(g)(A - \lambda I).$$

From Schur's first lemma Lemma 2.5.1, it then follows that $A - \lambda I = 0$, since $A - \lambda I$ is not invertible. Therefore $A = \lambda I$ for some $\lambda \in \mathbb{C}$. ■

2.5.3 Irreducible representations of abelian groups

A direct consequence of Schur's second lemma is the following theorem:

Theorem 2.5.3 Irreducible representations of finite abelian groups. *All irreducible representations of finite abelian groups are one-dimensional.*
Proof. Let $T : G \to GL(V)$ be an irreducible representation of a finite abelian group G. Then for any $a \in G$, $T(a)$ commutes with all $T(g)$, $g \in G$. Thus, by Schur's second lemma, $T(a) = \lambda I$ for some $\lambda \in \mathbb{C}$. This is true for all $a \in G$, and hence all $T(a)$ are scalar multiples of the identity matrix. It then follows that the only way that T can be irreducible is if it is one-dimensional. ■

Note that as usual, while we stated the result only for finite groups, it holds whenever Schur's second lemma holds. In particular, it is true for compact groups: all complex irreducible representations of compact abelian groups are one-dimensional.

2.6 The great orthogonality theorem

Objectives

You should be able to:

- Recall, state and prove the great orthogonality theorem.
- Deduce immediate consequences of the great orthogonality theorem.

We are now in a position to prove one of the most fundamental result in representation theory: the great orthogonality theorem. In essence, what the theorem says is that if we think of irreducible representations as being given by "vectors of matrices", where the "vectors" are indexed by the group elements, then they satisfy orthogonality relations. This is key to determine the possible irreducible representations of given group.

Theorem 2.6.1 The great orthogonality theorem. *Let $T : G \to GL(V)$ and $S : G \to GL(V')$ be two inequivalent irreducible unitary representations of a finite group G. Let $T(g)_{ij}$ and $S(g)_{ij}$ denote the matrix elements of the corresponding matrices, for all $g \in G$. Then*

$$\sum_{g \in G} T^\dagger(g)_{ij} S(g)_{kl} = 0.$$

For the matrices of a single irreducible unitary representation (or two equivalent ones), the relation is:

$$\sum_{g \in G} T^\dagger(g)_{ij} T(g)_{kl} = \frac{|G|}{d} \delta_{il} \delta_{jk},$$

where d is the dimension of the representation.

Proof. This theorem follows from Schur's lemmas. Let T and S be inequivalent irreducible unitary representations of dimensions d and d' respectively. Consider an arbitrary $d \times d'$ matrix X, and let

$$A = \sum_{g\in G} T(g)XS(g^{-1}) = \sum_{g\in G} T(g)XS^\dagger(g),$$

where the second equality follows since S is unitary. Now multiply by $T(h)$ on the left for any $h \in G$:

$$\begin{aligned} T(h)A &= \sum_{g\in G} T(h)T(g)XS(g^{-1}) \\ &= \sum_{g\in G} T(h)T(g)XS(g^{-1})S(h^{-1})S(h) \\ &= \sum_{g\in G} T(hg)XS^\dagger(g)S^\dagger(h)S(h) \\ &= \sum_{g\in G} T(hg)X(S(h)S(g))^\dagger S(h) \\ &= \left(\sum_{g\in G} T(hg)XS((hg)^{-1})\right) S(h). \end{aligned}$$

By the rearrangement theorem Theorem 1.4.2, multiplying all $g \in G$ by $h \in G$ does not change the sum over G, it simply rearranges the order of the terms in the sum. Thus we get:

$$T(h)A = AS(h).$$

But then, by Schur's first lemma Lemma 2.5.1, and the fact that T and S are assumed to be inequivalent, we conclude that $A = 0$. Therefore

$$\sum_{g\in G} T(g)XS^\dagger(g) = 0.$$

Now we can choose X to be any matrix. In particular, we can choose it to be a matrix with zero entries everywhere except in the j'th row and k'th column where it is one. Then we get:

$$\sum_{g\in G} T(g)_{ij}S^\dagger(g)_{kl} = 0,$$

for all i, j, k, l, which is the first orthogonality relation.

For the second one, we start with the matrix

$$A = \sum_{g\in G} T(g)XT(g^{-1}) = \sum_{g\in G} T(g)XT^\dagger(g).$$

Going through the same steps, we end up with the statement

$$T(h)A = AT(h).$$

Schur's second lemma Lemma 2.5.2 now implies that $A = \lambda_X I$ for some complex number $\lambda_X \in \mathbb{C}$. We wrote λ_X to remind ourselves that the constant depends on the choice of matrix X in the definition of A. Thus we have

$$\sum_{g\in G} T(g)XT(g^{-1}) = \lambda_X I.$$

Next, we need to fix the constants λ_X. We take the trace on both sides of the equation. On the left-hand-side, we get:

$$\begin{aligned}\sum_{g\in G} \mathrm{Tr}\left(T(g)XT(g^{-1})\right) =& \sum_{g\in G} \mathrm{Tr}\left(XT(g^{-1})T(g)\right)\\ =& \sum_{g\in G} \mathrm{Tr}\, X\\ =& |G|\, \mathrm{Tr}\, X.\end{aligned}$$

On the right-hand-side, we get:

$$\lambda_X \,\mathrm{Tr}\, I = d\lambda_X,$$

where d is the dimension of the representation. Therefore we conclude that

$$\lambda_X = \frac{|G|}{d}\,\mathrm{Tr}\, X,$$

and the relation becomes

$$\sum_{g\in G} T(g)XT(g^{-1}) = I\frac{|G|}{d}\,\mathrm{Tr}\, X.$$

To conclude the proof, we choose the matrices X as above, with zero entries everywhere except in position (j,k), where it is taken to be 1. Using index notation, if we write the left-hand-side as $\sum_{g\in G} T(g)_{ij}T^\dagger(g)_{kl}$, we can write $I = \delta_{il}$, and $\mathrm{Tr}\, X = \delta_{jk}$, since the trace is non-zero only if the entry 1 in X is on the diagonal. Therefore, we obtain the second orthogonality relation:

$$\sum_{g\in G} T(g)_{ij}T^\dagger(g)_{kl} = \frac{|G|}{d}\delta_{il}\delta_{jk}.$$

■

The great orthogonality theorem gives orthogonality relations between the matrices of the irreducible representations of any finite group G. Those are very useful, as will become even clearer when we introduce characters. But for the moment we can already deduce an immediate consequence of the orthogonality theorem:

Theorem 2.6.2 The number of irreducible representations of a finite group. *A finite group can only have a finite number of inequivalent irreducible representations.*

Proof. This is clear from the orthogonality theorem. We can think of irreducible representations as giving "vectors of matrices" $\left(T(g)_{ij}\right)_{g\in G}$ in a vector space of dimension $|G|$. The theorem tells us that those vectors must be orthogonal. But there are at most $|G|$ orthogonal vectors in a vector space of dimension $|G|$, and so the number of irreducible representations must be finite. In fact we will calculate the number of irreducible representations for any finite group explicitly when we introduce characters. ■

2.7 Characters

Objectives

You should be able to:

- Compute the character of a representation.
- Recognize that the character is a function of class.
- Calculate the character of direct sum and tensor product representations.

In this section we introduce the notion of the character of a representation. The idea is simple: we take the traces of the matrices in a given representation. Characters will give a clean way of studying and classifying irreducible representations, and determining whether a given representation is reducible or not.

2.7.1 Definition

Definition 2.7.1 The character of a representation. Let $T : G \to GL(V)$ be a representation of G. The **character** of this representation is given by the map $\chi : G \to \mathbb{C}$, which assigns to every group element the trace of the corresponding matrix in the representation T:

$$\chi^{(T)}(g) = \operatorname{Tr} T(g).$$

If T is irreducible, the character is called **simple** (or irreducible), while it is called **compound** otherwise. The **degree** of a character is the dimension of the corresponding representation. ♢

2.7.2 A few simple properties

The first thing to note about characters is that they are the same for all elements that are in the same conjugacy class. This is usually encapsulated in the statement that "character is a function of class":

Lemma 2.7.2 Character is a function of class. *The character of a representation is a* ***class function****, meaning that all elements of a group belonging to the same conjugacy class have the same character.*
Proof. Let $x, y \in G$ be conjugate. That is, there exists a $g \in G$ such that $x = gyg^{-1}$. Since a representation is a group homomorphism, this means that $T(x) = T(g)T(y)T(g^{-1})$. Taking the trace, we get

$$\operatorname{Tr} T(x) = \operatorname{Tr} \left(T(g)T(y)T(g^{-1})\right) = \operatorname{Tr} \left(T(y)T(g^{-1})T(g)\right) = \operatorname{Tr} T(y),$$

where we used the cyclic property of the trace of a product of matrices. ■

It is also important to note that equivalent representations have the same character:

Lemma 2.7.3 Equivalent representations have the same character. *If $T : G \to GL(V)$ and $S : G \to GL(V)$ are equivalent representations, then $\chi^{(T)}(g) = \chi^{(S)}(g)$ for all $g \in G$.*
Proof. Two representation T and S are equivalent if their matrices are related by a similarity transformation $T(g) = AS(g)A^{-1}$. Taking the trace, and using cyclicity again, we get:

$$\operatorname{Tr} T(g) = \operatorname{Tr} \left(AS(g)A^{-1}\right) = \operatorname{Tr} \left(S(g)A^{-1}A\right) = \operatorname{Tr} S(g).$$

■

What this tells us is that if two representations have different characters, then they must be inequivalent. We will see that the converse statement (that two inequivalent representations must have different characters) follows from the orthogonality theorem.

Another simple property of characters is how they behave for direct sums and tensor products of representations. We will leave the proof of the following lemma as an exercise:

Lemma 2.7.4 Character of a direct sum and a tensor product. *Let $T : G \to GL(V)$ and $S : G \to GL(V)$ be two representations of G. Then*

$$\chi^{(T\oplus S)}(g) = \chi^{(T)}(g) + \chi^{(S)}(g), \qquad \chi^{(T\otimes S)}(g) = \chi^{(T)}(g)\chi^{(S)}(g).$$

In particular, the character of a semisimple representation T can be written as

$$\chi^{(T)}(g) = \sum_{i=1}^{n} a_i \chi^{(T_i)}(g),$$

where the T_i are the irreducible representations appearing in the direct sum with multiplicity a_i.

2.8 Orthogonality for characters

Objectives

You should be able to:

- Recall and prove the orthogonality relations for characters.
- Calculate the number of irreducible representations of a finite group.
- Constrain the dimensions of the irreducible representations of a finite group.
- Constrain the character table of a group using orthogonality relations for characters.

Characters are very useful to, indeed, characterize representations. This becomes clear when we state the orthogonality relations in terms of characters. We know that equivalent representations have the same character. We will see that inequivalent representations have different characters, in fact orthogonal.

2.8.1 First orthogonality theorem

Recall that the character of a representation T is a complex-valued function $\chi^{(T)} : G \to \mathbb{C}$. We can think of it as a vector $\chi^{(T)}$ in a $|G|$-dimensional complex vector space. The standard inner product is then given by

$$\langle \chi^{(T)}, \chi^{(S)} \rangle = \sum_{g\in G} \left(\chi^{(T)}\right)^* (g)\chi^{(S)}(g).$$

In fact, we can do better. We know that the character is a class function, that is, all group elements in a conjugacy class have the same character. If c is the number of conjugacy classes in the group G, then we can write

$$\langle \chi^{(T)}, \chi^{(S)} \rangle = \sum_{i=1}^{c} n_i \left(\chi_i^{(T)}\right)^* \chi_i^{(S)},$$

where n_i is the number of elements in the i'th conjugacy class, and we use $\chi_i^{(T)}$ to denote the character of these elements. What this means is that we can think of the characters as vectors in a c-dimensional complex vector space. Note that we have reduced the dimension of the vector space from $|G|$ to c, the number of conjugacy classes in G.

What we will now see is that the simple characters (characters of irreducible representations) are in fact orthogonal vectors. In particular, inequivalent irreducible representations must have different (in fact orthogonal) characters.

Theorem 2.8.1 Orthogonality of characters. *The simple characters of a finite group G (the characters of its irreducible unitary representations) are orthogonal:*

$$\sum_{g\in G}\left(\chi^{(T)}\right)^*(g)\chi^{(S)}(g)=\sum_{i=1}^{c}n_i\left(\chi_i^{(T)}\right)^*\chi_i^{(S)}=|G|\delta_{TS},$$

where δ_{TS} is zero if T and S are inequivalent, and one if they are equivalent. Here n_i denotes the number of elements in the i'th conjugacy class.

Proof. This is a direct consequence of the great orthogonality theorem Theorem 2.6.1. Using the notation δ_{TS}, we can state Theorem 2.6.1 as

$$\sum_{g\in G}T^\dagger(g)_{ij}S(g)_{kl}=\frac{|G|}{d}\delta_{il}\delta_{jk}\delta_{TS}.$$

We look at the special case with $i=j$, $k=l$, and we sum over i and k. We get:

$$\sum_{i,k}\sum_{g\in G}T^\dagger(g)_{ii}S(g)_{kk}=\sum_{i,k}\frac{|G|}{d}\delta_{ik}\delta_{ik}\delta_{TS}.$$

The left-hand-side is simply $\sum_{g\in G}\left(\chi^{(T)}\right)^*(g)\chi^{(S)}(g)$. For the right-hand-side, we notice that

$$\sum_{i,k}\delta_{ik}\delta_{ik}=\sum_{i,k}\delta_{ik}=\sum_{i=1}^{d}1=d.$$

Therefore

$$\sum_{g\in G}\left(\chi^{(T)}\right)^*(g)\chi^{(S)}(g)=|G|\delta_{TS}.$$

■

We have shown that the characters are orthogonal vectors in the vector space of class functions, which has dimensions c, the number of conjugacy classes. It thus follows immediately that:

Corollary 2.8.2 Upper bound on the number of irreps. *The number ρ of inequivalent irreducible representations of a finite group is less or equal than the number c of conjugacy classes.*

2.8.2 Second orthogonality theorem

So far we looked at the character $\chi^{(T)}:G\to\mathbb{C}$ of an irreducible representation T as a vector in the c-dimensional space of class functions. Now we want to look at a different type of orthogonality. Let ρ be the number of irreducible representations of a finite group G, which we know is finite and bounded by c, and label the irreducible representations as $T^{(\alpha)}$. We denote their characters by $\chi^{(\alpha)}$. Now, for each conjugacy class $i\in\{1,\dots,c\}$, we can think of the $\chi_i^{(\alpha)}$

as being the components of a vector in the ρ-dimensional vector space of irreducible representations. We can also show that the characters are orthogonal from this point of view.

Theorem 2.8.3 Orthogonality of characters II. *Let $\chi_i^{(\alpha)}$ be the character of the irreducible representation $T^{(\alpha)}$ of a finite group G, in the i'th conjugacy class. Then:*

$$\sum_{\alpha=1}^{\rho} \left(\chi_i^{(\alpha)}\right)^* \chi_j^{(\alpha)} = \frac{|G|}{n_i}\delta_{ij}.$$

That is, the characters are orthogonal in the ρ-dimensional vector space of irreducible representations. Here n_i denotes the number of elements in the i'th conjugacy class.

We will skip the proof of this orthogonality theorem.

2.8.3 Some important consequences

A direct consequence of this theorem is that it completely fixes the number of irreducible representations of a finite group G:

Theorem 2.8.4 The number of irreps of a finite group. *The number ρ of inequivalent irreducible representations of a finite group is equal to the number c of conjugacy classes of the group.*

Proof. We have already seen that $\rho \leq c$. Since the characters are orthogonal in the vector space of irreducible representations, we also get that $c \leq \rho$. Thus $\rho = c$. ■

Remark 2.8.5 It follows from this statement that simple characters form an orthogonal basis on the space of class functions. Thus any class function can be written as a linear combination of simple characters.

Using orthogonality of characters we can also determine the possible dimensions of irreducible representations. Perhaps surprisingly, not only is the number of irreducible representations fixed, but their dimensions is also highly constrained. In other words, for a group of a certain size, its irreducible representations cannot be too large.

Theorem 2.8.6 Dimensions of irreducible representations. *Let d_α be the dimension of the irreducible representation $T^{(\alpha)}$ of a finite group G. Then*

$$\sum_{\alpha=1}^{c} d_\alpha^2 = |G|,$$

where the sum is over all inequivalent irreducible representations of G.

Proof. This follows from the second orthogonality relation Theorem 2.8.3. Set $i = j = 1$, where the first conjugacy class is identified with the conjugacy class which contains only the identity element. We get:

$$\sum_{\alpha=1}^{c} (\chi_1^{(\alpha)})^* \chi_1^{(\alpha)} = \sum_{\alpha=1}^{c} d_\alpha^2 = |G|.$$

■

Using this result, we can prove the converse of Theorem 2.5.3, which stated that all irreducible representations of finite abelian groups are one-dimensional.

Theorem 2.8.7 Irreducible representations of abelian groups. *A finite group is abelian if and only if all its irreducible representations are one-dimensional.*

Proof. We already know that the irreducible representations of a finite abelian group are one-dimensional. We need to show that if a finite group only has one-dimensional irreducible representations, then it is abelian.

Suppose that all irreducible representations are one-dimensional. Then from Theorem 2.8.6 we get that

$$\sum_{\alpha=1}^{c} 1 = |G|,$$

which means that the number of inequivalent irreducible representations c, which is equal to the number of conjugacy classes in G, must be equal to the order of the group $|G|$. In other words, all conjugacy classes of G must contain only one element, which implies that the group is abelian. ■

2.8.4 Character table

The character table of a finite group is a common way of summarizing the characters of the irreducible representations of a finite group G. It is very useful in physics, for instance to determine the direct sum decomposition of reducible representations using the decomposition theorem (see Theorem 2.9.2). In fact some textbook on group theory for physics compile character tables for various groups.

The rows of a character table are labeled by the inequivalent irreducible representations of G, while the columns are labeled by the conjugacy classes of G. Since the number of inequivalent irreducible representations is equal to the number of conjugacy classes, the character table is always square. Each entry in the table gives the character of the given irreducible representation in the corresponding conjugacy class. Orthogonality relations between characters are very useful to fill in the character table of a given group, even without knowing all details of its irreducible representations. We will see how this goes in the example of S_3 shortly (see Section 2.10).

2.9 Reducibility and decomposition

Objectives

You should be able to:

- Use the decomposition theorem to calculate the direct sum decomposition of a reducible representation.
- Use the character of a representation to determine whether it is reducible or not.

In the previous section we used characters to characterize inequivalent irreducible representations of finite groups. We found that the number of inequivalent irreducible representations of G is equal to the number of conjugacy classes in G, and that the dimensions of the irreducible representations are constrained by the order G. We saw that the orthogonality relations for characters can be neatly encoded in terms of a character table.

In this section we show how characters can also be used to study whether a given representation is reducible or not: the so-called “reducibility criterion”. In the case of a reducible (semisimple) representation, we also study how characters can be used to find the decomposition of the representation as a direct sum of irreducible representations: the so-called “decomposition theorem”.

2.9.1 A criterion for reducibility

Using orthogonality of characters we can find a simple criterion to determine whether a representation is reducible or not.

Let $T : G \to GL(V)$ be a semisimple representation. It can be decomposed as a sum of irreducible representations

$$T = \bigoplus_{\alpha=1}^{c} m_\alpha T^{(\alpha)},$$

where m_α is an integer that denotes the number of times that $T^{(\alpha)}$ appears in the decomposition. Taking the trace, we get a similar relation for the characters:

$$\chi^{(T)}(g) = \sum_{\alpha=1}^{c} m_\alpha \chi^{(\alpha)}(g).$$

Therefore, the character of a semisimple representation is a linear combination of simple characters with non-negative coefficients. (Any class function can be written as a linear combination of simple characters, but what is special here is that the coefficients are non-negative.)

A semisimple representation will then be irreducible if and only if only one of the coefficients m_α is non-zero and equal to one. Our goal is to find a simple criterion in terms of the characters of a representation to determine when this is the case, without first having to calculate the explicit direct sum decomposition of the representation.

Theorem 2.9.1 Criterion for reducibility. *A representation T of a finite group G is irreducible if and only if*

$$\sum_{i=1}^{c} n_i |\chi_i^{(T)}|^2 = |G|,$$

where n_i is the number of elements in the i'th conjugacy class. It it is reducible, then

$$\sum_{i=1}^{c} n_i |\chi_i^{(T)}|^2 > |G|.$$

Proof. As explained above, the characters of a semisimple representation are a linear combination of simple characters with non-negative coefficients:

$$\chi_i^{(T)} = \sum_{\alpha=1}^{k} m_\alpha \chi_i^{(\alpha)}.$$

The representation is irreducible if and only if only one of the coefficients m_α is non-zero and equal to one.

Taking the complex conjugate we also have

$$(\chi_i^{(T)})^* = \sum_{\alpha=1}^{k} m_\alpha (\chi_i^{(\alpha)})^*.$$

We take the product of these two equations, multiply by n_i (the number of elements in the i'th conjugacy class), and sum over i. We get:

$$\begin{aligned}\sum_{i=1}^{c} n_i \chi_i^{(T)} (\chi_i^{(T)})^* =& \sum_{\alpha=1}^{k}\sum_{\beta=1}^{k} m_\beta m_\alpha \left(\sum_{i=1}^{c} n_i \chi_i^{(\alpha)} (\chi_i^{(\beta)})^*\right) \\ =& |G| \sum_{\alpha=1}^{k} m_\alpha^2,\end{aligned}$$

where we used the first orthogonality theorem Theorem 2.8.1. In particular, T is irreducible if and only if all m_α are zero except one that is equal to one, so we get that it is irreducible if and only if

$$\sum_{i=1}^{c} n_i |\chi_i^{(T)}|^2 = |G|.$$

It thus follows that it is reducible if and only if

$$\sum_{i=1}^{c} n_i |\chi_i^{(T)}|^2 > |G|.$$

■

2.9.2 The decomposition theorem

We now introduce the decomposition theorem, which gives us a way of determining the irreducible representations that appear in the direct sum decomposition of a semisimple representation. The decomposition theorem uses orthogonality of characters to determine the coefficients m_α in the decomposition.

Theorem 2.9.2 The decomposition theorem. *Given a semisimple representation T of a finite group G with characters $\chi_i^{(T)}$, where i indexes the conjugacy classes of G, the coefficients m_α in the direct sum decomposition*

$$T = \bigoplus_{\alpha=1}^{c} m_\alpha T^{(\alpha)}$$

are given by

$$m_\alpha = \frac{1}{|G|} \sum_{i=1}^{c} n_i \chi_i^{(T)} (\chi_i^{(\alpha)})^*,$$

where n_i is the number of elements in the i'th conjugacy class, and $\chi_i^{(\alpha)}$ are the characters of the irreducible representation $T^{(\alpha)}$.

Proof. We know that

$$\chi_i^{(T)} = \sum_{\beta=1}^{c} m_\beta \chi_i^{(\beta)}.$$

We multiply by $n_i(\chi_i^{(\alpha)})^*$ and sum over i. We get:

$$\begin{aligned}\sum_{i=1}^{c} n_i \chi_i^{(T)} (\chi_i^{(\alpha)})^* =& \sum_{\beta=1}^{c} m_\beta \left(\sum_{i=1}^{c} n_i \chi_i^{(\beta)} (\chi_i^{(\alpha)})^* \right) \\ =& |G| m_\alpha,\end{aligned}$$

where we used the first orthogonality theorem Theorem 2.8.1. Therefore

$$n_\alpha = \frac{1}{|G|} \sum_{i=1}^{c} n_i \chi_i^{(T)} (\chi_i^{(\alpha)})^*.$$

■

The decomposition theorem is useful to determine the direct sum decomposition of reducible representations. If one knows the characters of all irreducible representations of a given group, then calculating the coefficients m_α

of the decomposition of a reducible representation becomes a simple algebraic calculation.

2.10 An example: S_3

Objectives

You should be able to:

- Apply the orthogonality theorems to calculate the character table of a finite group.
- Apply the reducibility criterion to find whether a representation is reducible.
- Apply the decomposition theorem to calculate the direct sum decomposition of a reducible representation.

Let us now study representations of our favourite symmetric group S_3 using all that we have seen about characters. This should make the various theorems and constructions more explicit.

2.10.1 The character table

First, recall that S_3 is the group of permutations of three objects. It has six elements. They are split into three conjugacy classes, corresponding to the three possible cycle structures. There is one permutation with three one-cycles (the identity permutation); 3 permutations with one two-cycle and one one-cycle (the transpositions); and 2 permutations with a three-cycle (the cyclic permutations of length 3). Let us denote the corresponding conjugacy classes as C_1, $C_2$and C_3, with number of elements respectively $n_1 = 1$, $n_2 = 3$ and $n_3 = 2$.

Since the number of inequivalent irreducible representations is equal to the number of conjugacy classes (Theorem 2.8.4), we know that S_3 has three inequivalent irreducible representations. Let us denote those by $T^{(\alpha)}$ with $\alpha = 1, 2, 3$. We have already found two of them: the identity representation ($T^{(1)}$) and the parity representation ($T^{(2)}$), which are both one-dimensional. We do not know yet the third one.

From Theorem 2.8.6, we know that the dimensions of the three irreducible representations must satisfy the equation

$$\sum_{\alpha=1}^{3} d_\alpha^2 = |S_3| = 6.$$

The only possibility is that two of the representations are one-dimensional (we know them already), and the third one is two-dimensional.

Even if we do not know the third two-dimensional irreducible representation explicitly, we can still construct the character table of S_3 using orthogonality of characters. The character table should look like:

Table 2.10.1 Constructing the character table for S_3

	C_1	C_2	C_3
$T^{(1)}$	1	1	1
$T^{(2)}$	1	-1	1
$T^{(3)}$	2	a	b

We have already filled many entries, and left two unknowns (a and b), which we will determine shortly. The first column is given by the trace of the matrices corresponding to the identity element of the group, which are just identity matrices. So these entries are just the dimension of the representations. This is true in general: the first column is always given by the dimensions of the representations.

The first row corresponds to the identity representation, which assigns 1 to all group elements. So the characters are always 1.

The second row corresponds to the parity representation, which assigns 1 to even permutations and -1 to odd permutations. The permutations in the conjugacy classes C_1 and C_3 are even, while those in C_2 are odd, so we can fill the second row appropriately.

There are two entries left, a and b. To find them, we use orthogonality of characters. We will use orthogonality of columns, which is just the statement of the second orthogonality relation Theorem 2.8.3. This relation says that

$$\sum_{\alpha=1}^{3} (\chi_i^{(\alpha)})^* \chi_j^{(\alpha)} = \frac{|G|}{n_i}\delta_{ij}.$$

Applying this to the first and second conjugacy classes (the first and second columns), that is, setting $i = 1$ and $j = 2$, the right-hand-side becomes zero, and we get:

$$(1)(1) + (1)(-1) + (2)(a) = 0.$$

Therefore $a = 0$. Similarly, using the first and third columns ($i = 1$ and $j = 3$), we get:

$$(1)(1) + (1)(1) + (2)(b) = 0,$$

that is $b = -1$. We have thus completed the character table of S_3:

Table 2.10.2 The character table for S_3

	C_1	C_2	C_3
$T^{(1)}$	1	1	1
$T^{(2)}$	1	-1	1
$T^{(3)}$	2	0	-1

For fun we can check that the other orthogonality relations are also satisfied. For instance, sticking with the second orthogonality relation Theorem 2.8.3, but applying it to the third column ($i = j = 3$), we get:

$$(1)(1) + (1)(1) + (-1)(-1) = \frac{|G|}{2} = \frac{6}{2} = 3,$$

which is indeed true.

The first orthogonality relation Theorem 2.8.1 correspond to orthogonality between rows. Recall that it says that:

$$\sum_{i=1}^{3} n_i (\chi_i^{(\alpha)})^* \chi_i^{(\beta)} = |G|\delta_{\alpha\beta}.$$

For instance, using it for the second and third irreducible representations ($\alpha = 2$ and $\beta = 3$), we get:

$$1(1)(2) + 3(1)(0) + 2(1)(-1) = 0,$$

which is indeed correct. Applying to the third row ($\alpha = \beta = 3$), we get:

$$1(2)(2) + 3(0)(0) + 2(-1)(-1) = |G| = 6,$$

which is again correct. Orthogonality relations are indeed satisfied!

We can also check that the reducibility criterion Theorem 2.9.1 is satisfied for the irreducible representations. For instance, looking at the two-dimensional irreducible representation $T^{(3)}$, we get:

$$\sum_{i=1}^{3} n_i|\chi_i^{(3)}|^2 = 1(2)^2 + 3(0)^2 + 2(-1)^2 = 6,$$

which is indeed equal to $|G| = 6$, therefore the representation is irreducible, as we know. In fact, we know this two-dimensional irreducible representation of S^3 very well: it is the two-dimensional representation constructed from symmetries of the equilateral triangle in Example 2.1.6. Indeed, one can check that the characters of this representation match with the third row of the character table for S^3.

2.10.2 The decomposition theorem

Let us now study the reducibility criterion and decomposition theorem. Let us consider the three-dimensional permutation representation T of S_3 from Example 2.1.4. Writing down the 3×3 matrices explicitly, it is easy to calculate the characters by taking the traces. The trace of the identity matrix is 3; the trace of the three matrices corresponding to the transpositions is 1; and the trace of the two matrices corresponding to the cyclic permutations are 0. Thus, the characters of T are:

$$\chi_1^{(T)} = 3, \qquad \chi_2^{(T)} = 1, \qquad \chi_3^{(T)} = 0.$$

Let us first show that this representation is reducible (we already know that since S_3 has only three irreducible representations of dimensions 1, 1 and 2.) We use the reducibility criterion Theorem 2.9.1. We calculate:

$$\sum_{i=1}^{3} n_i|\chi_i^{(T)}|^2 = 1(3)^2 + 3(1)^2 + 2(0)^2 = 12,$$

which is certainly greater than $|G| = 6$. Thus the representation is reducible, as expected.

We now want to find its direct sum decomposition. Recall the decomposition theorem Theorem 2.9.2. If we write the direct sum decomposition as

$$T = \bigoplus_{\alpha=1}^{c} m_\alpha T^{(\alpha)},$$

then the coefficients are given by

$$m_\alpha = \frac{1}{|G|}\sum_{i=1}^{c} n_i \chi_i^{(T)} (\chi_i^{(\alpha)})^*.$$

Let us calculate these coefficients using the character table Table 2.10.2. We get:

$$\begin{aligned} m_1 =& \frac{1}{6}\left(1(3)(1) + 3(1)(1) + 2(0)(1)\right) = 1, \\ m_2 =& \frac{1}{6}\left(1(3)(1) + 3(1)(-1) + 2(0)(1)\right) = 0, \\ m_3 =& \frac{1}{6}\left(1(3)(2) + 3(1)(0) + 2(0)(-1)\right) = 1. \end{aligned}$$

We thus obtain the direct sum decomposition:

$$T = T^{(1)} \oplus T^{(3)}.$$

What this means is that the three-dimensional permutation representation of S_3 is equivalent to the direct sum of the trivial one-dimensional representation $T^{(1)}$, and the two-dimensional irreducible representation $T^{(3)}$. This is something that you could show explicitly by finding a similarity transformation that brings all six 3×3 matrices in block diagonal form. But you see how much faster this was using characters! No need to find similarity transformations, only simple algebraic calculations are needed! Characters are really awesome!

2.11 An example: the regular representation

Objectives

You should be able to:

- Find the characters and the direct sum decomposition of the regular representation of any finite group.

As a last example in this section, let us study the regular representation for any finite group, which was introduced in Example 2.1.5. Recall that the regular representation is defined using Cayley's theorem. We think of a finite group G of order n as a subgroup of S_n, and keep only the matrices of the n-dimensional permutation representation of S_n that correspond to the elements of the subgroup to obtain the regular representation of G. Thus the regular representation is a n-dimensional representation.

It turns that for any finite group G, the characters of the regular representation are very easy to compute. First, the matrix corresponding to the identity element is the identity matrix, thus its trace is n, the dimension of the representation. What is the trace of the matrices corresponding to the other group elements in G? Recall from Cayley's theorem that one defines the mapping $G \to S_n$ by looking at the permutations of the group elements corresponding to the rows of the mulitiplication table. But by the rearrangement theorem Theorem 1.4.2, we know that each row must permute every single element (no element can be left fixed, since they cannot appear twice in the same column in the multiplication table). Thus, the corresponding n-dimensional permutation matrices will necessarily only have zeroes on the diagonal (otherwise it would leave some elements fixed). Their traces are then necessarily zero!

Therefore we conclude that for any finite group G of order n, its n-dimensional regular representation R will have characters $\chi_1^{(R)} = n$ for the conjugacy class containing the identity element and $\chi_i^{(R)} = 0$ for all other conjugacy classes $i > 1$.

With this we can show that for all finite groups with $n \geq 2$ the regular representation is reducible. Indeed, by the reducibility criterion Theorem 2.9.1, we get:

$$\sum_{i=1}^{c} n_i |\chi_i^{(R)}|^2 = 1 \cdot n^2 + 0 = n^2,$$

which is certainly $> |G| = n$ for all $n \geq 2$.

We can also find its decomposition into a direct sum of irreducible representations for all finite groups! Recall from the decomposition theorem Theorem 2.9.2 that, if we write the decomposition as

$$T^{(S)} = \bigoplus_{\alpha=1}^{c} m_\alpha T^{(\alpha)},$$

the coefficients are given by

$$m_\alpha = \frac{1}{|G|} \sum_{i=1}^{c} n_i \chi_i^{(S)} (\chi_i^{(\alpha)})^*.$$

In the case of the regular representation of a group of order n, since the only non-vanishing character is $\chi_1^{(S)} = n$, for any α we calculate:

$$m_\alpha = \frac{1}{n} \left(n(\chi_1^{(\alpha)})^* \right) = (\chi_1^{(\alpha)})^*.$$

But for any irreducible representation, the character of the identity element is the trace of the identity matrix, which is just the dimension of the representation:

$$\chi_1^{(\alpha)} = d_\alpha.$$

Therefore $m_\alpha = d_\alpha$, for all irreducible representations!

Therefore for any finite group the regular representation decomposes as the direct sum of irreducible representations:

$$T^{(S)} = \bigoplus_{\alpha=1}^{c} d_\alpha T^{(\alpha)},$$

whose coefficients are precisely the dimensions of the irreducible representations! In other words, in the decomposition each irreducible representation appears exactly the same number of times as its dimension. Isn't that cool?

For instance, the six-dimensional regular representation of S_3 has a direct sum decomposition:

$$T^{(S)} = T^{(1)} \oplus T^{(2)} \oplus T^{(3)} \oplus T^{(3)},$$

where as before $T^{(1)}$ is the one-dimensional identity representation, $T^{(2)}$ is the one-dimensional parity representation, and $T^{(3)}$ is the two-dimensional irreducible representation of S_3 (since the latter is two-dimensional it appears twice in the decomposition of the regular representation).

Remark 2.11.1 The existence of the regular representation for all finite groups also provide a simple proof of Theorem 2.8.6 about the dimensions of irreducible representations of finite groups. Indeed, the regular representation is n-dimensional, where $n = |G|$, and is equivalent to a block diagonal matrix with blocks of size d_α, where $d_\alpha$is the dimension of the irreducible representations, and such that each block appears exactly d_α times. Thus, simply comparing the dimension of the matrices, we conclude directly that

$$\sum_{\alpha=1}^{c} d_\alpha^2 = |G|.$$

2.12 Real, pseudoreal and complex representations

Objectives

You should be able to:

- Determine whether a given representation is real, pseudoreal or complex.

In this course we focus on group representations on complex vector spaces. However, sometimes such a representation may be real, just like complex numbers may be real. So a natural question is: given a representation on a complex vector space, can we determine easily when a representation is real or not?

First, how do we know whether a complex number z is actually real? Well, it is easy, you look at it, and see whether its imaginary part is zero! More precisely, the statement is that z is real if and only if $z = z^*$, that is, z is equal to is complex conjugate, which means that its imaginary part vanishes.

We can do the same thing with matrices. We say that a matrix M is real if and only if M^*. Note that this is really the complex conjugate, not the complex conjugate transpose. So this means that all the matrix entries are real numbers.

Now what about representations? When is a representation real? This is of great interest in physics; for instance, once one understands particles in terms of representations, then anti-particles transform in the complex conjugate representations. So understanding the relation between a representation and is complex conjugate is crucial.

2.12.1 Real, pseudoreal and complex

Let $T : G \to GL(V)$ be a representation of a group G on a complex vector space V. Let us first define the complex conjugate representation.

Lemma 2.12.1 Complex conjugate representation. *Let $T : G \to GL(V)$ be a representation of a group G on a complex vector space V. Then T^*, which is obtained by taking the complex conjugate matrices $T(g)^*$, is also a representation $T^* : G \to GL(V)$, and is called the* ***complex conjugate representation****. Note that its characters are $\chi_i^{(T^*)} = (\chi_i^{(T)})^*$.*

Proof. If T is a group homomorphism, then $T(gh) = T(g)T(h)$ for any $g, h \in G$. Taking the complex conjugate of the matrices, $T^*(gh) = T(g)^*T(h)^*$, and hence T^* is also a group homomorphism. ■

We would like to say that a representation is real if the entries of the corresponding matrices are real numbers. Or, more precisely, the matrices should be equivalent (by a similarity transformation) to matrices with only real numbers. We would also like to translate this condition into the statement that a real representation is equivalent to its complex conjugate. But this is subtle; real representations are indeed equivalent to their complex conjugate, but the converse is not true. Let us be a little more careful.

Definition 2.12.2 Complex representations. We say that T is **complex** if it is not equivalent to its complex conjugate representation T^*. We say that it is **non-complex** otherwise. That is, T is non-complex if there exists a matrix S such that for all $g \in G$,

$$T^*(g) = ST(g)S^{-1}.$$

◊

Remark 2.12.3 In terms of characters, it follows that if the characters $\chi_i^{(T)}$ are complex, i.e. not equal to their complex conjugate, then T is complex. The converse is not true however; if the characters are real, then we cannot conclude that T and T^* are equivalent.

We now show that we can divide non-complex representations further, when we focus on irreducible representations. We first need a lemma that tells us there are only two types of possible similarity transformations between an irreducible representation and its complex conjugate.

Lemma 2.12.4 Symmetric or anti-symmetric. *Let $T : G \to GL(V)$ be a non-complex unitary irreducible representation of a group G on a complex vector space V. Thus T is equivalent to T^*:*

$$T^*(g) = ST(g)S^{-1},$$

for some matrix S. Then S is either a symmetric matrix or an anti-symmetric matrix.

Proof. We start with $T^*(g) = ST(g)S^{-1}$. We take the transpose, to get:

$$T^\dagger(g) = (S^{-1})^T T(g)^T S^T.$$

Since T is unitary, the left-hand-side is equal to $T(g)^{-1} = T(g^{-1})$. Thus we get:

$$T(g^{-1}) = (S^{-1})^T T(g)^T S^T,$$

which shows that $T(g)$ and $T(g^{-1})$ are related. Now this equality is true for any $g \in G$, so we can write it for g^{-1} instead of g. We get:

$$T(g) = (S^{-1})^T T(g^{-1})^T S^T.$$

But $T(g^{-1})^T = ST(g)S^{-1}$, and hence, substituting back,

$$\begin{aligned} T(g) =& (S^{-1})^T \left(ST(g)S^{-1}\right) S^T \\ =& \left(S^{-1}S^T\right)^{-1} T(g) \left(S^{-1}S^T\right). \end{aligned}$$

What this means is that $S^{-1}S^T$ commutes with all matrices $T(g)$ of an irreducible representation. By Schur's lemma Lemma 2.5.2, we know that $S^{-1}S^T = \lambda I$ for some $\lambda \in \mathbb{C}$. Thus $S^T = \lambda S$. This means that

$$S = (S^T)^T = (\lambda S)^T = \lambda S^T = \lambda^2 S,$$

and hence $\lambda^2 = 1$. It follows that $S = \pm S$, that is, it is either symmetric or anti-symmetric. ■

Definition 2.12.5 Real and pseudoreal representations. Let $T : G \to GL(V)$ be a non-complex unitary irreducible representation of a group G on a complex vector space V. Thus T is equivalent to T^*:

$$T^*(g) = ST(g)S^{-1},$$

for some matrix S. We say that T is **real** if S is symmetric, and **pseudoreal** if S is anti-symmetric. ♢

Now if a representation is given by matrices $T(g)$ that only have real entries, then it is certainly real according to our definition, since S can be taken to be the identity matrix. But is it true that all real representations are equivalent to representations with matrices that only have real entries? This is what we show next.

Lemma 2.12.6 Real is real. *A real irreducible unitary representation T is equivalent to a representation whose matrices only have real entries.*

Proof. First, we notice that the similarity transformation S can be taken to be unitary. Indeed, since $T^*(g) = ST(g)S^{-1}$

$$ST(g) = T^*(g)S = \left(T(g^{-1})\right)^T S,$$

and so $S = T(g)^T ST(g)$. We can then calculate:

$$\begin{aligned} S^\dagger S =& T(g)^\dagger S^\dagger T(g)^* T(g)^T ST(g) \\ =& T(g)^\dagger S^\dagger ST(g), \end{aligned}$$

since T is unitary. Thus $T(g)S^\dagger S = S^\dagger ST(g)$. It follows that $S^\dagger S$ commutes with all $T(g)$, and hence by Schur's lemma Lemma 2.5.2, we know that $S^\dagger S = \lambda I$ for some $\lambda \in \mathbb{C}$.

But the similarity transformation S is defined only up to overall rescaling (and it is non-zero), so we can choose an appropriate rescaling so that $S^\dagger S = I$, that is, it is unitary.

Now that we know that S is unitary, we can prove the lemma. We assume that the representation is real, so that S is a symmetric unitary matrix. One can show that any symmetric unitary matrix S can be written as $S = W^2$ for some unitary symmetric matrix W (proof left as exercise). Thus $T^*(g) = ST(g)S^{-1}$ becomes $T^*(g) = W^2 T(g)(W^{-1})^2$. We multiply by W^{-1} on the left and W on the right to get:

$$W^{-1}T^*(g)W = WT(g)W^{-1}.$$

Using the fact that W is unitary symmetric, we have $W^{-1} = W^\dagger = W^*$, and $W = (W^{-1})^*$. So we can write:

$$WT(g)W^{-1} = W^{-1}T^*(g)W = W^*T^*(g)(W^{-1})^* = \left(WT(g)W^{-1}\right)^*.$$

It thus follows that $WT(g)W^{-1}$ has only real entries, thus T is equivalent to a representation with only real entries. ■

2.12.2 The Frobenius-Schur indicator

We have now defined when an irreducible representation on a complex vector space is real, pseudoreal, or complex. We understand that a real representation is equivalent to a representation whose matrices only have real entries. But in general, the representation may not be given in this form. How do we determine whether an irreducible representation is real, pseudoreal or complex?

Fortunately there is a neat criterion, known as the Frobenius-Schur indicator, which uses our beloved characters. Here goes!

Theorem 2.12.7 The Frobenius-Schur indicator. *Let T be an irreducible unitary representation of a finite group G, and define the Frobenius-Schur indicator as $\frac{1}{|G|}\sum_{g\in G}\chi(g^2)$. Then*

$$\frac{1}{|G|}\sum_{g\in G}\chi(g^2) = \begin{cases} 1 & \text{if } T \text{ is real} \\ -1 & \text{if } T \text{ is pseudoreal,} \\ 0 & \text{if } T \text{ is complex.} \end{cases}$$

We will leave the proof as an exercise.

Checkpoint 2.12.8 Prove Theorem 2.12.7.

So to determine whether an irreducible representation is real, pseudoreal or complex, we simply need to evaluate the Frobenius-Schur indicator. We also note here that if g and h are in the same conjugacy class, than g^2 and h^2 also are; thus the Frobenius-Schur indicator is also a class function, and could be written as a sum over conjugacy classes.

Remark 2.12.9 The classification between real, pseudoreal and complex irreducible representations also hold for compact groups. Moreover, those can be distinguished using the Frobenius-Schur indicator as above, but with the sum replaced by an integral over the compact group.

Example 2.12.10 Irreducible representations of S_3. As an example, let us look at the irreducible representations of S_3, as studied in Section 2.10. To evaluate the Frobenius-Schur indicator we will need the square of the group elements. We note that the square of the identity is still the identity. As for the transpositions (conjugacy class C_2), their square is the identity. For the cyclic permutations (conjugacy class C_3), we calculate. Take for instance (123). The square is $(123)(123) = (132)$, and hence it is still a cyclic permutation of length three. Thus the square of elements in the conjugacy class C_3 are still in C_3.

Let us now determine whether the irreducible representations are real, pseudoreal or complex. First, the two one-dimensional representations are manifestly real, since they map all group elements to real numbers. Let us, for fun, evaluate the Frobenius-Schur indicator for these two representations.

For the trivial representation, we get:

$$\frac{1}{|S_3|}\sum_{g\in S_3}\chi^{(1)}(g^2) = \frac{1}{6}\left((1)(1) + (3)(1) + (2)(1)\right) = 1,$$

and hence the representation is real, as expected (this is obviously always the case for the trivial representation).

For the non-trivial one-dimensional representation $T^{(2)}$ (using the notation in Section 2.10). The sum is:

$$\frac{1}{|S_3|}\sum_{g\in S_3}\chi^{(2)}(g^2) = \frac{1}{6}\left((1)(1) + (3)(1) + (2)(1)\right) = 1,$$

and hence this representation is also real, as expected.

Finally, we look at the two-dimensional representation. We get:

$$\frac{1}{|S_3|}\sum_{g\in S_3}\chi^{(2)}(g^2) = \frac{1}{6}\left((1)(2) + (3)(2) + (2)(-1)\right) = 1,$$

and thus this representation is also real! This is interesting, because depending on how you think about these 2×2 matrices, they may be written as matrices with complex entries. But in fact in Example 2.1.6 we had already seen one way to write these matrices with real entries, and thus the representation must be real indeed. □

Example 2.12.11 The two-dimensional representation of the quaternion group. Let us now look at the two-dimensional representation of the quaternion group given in Example 2.1.7. First, we should check that it is irreducible. Using the criterion for irreducibility Theorem 2.9.1, and calculating the traces of the matrices, we get:

$$(2)^2 + (-2)^2 + (0) + (0) + (0) + (0) + (0) + (0) = 8,$$

which is the order of the quaternion group. Thus it is irreducible. It is the only irreducible representation of dimension greater than one for the quaternion group.

Let us determine whether it is real, pseudoreal or complex. First, we note that for the quaternion group, $i^2 = (-i)^2 = j^2 = (-j)^2 = k^2 = (-k)^2 = -1$, and of course $1^2 = (-1)^2 = 1$. Thus the Frobenius-Schur indicator evaluates to:

$$\frac{1}{8}\left((2) + (2) + (-2) + (-2) + (-2) + (-2) + (-2) + (-2)\right) = -1.$$

We thus conclude that this representation is pseudoreal. That is, it is equivalent to its complex conjugate, but the similarity transformation involves an anti-symmetric matrix. This is an example where all the characters of the representation are real, but the representation is not, it is pseudoreal instead. In particular, the 2×2 matrices cannot be brought into matrices with only real entries by a similarity transformation. □

Example 2.12.12 The cyclic group $\mathbb{Z}_n$. As a last example, let us look at the one-dimensional representation for the cyclic group $\mathbb{Z}_4$ given by the complex numbers $\{1, -i, -1, i\}$. We calculate the square of the elements. We get that $1^2 = (-1)^2 = 1$, and $(-i)^2 = i^2 = -1$. Thus the Frobenius-Schur indicator is:

$$\frac{1}{4}\left(1 + (-1) + 1 + (-1)\right) = 0,$$

and hence the representation is complex, as expected. □

Chapter 3

Applications

In this section we study a few simple physical applications of group and representation theory, focusing on finite groups.

3.1 Crystallography

Objectives

You should be able to:

- Define what point and space groups are.
- Prove the crystallographic restriction theorem.

Group theory is really all about symmetries. In this section we will see how group theory can be applied to study objects with symmetries, such as molecules and crystals. We will see that symmetries are highly constraining, and that studying symmetries of physical objects such as crystals gives rise to highly non-trivial, and perhaps unexpected, results.

One should note that there is extensive literature on this subject, in fact there are whole books dedicated to the study of symmetries of crystals. Physicists and chemists have invented their own notation to denote the symmetry groups, and reading through this literature can quickly become confusing and even overwhelming. So what we will do here is just quickly go through some of the basic results that highlight how group theory is useful to understand the physics and chemistry of crystals.

3.1.1 Point groups

3.1.1.1 Definition

The starting point of the discussion involves the definition of **point groups**. Point groups are groups of geometric symmetries (or isometries) of objects that fix a point (usually set to be the origin of the coordinate system). The dihedral groups D_n that we have encountered previously are examples of discrete point groups in two dimensions.

One can think of points groups in N dimensions as being subgroups of the orthogonal group $O(N)$. Roughly speaking, the elements of points groups correspond to the symmetry operations of rotations, reflections, and "improper rotations" (i.e. elements with determinant equal to -1). There are infinite point groups (as an example, think of the group generated by a rotation in two dimensions by an irrational number of turns), and finite point groups. In the following we will mostly focus on finite point groups, which are the most interesting for crystallography.

Finite point groups can be classified.

- In one dimension, there are two finite points groups: the trivial group, and the reflection group about the origin, which is isomorphic to $\mathbb{Z}_2$.
- In two dimensions, there are two infinite families of finite point groups: the cyclic groups of rotations by angle $2\pi/n$, which are isomorphic to $\mathbb{Z}_n$, and the dihedral groups D_n. However, as we will, if we also impose consistency with the translational symmetries of crystals, only a finite (and very small) number of those survive, which is the essence of the crystallographic restriction theorem.
- In three dimensions, there is of course more possibilities. There are 7 infinite familes of finite point groups, and 7 additional point groups. For more information on those, see for instance the Wikipedia page on point groups and on point groups in 3D specifically. Again, imposing consistency with translational symmetries lead to the 32 crystallographic point groups.

3.1.1.2 Molecular point groups

Three-dimensional point groups are sometimes called **molecular point groups**, because they are important in the study of molecules, see for instance the Wikipedia page on molecular symmetry. For instance, the chemical compound of Xenon tetrafluoride XeF_4, which consists of four atoms of fluoride placed in a square planar configuration around an atome of xenon, has point group the dihedral group D_4 of symmetries of the square.

It turns out that representations and characters are crucial in the study of molecular symmetry. This is because the orbitals of the molecules transform according to irreducible representations of the point group. Thus, to understand the possible states of the system, it is essential to gain information about irreducible representations of the point group, which is encapsulated in its character table. In fact, a lot of the properties of molecules (such as optical activity, spectroscopy, dipole moments, electronic properties, what not - note that I have no idea what these terms really mean :-)) can be analyzed using irreducible representations of the point group. If you are not convinced, google orbitals irreducible representations, or something like that. You would be surprised how much group and representation theory appears in quantum chemistry! Representations and characters are so important that Wikipedia even has a list of character tables for chemically important 3D point groups...

3.1.2 Space groups

In the description of crystals, we need more than just point groups. We need **space groups**, also called **crystallographic groups**. Roughly speaking, space groups are symmetry groups of particular configurations of objects in space. We do not impose the restriction that a point (the origin) is fixed anymore. So we allow symmetries such as translational symmetries, which are crucial in the study of crystals.

In general, one can think of space groups as combinations of point groups with translational symmetries. More precisely, given a space group, there is an abelian subgroup of translations. Such a subgroup gives rise to the structure of a lattice of points that are invariant under these translations. For particular such lattices, there are sometimes extra symmetries (such as rotations, reflections, etc.), which, together with translations, form the structure of the space group.

Just as for point groups, space groups can be classified.

- In one dimension, there are two space groups (also known as “line groups”): the group of translations, which is isomorphic to the group of integers $\mathbb{Z}$, and the infinite dihedral group, which includes reflections.
- In two dimensions, there are 17 space groups, also known as wallpaper groups (have a look at this page, it’s super fun!). Those are obtained by combining the 5 types of two-dimensional lattices (known as “Bravais lattices”) specifying the translational symmetries with the two-dimensional point groups that are consistent with translational symmetries.
- In three dimensions, there is a total of 230 different space groups. Those are obtained by combining the 14 types of three-dimensional Bravais lattices with the 32 point groups that are consistent with translational symmetries.

3.1.3 Crystals

We are now ready to discuss symmetry groups of crystals, and the famous **crystallographic restriction theorem**, which shows how powerful the study of symmetries can be in physics and chemistry.

To start with: what is a crystal? In physics and chemistry, a crystal is a solid material which has a highly ordered microscopic structure: it consists in a lattice of atoms that extends in all directions and is invariant under translations. Examples include snowflakes, diamonds, table salt, etc.

Let us focus on crystals in three dimensions. Mathematically, the lattice of atoms in a crystal is a **Bravais lattice**: this is an infinite array of discrete points that are generated by a translation:

$$\vec{T} = n_1\vec{v_1} + n_2\vec{v_2} + n_3\vec{v_3},$$

where $n_1, n_2, n_3 \in \mathbb{Z}$ and $\vec{v_1}, \vec{v_2}, \vec{v_3}$ are primitive vectors in three dimensions which lie in different directions (not necessarily perpendicular) and span the lattice. Thus a crystal is invariant under translations generated by $\vec{T}$.

Furthermore, a given crystal may be invariant under symmetry operations that leave a point fixed, such as rotations and reflections. The group of such symmetry operations is a point group; in this context, it is known as a crystallographic point group. Together with translations, it forms the space group of the crystal.

We have seen that finite point groups in three dimensions can be classified: there are 7 infinite familes, and 7 additional point groups. But can all of those be point groups of crystals? In other words, can all of these point groups arise as symmetry groups of lattices? The answer to this question is the essence of the crystallographic restriction theorem, and may appear surprising: only a very small, finite, number of point groups are symmetry groups of crystals! For instance, there is no crystal that is invariant under rotations by an angle $2\pi/5$! Isn’t that surprising?

Theorem 3.1.1 Crystallographic restriction theorem. *Consider a crystal in three dimensions that is invariant under rotations by an angle $2\pi/n$ around an axis. Then the only allowed possibilities are $n = 1, 2, 3, 4, 6$, with the case $n = 1$ being the trivial case with no rotational symmetry.*

Note that the same result holds for a crystal in two dimensions invariant under rotations.

What? Really? Those are the only possibilities? Imposing translational symmetry is that constraining? Let's prove this! There are a number of different proofs of this theorem. We will first present a geometric proof, and then present a representation theoretic, or matrix based, proof, for fun.

Geometric proof. First, we forget about the case $n = 1$ since it is trivial, as it involves no rotational symmetry.

Now suppose that a lattice is invariant under rotations by an angle $2\pi/n$ about an axis. Consider two lattice points A and B in the plane perpendicular to the axis of rotation, and separated by a translation vector $\vec{v}$ of minimal length (by which I mean that there is no lattice point on the line between A and B). Let us rotate point A about point B by an angle $2\pi/n$ and call the resulting point A'. Similarly, let us rotate point B about A by the same angle $2\pi/n$ but in the opposite direction, and call the resulting point B'.

Let us call $\vec{v'}$ the vector joining A' and B'. We notice that $\vec{v'}$ is parallel to $\vec{v}$. Thus, if rotation by an angle $2\pi/n$ is a symmetry of the lattice, it follows that the new translation vector $\vec{v'}$ must be an integer multiple of $\vec{v}$, that is,

$$\vec{v'} = k\vec{v} \qquad \text{for some } k \in \mathbb{Z}.$$

This is because the two points A' and B' must be related by a translation generated by $\vec{v}$, and $\vec{v}$ was chosen to have minimal length.

Now we can relate the length of $\vec{v'}$ to the length of $\vec{v}$ and the angle $2\pi/n$. The four points A, B, A', B' form a trapezium, with angles $2\pi/n$ and $\pi - 2\pi/n$. By basic geometry, we get:

$$|\vec{v'}| = |\vec{v}| + 2|\vec{v}| \cos\left(\pi - \frac{2\pi}{n}\right) = |\vec{v}| \left(1 - 2\cos\frac{2\pi}{n}\right).$$

Since we must have $\vec{v'} = k\vec{v}$ for some $k \in \mathbb{Z}$, and hence $|\vec{v'}| = |k||vecv|$, we get the constraint:

$$|k| = 1 - 2\cos\frac{2\pi}{n},$$

that is, $1 - 2\cos\frac{2\pi}{n}$ must be a non-negative integer. Since $|\cos\alpha| \geq 1$ for any angle α, the only possible values of $|k|$ are $|k| = 0, 1, 2, 3$. Solving for n for each cases, we get, respectively, $n = 6, 4, 3, 2$. Those are the only possible rotational symmetries of a crystal, which concludes the proof. ■

Representation theoretic proof. Let us now give a representation theoretic, or matrix based, proof. A rotation by an angle $2\pi/n$ about an axis can be seen as a linear operator acting on the two-dimensional vector space perpendicular to the axis of rotation. After choosing a basis for this vector space, such a rotation can be written in terms of a 2×2 rotation matrix. In the language of group theory, this gives a two-dimensional representation of the cyclic rotation group $\mathbb{Z}_n$. On general grounds, we know how to write such a rotation matrix with respect to an orthonormal basis for $\mathbb{R}^2$:

$$\begin{pmatrix} \cos\frac{2\pi}{n} & \sin\frac{2\pi}{n} \\ -\sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{pmatrix}.$$

This matrix expression for the linear operator is only valid with respect to an orthonormal basis, but its trace (that is, the character of the representation), which is equal to $2\cos(2\pi/n)$, is invariant under changes of basis, as we know (that is, the characters are the same for equivalent representations related by similarity transformations). So in a different choice of basis, the 2×2 rotation matrix may look different, but its trace will always be equal to $2\cos(2\pi/n)$.

Now, in a lattice, it is appropriate to write down how rotations act with respect to the basis given by the primitive translation vectors $\vec{v_i}$. Those are in general not orthogonal, nor of length one, so the precise matrix representation may look ugly. However, if a rotation by an angle $2\pi/n$ is a symmetry of the lattice, then we know that it must map each translation vector into an integer linear combination of the other translation vectors. In other words, with respect to this lattice basis, the rotation matrix corresponding to rotation by an angle $2\pi/n$ must have integer entries. In particular, its trace must be integer. But since we know that the trace is invariant under changes of basis, its trace must also be equal to $2\cos(2\pi/n)$. Therefore we conclude that $2\cos(2\pi/n) = k$ for some $k \in \mathbb{Z}$. The only possibilities are $k = -2, -1, 0, 1, 2$, which correspond to, respectively, $n = 2, 3, 4, 6, 1$, with the last choice corresponding to the identity matrix (i.e. no rotational symmetry). These are the only possible rotational symmetries that are consistent with translational symmetries of a crystal. ■

This is a highly non-trivial result, and is perhaps unexpected. Why can't we build a lattice that is invariant under five-fold rotations? Or seven-fold rotations? This is related to the fact that we cannot combine objects with a five-fold or seven-fold apparent symmetry in such a way that they completely fill the plane. See for instance this webpage for pretty pictures.

Using the crystallographic restriction theorem, we can enumerate the possible crystallographic point groups in low dimensions.

- In two dimensions, we had two infinite families of point groups: the cyclic rotations groups ($\mathbb{Z}_n$) and the dihedral groups (D_n). Since crystals can only be invariant under rotations by an angle $2\pi/n$ with $n \in \{1, 2, 3, 4, 6\}$, we see that the only crystallographic point groups are the five cyclic rotation groups $\mathbb{Z}_n$ and five dihedral groups D_n with $n \in \{1, 2, 3, 4, 6\}$, for a total of 10 crystallographic point groups in two dimensions.

- In three dimensions, we started with 7 infinite families and 7 additional point groups. Looking at the rotations that are elements of these point groups, one can conclude that there are only 32 crystallographic point groups in three dimensions.

To summarize, the possible symmetries of crystals are highly constrained! This is neat, and is a nice, direct, application of group theory to physics and chemistry. Crystallography is all about group and representation theory!

To end this section, note that the crystallographic restriction theorem is not the end of the story. The story continues with aperiodic tilings and quasicrystals. It is truly fascinating (and has led to a Nobel prize). Have a look at these Wikipedia pages and further resources if you are interested!

3.2 Quantum Mechanics

Objectives

You should be able to:

- Explain why the states of a quantum mechanical system can be labeled using irreducible representations of the symmetry group, and how the degeneracy of states is computed by looking at the dimensions of the irreducible representations.

- Use representation theory to determine the form of the states in the case of a cyclic periodic potential (Bloch's theorem).

In this section we give a very brief introduction to the power of group and representation theory in quantum mechanics.

3.2.1 Schrodinger equation

The starting point of quantum mechanics is Schrodinger equation. The state of a physical system in quantum mechanics is determined by a wave-function Ψ, which is a solution of a differential equation known as the **Schrodinger equation** of the system:

$$i\hbar \frac{\partial}{\partial t}\Psi = H\Psi, \tag{3.2.1}$$

where H is the **Hamiltonian**, which is a linear differential operator. It can be seen as a quantization of the Hamiltonian of classical mechanics, as we will see later on.

Remark 3.2.1 There is something important here that is not often mentioned: quantum mechanics is in many ways easier than classical mechanics. That is because in quantum mechanics, we are solving a linear differential equation to determine the wave-function Ψ. So the superposition principle applies; given two solutions Ψ_1 and Ψ_2 of the Schrodinger equation, any linear combination of those is also a solution. So we can really think of the vector space of solutions (the Hilbert space of states), and think of the Hamiltonian as an operator on that space (if it is finite-dimensional, we can choose a basis and represent it as a matrix).

On the contrary, in classical mechanics, to solve a problem one needs to solve the equations of motion for, say, the position function $x(t)$. Newton's law can be rewritten as

$$m\frac{d^2}{dt^2}\vec{x}(t) = -\nabla V(\vec{x}).$$

These equations of motion for $\vec{x}(t)$are generally non-linear, except for the case of harmonic motion, i.e. masses connected by springs, in which case $\nabla V(\vec{x})$ is linear in $\vec{x}$. So in general, the superposition principle does not apply in classical mechanics, and linear combinations of solutions are not solutions of the equations of motion.

This is why group and representation theory are much more useful and powerful in quantum mechanics than in classical mechanics.

To solve Schrodinger equation we do a separation of variables. We write $\Psi = \psi e^{-iEt/\hbar}$, where ψ does not depend on time. Substituting back in (3.2.1), we get:

$$H\psi = E\psi,$$

which is an eigenvalue problem for the Hamiltonian of the system. E here is the energy of the system. This is the key equation that we need to solve.

3.2.2 The symmetry group of the Hamiltonian

The power of group and representation theory is that we will be allowed to characterize the form of the space of solutions for a given quantum mechanical problem without ever solving the eigenvalue problem. All that we will use are the symmetry properties of the system.

Suppose that R is another operator acting on the space of states. We know how an operator transforms: it transforms by a similar transformation. Thus, the Hamiltonian transforms as

$$H' = R^{-1}HR.$$

We say that the Hamiltonian is **invariant** under R if $H' = H$, that is,

$$H = R^{-1}HR,$$

or, equivalently,

$$RH = HR.$$

We can look at all such operators R that commute with H: those form a group G, which we call the **symmetry group of the Hamiltonian**.

More precisely, we could start with the symmetry group G, realized in terms of properties of the Hamiltonian (for instance some explicit symmetry of the system, such as periodicity). Then the operators R would form a representation of the group G, acting on the space of states of the system.

3.2.3 Irreducible representations and degeneracy of states

Now, the key is to realize that two eigenfunctions that are mapped into each other by a symmetry transformation R must have the same energy eigenvalue E. Indeed, if $H\psi = E\psi$, then

$$H(R\psi) = R(H\psi) = R(E\psi) = E(R\psi).$$

So what this means is that we can group the eigenfunctions of the systems into subspaces that are invariant under the action of the symmetry group; all eigenfunctions in each of these subspaces share the same eigenvalue.

This statement follows neatly from representation theory. If R is an irreducible presentation, then, since $HR = RH$, by Schur's lemma we must have $H = EI$ for some number E. What this means is that all states have the same energy eivenvalue. However, in general R is not irreducible. But one can choose a basis for the space of states wuch that R decomposes as direct sum of irreducible representations $R = \bigoplus_\alpha R^{(\alpha)}$, and by Schur's lemma, it follows that with respect to this choice of basis the Hamiltonian also decomposes as a direct sum of diagonal matrices $H = \bigoplus_\alpha E^{(\alpha)} I$, where I is the identity matrix of the size of the corresponding irreducible representation $R^{(\alpha)}$.

To summarize, what this means is that *we can choose a basis for the space of states, composed of eigenfunctions of the Hamiltonian, in which the state space decomposes as a direct sum of subspaces that transform according to irreducible representations of the symmetry group. All states in each such subspace are eigenfunctions of the Hamiltonian with the same energy eigenvalue.*

This leads to the very important physical concept of degeneracy.

Definition 3.2.2 Degeneracy of states. We say that two eigenfunctions of the Hamiltonian are **degenerate** if they share the same energy eigenvalue. ♢

At first, in the history of physics, when eigenstates that shared the same energy eigenvalues were discovered, it was very puzzling to physicists. What should states have the same energy eigenvalue? When you diagonalize a generic matrix, you expect distinct eigenvalues.

But from the viewpoint of the representation theory, degeneracy is expected, and in fact can be constrained using representation theory. Indeed, we know that the "degree of degeneracy" of an eigenfunction (that is, the number of degenerate states) is at least as large as the dimension of the irreducible representation that it belongs to. It can be larger if two eigenspaces corresponding to different irreducible representations have the same energy eigenvalue; we call such degeneracy "accidental", since it is not determined by symmetry. However, most degeneracies are determined by symmetry.

This provides a beautiful interplay between mathematics and physics. On the one hand, if we know the group of symmetry of a system, we can determine the possible degeneracies, simply by enumerating its irreducible representations and their dimensions. On the other hand, if we observe experimentally that say 8 states are degenerate (as was observed in the 50s when eight baryons were observed with approximately the same mass), then we can use this information to restrict the possible groups of symmetries: assuming that the degeneracy is not accidental, the group of symmetries must be such that it has an 8-dimensional irreducible representation.

The upshot of this discussion could be summarized as follows:

- To each energy eigenvalue we can associate a corresponding irreducible representation of the symmetry group of the Hamiltonian. The degeneracy of the eigenvalue is the dimension of the irreducible representation (it could be larger if there is accidental degeneracy).
- In the basis for the space of states given by eigenfunctions of the Hamiltonian, the state space decomposes as a direct sum of subspaces that transform according to irreducible representations of the symmetry group. All eigenfunctions in a given subspace share the same energy eigenvalue. Thus, we can label the eigenfunctions of the system using the irreducible representations of the symmetry group, which we call in physics the **quantum numbers**.

3.2.4 Examples

We will look at two simple examples of the construction. But most relevant examples to quantum mechanics involve continuous groups and their representations, thus for this we will need to wait until later.

Example 3.2.3 Parity. Let us start with a very simple one-dimensional example. The Hamiltonian of a one-dimensional system takes the form $H = -\frac{1}{2}\frac{d^2}{dx^2} + V(x)$ for some potential function $V(x)$. Now suppose that $V(-x) = V(x)$. Then the mapping $x \mapsto -x$ leaves the Hamiltonian H invariant. This symmetry group is the order two group $\mathbb{Z}_2 = \{e, a\}$.

This means that we can label the eigenfunctions of the system according to irreducible representations of the symmetry group. Here, $\mathbb{Z}_2$ has only two one-dimensional irreducible representations: the trivial representation (given by $T(e) = T(a) = 1$), and the parity representation (given by $P(e) = 1$, $P(a) = -1$). Those act on states of the system (eigenfunctions). States that transform according to the trivial representation satisfy:

$$\psi(-x) = T(a)\psi(x) = \psi(x),$$

that is, they are even. States that transform according to the parity representation transform as:

$$\psi(-x) = S(a)\psi(x) = -\psi(x),$$

that is, they are odd. The upshot of the representation theory is that we can label the states according to whether they are even or odd.

We also know that there is no degeneracy here (barring accidental degeneracy), since all irreducible representations of $\mathbb{Z}_2$ are one-dimensional.

This example is of course almost trivial: the goal was just to get familiar with how to use the language of representation theory to understand states of a quantum mechanical system. □

Example 3.2.4 Bloch's theorem. The second example is more interesting, and in fact is of fundamental importance in solid state physics. We stick with a one-dimensional system, but we consider a potential $V(x)$ that is periodic: $V(x+a) = V(x)$ for some constant a. We consider the mapping $T : x \to x+a$. This generates a discrete, infinite, abelian group G.

Since G is abelian, it only has one-dimensional irreducible representations. So as before, there is no degeneracy of states. We can label states using irreducible representations. Those representations are unitary. So for a one-dimensional irreducible representation S_k we get the transformation:

$$\psi(x+a) = S_k\psi(x) = e^{ika}\psi(x),$$

for some real number k. Since $e^{ika} = e^{ika+i2\pi}$, the representations are indexed by a real number k such that

$$-\frac{\pi}{a} \le k < \frac{\pi}{a}.$$

So the states of the system can be labeled by such a real number. This is an example of a Brillouin zone.

We can look at the special case of a one-dimensional crystal (note that it can be easily generalized to more than one dimension) with a finite number of sites, say N. We can impose a periodic boundary condition, meaning that if we translate from the Nth site, we go back to the first one. Thus the group of symmetry is generated by the translation T, but with the condition $T^n = e$. This group is the cyclic group $\mathbb{Z}_N$.

The abelian group $\mathbb{Z}_N$ has N one-dimensional irreducible representations S_k. Those are given by representing the generator T by the roots of unity $S_k(T) = w^k$, for $k = 0, 1, \ldots, N-1$, where $w = e^{2\pi i/N}$. So we can label the states according to how they transform. For instance, a state ψ_k that transforms in the kth irreducible representation would transform as

$$\psi_k(x+a) = S_k(T)\psi_k(x) = e^{2\pi ik/N}\psi_k(x). \tag{3.2.2}$$

Now let us write the state $\psi_k(x)$ as

$$\psi_k(x) = e^{i\phi_k(x)}u_k(x), \tag{3.2.3}$$

where $u_k(x+a) = u_k(x)$ and $\phi_k(x)$ is some phase factor. Because of (3.2.2), we must have:

$$\psi_k(x+a) = e^{i\phi_k(x+a)}u(x) = e^{2\pi ik/N}\psi_k(x) = e^{2\pi ik/N + i\phi_k(x)}u(x).$$

Thus we must have

$$\phi_k(x+a) = \frac{2\pi k}{N} + \phi_k(x), \tag{3.2.4}$$

and, iterating, $\phi_k(x+ma) = \frac{2\pi km}{N} + \phi_k(x)$. Thus ϕ_k is a linear function of m, and hence of $x+ma$. Therefore we can write $\phi_k(x) = Ax + B$ for some A and B. Substituting back in (3.2.4), we get:

$$A = \frac{2\pi k}{Na}.$$

Thus $\phi_k(x) = \frac{2\pi k}{Na}x + B$. Subsituting back in (3.2.3), and redefining $u_k(x)$ to include in it the arbitrary constant B, we get:

$$\psi_k(x) = e^{\frac{2\pi ikx}{Na}}u_k(x).$$

This is known as a Bloch wave: it is a periodic function $u_k(x)$ multiplied by a plane wave. Representation theory tells us that the quantum mechanical states of a periodic lattice have this form. This is pretty much the statement of Bloch's theorem: that a basis of wavefunctions is given by Bloch waves, and that these wavefunctions are energy eigenstates of the system.

The amazing thing is that this result follows directly from representation theory, without ever solving the Schrodinger equation! All that we used is the periodic symmetry of the lattice. The particular form of the potential would then be needed to solve for the periodic functions $u_k(x)$ and identity the particular energy eigenvalues corresponding to the eigenstates of the system. □

3.3 Coupled harmonic oscillators

Objectives

You should be able to:

- Determine the form of the normal modes of a system of masses tied by springs with particular symmetries using representation theory.

In the previous section we argued that quantum mechanics was somewhat easier than classical mechanics, because Schrodinger equation is linear, and hence representation theory naturally plays a role. But there is one particular context in classical mechanics where the equation of motion is linear: it is the case of a system of coupled harmonic oscillators. As Michael Peskin, a famous physicist, said: "Physics is that subset of human experience which can be reduced to coupled harmonic oscillators." The dynamics of systems of coupled harmonic oscillators is crucial in physics, and in fact forms the basis of much of quantum mechanics and quantum field theory.

In classical mechanics, one can think of a system of coupled harmonic oscillators as a system of masses connected by springs. But in fact what we will do in this section applies more generally to study oscillations of any physical system about a stable equilibrium point, as you may have studied in your classical mechanics course.

Going back to the spring-mass system, the system is governed by Hooke's law, which says that the force on the mass by the spring is proportional to its displacement from equilibrium. For instance, if we consider the very simple system of a mass moving in one dimension, and attached by a spring to a fixed wall, then the equation of motion would be

$$m \frac{d^2}{dt^2} x = -kx,$$

where m is the mass, k is Hooke's constant, and x is the displacement of the mass from its equilibrium position. The goal of classical mechanics is to solve the equation of motion to find the displacement x as a function of time. In this case, we see that the equation of motion is linear. We know that the general solution will take the form $x(t) = C\cos(\omega t + \delta)$, with $\omega = \sqrt{k/m}$.

For a more general system of N masses moving in D dimensions, each mass can move in D independent directions, and hence the displacement of the whole system is given by N D-dimensional displacement vectors. All-in-all, this can be packaged in a ND-dimensional real vector, which we call $\vec{\eta}(t)$. The equation of motions will look like (in vector form):

$$\frac{d^2}{dt^2}\vec{\eta} = -H\vec{\eta}, \tag{3.3.1}$$

for some $ND \times ND$ matrix of coefficients H. The key point here is that the equation is still linear. But given a particular spring-mass system, the precise form of the matrix H may be mess, and in fact not so easy to write down. The goal of this section will be to extract a lot of information about the motion of the system without ever having to write down H explicitly. As you could guess, what we will use is simply the symmetries of the system, and representation theory.

3.3.1 Normal modes

To analyze the behaviour of spring-mass systems in more than one dimension, one needs to understand the concept of ***normal modes*** of the system.

Definition 3.3.1 Normal modes. We say that a system is in a ***normal mode*** if all its parts are oscillating with the same frequency. ◊

Let us go back to (3.3.1). To solve this equation, we will assume that normal modes exist, and see what they look like. We take the ansatz

$$\vec{\eta}(t) = \vec{a}\cos(\omega t + \delta),$$

for some positive constant ω and a vector of coefficients $\vec{a}$ (for the case $\omega = 0$, we would take the ansatz $\vec{\eta}(t) = \vec{a}t + \vec{b}$). Plugging back in (3.3.1), we obtain the equation

$$H\vec{a} = \omega^2\vec{a},$$

which is nothing but an eigenvalue problem for the $ND \times ND$ matrix of coefficients H!

In other words, what we have found is that the normal modes are given by eigenvectors of H, with oscillating frequency given by the square root of their eigenvalues. Because the equations of motion are linear, we can conclude that the general motion of the system will be given by a linear combination (superposition) of these ND normal modes.

A few things to note: for the motion to be oscillatory, all eigenvalues of H must be non-negative (otherwise we would have an unstable equilibrium point). For a normal mode with positive ω, all masses are oscillating with the same frequency ω. For a normal mode with a zero eigenvalue, the corresponding motion is uniform translation of all masses.

The upshot of this discussion is that to understand the motion of the system, we need to find its normal modes, and to find its normal modes, we need to find the eigenvectors of the $ND \times ND$ matrix H. This is where representation theory comes into play!

3.3.2 Representation theory to the rescue

Our goal is now to find the eigenvectors and eigenvalues of H. As mentioned previously, for large systems of masses, H may be a mess. Can we get information about the normal modes without writing H explicitly?

The idea is to use symmetries of the system. We consider the group G of symmetries that leave the system invariant. Concretely, this means that it leaves H invariant. As before, this means that G acts on the space of solutions as an ND-dimensional representation that commutes with H. This representation is in general reducible, but one can choose a basis for the space of states such that it decomposes as a direct sum of irreducible representations. By Schur's lemma, then we know that in this basis, the matrix H becomes diagonal. Thus the basis vectors are eigenvectors of H. Moreover, the space of states also decomposes as a direct sum of subspaces that transform according to the irreducible representations in the decomposition, and any two states in the same subspace share the same eigenvalue.

So given a particular spring-mass system, to understand the different normal modes, or eigenstates of the system, one proceeds as follows:

- We find the symmetry group G of the system, and how it acts on the space of states as an ND-dimensional (reducible) representation (in fact we only need its characters).
- We use Theorem 2.9.2 to calculate how it decomposes as a direct sum of irreducible representations.
- We know that in this basis, the matrix H becomes diagonal, with eigenvalues appearing with degeneracy equal to the dimension of the corresponding irreducible representation. The different eigenvalues give us the possible different normal modes of the system.
- We use physical insight or further calculations to determine what these normal modes are.

3.3.3 Examples

In this section we consider two examples. The first one is very simple, and is studied simply to set the stage; the second one is more involved.

Example 3.3.2 Two masses connected by a spring in one dimension. We start with the simple example of two masses connected by a spring and restricted to move in one dimension. We want to understand the normal modes of the system. In this case we can do everything explicitly, so let us start by doing that. But the goal of this example is to show how representation theory can be used to gain information on the system without solving explicitly.

Let $\eta_1(t)$ and $\eta_2(t)$ be the displacements of the two masses. The equations of motion are (setting the constants $m_1 = m_2 = k = 1$ for simplicity):

$$\frac{d^2}{dt^2}\eta_1 = -(\eta_1 - \eta_2), \qquad \frac{d^2}{dt^2}\eta_2 = -(\eta_2 - \eta_1).$$

In matrix form, those reads:

$$\frac{d^2}{dt^2}\vec{\eta} = -\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\vec{\eta}.$$

We could of course solve this equation explicitly here. But let us use representation theory instead. To get the normal modes we want to find eigenvectors for the matrix

$$H = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{3.3.2}$$

We want to do so using symmetries of the system.

The symmetry group G here is rather simple. It has two elements: the identity, and the exchange of the two masses. Thus $G \cong \mathbb{Z}_2$. It acts on the two-dimensional space of states in a two-dimensional representation T. We can write down the two-dimensional representation explicitly (we will do that below for completeness), but to find its decomposition in terms of irreducible representations we only need its characters. The identity element has character $\chi(e) = 2$, while the non-trivial element does not leave any mass fixed, hence it must have character $\chi(a) = 0$. For completeness, the two-dimensional representation is given by the matrices:

$$T(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T(a) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now let us decompose T as a direct sum of irreducible representations. Using the character table for $\mathbb{Z}_2$, and the decomposition theorem Theorem 2.9.2, we can find the decomposition of T. If we denote by $T^{(1)}$ the trivial representation, and $T^{(2)}$ the parity representation (the other irreducible representation of $\mathbb{Z}_2$), then we can write $T = m_1 T^{(1)} \oplus m_2 T^{(2)}$. From Theorem 2.9.2, we find the coefficients to be:

$$m_1 = \frac{1}{2}((2)(1) + (0)(1)) = 1,$$
$$m_2 = \frac{1}{2}((2)(1) + (0)(-1)) = 1.$$

Thus $T = T^{(1)} \oplus T^{(2)}$. In this basis the representation is given by diagonal matrices.

By Schur's lemma, we know that in this basis H is also diagonal, i.e. the basis vectors are eigenvectors of H, and hence give normal modes of the system. The eigenvalues here are non-degenerate since the irreducible representations are one-dimensional. We do not know from symmetry alone however what the diagonal entries of H should be; those correspond to the eigenvalues (of frequencies of the normal modes).

But since here we know the specific form of H (see (3.3.2)), and see that the trace is 2 and the determinant is 0, the diagonal form of H, must be

$$H = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

So there is one normal mode with zero frequency, and one normal mode with a non-zero frequency. It is easy to identify these normal modes physically. The first mode corresponds to linear translation of the whole system (the spring is not stretched at all). The second mode corresponds to the masses oscillating with the same frequency in opposite directions (the "breathing mode").

Of course this example was rather trivial. We could have solved the equations of motion explicitly. But in more complicated systems, H may be difficult to write down, and studying symmetries provides valuable information on the normal modes of the system without having to solve any equation of motion. □

Example 3.3.3 The triangular molecule. We now consider the more interesting example of 3 masses in two dimensions connected by three springs in the form of an equilateral triangle. Each mass has two displacement coordinates (horizontal and vertical directions), and hence the displacement vector is $3 \times 2 = 6$-dimensional.

Writing down the matrix H for this system would be pretty annoying. So let us not do that. We will use symmetry instead to gain information on the normal modes of the system.

The symmetry group here is the dihedral group D_3 which is isomorphic to the symmetric group S_3. It acts on the space of states in the form of a 6-dimensional representation T. We will not write down the representation explicitly, but simply deduce its characters on the three conjugacy classes of S_3. We will use the notation in Section 2.10.

In the conjugacy class C_1 containing the identity element, the character of

T is $\chi_1 = 6$. In the conjugacy class C_2 corresponding to the transpositions, the action should leave one mass fixed, and hence its two displacement coordinates invariant. Therefore, the character must be $\chi_2 = 2$. For the class C_3 corresponding to the cyclic permutations, nothing is left invariant, and hence $\chi_3 = 0$.

From this we can get the decomposition of T. Using Theorem 2.9.2 and the character table Table 2.10.2 (with the notation that $T^{(1)}$ is the trivial irrep, $T^{(2)}$ is the parity one-dimensional irrep, and $T^{(3)}$ is the two-dimensional irrep) we get the coefficients:

$$m_1 = \frac{1}{6}(1(6)(1) + 3(2)(1) + 2(0)(1)) = 2,$$
$$m_2 = \frac{1}{6}(1(6)(1) + 3(2)(-1) + 2(0)(1)) = 0,$$
$$m_3 = \frac{1}{6}(1(6)(2) + 3(2)(0) + 2(0)(-1)) = 2.$$

Thus $T = T^{(1)} \oplus T^{(1)} \oplus T^{(3)} \oplus T^{(3)}$. So in this basis the matrix H takes the diagonal form:

$$H = \text{diag}(\omega_1, \omega_2, \omega_3, \omega_3, \omega_4, \omega_4),$$

for some eigenvalues $\omega_1, \omega_2, \omega_3, \omega_4$. Thus there should be four types of normal modes, two of which would have a two-dimensional eigenspace. At least that's what we get from symmetry; there could be additional accidental degeneracies (i.e. some of the ω_i could be identical).

In fact one can identify the zero modes physically. Two of them correspond to horizontal and vertical translations of the whole system; those have zero frequency. This corresponds to one of the two-dimensional irreps. There is in fact an accidental degeneracy here; there is another mode with zero frequency, corresponding to cyclic rotation of the whole system. This corresponds to one of the one-dimensional irreps. So there are three zero modes for this system.

We still have one one-dimensonal irrep and one two-dimensional irrep. There is one "breathing mode", where all masses oscillate inwards and outwards simultaneously. This corresponds to the remaining one-dimensional irrep. The two other normal modes are not so easy to guess, but one could calculate the eigenvectors from linear algebra, since they should be orthogonal to the four eigenvectors that we have already determined, and then see how they act on the three masses. See for instance Section III.2 in Zee's book. □

3.4 The Lagrangian

Objectives

You should be able to:

- Use symmetries to specify the form of a Lagrangian, and explain how symmetries are related to conserved quantities.

Symmetries also play a very important role in determining the physical laws of a system. In modern physics, we usually formulate classical and quantum mechanics in terms of either a Lagrangian or Hamiltonian. In this section we briefly review the Lagrangian formalism, and show how symmetries play an important role.

3.4.1 The Lagrangian and classical mechanics

In the Lagrangian formulation of classical mechanics, one starts with the **Lagrangian** L of the system, which is defined as $L = T - V$, where T is the kinetic energy of the system and V is its potential energy. We then construct the **action** of the system from time t_1 to t_2:

$$S = \int_{t_1}^{t_2} L dt.$$

Hamilton's principle then states that the motion of the system from time t_1 to t_2 is such that the action has a stationary value (is extremized). This is a very profound statement. Out of all possible paths in configuration space, Nature chooses the path which extremizes the action. The physics is entirely encoded in its Lagrangian; the equations of motion can be obtained by extremizing the action. Isn't that beautiful? All of classical mechanics can be summarized in this neat variational principle.

Using variational calculus, one finds that the action S has a stationary value if and only if the **Euler-Lagrange** equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

are satisfied, for $i = 1, \ldots, s$, where the q_i are coordinates for the system, and $\dot{q}_i$ denotes time derivative of the coordinates.

Thus the physics of a system in classical mechanics is encoded in a single object: its Lagrangian. This makes understanding symmetries of the theory much easier. In terms of the equations of motion, one would say that a system has a symmetry if both sides of the equations of motion transform in the same way, so that the equations are invariant under the symmetry operation. But here the statement is much simpler. A system if symmetric if its action is invariant under the symmetry operation. Or, in the language of representation theory, the action must transform in the trivial representation of the symmetry group. In fact, in most cases, invariance of the action comes from invariance of the Lagrangian itself.

In fact, in modern physics one often starts with the desired symmetries of a given physical system, and use this to determine the form of the Lagrangian, by writing down the most general action which is invariant under these symmetries. Let me give a very simple example. Consider a free particle in three-dimensional space. Homogeneity of space and time implies that L can only depend on $|\vec{v}|^2$: it cannot depend explicitly on $\vec{x}$ or t, nor can it depend on the direction of $\vec{v}$. Then, one can show that L must be proportional to $|\vec{v}|^2$ by requiring that it is invariant under Galilean transformations of the frame of reference. This constant of proportionality defines the mass of the particle, which can be shown to be positive, since otherwise the action would have no extremum. Thus, simply by requiring invariance under given symmetries, we are able to recover the Lagrangian of the system! The same can be done for a system of N particles, see "Mechanics" by Landau and Lifschitz, Sections 1.3 and 1.4.

This approach is what is done for instance in particle physics. Hamilton's principle (and its quantum version) is postulated. We then study the symmetries of the system, and figure out the most general Lagrangian that is compatible with these symmetries. This Lagrangian will involve a certain number of constants; we then use experiments to fix the value of these constants.

3.4.2 The Lagrangian and quantum physics

Talking about particle physics, I should mention here that quantum mechanics and quantum field theory can also be obtained from the Lagrangian of the physics. This beautiful approach is due to Feynman.

While classical mechanics can be summarized by saying that the system takes the path in configuration space which extremizes the action, quantum mechanics can be summarized by saying that the systems takes *all paths* in configuration space, each of which is weighted by the exponential of the action. A simple argument for this goes as follows. Consider the standard double-slit experiment. A particle emitted from a source at time t_1 passes through one or the other of two holes, drilled in a screen, and is detected at time t_2 by a detector located on the other side of the screen. In quantum mechanics, the amplitude for detection is given by the sum of the amplitudes corresponding to the two paths that the particle can take. But then Feynman asked the following. First, what happens if there are three holes in the screen? Well, then the amplitude is the sum of the three paths. What if there are four holes? Well, it must be the sum of the four paths. And what if I put another screen with some holes in it? Then the amplitude becomes the sum of all possible paths. But then, what if I have an infinity of such screens with an infinity of holes in them, such that the screens are actually no longer there??? The only logical conclusion is that the amplitude then must be the sum of all paths from the source to the detector. It is easy to show (in quantum mechanics) that each path must be weighted by a factor of

$$e^{\frac{i}{\hbar}S},$$

where S is the action of the system. This infinite sum over all paths is what is called in quantum physics a **path integral**. When $\hbar \to 0$, the classical limit is recovered, since the path integral "localizes" on the classical configuration given by the extremum of S. This gives a beautiful conceptual explanation of the difference between quantum and classical physics.

We have just seen that quantum physics can be understood neatly in terms of the action of a system. In fact, all of the fundamental laws of physics can be written in terms of an action principle. This includes electromagnetism, general relativity, the Standard Model of particle physics, and string theory. For instance, almost everything we know about Nature can be captured in the Lagrangian

$$L = \sqrt{g}(R + \frac{1}{2}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\not{D}\psi),$$

where the first term is the Einstein term, the second the Yang-Mills (or Maxwell) term, and the last the Dirac term. Those describe gravity, the forces of Nature (like electromagnetism and the nuclear forces), and the dynamics of particles like electrons and quarks. You are welcome to try to understand what these terms really mean! :-)

3.4.3 Symmetries and conservation laws

Going back to classical mechanics (although the discussion here has a parallel in quantum physics), from the Lagrangian of a system one can extract the equations of motion. However, for most systems in Nature, these equations are not integrable, which means that they can only be solved numerically (with a big enough computer).

In many cases however a great deal of information can be obtained by studying **conserved quantities**, which are quantities whose values remain constant during motion. These conserved quantities can often be found relatively easily and then used to reduce the differential system to a simpler one which is easier to solve. In fact, in many cases these conserved quantities are even more interesting than a full solution to the equations of motion. It is therefore of interest to learn how to find conserved quantities.

The key result here is that we can associate conserved quantities to symmetries of the system! This is the fundamental statement of **Noether's theorem**: For any continuous symmetry of the action of a system, there is a corresponding conserved quantity. And this is not just an abstract statement; we can construct this conserved quantity explicitly. For instance, invariance under time translations gives rise to conservation of energy. Invariance under space translations gives rise to conservation of momentum. Invariance under space rotations gives rise to conservation of angular momentum. Once again, group theory is fundamental!

Chapter 4

Lie groups and Lie algebras

In this section we introduce Lie groups and Lie algebras.

4.1 Lie groups

Objectives

You should be able to:

- Recall the definition of Lie groups and matrix Lie groups.
- Relate the interpretation of Lie groups as groups of matrices, groups of continuous transformations of a target space, and in terms of the manifold structure of their parameter spaces.
- Identify the fundamental representation of a matrix Lie group.

In the first few sections we discussed at length finite groups and their representations. But in the beginning we also introduced examples of infinite and continuous groups. In this section we go back to the study of continuous groups, and focus on those that also satisfy a differentiability criterion, known as Lie groups. But in fact, all examples of continuous groups that we have seen previously are also examples of Lie groups.

4.1.1 Continuous (topological) groups and Lie groups: abstract definition

Let us go back to discrete groups for a moment. What is a discrete group? It is a discrete set G, with a binary operation (that we call group multiplication) satisfying a bunch of axioms. In particular, the binary operation closes, so we can think of it as a mapping $\alpha : G \times G \to G$, given by $\alpha : (a, b) \mapsto ab$, where $a, b \in G$ and on the right-hand-side we mean group multiplication of a and b. We can also define another mapping $\beta : G \to G$, which maps elements to their inverses: $\beta : g \mapsto g^{-1}$ for $g \in G$. So we encode the structure of a group in terms of the mappings α and β.

Now, what is a continuous group? The idea is to replace the discrete set G by a continuous space M of a certain dimension, which we call the **dimension** of the continuous group. More precisely, to define a continuous group we need to define the notion of continuity. The precise statement is that we add to the set M the extra structure of a topology (i.e. a notion of open sets, or "proximity"), so that it becomes a topological space. Then, we require that the mappings α and β be continuous with respect to this topology. This defines the concept of a **topological group**, which is what we really mean by "continuous group".

The idea of Lie groups is to impose the extra structure of differentiability. To define what it means to be "differentiable", we introduce an extra structure on the set M: we require that M be a smooth (i.e. C^∞) manifold. Then, we require that the mappings $\alpha : M \times M \to M$ and $\beta : M \to M$ be smooth maps.

Definition 4.1.1 Lie groups. A **Lie group** is a real smooth (C^∞) manifold M that is also a group, in which the group operations of multiplication $\alpha : M \times M \to M$ and inversion $\beta : M \to M$ are smooth maps. ◊

This is a nice, abstract definition, but it requires an understanding of manifolds and differential geometry. Fortunately, as we will see in practice we will not need to know much about differential geometry, as the Lie groups that we will work with will all be realized as groups of matrices that depend continuously on a set of parameters. But let us not get ahead of ourselves: let us first work through a few simple examples.

Example 4.1.2 A one-dimensional Lie group. Consider the set $\mathbb{R}^*$, with binary operation given by multiplication. It is easy to see that it satisfies the axioms of a group. Moreover, $\mathbb{R}^*$ is a smooth manifold, and the group operations of multiplication and inversion are given by the mappings $\alpha : (a, b) \mapsto ab$ and $\beta : a \to \frac{1}{a}$, for $a, b \in \mathbb{R}^*$, which are smooth maps. Thus $\mathbb{R}^*$ is a one-dimensional Lie group. □

Example 4.1.3 The circle group. Consider the one-dimensional circle S^1, which can be parameterized by an angle $\theta \in [0, 2\pi)$. This is a group under the operation given by addition of angles. Again, S^1 is a smooth manifold, and the group operations, given by $\alpha : (\theta_1, \theta_2) \mapsto \theta_1 + \theta_2 \pmod{2\pi}$ and $\beta : \theta_1 \mapsto -\theta_1 \pmod{2\pi}$ are smooth maps. Thus S^1 is a one-dimensional Lie group. □

Example 4.1.4 The affine group of one dimension. Let us now consider the set of matrices of the form

$$T = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},$$

with $a \in \mathbb{R}_{>0}$ and $b \in \mathbb{R}$. This set forms a group under matrix multiplication. Its underlying manifold is given by $\mathbb{R}_{>0} \times \mathbb{R}$, which is a smooth manifold. Using matrix multiplication, we see that

$$T_1 T_2 = \begin{pmatrix} a_1 & b_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 a_2 & a_1 b_2 + b_1 \\ 0 & 1 \end{pmatrix},$$

and hence the group operation on $\mathbb{R}_{>0} \times \mathbb{R}$ can be written as:

$$\alpha : ((a_1, b_1), (a_2, b_2)) \mapsto (a_1 a_2, a_1 b_2 + b_1),$$

which is smooth. As for inverses, given a matrix

$$T = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},$$

its inverse is

$$T = \begin{pmatrix} \frac{1}{a} & -\frac{b}{a} \\ 0 & 1 \end{pmatrix}.$$

Thus the group inversion on $\mathbb{R}_{>0} \times \mathbb{R}$ can be written as:

$$\beta : (a, b) \mapsto \left(\frac{1}{a}, -\frac{b}{a} \right),$$

which is also differentiable since $a \in \mathbb{R}_{>0}$. Therefore, this is a two-dimensional Lie group. □

Now that we have endowed our groups with the extra structure of a topology, we can define what it means for a group to be compact.

Definition 4.1.5 Compact groups. We say that a topological group is **compact** if its topology is compact. ♢

Of the three examples that we saw above, only the circle group S^1 is compact.

4.1.2 Matrix Lie groups

The abstract definition of Lie groups as groups that are also smooth manifolds is nice, but in practice it is a bit too abstract for our purposes. All the Lie groups that we will study in this class have very explicit realizations as matrix groups. To study those, we need to introduce the mother of all Lie groups: the general linear group, which we have already encountered many times.

Example 4.1.6 $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$. Consider the group $GL(n, \mathbb{R})$ of real $n \times n$ invertible matrices. Real invertible matrices have n^2 real independent entries. One can show that this is a Lie group of dimension n^2. The details of the proof are beyond the scope of this class, but let us briefly sketch the argument. Let $M_n(\mathbb{R})$ be the set of $n \times n$ matrices. Then the subset of real invertible matrices $GL(n, \mathbb{R})$ is obtained by requiring that the determinant is non-zero. But since the determinant is a polynomial map, one can show that the subset $GL(n, \mathbb{R})$ forms an open set in $M_n(\mathbb{R})$ (with respect to the Zarisky topology). Using this, one can conclude that $GL(n, \mathbb{R})$ is a Lie group of dimension n^2.

Using a similar line of reasoning, one can show that $GL(n, \mathbb{C})$ is also a Lie group, with real dimension $2n^2$ (since complex invertible matrices have n^2 independent complex entries, that is, $2n^2$ independent real parameters). □

With this example, we can construct many Lie groups, because of the following crucial theorem, known as the "closed subgroup theorem":

Theorem 4.1.7 The closed subgroup theorem. *If H is a closed subgroup of a Lie group G, then H is also a Lie group (with the smooth structure agreeing with the embedding). In particular, all closed subgroups of $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$ are Lie groups. Those are known as **matrix Lie groups**.*

We will not prove this theorem here, as it is beyond the scope of this class. But the key is that it implies that all our favourite continuous groups are Lie groups.

Example 4.1.8 $SL(n, \mathbb{R})$ and $SL(n, \mathbb{C})$. The group $SL(n, \mathbb{R})$ of real matrices with determinant one is a closed subgroup of $GL(n, \mathbb{R})$, and hence a Lie group. Imposing that the determinant is equal to one removes one degree of freedom, and hence the dimension of $SL(n, \mathbb{R})$ is $n^2 - 1$.

Similarly, the group $SL(n, \mathbb{C})$ of complex-valued matrices with determinant one is a closed subgroup of $GL(n, \mathbb{C})$, and hence a Lie group. Again, imposing that the determinant is equal to one fixes one constraint, but in this case it reduces the number of complex degrees of freedom by one, and hence the number of real degrees of freedom by two. So the real dimension of $SL(n, \mathbb{C})$

is $2(n^2 - 1)$.

Note that $SL(n, \mathbb{R})$ and $SL(n, \mathbb{C})$ are non-compact. □

Example 4.1.9 $O(n)$ **and** $SO(n)$**.** The group $O(n)$ of real orthogonal matrices is a closed subgroup of $GL(n, \mathbb{R})$, and hence a Lie group. The number of independent real parameters in an $n \times n$ orthogonal matrix can be calculated to be $\frac{1}{2}n(n-1)$, which is the dimension of the Lie group.

The group $SO(n)$ of special orthogonal matrices (with determinant one) is also a closed subgroup of $GL(n, \mathbb{R})$, and hence a Lie group. In this case, imposing that the determinant of an orthogonal matrix A is equal to one is a discrete condition, as any orthogonal matrix must satisfy $\det A = \pm 1$ to start with. Thus it does not reduce the number of degrees of freedom, and hence the dimension of $SO(n)$ is also $\frac{1}{2}n(n-1)$.

The orthogonal groups $O(n)$ and special orthogonal groups $SO(n)$ are compact Lie groups. □

Example 4.1.10 $U(n)$ **and** $SU(n)$**.** The group $U(n)$ of unitary matrices is a closed subgroup of $GL(n, \mathbb{C})$, and hence a Lie group. The number of real independent parameters in an $n \times n$ unitary matrix can be calculated to be n^2, which is the dimension of $U(n)$.

The special unitary group $SU(n)$ is also a closed subgroup of $GL(n, \mathbb{C})$, and hence a Lie group. Imposing the condition that the determinant is equal to one reduces the number of real degrees of freedom by one. Indeed, any unitary matrix A has $|\det A|^2 = 1$, which means that its determinant can be written as $\det A = e^{i\theta}$ for some real number θ. Imposing that $\det A = 1$ fixes this real parameter θ. As a result, the real dimension of $SU(n)$ is $n^2 - 1$.

Just as for orthogonal groups, both $U(n)$ and $SU(n)$ are compact Lie groups. □

In this course we will focus on matrix Lie groups, such as the orthogonal and unitary groups, which are realized as closed subgroups of the general linear groups. This is why we will rarely need to go back to the formal abstract definition of Lie groups.

4.1.3 Lie groups as groups of transformations of a target space

We first introduced Lie groups abstractly, as groups that are also smooth manifolds. We then studied an important class of Lie groups, known as matrix Lie groups, that are realized as closed subgroups of the general linear groups. But concretely, we often think of Lie groups in a third way, as groups of transformations of a target space. For instance, we often think of $SO(n)$ as the group of rotations in $\mathbb{R}^n$. Let us make these statements a little more precise.

Suppose that we are interested in a group of continuous transformations of some space, which we call the **target space** T:

$$L : T \to T.$$

If the target space has dimension d, and we use coordinates $x_1, \ldots, x_d$ on T, we can write these transformations as:

$$x_i' = g_i(x_1, \ldots, x_d; a_1, \ldots, a_n), \qquad i = 1, \ldots, d,$$

where the $a_1, \ldots, a_n$ are parameters for the continuous transformations. From this point of view, the abstract continuous group is the **parameter space** of the transformations, which is parameterized by the $a_1, \ldots, a_n$. In other words, to each choice of real numbers $a_1, \ldots, a_n$ in a certain set M, we assign a corresponding continuous transformation of the target space as above.

From the point of view of transformations, the group operation consists in composition of transformations. How does this relate to the group operation on parameter space? Given two transformations of the target space, corresponding to two points in parameter space, composing the transformations gives a new transformation of the target space, which is assigned to a new point in parameter space. Thus, from the point of view of the parameter space, composing transformation of the target space gives a mapping $\alpha : M \times M \to M$ on parameter space, which specifies the group operation. If M is a smooth manifold, we recover the structure of Lie group.

Remark 4.1.11 It is important however to distinguish between the dimension of the target space (d), and the dimension of the Lie group, which is the number of parameters n on which the transformations depend. In other words, the dimension of the Lie group is the dimension of the parameter space M, which is the underlying smooth manifold of the Lie group.

All the Lie groups that we have encountered above can be understood as groups of transformations.

Example 4.1.12 Rescalings in one dimension. Let $x \in \mathbb{R}$, and consider the group of rescalings

$$x' = ax$$

parameterized by a non-zero real number $a \in \mathbb{R}^*$. These transformations form an abelian group, with the operation being composition. The parameter space is $\mathbb{R}^*$. Composing two transformations with parameters a and b, we get:

$$x'' = ax' = abx,$$

and hence the group operation on parameter space is the mapping $\alpha : \mathbb{R}^* \times \mathbb{R}^* \to \mathbb{R}^*$ given by

$$\alpha : (a, b) \mapsto ab.$$

The underlying parameter space is thus the group $\mathbb{R}^*$ with operation given by multiplication, which is a one-dimensional Lie group. It is Example 4.1.2. □

Example 4.1.13 Affine transformations in one dimension. As a variation of the previous example, we consider the group of transformations:

$$x' = a_1 x + b_1,$$

with $a_1 \in \mathbb{R}_{>0}$ and $b_1 \in \mathbb{R}$. These transformations form a non-abelian group. Its parameter space is two-dimensional, and given by $M = \mathbb{R}_{>0} \times \mathbb{R}$. Composing two transformations with parameters (a_1, b_1) and (a_2, b_2), we get:

$$x'' = a_1 x' + b_1 = a_1(a_2 x + b_2) + b_1 = a_1 a_2 x + (a_1 b_2 + b_1).$$

So the group operation corresponds to the mapping $\alpha : M \times M \to M$ on parameter space given by:

$$\alpha : ((a_1, b_1), (a_2, b_2)) \mapsto (a_1 a_2, a_1 b_2 + b_1). \tag{4.1.1}$$

This is nothing else than the group operation on the two-dimensional Lie group constructed in Example 4.1.4. □

We also introduced matrix Lie groups above as groups of matrices. But it is also customary to think of them as groups of transformations of target space.

Example 4.1.14 General linear transformations. Let the target space be an n-dimensional real vector space V. The linear invertible transformations $L : V \to V$ form the general linear group $GL(V)$, which we have already studied. If $V = \mathbb{R}^n$, then we get the general linear group $GL(n, \mathbb{R})$. Note that, while the dimension of the target space is n, the dimension of the Lie group (which is the dimension of the parameter space for $n \times n$ real invertible matrices) is n^2.

If we choose the complex vector space $V = \mathbb{C}^n$ for target space, then we get the group $GL(n, \mathbb{C})$. The real dimension of the target space here is $2n$, while the dimension of the Lie group is $2n^2$. □

Remark 4.1.15 When we think of Lie groups as groups of transformations, or as groups of matrices, the underlying manifold structure of the parameter space many not be obvious. We will see that in the following examples.

Example 4.1.16 Orthogonal transformations and rotations in $\mathbb{R}^n$. One can restrict to general linear transformations in $\mathbb{R}^n$ that preserve length and angles. Those are given by orthogonal transformations, and form the Lie group $O(n)$, studied in Example 4.1.9. One can also think of orthogonal transformations as symmetries of a real n-dimensional sphere, since they preserve its defining equation

$$x_1^2 + x_2^2 + \ldots + x_{n+1}^2 = R^2.$$

Note that while the dimension of the target space is n, the dimension of the Lie group $O(n)$ is $\frac{1}{2}n(n-1)$, as orthogonal $n \times n$ matrices depend on $\frac{1}{2}n(n-1)$ real independent parameters.

We can also further restrict to special orthogonal transformations, which form the group $SO(n)$. These consists in all elements of $O(n)$ that are continuously connected to the identity. Geometrically, it can be understood as the group of rotations in $\mathbb{R}^n$.

We note that the underlying manifold structure of $SO(n)$ (that is, of its parameter space) is not obvious to see in general. But for $SO(2)$ it is clear. Rotations in two dimensions are parameterized in terms of a single angle $\theta \in [0, 2\pi)$. Thus the parameter space is the circle S^1, with the group operation given by addition of angles. We recover the one-dimensional Lie group of Example 4.1.3! So instead of thinking of $SO(2)$ as the group of 2×2 special orthogonal matrices, or as the group of rotations in $\mathbb{R}^2$, we can also think of it as the circle S^1 with group operation given by addition of angles. This also makes it clear that it is a compact group, since the circle is compact.

For $SO(3)$, things are already more complicated. Rotations in $\mathbb{R}^3$ are parameterized by three angles. So its parameter space is three-dimensional, which is indeed the dimension of the Lie group. But if you remember how to parameterize rotations in three dimensions with angles, it is not so straightforward. It turns out that the manifold structure of the parameter space here is $\mathbb{RP}^3$, the real projective three-dimensional space. This example already higlights the fact that finding the manifold structure of the underlying parameter space of a matrix Lie group is not in general straightforward. □

Example 4.1.17 Unitary transformations in $\mathbb{C}^n$. The unitary group $U(n)$, studied in Example 4.1.10 can be realized as the group of transformations in $\mathbb{C}^n$ that preserve the inner product. One could also think $U(n)$ as the group of symmetries of a "complex n-dimensional sphere".

As for the the orthogonal group, the underlying manifold structure of $U(n)$ and $SU(n)$ is not so obvious to see. But for $SU(2)$ we can find the parameter space directly. Any 2×2 special unitary matrix A can be parametrized as:

$$A = \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix}$$

for $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$. Writing $\alpha = a + ib$ and $\beta = c + id$, the equation $|\alpha|^2 + |\beta|^2 = 1$ becomes

$$a^2 + b^2 + c^2 + d^2 = 1,$$

which is the equation of a three-sphere S^3 in $\mathbb{R}^4$. Thus, the parameter space (or underlying manifold structure) of $SU(2)$ is the three-sphere S^3, which is a compact manifold. □

Remark 4.1.18 When we think of Lie groups as groups of transformations, or equivalently as subgroups of $GL(n, \mathbb{C})$, formally what we are doing is constructing a representation of the underlying abstract group. Indeed, to every group element (i.e. points in parameter space), we assign a unique matrix (or a transformation of the target space). So we are building an explicit group homomorphism between the parameter space and $GL(n, \mathbb{C})$. We call this representation the **fundamental** (or **defining**) representation of the matrix Lie group.

For instance, the interpretation of $SO(2)$ as 2×2 special orthogonal matrices is the fundamental representation of the abstract Lie group, which is isomorphic to the circle S^1. Similarly, the interpretation of $SU(2)$ as 2×2 special unitary matrices is the fundamental representation of the abstract Lie group, which is isomorphic to the three-sphere S^3.

4.2 Rotations in two and three dimensions

Objectives

You should be able to:

- Calculate the generators of two and three dimensional rotations.
- Recover finite rotations by exponentiating the generators.

The very cool thing about Lie groups is that almost all of their structures is encoded in their linearization. This is the key insight of Lie. In calculus, you can approximate an analytic function $f(x)$ by looking at its linearization (or tangent line) at a point, say $x = 0$. In fact, if the function is analytic, Taylor tells you that you can completely reconstruct the function locally if you know all its derivatives at $x = 0$. But you need to know all derivatives. A similar statement holds in more than one dimensions; you can approximate a manifold at a point by looking at its tangent space. But this is only a linearization of the space.

For Lie groups the situation is different. It turns out that you only need to know the linearization of a group near the origin to recover the whole group locally! This is because of the group structure. If you know the first derivative of the group elements (thought of as matrices) at the origin, you can recover the group elements through exponentiation. So we can study the structure of Lie groups by looking at their linearizations at the origin, which are called **Lie algebras**. In other words, the Lie algebra of a Lie group can be thought of as the linearization, or tangent space, of the group manifold at the origin.

To introduce the concept of Lie algebras, we will first study rotations in two and three dimensions, and then define the formal idea that applies to all Lie groups.

4.2.1 Rotations in two dimensions

Let us start by looking at the Lie group $SO(2)$ of rotations in two dimensions. Those are given by 2×2 special orthogonal matrices. We know that we can write any such rotation as a matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

in terms of an angle of rotation $\theta \in [0, 2\pi)$. The deep idea of Lie is that we can recover all rotations by doing a rotation by a very small angle many many times. Doesn't sound very deep, but it is. :-) More precisely, we think of infinitesimal rotations. If we take the angle θ to be small, we can approximate $\cos\theta \cong 1 + \mathcal{O}(\theta)^2$ and $\sin\theta \cong \theta + \mathcal{O}(\theta)^3$. Keeping only terms of first order in θ, we can thus approximate the rotation matrix as

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cong \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} = I + \theta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Conversely, we can do a similar linear approximation for special orthogonal matrices. If A is a 2×2 special orthogonal matrix A which is very close to the identity matrix, we can write $A \cong I + M$ for some "infinitesimal" M. The condition that $A^T A = I$ implies that

$$(I + M^T)(I + M) \cong I + M + M^T = I,$$

where we kept only terms of first order in M. Thus $M = -M^T$, that is, M is an antisymmetric matrix. But in two dimensions any real antisymmetric matrix is a multiple of $X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, thus we can write

$$A \cong I + \theta X$$

for some real parameter θ as before. So we have recovered the same first order expansion as for rotations. We call X the **generator** of the Lie group (more precisely, it is a matrix representation of the generator of the Lie group).

Checkpoint 4.2.1 Starting from orthogonal matrices, we recovered the first order expansion of rotations. But this should only work for special orthogonal matrices. It looks like we never imposed the condition that $\det A = 1$. Why?

The key insight of Lie is that we can now recover finite rotations from the knowledge of the generator X alone. Let $R(\theta)$ be a rotation by a finite angle θ. Pick a large integer N. We can recover $R(\theta)$ by doing $R(\theta/N)$ N times. In the limit as $N \to \infty$, $R(\theta/N)$ becomes an infinitesimal rotation. Thus we get:

$$R(\theta) = \lim_{N\to\infty} R(\theta/N)^N = \lim_{N\to\infty} \left(I + \frac{\theta}{N} X \right)^N.$$

If we naively use the relation $e^x = \lim_{N\to\infty}(1 + \frac{x}{N})^N$ for matrices, we would conclude that

$$R(\theta) = e^{\theta X}.$$

In other words, we can recover finite rotations by exponentiating the infinitesimal generator! That is cool. This is due to the group structure, which says that we can recover rotations by successively repeating rotations of smaller angles.

Let us be a little more precise and prove this explicitly, which makes the relation with the group structure more explicit.

Lemma 4.2.2 The generator of two-dimensional rotations. *Let*

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We define the exponential of a matrix A through the series:

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

Then any two-dimensional rotation $R(\theta)$ can be written in terms of the generator X via ***exponentiation****:*

$$R(\theta) = e^{\theta X}.$$

Proof. Let us start with a two-dimensional rotation $R(\theta)$. Since it is a Lie group, we can Taylor expand near the identity:

$$R(\theta) = I + \frac{dR}{d\theta}\Big|_{\theta=0} \theta + \frac{1}{2}\frac{d^2R}{d\theta^2}\Big|_{\theta=0} \theta^2 + \dots \tag{4.2.1}$$

(By derivative of a matrix here we mean derivative of its entries.) Let us now identity the derivatives of the rotation matrix. Rotations form a Lie group. The group structure is given by composition of rotations, which can be written as the requirement that

$$R(\theta_1 + \theta_2) = R(\theta_1)R(\theta_2).$$

Taking the derivative on both sides with respect to θ_1, and then setting $\theta_1 = 0$, we get

$$\frac{dR(\theta_1 + \theta_2)}{d\theta_1}\Big|_{\theta_1=0} = R(\theta_2)\frac{dR(\theta_1)}{d\theta_1}\Big|_{\theta_1=0}. \tag{4.2.2}$$

We can calculate the left-hand-side of (4.2.2) via the chain rule:

$$\left(\frac{dR(\theta_1 + \theta_2)}{d(\theta_1 + \theta_2)}\frac{d(\theta_1 + \theta_2)}{d\theta_1}\right)\Big|_{\theta_1=0} = \frac{dR(\theta_2)}{d\theta_2}.$$

For the right-hand-side of (4.2.2), we define the matrix

$$X := \frac{dR(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Thus (4.2.2) becomes

$$\frac{dR(\theta)}{d\theta} = XR(\theta). \tag{4.2.3}$$

In particular,

$$\frac{dR(\theta)}{d\theta}\Big|_{\theta=0} = X.$$

In fact, taking repeated derivatives of (4.2.3), we get:

$$\frac{d^nR(\theta)}{d\theta^n} = X\frac{d^{n-1}R(\theta)}{d\theta^{n-1}}.$$

Evaluating at $\theta = 0$, we get, by induction on n,

$$\frac{d^n R(\theta)}{d\theta^n}\Big|_{\theta=0} = X^n.$$

Thus all derivatives of the rotation matrix at the origin are determined by the generator X! As is clear from the proof, this follows because of the group structure of rotations.

Plugging this back into (4.2.1), with $X^0 := I$, we get

$$R(\theta) = \sum_{n=0}^{\infty} \frac{1}{n!}\theta^n X^n = e^{\theta X},$$

which concludes the proof. ■

The cool thing here is that all two-dimensional rotations can be recovered by exponentiating the infinitesimal generator X. So instead of studying the Lie group $SO(2)$ directly, we could instead study the properties of the generator X. This is our first example of a Lie algebra. We say that X is an element of the **Lie algebra** $\mathfrak{so}(2)$: we use the weird font to distinguish the Lie algebra from the Lie group. The algebra is rather trivial however here, since it has only one generator. Our next example will be less trivial.

4.2.2 Rotations in three dimensions

Let us now consider the Lie group $SO(3)$ consisting of three-dimensional rotations. We think of those as 3×3 special orthgonal matrices A. We do an infinitesimal expansion $A \cong I + M$ for an infinitesimal M. Then the orthogonality condition becomes

$$A^T A = (I + M^T)(I + M) \cong I + M^T + M = I.$$

Therefore, as for two-dimensional rotations, we conclude that M is a real antisymmetric 3×3 matrix. In fact, it is customary in physics to introduce a factor of i in our linearization, and define instead $A \cong I + iL$. Then what we have shown is that $(iL)^T = -iL$. Since $(iL)^T$ is a real matrix, then it is equal to its complex conjugate $-iL^\dagger$. Thus the condition that $(iL)^T = -iL$ can be rewritten as $L^\dagger = L$, that is, L is a purely imaginary Hermitian 3×3 matrix.

Any 3×3 purely imaginary Hermitian matrix can be written as a linear combination of three matrices:

$$L_1 = -i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad L_2 = -i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad L_3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.2.4}$$

Those are the infinitesimal generators of the Lie group $SO(3)$. We can write an arbitrary infinitesimal rotation as $i\theta_1 L_1 + i\theta_2 L_2 + i\theta_3 L_3$ for three real parameters $\theta_1, \theta_2, \theta_3$ (we thus see that $SO(3)$ is a three-dimensional Lie group). Following the same argument as for rotations in two dimensions, we can also conclude that we can write an arbitrary finite rotation in three dimensions (i.e. an element of the Lie group $SO(3)$) by exponentiation:

$$R(\theta) = e^{i\sum_{i=1}^{3}\theta_i L_i}.$$

The algebra generated by the generators L_1, L_2, L_3 is the Lie algebra $\mathfrak{so}(3)$. Now that we have more than one generator, this is more interesting. How do we define the abstract properties of this algebra? (Instead of writing down an explicit representation in terms of three-dimensional matrices.) For this, we need one more element: the notion of a binary operation on the algebra, which, in matrix form, will be given by the commutator. Let us see how this comes about.

4.2.3 Commutation and the commutator

What we have seen so far is that for rotations in two and three dimensions, we can reconstruct the group elements by exponentiating the infinitesimal generators. This is in fact a general statement for all Lie groups, as we will see in the next section. In the context of three-dimensional rotations, we found an explicit representation for the generators in terms of 3×3 matrices. But just as when we defined abstract groups, we would like to obtain an abstract definition of the algebra of generators of a Lie group. For this, there is one element missing. In general, rotations do not commute. How can we see that from the point of view of the infinitesimal generators?

Let $R \cong I + M$ and $R' \cong I + M'$ be infinitesimal rotations. Then

$$RR'R^{-1} \cong (I + M)(I + M')(I - M) = I + M' + (MM' - M'M),$$

where we neglected terms of higher order. If we define the **commutator** $[M, M']$ as

$$[M, M'] := MM' - M'M,$$

then non-commutativity of R and R' is encapsulated in the statement of whether the commutator of their infinitesimal generators vanishes or not.

Thus, to encode the group structure of a Lie group in terms of the abstract notion of the algebra of its infinitesimal generators, we need to specify the commutation relations between the generators. This will give an abstract definition of a Lie algebra, from which a Lie group can be obtained by exponentiation.

For three-dimensional rotations, looking at the representation of the generators L_1, L_2, L_3 in terms of 3×3 matrices that we found, it is easy to compute that

$$[L_1, L_2] = iL_3, \qquad [L_2, L_3] = iL_1, \qquad [L_3, L_1] = iL_2.$$

Checkpoint 4.2.3 Check that the generators of three-dimensional rotations L_1, L_2, L_3 satisfy these commutation relations.

This can be encoded neatly using the Levi-Civita symbol ϵ_{ijk}, which is defined as:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i, j, k) \text{ is an even permutation of } (1, 2, 3); \\ -1 & \text{if } (i, j, k) \text{ is an odd permutation of } (1, 2, 3); \\ 0 & \text{if two indices are repeated.} \end{cases}$$

Then:

$$[L_i, L_j] = i \sum_{k=1}^{3} \epsilon_{ijk} L_k. \tag{4.2.5}$$

As we can see, the commutator closes, since the right-hand-side is a linear combination of the generators. Thus it provides a bilinear operation $L \times L \to L$ on the vector space L spanned by L_1, L_2, L_3. This is generally true, as we will see. For any Lie algebra, we are given a bilinear operation, which we write as a commutator, such that

$$[L_i, L_j] = \sum_k c_{ijk} L_k.$$

We call the c_{ijk} the **structure constants** of the Lie algebra.

Abstractly, we can define the Lie algebra $\mathfrak{so}(3)$ as being the three-dimensional vector space V of real linear combinations of the generators L_1, L_2, L_3, with a bilinear operation $[\cdot, \cdot] : V \times V \to V$ specified by (4.2.5).

4.2.4 Differential representation

So far we have worked exclusively with matrix representations of the rotation groups. In fact, we defined the rotation groups in terms of their fundamental, or defining, representations, as subgroups of $GL(n, \mathbb{R})$.

But there are other types of representations that are very useful. Let us focus on three-dimensional rotations as an example. We can represent the infinitesimal generators of rotations as differential operators acting on functions $f(x, y, z)$ on $\mathbb{R}^3$. To do that, what we need to do is find differential operators that satisfy the commutation relations (4.2.5). It is not too difficult to check that the following differential representation works:

$$L_1 = -i \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \qquad L_2 = -i \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \qquad L_3 = -i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \tag{4.2.6}$$

Checkpoint 4.2.4 Check that the differential operators L_1, L_2, L_3 satisfy the commutation relations (4.2.5).

You may recognize those operators as the angular momentum operators in quantum mechanics (up to a factor of $\hbar$). This is not a coincidence! The angular momentum operators in quantum mechanics are a representation of the infinitesimal generators of the group of rotations in three dimensions. In fact, the possibility of going back and forth between differential representations and matrix representations of the Lie algebra is the essence of the duality between the Schrodinger and Heisenberg pictures of quantum mechanics.

4.3 Lie algebras

Objectives

You should be able to:

- Recall the definition of Lie algebras.
- Calculate the Lie algebra corresponding to a given Lie group.
- Reconstruct the group elements from the Lie algebra generators by exponentiation.

In this section we generalize the discussion about two-dimensional and three-dimensional rotations to more general Lie groups. We will try to come up with an abstract definition for the algebra of generators of a Lie group.

We will proceed in three steps. First, we will study how to extract the Lie algebra corresponding to a given matrix Lie group. Second, we will define the abstract notion of a Lie algebra. Third, we will show how we can recover the group structure corresponding to a Lie algebra through exponentiation.

4.3.1 The Lie algebra of a Lie group

Let us first describe how, given a matrix Lie group, one can extract the corresponding Lie algebra. The discussion here is very similar to what was done for two-dimensional and three-dimensional rotations in Section 4.2.

Geometrically, the Lie algebra is the linearization of the Lie group at the origin. In other words, we replace the group manifold locally near the origin by its tangent space. Since we are working with matrix Lie groups, we can calculate this linearization explicitly by expanding the matrices of the fundamental representation near the identity, and keeping only terms of first order. Let us be a little more precise.

We start with a matrix Lie group. That is, we start with its defining representation, so we can think of the group elements as matrices A that depend on n real parameters. We want to construct the associated Lie algebra. We proceed in three steps.

1. We expand the group elements near the identity, to first order:
$$A = I + iL.$$
We determine the properties of the matrices L by imposing the appropriate condition on A. For instance, for $SO(n)$, imposing that A is a real orthogonal matrix constrains L to be a purely imaginary Hermitian matrix.

2. We find a basis $L_1, \ldots, L_n$ for the n-dimensional vector space of matrices L satisfying the appropriate constraint. We call the L_i the infinitesimal generators of the Lie group. We write L as a linear combination of the L_i:
$$L = \sum_{i=1}^{n} \theta_i L_i.$$

3. For any two group elements A, A' with first order expansions $A = I + iL$, $A' = I + iL'$, the commutator of matrices $[L, L']$ encodes whether the group elements commute. Since L and L' are linear combinations of the generators L_i, to know all such commutators we only need to calculate the commutators of the generators L_i. Because of the group structure we know that the commutator closes, that is, the commutator of two generators will be itself a linear combination of the generators. So we can write:
$$[L_i, L_j] = \sum_{k=1}^{d} c_{ijk} L_k \tag{4.3.1}$$
for some structure constants c_{ijk}. We calculate these structure constants from the form of the generators of the group.

This determines the Lie algebra associated to a Lie group. It is the vector space V of real linear combinations of the generators L_i, with a binary operation $[\cdot, \cdot] : V \times V \to V$ specified by (4.3.1).

Remark 4.3.1 Note that in this approach, it is crucial that we start with a matrix representation of the group. This allows us to do an expansion near the identity matrix, and to determine the bracket $[\cdot, \cdot]$ by evaluating the commutator of the matrices corresponding to the infinitesimal generators. But in the end, we end up with an algebra that can be defined abstractly in terms of the generators and the relation (4.3.1). It does not depend on the matrix representation anymore. And in fact, had we started with a different matrix representation of the Lie group, we would have obtained a different matrix representation of the generators L_i, but the resulting algebra would have been the same, with the same relation (4.3.1).

4.3.2 Abstract definition of a Lie algebra

In the previous section we have seen how we can determine the linearization of a matrix Lie group at the origin (that is, the associated Lie algebra). We obtained a vector space spanned by the infinitesimal generators of the Lie group, equipped with a bracket operation. We calculated the Lie algebra associated to a Lie group using a specific representation of the group (usually the defining representation). But the end result is an algebraic structure that can be defined entirely abstractly. Let us now define the abstract concept of a Lie algebra.

Definition 4.3.2 Lie algebra. A **Lie algebra** is a vector space $\mathfrak{g}$ over some field F (generally taken to be the real numbers in this class), together with a binary operation $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the **Lie bracket**, that satisfies the following axioms:

1. Bilinearity:

$$[ax + by, z] = a[x, z] + b[y, z],$$
$$[z, ax + by] = a[z, x] + b[z, y],$$

for all $a, b \in F$ and all $x, y, z \in \mathfrak{g}$.

2. Antisymmetry:

$$[x, y] = -[y, x],$$

for all $x, y \in \mathfrak{g}$.

3. The Jacobi identity:

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0,$$

for all $x, y, z \in \mathfrak{g}$.

◊

This is the abstract definition of a Lie algebra. Concretely, in this course the Lie algebras that we will study will be obtained as linearizations of matrix Lie groups. That is, they will be constructed as in Subsection 4.3.1, starting from a matrix representation of a Lie group. Then the vector space $\mathfrak{g}$ is constructed as real linear combinations of the infinitesimal generators of the group, and the Lie bracket is realized explicitly as a commutator of the matrix representation for these generators. We know that the Lie bracket closes, that is, for any two $x, y \in \mathfrak{g}$, $[x, y] \in \mathfrak{g}$, as required for a Lie algebra.

In this context, the three axioms in the definition of Lie algebras are clear. The commutator of matrices is certainly bilinear and antisymmetric, by definition. As for the Jacobi identity, one can check that given any three matrices A, B, C, the commutator $[A, B] = AB - BA$ satisfies the Jacobi identity $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

Checkpoint 4.3.3 Check that for any three matrices A, B, C, the commutator $[A, B] = AB - BA$ satisfies the Jacobi identity $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

Thus the procedure of Subsection 4.3.1 always produces a Lie algebra. The structure of the algebra is encoded in the structure constants as in (4.3.1).

4.3.3 Recovering a Lie group from a Lie algebra

Now suppose that you are given a Lie algebra $\mathfrak{g}$. How can you reconstruct the group structure from the Lie algebra?

The procedure is simple: exponentiation. Start with the vector space $\mathfrak{g}$. We assume that there is a matrix representation for the Lie algebra, so that we can represent elements of the vector spaces in terms of matrices. Then:

1. We pick a basis L_i, $i = 1, \ldots, n$ for $\mathfrak{g}$. Using the matrix representation we can think of the elements L_i as matrices. A general element of $\mathfrak{g}$ can be written as a linear combination $\sum_{i=1}^{n} t_i L_i$.

2. We construct a matrix representation for the Lie group by exponentiation:
$$g(\vec{t}) = e^{i \sum_{i=1}^{n} t_i L_i}.$$
The matrices L_i form a representation of the infinitesimal generators of the Lie group.

The claim is that the matrices $g(\vec{t})$ thus constructed form a group under matrix multiplication. Associativity is clear. We need to check that the identity matrix is in there; that inverses exist; and the matrix multiplication closes.

Exponentiation maps the origin of the vector space $\vec{0} \in \mathfrak{g}$ to the matrix

$$g(\vec{0}) = e^0 = I,$$

which is the identity matrix. So this is good.

As for inverses, for any $\sum_i t_i L_i \in \mathfrak{g}$, $-\sum_i t_i L_i$ is also in $\mathfrak{g}$, and by exponentiation those are mapped to

$$g(\vec{t})g(-\vec{t}) = e^{i \sum_{i=1}^{n} t_i L_i} e^{-i \sum_{i=1}^{n} t_i L_i} = I.$$

This follows because for any matrix A, we have $e^A e^{-A} = I$, which can be checked by expanding the Taylor series for the exponential.

The tricky statement is whether the group multiplication closes. In other words, we need to show that given any two $\sum_i t_i L_i$ and $\sum_i s_i L_i$ in $\mathfrak{g}$, then there exists a $\sum_i u_i L_i \in \mathfrak{g}$ such that

$$g(\vec{t})g(\vec{s}) = g(\vec{u}).$$

In other words, for any two $A, B \in \mathfrak{g}$, we want to find a $C \in \mathfrak{g}$ such that

$$e^A e^B = e^C.$$

To show such a thing, we would like to construct C explicitly from A and B. For C to be in $\mathfrak{g}$, this means that we should be able to write it as a linear combination of A, B and commutators of those.

First, if A and B commute, then it straightforward to show that

$$e^A e^B = e^{A+B}.$$

Thus, in this case, the statement is certainly true since $C = A + B$, and exponentiation produces a (abelian) group.

Checkpoint 4.3.4 Prove that $e^A e^B = e^{A+B}$ for any two commuting matrices A and B.

However, $e^A e^B$ is not equal to e^{A+B} anymore when A and B do not commute. If we define C by $e^C = e^A e^B$, expand the formal series on both sides of the equality, and determine C order by order in A and B, we obtain the expansion:

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] + \dots,$$

where ... stands for an infinite formal series of terms that are expressed as linear combinations of commutators of A and B. This is essentially the Baker-Campbell-Haussdorff formula. The key point here is that C is a formal series of terms in $\mathfrak{g}$. And if A and B are close enough to the identity (this can be stated formally), then the formal series converges to a Lie algebra element $C \in \mathfrak{g}$. Thus exponentiation constructs a group, at least locally near the identity.

To summarize, from a Lie algebra, one can recover the local behaviour of a Lie group through exponentiation. Nice!

4.4 $SU(2)$

Objectives

You should be able to:

- Determine the Lie algebra $\mathfrak{su}(2)$.
- Explain the statement that $SU(2)$ is a double cover of $SO(3)$.

In this course we will mostly focus on the two most important Lie groups in physics, namely $SO(3)$ and $SU(2)$. We have already calculated the Lie algebra $\mathfrak{so}(3)$ associated to the Lie group $SO(3)$ in Section 4.2. We obtained that $\mathfrak{so}(3)$ is the three-dimensional vector spanned by elements L_1, L_2, L_3 with the bracket

$$[L_i, L_j] = i \sum_{k=1}^{3} \epsilon_{ijk} L_k.$$

4.4.1 The Lie algebra $\mathfrak{su}(2)$

Let us now study the Lie algebra $\mathfrak{su}(2)$ associated to the special unitary group $SU(2)$. The fundamental representation of $SU(2)$ consists in 2×2 unitary matrices U with determinant one:

$$U^\dagger U = I, \qquad \det U = 1.$$

To obtain the Lie algebra, we do a first order expansion $U = I + iL$. To first order, the unitarity condition becomes

$$U^\dagger U \cong (I + iL)^\dagger (I + iL) \cong I - iL^\dagger + iL = I.$$

Therefore $L^\dagger = L$, which says that L is a Hermitian matrix. We also need to impose the condition that $\det U = 1$. We can write a general 2×2 complex-valued matrix as

$$L = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

for $a, b, c, d \in \mathbb{C}$. Then

$$\det U \cong \det(I + iL) = \det \begin{pmatrix} 1 + a & b \\ c & 1 + d \end{pmatrix} = (1 + a)(1 + d) - bc.$$

Since we keep only terms of first order in L, we can ignore any terms that are order two in the complex numbers a, b, c, d. Thus, to first order, we have

$$\det U \cong 1 + a + d.$$

For this to be equal to one we must have $a = -d$, that is, the matrix L is traceless (has vanishing trace).

The conclusion is that L must be a 2×2 traceless Hermitian matrix. Any such matrix can be written as

$$L = \begin{pmatrix} t_1 & t_2 + it_3 \\ t_2 - it_3 & -t_1 \end{pmatrix}$$

for some real numbers $t_1, t_2, t_3 \in \mathbb{R}$. Indeed, since the matrix is Hermitian, the diagonal terms must be real, and since it is traceless they must add up to zero. The off-diagonal terms must be complex conjugate since the matrix is Hermitian.

It thus follows that the Lie algebra $\mathfrak{su}(2)$ is three-dimensional. A natural basis for the vector space consists of the matrices:

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.4.1}$$

Those are known as **Pauli matrices**. By calculating the commutators, one finds:

$$[T_1, T_2] = iT_3, \qquad [T_2, T_3] = iT_1, \qquad [T_3, T_1] = iT_2.$$

In other words,

$$[T_i, T_j] = i \sum_{k=1}^{3} \epsilon_{ijk} T_k.$$

What? This is the same abstract algebra as $\mathfrak{so}(3)$! Indeed, it turns out that the Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ are isomorphic. Really, they are the same. This is absolutely fundamental, and, as we will see, is the reason for the appearance of fermions (particles with half-integer spin) in non-relativistic physics.

This does not mean however that the Lie groups $SO(3)$ and $SU(2)$ are isomorphic: in fact they are not. But they are certainly closely related. More precisely, there exists a two-to-one group homomorphism from $SU(2)$ to $SO(3)$. We say that $SU(2)$ is a "double covering" of $SO(3)$.

4.4.2 $SU(2)$ vs $SO(3)$

We have seen in the previous section that the Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic. What does this mean for the Lie groups $SU(2)$ and $SO(3)$?

The fact that their Lie algebras are the same means that the two groups are "locally isomorphic", which is a rather striking statement. What this means is that, using exponentiation of the Lie algebra, we would recover the same group structure locally near the identity. This does not mean however that the two groups are globally isomorphic: in fact they are not.

But they are certainly closely related. The precise statement is that there exists a group homomorphism from $SU(2)$ to $SO(3)$ that is two-to-one. This is what we show next. We say that $SU(2)$ is a "double covering" of $SO(3)$.

Let us now construct the group homomorphism $f : SU(2) \to SO(3)$.

Theorem 4.4.1 **$SU(2)$ covers $SO(3)$.** *There exists a two-to-one group homomorphism $f : SU(2) \to SO(3)$, which means that the Lie groups $SU(2)$*

and $SO(3)$, which share the same Lie algebras $\mathfrak{su}(2) \cong \mathfrak{so}(3)$, are only locally isomorphic.

We provide a formal proof of this statement using the adjoint representation of the Lie group $SU(2)$ below. But before we do that, let us write down an informal argument, which may be more enlightening.

Both $SU(2)$ and $SO(3)$ share the same Lie algebra. For $SU(2)$, we obtained the Lie algebra by constructing the generators of $SU(2)$, which is given by the vector space spanned by the Pauli matrices (4.4.1), which we reproduce here for convenience:

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

For $SO(3)$, we obtained the algebra by looking at the generators of rotations in three dimensions (4.2.4), which we also reproduce here:

$$L_1 = -i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad L_2 = -i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad L_3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One can show that all special unitary matrices in $SU(2)$ can be obtained by exponentiation, and similarly for special orthogonal matrices in $SO(3)$ (this is because $SU(2)$ and $SO(3)$ are compact and connected). Thus we can write an arbitrary $U \in SU(2)$ as

$$U(\theta_1, \theta_2, \theta_3) = \exp\left(i \sum_{k=1}^{3} \theta_k T_k\right),$$

and an arbitrary $R \in SO(3)$ as

$$R(\theta_1, \theta_2, \theta_3) = \exp\left(i \sum_{k=1}^{3} \theta_k L_k\right).$$

The mapping $f : SU(2) \to SO(3)$ is then given by $U(\theta_1, \theta_2, \theta_3) \mapsto R(\theta_1, \theta_2, \theta_3)$.

How do we see that the mapping is two-to-one? Let us argue that it is two-to-one by looking at the special case with $\theta_1 = \theta_2 = 0$, but the argument is general (although a bit more tedious) and can be written explicitly for any choice of parameters.

We know that $R(0, 0, \theta) = R(0, 0, \theta + 2\pi)$, since $R(0, 0, \theta)$ corrrespond to a rotation in three dimensions by angle θ about an axis, hence rotating further by 2π does not do anything. On the other hand, since T_3 is diagonal, we can write

$$U(0, 0, \theta) = \exp(i\theta T_3) = \begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix}.$$

Thus,

$$U(0, 0, \theta + 2\pi) = \begin{pmatrix} e^{i\frac{\theta}{2} + i\pi} & 0 \\ 0 & e^{-i\frac{\theta}{2} - i\pi} \end{pmatrix} = -U(0, 0, \theta).$$

So shifting θ by 2π does not send the unitary matrix $U(0, 0, \theta)$ to itself; rather it sends it to minus itself. Shifting by 4π would bring back $U(0, 0, \theta)$ to itself.

As a result, we see that both $U(0, 0, \theta)$ and $U(0, 0, \theta + 2\pi) = -U(0, 0, \theta)$ are mapped to the same rotation matrix $R(0, 0, \theta)$, and hence the mapping is two-to-one. The point here is that rotations have period 2π, while unitary matrices as written in terms of Pauli matrices have period 4π. This is why $SU(2)$ covers $SO(3)$ twice.

Let us now write down a formal proof of the theorem.

Proof. We will construct the group homomorphism explicitly. In fact, it will be given by the adjoint representation of the group $SU(2)$, so let us construct this explicitly.

The adjoint representation of a Lie group G is given by a representation of the group elements of G as linear operators acting on the Lie algebra $\mathfrak{g}$ of G. Since the Lie algebra $\mathfrak{g}$ is a n-dimensional vector space, where n is the dimension of the Lie group, the adjoint representation maps the group elements to $n \times n$matrices acting on $\mathfrak{g}$. So it gives a representation $T : G \to GL(\mathfrak{g})$ of the same dimension as the dimension of the Lie group G. In the case of $G = SU(2)$, which has dimension 3, the adjoint representation is three-dimensional, and represents the group of elements of $SU(2)$ as 3×3 matrices in $GL(\mathfrak{su}(2))$. As we will show, it turns out that the image of the adjoint representation is $SO(3) \subset GL(\mathfrak{su}(2))$, and hence it provides the group homomorphism $SU(2) \to SO(3)$ that we are looking for.

Let us now construct the adjoint representation explicitly. We start with an element $X \in \mathfrak{su}(2)$, which we think of as a 2×2 traceless Hermitian matrix. We take an element $U \in SU(2)$, which is a 2×2 special unitary matrix. Then we define:

$$Ad_U(X) = UXU^\dagger,$$

which is basically a similarity transformation of X by U.

Let us now show that $Ad_U(X) \in \mathfrak{su}(2)$, i.e. $Ad_U(X)$ is traceless Hermitian. The tracelessness property follows since

$$\operatorname{Tr} Ad_U(X) = \operatorname{Tr}(UXU^\dagger) = \operatorname{Tr}(XU^\dagger U) = \operatorname{Tr} X = 0,$$

by the cyclic property of traces. As for being Hermitian, we find:

$$(Ad_U(X))^\dagger = (UXU^\dagger)^\dagger = UX^\dagger U^\dagger = UXU^\dagger = Ad_U(X),$$

since $X = X^\dagger$. Therefore, $Ad_U(X)$ is Hermitian and traceless. In other words, we have a mapping $Ad_U : \mathfrak{su}(2) \to \mathfrak{su}(2)$, which takes $X \mapsto Ad_U(X) = UXU^\dagger$. That is, $Ad_U \in GL(\mathfrak{su}(2))$.

We can now define the adjoint representation of $SU(2)$. It is given by the group homomorphism $Ad : SU(2) \to GL(\mathfrak{su}(2))$, with $U \mapsto Ad_U$. In other words, it takes a special unitary matrix $U \in SU(2)$, and maps it to the linear operator Ad_U on $\mathfrak{su}(2)$. Since $\mathfrak{su}(2)$ is a three-dimensional vector space, this gives a three-dimensional representation of $SU(2)$. Note that it is straightforward to show that Ad is a group homomorphism, since $Ad : U_1U_2 \mapsto Ad_{U_1U_2} = Ad_{U_1} \circ Ad_{U_2}$ by definition of Ad_U.

The image of Ad is in $GL(\mathfrak{su}(2))$. We can write any 2×2 Hermitian traceless matrix X as a real linear combination of the Pauli matrices (4.4.1):

$$X = x_1T_1 + x_2T_2 + x_3T_3.$$

Thus we can identify $\mathfrak{su}(2) \cong \mathbb{R}^3$, where we map an Hermitian traceless matrix X to the vector of coefficients $\vec{x}$ in the basis for $\mathfrak{su}(2)$ given by Pauli matrices. With this identification we can see $Ad : SU(2) \to GL(3, \mathbb{R})$. But in fact, the image is in $SO(3) \subset GL(3, \mathbb{R})$, as we now show.

We notice that for any Hermitian matrix $X = x_1T_1 + x_2T_2 + x_3T_3$, the determinant is calculated as:

$$4 \det X = (x_3)(-x_3) - (x_1 - ix_2)(x_1 + ix_2) = -|\vec{x}|^2.$$

So we can identify the determinant of $X \in \mathfrak{su}(2)$ with the length of the vector $\vec{x} \in \mathbb{R}^3$. Since the determinant is preserved by similarity transformations, it follows that

$$\det X = \det UXU^\dagger = \det Ad_U(X).$$

This means that the vectors in $\mathbb{R}^3$ corresponding to X and $Ad_U(X)$ have the same length. In other words, Ad_U is a linear operator on $\mathbb{R}^3$ that preserves the length of vectors: it is an orthogonal transformation in $O(3) \subset GL(3,\mathbb{R})$. Thus $Ad : SU(2) \to O(3) \subset GL(3,\mathbb{R})$. But in fact, we know even more. From Example 4.1.10, we know that the underlying manifold structure of $SU(2)$ is S^3. In particular, it is connected (in fact, it is also simply connected). Therefore, the image of the continuous mapping $Ad : SU(2) \to O(3)$ must be connected to the identity in $O(3)$, which means that it must be a subgroup of $SO(3) \subset O(3)$. Thus $Ad : SU(2) \to SO(3)$.

What remains to be shown is that the group homomorphism $Ad : SU(2) \to SO(3)$ is surjective, and that it is two-to-one.

We will leave the proof that $Ad : SU(2) \to SO(3)$ is surjective as an exercise. (Note that there are many ways that this can be done.)

To show that $Ad : SU(2) \to SO(3)$ is two-to-one, pick any unitary matrix U_1, and set $U_2 := -U_1$. Then

$$Ad(U_1) = Ad_{U_1} = Ad_{U_2} = Ad(U_2),$$

since for any Hermitian matrix X,

$$Ad_{U_1}(X) = U_1 X U_1^\dagger = (-U_1)X(-U_1)^\dagger = U_2 X U_2^\dagger = Ad_{U_2}(X).$$

Thus the mapping $Ad : SU(2) \to SO(3)$ is two-to-one.

Equivalently, we could show that the kernel is non-trivial. The kernel of Ad corresponds to 2×2 unitary matrices U that are mapped to Ad_U being the identity element of $SO(3)$. There are two such matrices,

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

since in both cases $Ad_U(X) = X$, i.e. Ad_U is the identity in $SO(3)$. Thus $\ker(f) \cong \mathbb{Z}_2$, where $\mathbb{Z}_2$ is realized as ± 1 times the identity matrix. Since the kernel has two elements, this means that $Ad : SU(2) \to SO(3)$ is two-to-one. ■

Remark 4.4.2 We have shown that there is a two-to-one group homomorphism $Ad : SU(2) \to SO(3)$. The kernel is $\mathbb{Z}_2$, realized as ± 1 times the identity matrix in $SU(2)$. Thus, by the first isomorphism theorem, we can write

$$SO(3) \cong SU(2)/\mathbb{Z}_2.$$

The groups however are said to be "locally isomorphic", since they have isomorphic Lie algebras. Equivalently, this can be seen since if $U \in SU(2)$ is near the identity, then $-U \in SU(2)$ is not. And hence in a neighborhoud of the identity the map is a local isomorphism.

4.5 General remarks

The presentation of Lie groups and Lie algebras in this course is necessarily unsystematic, and does not do justice to this beautiful area of mathematics. But let us conclude this section with a few general remarks about Lie groups and Lie algebras.

Remark 4.5.1 *To every Lie group is associated a unique Lie algebra.* The Lie algebra is uniquely determined as being the tangent space of the Lie group at the origin, or, in the language of matrix Lie groups, as being the algebra of the infinitesimal generators of the group.

Remark 4.5.2 *Two Lie groups that have isomorphic Lie algebras are not necessarily isomorphic.* Different Lie groups may share the same Lie algebras, such as $SO(3)$ and $SU(2)$. However, Lie groups that have the same Lie algebras are locally isomorphic near the identity. Globally, they are generally related by surjective group homomorphisms, i.e. covering maps. For instance, $SU(2)$ covers $SO(3)$ twice, as we have seen. The "largest" of the groups associated to a Lie algebra is called the **universal covering group**. It is simply connected (i.e. it consists of one piece and has no "holes" in it, that is, its fundamental group is trivial). In the case of $SO(3)$, its universal covering is $SU(2)$, which is simply connected (since it is diffeomorphic to S^3, see Example 4.1.10).

Remark 4.5.3 While different Lie groups may share the same Lie algebra, there is nonetheless a uniqueness result that is crucial. Lie's third theorem states that *every finite-dimensional real Lie algebra is the Lie algebra of some simply connected Lie group.* Moreover, if two Lie groups are simply connected and have isomorphic Lie algebras, then the groups themselves are isomorphic. What this means is that *to every finite-dimensional real Lie algebra one can associate a unique Lie group that is simply connected.* $SU(2)$ plays this role for the Lie algebras $\mathfrak{so}(3) \cong \mathfrak{su}(2)$.

Our goal for the remaining of this class is to study representation theory of Lie groups such as $SO(3)$ and $SU(2)$. More precisely, from physics we are particularly interested in representation theory of $SO(3)$, which corresponds to non-relativistic particles.

Representation theory of Lie algebras is generally easier than representation theory of Lie groups, because Lie algebras are linear. We have already seen that given a representation of a Lie algebra, one can reconstruct a representation of a Lie group using exponentiation. So the goal is to study representation theory of Lie groups by studying representation theory of Lie algebras. However, one has to be careful, since different Lie groups may share the same Lie algebra. If we start with a representation of a Lie algebra, and define a representation via exponentiation, how do we know what Lie group we are talking about?

The main result here is that *every representation of a simply connected Lie group comes from a representation of its corresponding Lie algebra.* So by starting with representations of the Lie algebra, what we are constructing is all representations of the unique simply connected Lie group associated to the algebra (the universal covering). The simply connected property is crucial.

For instance, representation theory of the Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ constructs all representations of the universal covering $SU(2)$. However, some of those will not be honest representations of the non-simply connected $SO(3) \cong SU(2)/\mathbb{Z}_2$. As we will see, they are so-called "projective", or "spin", representations. Those are mappings that preserve group multiplication but only up to a constant. So from the point of view of $SO(3)$, representation theory of the Lie algebra $\mathfrak{so}(3)$ constructs not only the honest representations of $SO(3)$, but also some spin representations.

Fortunately, in quantum mechanics it is sufficient to preserve only group operation up to a constant (since wave-functions are only defined up to phase), so we actually do care about spin representations, and should include them. Indeed, as we will see, the odd-dimensional representations (corresponding in physics to particles with integer spin) coming from $\mathfrak{so}(3)$ are honest representations of $SO(3)$, while the even-dimensional ones (corresponding in physics to particles with half-integer spin) are spin representations of $SO(3)$ (they are however true representations of the simply connected $SU(2)$).

Let us now move on to representation theory!

Chapter 5

Representation theory of Lie groups

5.1 Tensor representations of $SO(3)$

Objectives

You should be able to:

- Recognize tensors as objects that transform according to representations of $SO(3)$.
- Determine the dimension of the irreducible representations associated to symmetric traceless tensors.

In this section we study intuitively what representations of $SO(3)$ look like. This will give us an intuitive understanding of what the representation theory of $SO(3)$ should be, which we will then construct explicitly using representations of its Lie algebra in the following sections.

5.1.1 Representations of $SO(2)$

Let us start by giving a brief and naive construction of the irreducible representations of $SO(2)$.

We start with the cyclic group $\mathbb{Z}_N$, which we can think of a discrete set of rotations in two dimensions by angle $2\pi/N$. We know its irreducible representations: there are N of them, all one-dimensional, given by sending the generator of $\mathbb{Z}_N$ to $e^{2\pi i k/N}$ for some $k \in \{0, 1, \ldots, N-1\}$.

Well, as N goes to infinity, we can think of the group $\mathbb{Z}_N$ as turning into the continuous group $SO(2)$ of rotations in two dimensions. So we could expect the irreducible representations of $SO(2)$ to be all one-dimensional (which is true since $SO(2)$ is abelian), and to be given by the complex numbers $e^{ik\theta}$ for $\theta \in [0, 2\pi)$, and with k a non-negative integer. So there would be an infinite number of them, indexed by k.

While this construction is of course very naive (since it is not clear what precisely is meant by the limit $N \to \infty$ here), the intuition is correct. Those are the irreducible representations of $SO(2)$, and they are indexed by a non-negative integer k.

5.1.2 Representations of $SO(3)$

Let us now move on to the more interesting case of rotations in three dimensions, that is, $SO(3)$. What are the irreducible representations? After looking at $SO(2)$, we expect an infinite number of them. But here we do not expect them to be all one-dimensional, since $SO(3)$ is not abelian.

The simplest representation is of course the trivial one, where we sent every element $R \in SO(3)$ to $T(R) = 1$. This is as usual a rather boring one. It is sometimes called the **scalar representation** in physics, since, if you think of it as acting on a one-dimensional vector space V, then $T(R)v = v$ for all $v \in V$, that is, it leaves everything invariant. So a scalar object (which is an object that is invariant under rotations) transforms according to the trivial representation of $SO(3)$.

We have already studied another representation of $SO(3)$: its defining representation, which associates to every $R \in SO(3)$ a three-dimensional rotation matrix $T(R)$. In physics, we sometimes call this representation the **vector representation**. This is because it defines how vectors in $\mathbb{R}^3$ transform under rotations. Indeed, if $v \in \mathbb{R}^3$, then under a rotation we get $v' = Rv$, where R is a 3×3 rotation matrix. In index notation, we would write

$$v'_a = \sum_{j=1}^{3} R_{aj} v_j.$$

Turing this around, one could think of vectors as being objects that transform according to the three-dimensional defining representation of $SO(3)$.

Remark 5.1.1 This is a good time to introduce a notation that is standard in physics, which is to denote the representation of a Lie group by its dimension in bold face. For instance, we would write **3** for the defining representation of $SO(3)$. The notation is slightly ambiguous though, because there may be more than one representation of the same dimension. For instance, if a representation is complex, then its complex conjugate has the same dimension (we would then write, say, **5** and $\overline{\mathbf{5}}$ for the complex conjugate representations). Nevertheless, this is common notation, and it is usually clear from context what representation is being discussed.

Now you probably see a pattern. How can we get more representations of $SO(3)$? A natural way to think about representations is to think about how they act. Namely, we think of objects that transform in certain ways under rotations. We have already studied how scalars and vectors transform, which gave us the one-dimensional and three-dimensional representation of $SO(3)$. What next?

The next objects to look at are matrices. How do matrices tranform under rotations? Think of a matrix M as being a linear operator on $\mathbb{R}^3$. Then under a rotation R (which is a change of basis for $\mathbb{R}^3$) the matrix M will change by a similarity transformation $M' = RMR^T$, since R is orthogonal. In index notation, this becomes

$$M'_{ab} = \sum_{i,j=1}^{3} R_{ai} M_{ij} R^T_{jb} = \sum_{i,j=1}^{3} R_{ai} R_{bj} M_{ij}.$$

This defines a new representation of $SO(3)$. Indeed, we could place the nine components of the matrix M in a vector, and then the transformation rule above would give us a nine-dimensional representation of $SO(3)$, where we associate to every group element of $SO(3)$ the 9×9 matrix with components given by $R_{ai} R_{bj}$, where R is the corresponding 3×3 rotation matrix. In other words, if you recall the construction of tensor product representations, what we are constructing here is the nine-dimensional representation that is the tensor product of the defining representation with itself:

$$\mathbf{9} = \mathbf{3} \otimes \mathbf{3}.$$

We know that matrices transform according to this representation of $SO(3)$.

We could keep going and define higher-dimensional representations by looking at how objects with more than two indices transform under rotations. Those objects are known as **tensors**, and the corresponding representations are known as **tensor representations**. The rank of a tensor is the number of indices. For instance, a rank 3 tensor T_{ijk} transforms as

$$T'_{abc} = \sum_{i,j,k=1}^{3} R_{ai} R_{bj} R_{ck} T_{ijk},$$

which defines a 27-dimensional representation of $SO(3)$, namely $\mathbf{27} = \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$. In fact, we can think of tensors as being objects that transform according to representations of $SO(3)$.

5.1.3 Irreducible representations of $SO(3)$

In the previous section we constructed an infinite class of representations of $SO(3)$ by looking at how tensors transform under rotations. But are these representations irreducible? We will not answer this question rigorously here. We will instead look at the rank 2 case to get an intuition about what to expect, and then state the general result without proof.

Let us look at the rank 2 tensor representation (the matrix one) as a starting point. This 9-dimensional representation maps elements of $SO(3)$ to operators that act on matrices as

$$M'_{ab} = \sum_{i,j=1}^{3} R_{ai} R_{bj} M_{ij}.$$

This representation will be reducible if there exists a proper subspace of the space of matrices that is invariant under the action of the representation.

The key here is to notice that the transformation preserves the symmetry property of tensors under permutation of indices. In the case of a matrix, we can always write a matrix as a sum of its symmetric and antisymmetric parts:

$$M_{ij} = \frac{1}{2}(M_{ij} - M_{ji}) + \frac{1}{2}(M_{ij} + M_{ji}) := A_{ij} + S_{ij},$$

where $A_{ij} = -A_{ji}$ and $S_{ij} = S_{ji}$. The subspaces given by antisymmetric and symmetric matrices respectively are invariant under the action of the representation. Indeed, we get that

$$\begin{aligned}
A'_{ab} =& \frac{1}{2}(M'_{ab} - M'_{ba}) \\
=& \frac{1}{2} \sum_{i,j=1}^{3} \left(R_{ai} R_{bj} M_{ij} - R_{bi} R_{aj} M_{ij} \right) \\
=& \frac{1}{2} \sum_{i,j=1}^{3} R_{ai} R_{bj} \left(M_{ij} - M_{ji} \right)
\end{aligned}$$

$$= \sum_{i,j=1}^{3} R_{ai} R_{bj} A_{ij}.$$

A similar argument holds for S_{ij}. Thus the subspaces of symmetric and anti-symmetric rank 2 tensors are invariant subspaces, and thus provide subrepresentations. The number of independent components in a rank 2 anti-symmetric tensor in three dimensions is 3, while the number of independent components in a rank 2 symmetric tensor in three dimensions is 6. We thus obtain the decomposition $\mathbf{9} = \mathbf{6} \oplus \mathbf{3}$ for the nine-dimensional tensor representation.

Now, are the **6** and **3** representations irreducible? It turns out that the three-dimensional one is, and is in fact dual to the vector representation (since in three dimensions rank 2 anti-symmetric tensors are dual to vectors, see Zee's book, section IV.1, for more on this). However the six-dimensional representation is still reducible.

Indeed, consider the subspace consisting of the trace $S = \sum_{i=1}^{3} S_{ii}$. Then one sees that

$$\begin{aligned}
S' &= \sum_{a=1}^{3} S'_{aa} \\
&= \sum_{a=1}^{d} \sum_{i,j=1}^{3} R_{ai} R_{aj} S_{ij} \\
&= \sum_{i,j=1}^{3} \left(\sum_{a=1}^{d} R^T_{ia} R_{aj} \right) S_{ij} \\
&= \sum_{i,j=1}^{3} \delta_{ij} S_{ij} \\
&= S,
\end{aligned}$$

where we used the fact that R is an orthogonal matrix, and hence $R^T R = I$. Therefore the trace is invariant! Thus the subspace spanned by the trace is a one-dimensional invariant subspace, which transforms according to the trivial one-dimensional representation. We thus get the further decomposition $\mathbf{6} = \mathbf{5} \oplus \mathbf{1}$. Overall, the nine-dimensional tensor representation decomposes as $\mathbf{9} = \mathbf{5} \oplus \mathbf{3} \oplus \mathbf{1}$. It turns out that these summands are now all irreducible. The objects that transform according to the five-dimensional representation are the symmetric traceless tensors, which can be written as

$$S_{ij} - \delta_{ij} \frac{S}{3},$$

and are indeed traceless.

A similar decomposition, albeit slightly more complicated, is possible for higher rank tensors. For $SO(3)$, it turns out that there is a special duality that can be used to show that for higher rank tensors, only the symmetric traceless tensors give rise to new irreducible representations (see Zee's book, section IV.1). The end result, which we will not prove here, is that *all irreducible representations of $SO(3)$ can be constructed by looking at how symmetric traceless tensors transform.*

Our last question for this section is: what is the dimension of these irreducible representations? To find the dimension we need to find the number of independent components of symmetric traceless tensors. This will give the dimensions of the irreducible representations of $SO(3)$.

Lemma 5.1.2 *Let $S_{k_1\cdots k_j}$ be a rank j tensor that is fully symmetric under permutations of indices, and such that it is traceless:*

$$\sum_{k_1,k_2=1}^{3} \delta_{k_1 k_2} S_{k_1 k_2 \ldots k_j} = 0.$$

(Note that it does not matter what pair of indices we choose here, since S is fully symmetric.) Then S has $2j+1$ independent components.

Proof. We need to count the number of independent components of S. First, let us count the number of components of a symmetric rank j tensor. Suppose first that each index can only take values 1 and 2. There the possibilities are $2\cdots 2$, $2\cdots 21$, $2\cdots 211$, and so on. This gives $j+1$ possibilities. Then we add a 3. If we have a 3 for the first index, then we have j possibilities for 2s and 1s in the remaining $j-1$ indices. If we have two 3s, then we have $j-1$ possibilities for the remaining indices. And so on. Overall we get: $\sum_{k=0}^{j}(k+1) = \left(\sum_{k=0}^{j} k\right) + (j+1) = \frac{1}{2}j(j+1) + (j+1) = \frac{1}{2}(j+1)(j+2)$ possibilities. So a symmetric rank j tensor has $\frac{1}{2}(j+1)(j+2)$ independent components.

The tracelessness condition $\sum_{k_1,k_2=1}^{3} \delta_{k_1 k_2} S_{k_1 k_2 \ldots k_j} = 0$ consists in a number of independent conditions. The number of conditions here is the number of values that the indices $k_3 \cdots k_n$ can take. From the previous paragraph, this is the number of independent components of a rank $j-2$ symmetric tensor, which is $\frac{1}{2}j(j-1)$. Therefore, the total number of independent components of a symmetric traceless rank j tensor is

$$\frac{1}{2}(j+1)(j+2) - \frac{1}{2}j(j-1) = 2j+1.$$

■

The end result is that *there is an infinite number of irreducible representations for $SO(3)$, indexed by a non-negative integer j, with dimensions $2j+1$. The objects that transform according to these representations are symmetric traceless rank j tensors.*

This is the key result of this section. And, if you have done some physics, you may have encountered the formula $2j+1$ before. It is rather famous in the history of quantum mechanics and atomic physics. It corresponds to the degeneracy of states of the hydrogen atom (or spherical harmonics). It also corresponds to the multiplicity, or quantum states, of particles with integer spin. This is certainly not a coincidence, as we will see!

5.2 Representations of Lie algebras and the adjoint representation

Objectives

You should be able to:

- Recall the definition of representations of Lie algebras, and distinguish with representations of Lie groups.
- Determine the adjoint representation of a Lie algebra.

The irreducible representations of a Lie algebra are in one-to-one correspondence with the irreducible representations of its associated simply connected Lie group (the universal covering). Concretely, given a representation of a Lie algebra, one may obtain the corresponding representation of the Lie group by exponentiation. Since representations of Lie algebras are easier to construct than for the corresponding Lie groups, we now move on to the construction of irreducible representations of Lie algebras.

In this section we will define more precisely what representations of Lie algebras are, and then look at one important representation that exists for all Lie algebras: the adjoint representation.

5.2.1 Definition

Recall that a Lie algebra $\mathfrak{g}$ is a vector space with a bracket $[\cdot,\cdot]$ that satisfies a bunch of axioms. Suppose that $\mathfrak{g}$ is n-dimensional, and pick a basis $L_1,\ldots,L_n$ for $\mathfrak{g}$. Then we can write

$$[L_i, L_j] = \sum_{k=1}^{n} c_{ijk} L_k, \tag{5.2.1}$$

for some choice of structure constant c_{ijk}.

A representation of a Lie algebra is simply a mapping that represents the elements of $\mathfrak{g}$ as matrices, with the bracket $[\cdot,\cdot]$ realized as commutator of matrices. Concretely, one maps the basis $L_1,\ldots,L_n$ to n matrices that satisfy the commutation relations (5.2.1). As usual, if the matrices are $d\times d$, we call d the dimension of the representation.

More formally:

Definition 5.2.1 Representation of a Lie algebra. Let $\mathfrak{g}$ be a Lie algebra, and let V be a vector space. Let $\mathfrak{gl}(V)$ be the Lie algebra of all linear endomorphisms (i.e. linear maps $V\to V$) with bracket given by commutator of matrices $[X,Y]=XY-YX$. A **representation** Γ of $\mathfrak{g}$ is a mapping

$$\Gamma:\mathfrak{g}\to\mathfrak{gl}(V)$$

that is compatible with the brackets (this is called a **Lie algebra homomorphism**):

$$\Gamma([x,y]) = [\Gamma(x),\Gamma(y)] := \Gamma(x)\Gamma(y)-\Gamma(y)\Gamma(x),$$

for all $x,y\in\mathfrak{g}$. The **dimension** of the representation is the dimension of V. ♢

As for groups, we can talk about **irreducible representations**, which are representations such that V has no proper invariant subspace, that is, no proper subrepresentation.

Remark 5.2.2 When we define representations of groups, we map group elements to linear operators in $GL(V)$, that is, invertible matrices. Invertibility is required because of the group properties, since we map the group operation to matrix multiplication. For Lie algebras, we map elements of the vector space to linear operators in $\mathfrak{gl}(V)$, i.e. matrices that are not necessarily invertible. This is an important distinction.

We have already seen examples of Lie algebra representations.

Example 5.2.3 Trivial representation. As for groups, the trivial representation Γ always exists for Lie algebra. It is defined by sending all elements of $\mathfrak{g}$ to the zero endomorphism 0 of a one-dimensional vector space V, which maps all elements of V to the origin. Indeed, if $\Gamma(L_i)=0$ for $i=1,\ldots,n$, where $L_1, \ldots, L_n$ is a basis for $\mathfrak{g}$, then it certainly preserves the bracket, since

$$[\Gamma(L_i), \Gamma(L_j)] = \sum_{k=1}^{n} c_{ijk} \Gamma(L_k),$$

for any choice of structure constants. □

Example 5.2.4 The defining representation for $SU(2)$ and $SO(3)$. We constructed the Lie algebras $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ by looking at the infinitesimal generators of $SU(2)$ and $SO(3)$ in their fundamental, or defining, representations. This gives rise to a corresponding representation of the Lie algebras.

For $SU(2)$, we started with the 2-dimensional representation of $SU(2)$ consisting of 2×2 special unitary matrices. We found in (4.4.1) that the generators are the Pauli matrices:

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.2.2}$$

Those satisfy the commutation relations

$$[T_i, T_j] = i \sum_{k=1}^{3} \epsilon_{ijk} T_k,$$

which defines the three-dimensional Lie algebra $\mathfrak{su}(2)$. More precisely, the Pauli matrices (5.2.2) furnish a two-dimensional representation of the Lie algebra $\mathfrak{su}(2)$, since they map the generators to 2×2 matrices that obey the appropriate commutation relations.

In the case of $SO(3)$, we started with the 3-dimensional representation of $SO(3)$ consisting of 3×3 special orthogonal matrices. We found in (4.2.4) that the infinitesimal generators are:

$$L_1 = -i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad L_2 = -i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad L_3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.2.3}$$

Those also satisfy the commutation relations:

$$[L_i, L_j] = i \sum_{k=1}^{3} \epsilon_{ijk} L_k.$$

Thus (5.2.3) defines another representation of the same Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$, this time a three-dimensional one. □

5.2.2 The adjoint representation

Before we move on to the study of irreducible representations of the Lie algebra $\mathfrak{su}(2)$, let us define an important representation that exists for all Lie algebras $\mathfrak{g}$, which is called the adjoint representation.

Lemma 5.2.5 The adjoint representation. *Let $\mathfrak{g}$ be a n-dimensional Lie algebra, and pick a basis $L_1, \ldots, L_n$, with bracket*

$$[L_i, L_j] = \sum_{k=1}^{n} c_{ijk} L_k.$$

The structure constants define a n-dimensional representation of $\mathfrak{g}$, called the

adjoint representation*, as follows. One defines* n *matrices* $\Gamma_1, \ldots, \Gamma_n$ *by*

$$(\Gamma_i)_{jk} := -c_{ijk}.$$

What this equation means is that the jk*'th component of the matrix* Γ_i *is given by the structure constant* $-c_{ijk}$*. Then these* $\Gamma_1, \ldots, \Gamma_n$ *form a* n*-dimensional representation of* $\mathfrak{g}$*.*

Proof. To show that it is a representation, we need to check that the mapping respects the bracket, that is, that the matrices Γ_i satisfy the commutation relations

$$[\Gamma_i, \Gamma_j] = \sum_{k=1}^{n} c_{ijk}\Gamma_k.$$

Well, it turns out that this is a direct consequence of the Jacobi identity. Recall that for a Lie algebra $\mathfrak{g}$, the bracket $[\cdot, \cdot]$ satisfies the Jacobi identity

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0,$$

for all $x, y, z \in \mathfrak{g}$. In particular, if we pick $x = L_i, y = L_j, z = L_k$, we get

$$[L_i, [L_j, L_k]] + [L_j, [L_k, L_i]] + [L_k, [L_i, L_j]] = 0. \tag{5.2.4}$$

But $[L_i, L_j] = \sum_{k=1}^{n} c_{ijk} L_k$. Thus we can rewrite (5.2.4) as

$$\sum_{a=1}^{n} (c_{jka}[L_i, L_a] + c_{kia}[L_j, L_a] + c_{ija}[L_k, L_a]) = 0.$$

Doing the same thing one more time we get:

$$\sum_{a=1}^{n}\sum_{b=1}^{n} (c_{jka}c_{iab} + c_{kia}c_{jab} + c_{ija}c_{kab})\, L_b = 0.$$

Since the L_b are linearly independent, this means that the coefficients of the linear combination must be identically zero, so we obtain the following condition on the structure constants:

$$\sum_{a=1}^{n} (c_{jka}c_{iab} + c_{kia}c_{jab} + c_{ija}c_{kab}) = 0.$$

We note that $c_{ijk} = -c_{jik}$, so we can write:

$$\sum_{a=1}^{n} (c_{ika}c_{jab} - c_{jka}c_{iab}) = -\sum_{a=1}^{n} c_{ija}c_{akb}.$$

Now if we use the definition of the matrices of the adjoint representation $(\Gamma_i)_{jk} = -c_{ijk}$, we can rewrite this equation as

$$\sum_{a=1}^{n} ((\Gamma_i)_{ka}(\Gamma_j)_{ab} - (\Gamma_j)_{ka}(\Gamma_i)_{ab}) = \sum_{a=1}^{n} c_{ija}(\Gamma_a)_{kb}.$$

This is an equation for the kb'th component of a matrix, where the sum over a comes from matrix multiplication. Thus it can be rewritten in matrix form as the equation

$$\Gamma_i\Gamma_j - \Gamma_j\Gamma_i = \sum_{a=1}^{n} c_{ija}\Gamma_a,$$

which is precisely the commutation relation of the Lie algebra. ■

We can define the adjoint representation a little more abstractly, for those who like this stuff. The adjoint representation is a representation that maps elements of the vector space $\mathfrak{g}$ to linear operators on the vector space $\mathfrak{g}$ itself. That is, we think of elements of the Lie algebra as operators acting on the algebra itself. The formal definition is the following:

Definition 5.2.6 The adjoint representation (bis). Let $\mathfrak{g}$ be a Lie algebra, and let $\mathfrak{gl}(\mathfrak{g})$ be the Lie algebra of linear endomorphisms of $\mathfrak{g}$. The adjoint representation is defined as the mapping

$$\begin{aligned} ad :& \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \\ & x \mapsto ad_x := [x, \cdot], \end{aligned}$$

where $ad_x = [x, \cdot]$ is the operator on $\mathfrak{g}$ defined by

$$\begin{aligned} ad_x :& \mathfrak{g} \to \mathfrak{g} \\ & y \mapsto [x, y]. \end{aligned}$$

◊

In this more abstract formulation, we send an element $x \in \mathfrak{g}$ to the operator that acts on $\mathfrak{g}$ by evaluating the bracket of any $y \in \mathfrak{g}$ with the fixed $x \in \mathfrak{g}$. It may not be obvious at first that this is equivalent to the definition in terms of structure constants, but it is. We will leave the comparison as an exercise.

Checkpoint 5.2.7 Show that the abstract definition of the adjoint representation is equivalent to the definition in terms of structure constants, by calculating the components of the matrices corresponding to the linear operators ad_{L_i}.

Example 5.2.8 The adjoint representation of $\mathfrak{su}(2)$. Let us now construct the adjoint representation of $\mathfrak{su}(2)$. We recall the commutation relations

$$[T_i, T_j] = i \sum_{k=1}^{3} \epsilon_{ijk} T_k.$$

The structure constants are thus $c_{ijk} = i\epsilon_{ijk}$. We construct the adjoint representation by defining the three matrices Γ_i, $i = 1, 2, 3$ with components given by

$$(\Gamma_i)_{jk} = -i\epsilon_{ijk}.$$

Recalling the definition of the Levi-Civita symbol, we see that these matrices take the form

$$\begin{aligned} \Lambda_1 =& -i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \\ \Lambda_2 =& -i \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ \Lambda_3 =& -i \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Those are precisely the generators of $SO(3)$, see (5.2.3)! Therefore, the adjoint representation of $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ constructs the infinitesimal generators of $SO(3)$, and by exponentiation the defining representation of $SO(3)$. Note that this is particular to $SO(3)$ though; in general, the adjoint representation of a Lie algebra does not construct the defining representation of its associated Lie group. For instance, for $SO(n)$, the defining representation is n-dimensional (consisting of $n \times n$ rotation matrices), while the adjoint representation of the Lie algebra $\mathfrak{so}(n)$ is $\frac{1}{2}n(n-1)$-dimensional (the dimension of the Lie algebra). These dimensions match only when $n = 3$. □

5.3 The highest weight construction for $\mathfrak{su}(2)$

Objectives

You should be able to:

- Construct irreducible representations of $\mathfrak{su}(2)$ using the highest weight construction.
- Determine the matrices of an irreducible representation in the orthonormal basis of states given by the highest weight construction.
- Relate these representations to spherical harmonics.

Let us now do a systematic construction of irreducible representations of the Lie algebra $\mathfrak{su}(2)$ using the highest weight construction. This approach is rather general and can be suitably generalized to construct irreducible representations of Lie algebras in general.

The idea is to construct representations by looking at how they act on vector spaces. In other words, we will construct first the space on which a representation acts, and from there extract the matrices of the representation in an appropriate basis of states.

5.3.1 The highest weight construction

Let L_1, L_2, L_3 be a basis for the three-dimensional Lie algebra $\mathfrak{su}(2)$, with bracket

$$[L_i, L_j] = i \sum_{k=1}^{3} \epsilon_{ijk} L_k.$$

We want to construct representations of this algebra. To construct a representation, we are looking for three matrices $\Gamma_1, \Gamma_2, \Gamma_3$ with commutator

$$[\Gamma_i, \Gamma_j] = i \sum_{k=1}^{3} \epsilon_{ijk} \Gamma_k. \tag{5.3.1}$$

Since we are constructing representations up to equivalence, we can always use similarity transformations to make the representation as simple as possible. Our goal is to use a similarity transformation to simultaneouly diagonalize as many of the matrices Γ_is as possible. But since two matrices that can be diagonalized by the same similarity transformation must commute, we can only diagonalize one of the three Γ_is. Its is conventional to choose Γ_3 as being diagonal. In other words, we will be constructing representations of $\mathfrak{su}(2)$ such that Γ_3 is a diagonal matrix.

The idea is to construct the vector space on which the representation acts. We will call a **state** an element of the vector space on which the representation acts and that is an eigenvector of Γ_3. We use the standard Dirac "bra-ket" notation. We write the eigenstate of Γ_3 with eigenvalue m as $|m\rangle$:

$$\Gamma_3 |m\rangle = m|m\rangle.$$

At this stage, all that we know is that m is a real number, since representations of $\mathfrak{su}(2)$ are composed of Hermitian matrices (and hence with real eigenvalues).

Remark 5.3.1 The Dirac bra-ket notation. The Dirac bra-ket notation is conventional in the treatment of Lie algebras in both physics and mathematics. In this notation, we represent a complex vector v as $|v\rangle$. We represent its complex conjugate transpose $(v^*)^T$ as $\langle v|$. Thus the norm square of v can be written as

$$\langle v|v\rangle = (v^*)^T v = \sum_{i=1}^{n} |v_i|^2.$$

In general, for two vectors u, v,

$$\langle u|v\rangle = (u^*)^T v = \sum_{i=1}^{n} u_i^* v_i.$$

We are constructing finite-dimensional representations. So we know that Γ_3 only has a finite number of eigenvectors. We pick the state with the largest eigenvalue, and call this state the **highest weight state**. We will denote it by $|j\rangle$, with the letter j reserved to the largest eigenvalue of Γ_3 in a given representation. We normalize this state so that

$$\langle j|j\rangle = 1.$$

We now define a **lowering operator** Γ_- and a **raising operator** Γ_+ by

$$\Gamma_\pm = \frac{1}{\sqrt{2}}(\Gamma_1 \pm i\Gamma_2).$$

It will be convenient, for reasons that will become clear shortly, to use $\Gamma_-, \Gamma_+, \Gamma_3$ as basis for our representation, instead of $\Gamma_1, \Gamma_2, \Gamma_3$. So we should transform the commutation relations (5.3.1) into commutations relations for the new basis. A straightforward calculation shows that

$$[\Gamma_3, \Gamma_\pm] = \pm\Gamma_\pm, \qquad [\Gamma_+, \Gamma_-] = \Gamma_3. \tag{5.3.2}$$

So instead of constructing matrices $\Gamma_1, \Gamma_2, \Gamma_3$ satisfying (5.3.1), we are now trying to construct a diagonal matrix Γ_3 and two matrices $\Gamma_\pm$ satisfying (5.3.2).

$\Gamma_\pm$ are called lowering and raising operator for the following reason. Pick a state $|m\rangle$, and act with $\Gamma_3\Gamma_\pm$. We get:

$$\begin{aligned}\Gamma_3\Gamma_\pm|m\rangle =& (\Gamma_\pm\Gamma_3 \pm \Gamma_\pm)|m\rangle \\ =& (m \pm 1)\Gamma_\pm|m\rangle.\end{aligned}$$

In other words, $\Gamma_\pm|m\rangle$ is also an eigenstate of Γ_3, with eigenvalue $(m \pm 1)$. So acting with Γ_- on an eigenstate of Γ_3 produces a new eigenstate with eigenvalue lowered by 1 (hence the name lowering operator), while acting with Γ_+ produces a new eigenstate with eigenvalue raised by 1 (hence the name raising operator).

We will fix the normalization of the eigenstate of Γ_3 using Γ_-:

$$|m-1\rangle := \Gamma_-|m\rangle.$$

Then, acting on the highest weight state $|j\rangle$ with Γ_-, we produce a tower of states:

$$|j\rangle, \qquad |j-1\rangle, \qquad \ldots, \qquad |q\rangle.$$

We know that the tower must end at some point, since we are constructing a finite-dimensional representation. Thus there must be an eigenstate $|q\rangle$ of Γ_3 such that

$$\Gamma_-|q\rangle = 0.$$

The next step is to figure out what q is for a given j. This will tell us the dimension of the corresponding representation. To this end, we notice that we have not fully determined how Γ_+ acts on an eigenstate of Γ_3. All that we know is that

$$\Gamma_+|m\rangle = N_m|m+1\rangle$$

for some normalization constant N_m.

Lemma 5.3.2 *Let $|m\rangle$ be eigenstates of Γ_3 such that $\Gamma_-|m\rangle = |m-1\rangle$, with highest weight state $|j\rangle$ (that is, $\Gamma_+|j\rangle = 0$). Then*

$$\Gamma_+|m\rangle = N_m|m+1\rangle$$

with

$$N_m = \frac{1}{2}j(j+1) - \frac{1}{2}m(m+1).$$

Proof. We can compute N_m as follows:

$$\begin{aligned}
N_m|m+1\rangle =&\Gamma_+|m\rangle\\
=&\Gamma_+\Gamma_-|m+1\rangle\\
=&(\Gamma_-\Gamma_+ + \Gamma_3)|m+1\rangle\\
=&(N_{m+1} + (m+1))|m+1\rangle.
\end{aligned}$$

Therefore, we get a recursion

$$N_m = N_{m+1} + (m+1).$$

We can solve this recursion as follows. First, since $|j\rangle$ is the highest weight state, we know that $N_j = 0$. Thus

$$N_{j-1} = j.$$

Then

$$N_{j-2} = N_{j-1} + (j-1) = j + (j-1).$$

By induction, we get

$$N_{j-s} = \sum_{k=0}^{s-1}(j-k) = sj - \frac{1}{2}s(s-1) = \frac{s}{2}(2j-s+1),$$

for $s = 0, \ldots, 2j$. We redefine the index as $m = j - s$ to get:

$$N_m = \frac{j-m}{2}(j+m+1) = \frac{1}{2}j(j+1) - \frac{1}{2}m(m+1).$$

■

With this result we can determine what the lowest weight state $|q\rangle$ is. By definition, it is the end of the tower, so we must have $\Gamma_-|q\rangle = 0$. Then:

$$\begin{aligned}
0 =&\Gamma_+\Gamma_-|q\rangle\\
=&(\Gamma_-\Gamma_+ + \Gamma_3)|q\rangle\\
=&(N_q + q)|q\rangle,
\end{aligned}$$

with $N_q = \frac{1}{2}j(j+1) - \frac{1}{2}q(q+1)$. Thus we are looking for a solution to the equation

$$\frac{1}{2}j(j+1) - \frac{1}{2}q(q+1) + q = \frac{1}{2}j(j+1) - \frac{1}{2}q(q-1) = 0.$$

There are two solutions: $q = j+1$ and $q = -j$. The first one does not make sense, since $|j\rangle$ is a highest weight state. Therefore $q = -j$.

What this means is that we have constructed the following tower of eigenstates for Γ_3:

$$|j\rangle, \qquad |j-1\rangle, \qquad \ldots, \qquad |-j+1\rangle, \qquad |-j\rangle.$$

There are $2j+1$ states. Since these states span the vector space on which our representation acts, $2j+1$ must be a positive integer. That is, j must be a non-negative half-integer.

The outcome of this construction is the following important result:

Theorem 5.3.3 Irreducible representations of $\mathfrak{su}(2)$. *There is an infinite tower of irreducible representations for $\mathfrak{su}(2)$, indexed by a non-negative half-integer j, with dimensions $2j+1$. We call the $(2j+1)$-dimensional representation the **spin-j representation**, and denote its states by $|j,m\rangle$ (those are eigenstates of Γ_3 with eigenvalues $m = -j, -j+1, \ldots, j-1, j$).*

We have not proved here that these representations are irreducible, and that we have obtained all irreducible representations: this is beyond the scope of this class.

5.3.2 Normalization of states

In the construction above we define the states $|j,m\rangle$ as:

$$\Gamma_3|j,m\rangle = m|j,m\rangle, \qquad \Gamma_-|j,m\rangle = |j,m-1\rangle, \qquad \Gamma_+|j,m\rangle = N_m|j,m+1\rangle,$$

with $N_m = \frac{1}{2}j(j+1) - \frac{1}{2}m(m+1)$. One can show that these states are orthogonal, that is,

$$\langle j,m'|j,m\rangle = 0 \qquad \text{if } m' \neq m.$$

Checkpoint 5.3.4 Show that $\langle j,m'|j,m\rangle = 0$ if $m' \neq m$.

However, the states $|j,m\rangle$ are not normalized. We normalized the highest weight state by requiring that $\langle j,j|j,j\rangle = 1$, but $\langle j,m|j,m\rangle \neq 1$ for $m \leq j-1$. Let $C_m^2 = \langle j,m|j,m\rangle$. We can construct a new basis of states that are orthonormal as

$$\widetilde{|j,m\rangle} = \frac{1}{C_m}|j,m\rangle.$$

Lemma 5.3.5 *Let $\widetilde{|j,m\rangle}$ be a tower of eigenstates $\Gamma_3\widetilde{|j,m\rangle} = m\widetilde{|j,m\rangle}$, for $m = -j, -j+1, \ldots, j-1, j$ that are orthonormal:*

$$\widetilde{\langle j,m'|}\widetilde{j,m\rangle} = \delta_{m',m}.$$

Then

$$\Gamma_+\widetilde{|j,m\rangle} = \sqrt{\frac{1}{2}(j+m+1)(j-m)}\,\widetilde{|j,m+1\rangle},$$
$$\Gamma_-\widetilde{|j,m\rangle} = \sqrt{\frac{1}{2}(j+m)(j-m+1)}\,\widetilde{|j,m-1\rangle}.$$

Proof. We can construct a basis of orthonormal states from our previous basis as

$$\widetilde{|j,m\rangle} = \frac{1}{C_m}|j,m\rangle.$$

Clearly, this does not change how Γ_3 acts, since

$$\Gamma_3\widetilde{|j,m\rangle} = m\widetilde{|j,m\rangle}.$$

We want to find how Γ_+ and Γ_- act on this new basis.

Since $\Gamma_\pm = \frac{1}{\sqrt{2}}(\Gamma_1 \pm i\Gamma_2)$, and Γ_1, Γ_2 are Hermitian, we find $\Gamma_+^\dagger = \Gamma_-$ and $\Gamma_-^\dagger = \Gamma_+$. Thus we can write

$$\langle j, m-1|j, m-1\rangle = \langle j,m|\Gamma_+\Gamma_-|j,m\rangle = N_{m-1}\langle j,m|j,m\rangle,$$

where for the second equality we applied $\Gamma_+\Gamma-$ on the ket $|j,m\rangle$. Thus, if we write $C_m^2 = \langle j,m|j,m\rangle$, we get the relation

$$C_{m-1}^2 = N_{m-1}C_m^2 = \frac{1}{2}(j+m)(j-m+1)C_m^2. \tag{5.3.3}$$

Now we know that

$$\Gamma_-|j,m\rangle = |j,m-1\rangle.$$

In terms of the normalized basis, this becomes

$$\Gamma_-\widetilde{|j,m\rangle} = \frac{C_{m-1}}{C_m}\widetilde{|j,m-1\rangle}$$

Substituting (5.3.3), we get:

$$\Gamma_-\widetilde{|j,m\rangle} = \sqrt{\frac{1}{2}(j+m)(j-m+1)}\widetilde{|j,m-1\rangle}.$$

As for Γ_+, we know that

$$\Gamma_+|j,m\rangle = N_m|j,m+1\rangle = \frac{j-m}{2}(j+m+1)|j,m+1\rangle.$$

In terms of the normalized basis, we get:

$$\Gamma_+\widetilde{|j,m\rangle} = \frac{C_{m+1}}{C_m}\frac{j-m}{2}(j+m+1)\widetilde{|j,m+1\rangle}.$$

Substituting (5.3.3):

$$\begin{aligned}\Gamma_+\widetilde{|j,m\rangle} =& \sqrt{\frac{2}{(j+m+1)(j-m)}}\frac{j-m}{2}(j+m+1)\widetilde{|j,m+1\rangle}\\ =& \sqrt{\frac{1}{2}(j+m+1)(j-m)}\widetilde{|j,m+1\rangle}.\end{aligned}$$

■

Remark 5.3.6 For clarity, we will now drop the tilde symbol on the normalized states. From now on all states will be assumed to be normalized.

5.3.3 Matrix representations

We have constructed the irreducible representations of $\mathfrak{su}(2)$ using the highest weight construction. We can now write down explicit matrices for these representations, using the orthonormal basis for the vector space on which they act. We will do explicitly the most common representations in physics, the spin 1/2 and spin 1 representations.

Example 5.3.7 The spin 1/2 representation. We consider the spin 1/2 representation, with $j = 1/2$. There are two states: $|1/2, 1/2\rangle$ and $|1/2, -1/2\rangle$. Those are normalized eigenstates of Γ_3 with eigenvalues 1/2 and −1/2.

Using Lemma 5.3.5, we get:

$$\Gamma_-|1/2,1/2\rangle = \frac{1}{\sqrt{2}}|1/2,-1/2\rangle, \qquad \Gamma_-|1/2,-1/2\rangle = 0,$$

and

$$\Gamma_+|1/2,1/2\rangle = 0, \qquad \Gamma_+|1/2,-1/2\rangle = \frac{1}{\sqrt{2}}|1/2,1/2\rangle.$$

To write down a matrix representation, we think of the states as basis vector of a two-dimensional vector space. We define

$$\begin{pmatrix}1\\0\end{pmatrix} = |1/2,1/2\rangle, \qquad \begin{pmatrix}0\\1\end{pmatrix} = |1/2,-1/2\rangle.$$

Then the action of the operators Γ_3 and $\Gamma_\pm$ can be rewritten in matrix form as

$$\Gamma_3 = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \qquad \Gamma_+ = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \qquad \Gamma_- = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}.$$

We are usually interested in going back to the $\Gamma_1, \Gamma_2, \Gamma_3$ basis. Since $\Gamma_\pm = \frac{1}{\sqrt{2}}(\Gamma_1 \pm i\Gamma_2)$, we have

$$\Gamma_1 = \frac{1}{\sqrt{2}}(\Gamma_+ + \Gamma_-), \qquad \Gamma_2 = -\frac{i}{\sqrt{2}}(\Gamma_+ - \Gamma-). \tag{5.3.4}$$

Thus

$$\Gamma_1 = \frac{1}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad \Gamma_2 = -\frac{i}{2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \qquad \Gamma_3 = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Those are none other than the Pauli matrices giving the defining representation of $SU(2)$! See Example 5.2.4. Thus we have recovered the two-dimensional representation of $\mathfrak{su}(2)$ that is obtained from the infinitesimal generators of $SU(2)$. □

Example 5.3.8 The spin 1 representation. Consider now the spin 1 representation with $j = 1$, with the three normalized states $|1, 1\rangle$, $|1, 0\rangle$ and $|1, -1\rangle$. The action of $\Gamma_\pm$ can be calculated to be:

$$\Gamma_-|1,1\rangle = |1,0\rangle, \qquad \Gamma_-|1,0\rangle = |1,-1\rangle, \qquad \Gamma_-|1,-1\rangle = 0,$$

and

$$\Gamma_+|1,1\rangle = 0, \qquad \Gamma_+|1,0\rangle = |1,1\rangle, \qquad \Gamma_+|1,-1\rangle = |1,0\rangle.$$

Using the basis vectors

$$\begin{pmatrix}1\\0\\0\end{pmatrix} = |1,1\rangle, \qquad \begin{pmatrix}0\\1\\0\end{pmatrix} = |1,0\rangle, \qquad \begin{pmatrix}0\\0\\1\end{pmatrix} = |1,-1\rangle,$$

we get the matrices

$$\Gamma_+ = \begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}, \qquad \Gamma_- = \begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix}, \qquad \Gamma_3 = \begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}.$$

Using (5.3.4), we get:

$$\Gamma_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}, \qquad \Gamma_2 = -\frac{i}{\sqrt{2}}\begin{pmatrix}0&1&0\\-1&0&1\\0&-1&0\end{pmatrix}, \qquad \Gamma_3 = \begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}.$$

One can check that these matrices indeed satisfy the commutation relations of the $\mathfrak{su}(2)$ Lie algebra. In fact, they are equivalent to the defining representation of $SO(3)$ as presented in Example 5.2.4. Indeed, one can find a similarity transformation that brings the matrices of the defining representation of $SO(3)$ into the matrices above (this is the similarity trasnformation that diagonalizes Γ_3). □

5.3.4 Casimir invariant

I will not develop the general theory of Casimir elements here, but let me just say a few words in the context of $\mathfrak{su}(2)$. Roughly speaking, a Casimir element is a combination of the elements of the Lie algebra that commutes with all elements of the algebra. However, to construct a Casimir element we may take products of matrices in a given representation, and hence we actually leave the algebra (more precisely, a Casimir element lives in the "universal enveloping algebra" of a Lie algebra).

An important Casimir element is the quadratic Casimir invariant. In the context of $\mathfrak{su}(2)$, this is the operator

$$\Gamma^2 := \Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2 = \Gamma_+\Gamma_- + \Gamma_-\Gamma_+ + \Gamma_3^2.$$

Using simple properties of commutators, it is easy to see that

$$[\Gamma^2, \Gamma_1] = [\Gamma^2, \Gamma_2] = [\Gamma^2, \Gamma_3] = 0,$$

and hence Γ^2 commutes with all elements of the Lie algebra $\mathfrak{su}(2)$.

Since Γ^2 commutes with all elements of the Lie algebra, it follows that $\Gamma^2|j,m\rangle$ can only depend on j; it cannot depend on m. (That's because it must be a multiple of the identity matrix, as an operator acting on the vector space spanned by the states $|j,m\rangle$.) Thus we can calculate $\Gamma^2|j,m\rangle$ by acting on the highest weight vector:

$$\begin{aligned}\Gamma^2|j,j\rangle &= \left(\Gamma_+\Gamma_- + \Gamma_-\Gamma_+ + \Gamma_3^2\right)|j,j\rangle\\ &=(j+j^2)|j,j\rangle.\end{aligned}$$

Therefore

$$\Gamma^2|j,m\rangle = j(j+1)|j,m\rangle.$$

This Casimir invariant in physics corresponds to the total angular momentum, which you may have seen in your quantum mechanics class. It is the same for all states in a given irreducible representation of $\mathfrak{su}(2)$.

5.3.5 Differential representation and spherical harmonics

In most of this class we work with matrix representations of Lie groups and Lie algebras. But we have already seen in Subsection 4.2.4 that we can also represent operators as differential operators acting on the space of functions of some variables. For instance, in (4.2.6) we wrote down a representation for the basis vectors L_1, L_2, L_3 of the Lie algebra $\mathfrak{su}(2)$ as differential operators acting on functions $f(x, y, z)$, which we reproduce here for convenience:

$$L_1 = -i\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \qquad L_2 = -i\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \qquad L_3 = -i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).$$

As mentioned there, these operators correspond to the angular momentum operators in quantum mechanics.

Now comes the power of abstract mathematics. We can apply everyting that we have found about representation theory of $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ to this particular representation of the Lie algebra. We know that we can decompose the space of functions of (x, y, z) in terms of how they transform under the irreducible representations of $SO(3)$. It is easier to work in spherical coordinates (r, θ, ϕ), since the group of symmetries here is rotation, and so functions that are invariant under rotations will depend only on (θ, ϕ). Then we can write a basis for spherically symmetric functions $f(\theta, \phi)$ given by functions $Y_j^m(\theta, \phi)$ that are eigenfunctions of L_3:

$$L_3 Y_j^m(\theta, \phi) = m Y_j^m(\theta, \phi).$$

Moreover, by the construction of the Casimir invariant we know that these functions satisfy:

$$L^2 Y_j^m(\theta, \phi) = j(j+1) Y_j^m(\theta, \phi).$$

In other words, we decompose the space of spherically symmetric functions as a direct sum of subspaces that transform according to the irreducible representations of $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. Here we want true representations of $SO(3)$, the group of rotations, so we keep only the representations with j a non-negative integer. As a result, any spherically symmetric function can be written as a linear combination of the functions $Y_j^m(\theta, \phi)$ satisfying the properties above, with $j = 0, 1, 2, 3, \ldots$ and $m = -j, -j+1, \ldots, j-1, j$. The functions $Y_j^m(\theta, \phi)$ are known as **spherical harmonics**, and will certainly show up when you study for instance the hydrogen atom in quantum mechanics. In our context, this is just a different point of view on representation theory of the group of rotations $SO(3)$.

5.4 Projective and spin representations

Objectives

You should be able to:

- Define projective and spin representations of a group.
- Recognize the spin representations of $SO(3)$ from the representations of its Lie algebra $\mathfrak{so}(3)$.

Now that we understand the irreducible representations of the Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$, let us now go back to the representations of the Lie group $SO(3)$. We know that all irreducible representations of $\mathfrak{su}(2)$ correspond to irreducible representations of the universal cover $SU(2)$ by exponentiation. But what do they correspond to in terms of the Lie group $SO(3)$?

To answer this question we need to define projective and spin representations, to which we now turn to.

5.4.1 Projective representations

One motivation for the definition of projective representations comes from quantum mechanics. In quantum mechanics, the state of a system is specified by a vector in some vector space (or Hilbert space) of possible states. This state is called the "wave-function" of the system.

But this is not quite precise. The physical probability of finding the system in a state specified by a wave-function ψ is equal to the norm square $|\psi|^2 = \psi^* \psi$. In particular, two wave-functions that only differ by a phase give the same physical probability. So the wave-function is only really defined up to a phase factor: mathematically, the statement is that the state of a system is a vector in a projective vector space (or projective Hilbert space), where two vectors that differ by overall rescaling are identified.

Now suppose that the system has a group of symmetries. As we have already seen, the group of symmetries act on the space of states of the system. In other words, we get a representation of the group as acting on the space of states. But in fact, since states are only defined up to overall rescaling, we do not really need an honest representation of the group; it is sufficient for the symmetry operations to transform the wavefunctions only up to rescaling. This is the idea behind projective representations.

Definition 5.4.1 Projective representations. A **projective representation** of a group G is a collection of matrices $T(g)$, $g \in G$, such that it preserves the group structure up to a constant:

$$T(g)T(h) = c(g,h)T(gh)$$

for some constants $c(g,h)$.

More mathematically, a projective representation of a group G on a vector space V over a field F (such as $\mathbb{C}$ or $\mathbb{R}$) is a group homomorphism from G to the projectivization of $GL(V)$:

$$T : G \to GL(V)/F^*,$$

where elements of $GL(V)/F^*$ are equivalence classes of invertible linear transformations of V that differ by overall rescaling. ◊

This is precisely what we want in quantum mechanics, since overall rescaling of the wave-function by a phase factor does not change the physics. In other words, given a group of symmetry for a quantum mechanical system, we know that it acts on the space of states of the system as a projective representation. So our goal to understand the system is to classify projective representations of the group of symmetries.

But how do we construct projective representations of a group G? We have already seen that the irreducible representations of a Lie algebra are in one-to-one correspondence with the irreducible ordinary representations of the associated simply-connected Lie group. As for projective representations, the key result is **Bargmann's theorem**, which tells us that for a large class of groups (which includes $SO(3)$, the Lorentz group and the Poincare group), *every projective unitary representation of G comes from an ordinary representation of the universal cover*. Great! In fact, for finite-dimensional representation this is always true regardless of the group. But in quantum mechanics the Hilbert space is often infinite-dimensional, so one has to be careful, but Bargmann's theorem holds for the groups of interest such as rotations, translations and Lorentz transformations.

In the context of $SO(3)$, we know that its universal cover is $SU(2)$. So what this means is that all irreducible projective representations of $SO(3)$ come from ordinary irreducible representations of $SU(2)$, which in turn correspond to irreducible representations of the Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$, which we have already constructed. So we already know all projective representations of $SO(3)$!

Before we look at those representations more closely, let us define a particular type of projective representations that is very important in physics.

5.4.2 Spin representations

There is a particular type of projective representations that is fundamental in physics. Those are called **spin representations**. They are defined for the special orthogonal groups $SO(n)$ (and more generally for $SO(p, q)$ - the discussion below can be easily adapted to this more general setup).

Roughly speaking, spin representations are representations of the elements of $SO(n)$ (which you can think of as rotations) that pick a sign when you rotate by an angle of 2π about an axis. Those are particular types of projective representations, since they preserve the group law only up to a sign. We call **spinors** the objects that transform according to spin representations of $SO(n)$, which should be contrasted with **tensors**, which are objects that transform according to ordinary representations of $SO(n)$.

The more precise definition of spin representations is in terms of the double cover of $SO(n)$. We have already seen that $SO(3)$ has a double cover, which is given by the group $SU(2)$. It turns out that in general, the group $SO(n)$ has a unique connected double cover. We call this double cover the **spin group**, and denote it by $Spin(n)$. Then there is a group homomorphism $Spin(n) \to SO(n)$ whose kernel is $\mathbb{Z}_2$. Spin representations of $SO(n)$ are defined as being the ordinary representations of the spin group $Spin(n)$ that do not come from ordinary representations $SO(n)$.

In the case of $SO(3)$, its unique connected double cover is $SU(2)$, and hence $Spin(3) \cong SU(2)$. Therefore, we know that all ordinary representations of $SU(2)$ that do not come from ordinary representations of $SO(3)$ are spin representations.

5.4.3 Spin representations for $SO(3)$

With this under our belt, we can now go back to representation theory of $SO(3)$. In the previous section we constructed all irreducible representations of the Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ using the highest weight construction. We obtained an infinite tower of irreducible representations, indexed by a non-negative half-integer j, with dimensions $2j + 1$.

We know that these representations are in one-to-one correspondence with the irreducible representations of the simply connected Lie group $SU(2)$. But what we are often interested in in physics is the group of rotations $SO(3)$, not $SU(2)$. What do the representations of the Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ correspond to in terms of the non-simply connected Lie group $SO(3)$?

From the previous discussion, from Bargmann's theorem we know that exponentiating the representations of $\mathfrak{su}(2)$ will generally give projective representations of $SO(3)$. Moreover, since $SU(2)$ is the unique connected double cover of $SO(3)$, we know that they will all be either ordinary or spin representations of $SO(3)$. How do we distinguish between the two?

Let us first work out an example. Consider the spin-1/2 representation of $\mathfrak{su}(2)$, with $j = 1/2$. The representation is given by the Pauli matrices (4.4.1), reproduced here for convenience:

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Now consider two elements of the Lie algebra given by $\theta_1 T_3$ and $\theta_2 T_3$ for some $\theta_1, \theta_2 \in \mathbb{R}$. By exponentiation, these give rise to matrices

$$R(\theta_1) = e^{i\theta_1 T_3}, \qquad R(\theta_2) = e^{i\theta_2 T_3}.$$

Since exponentiating a diagonal matrix means exponentiating the diagonal components, we get:

$$R(\theta_1) = \begin{pmatrix} e^{i\theta_1/2} & 0 \\ 0 & e^{-i\theta_1/2} \end{pmatrix}, \qquad R(\theta_2) = \begin{pmatrix} e^{i\theta_2/2} & 0 \\ 0 & e^{-i\theta_2/2} \end{pmatrix}.$$

Now suppose that $\theta_1 + \theta_2 = 2\pi$. Then we get:

$$\begin{aligned} R(\theta_1)R(\theta_2) &= \begin{pmatrix} e^{i(\theta_1+\theta_2)/2} & 0 \\ 0 & e^{-i(\theta_1+\theta_2)/2} \end{pmatrix} \\ &= -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= -I, \end{aligned}$$

where I is the 2×2 identity matrix. But in $SO(3)$, rotating about an axis by 2π should be the identity transformation, not minus the identity transformation! Thus we conclude that this representation does not quite preserve the group structure: it preserves it only up to sign. That is, it is a spin representation of $SO(3)$!

This was somehow expected, because when we constructed representations of $SO(3)$ in terms of how tensors transform, we only obtained $(2j+1)$-dimensional representations with j a non-negative integer: we never saw the representations with j a half-integer.

But what about the other representations of $SO(3)$ with half-integer j? Are they ordinary or spin representations?

Consider the highest weight construction in the previous section. In this construction, Γ_3 was always a diagonal matrix, whose diagonal entries corresponded to the values of m between $-j$ and j. If j is a half-integer, then those diagonal entries are all half-integers. Then if we construct matrices $R(\theta)$ by exponentiating the matrices $\theta\Gamma_3$, the same argument above will go through, and we will always end up with the statement that a rotation by 2π gives minus the identity matrix. (Note that this does not happen if the diagonal values of Γ_3 are integers, which is the case when j is an integer.)

Therefore we obtain the very important result that finite-dimensional irreducible projective representations of $SO(3)$ come in two classes:

- An infinite class of ordinary representations, indexed by a non-negative integer j, with dimensions $2j+1$. Those are **tensor representations**, and we call objects that transform according to these representations **tensors** (for $j = 0$, we call them **scalars**, and for $j = 1$, we call them **vectors**).

- An infinite class of spin representations, indexed by a positive half-integer j (which is not an integer), with dimensions $2j+1$. Those are **spin representations**, and we call objects that transform according to these representations **spinors**.

Isn't that cool? Now you see how spinors come about in physics. As you may already know, in physics we call particles that transform according to ordinary representations **bosons** (they have integer spin), and particles that transform according to spin representations **fermions** (they have half-integer spins). Those have very different physical properties, due to the difference between Bose-Einstein and Fermi-Dirac statistics (the Pauli exclusion principle). For instance, particles that exchange forces, such as photons, gluons, the W and Z bosons, and the Higgs, are bosons. On the other hand, particles that make up matter, such as electrons, protons, muons, quarks, etc., are fermions. Cool, hey?

5.5 Tensor representations of $SU(N)$

Objectives

You should be able to:

- Recognize tensors as objects that transform according to representations of $SU(N)$.
- Determine the dimension of the irreducible representations of $SU(N)$ by looking at the corresponding tensors.
- Show that the defining representation of $SU(2)$ is pseudo-real.
- List the irreducible representations of $SU(3)$ as tensor representations.

In Section 5.1 we studied representations of $SO(3)$ by looking at how tensors transform under rotations. In this section we use a similar approach to study tensor representations of $SU(N)$.

5.5.1 Tensor representations of $SO(N)$

Let us start with a brief recap of the construction of tensor representations of $SO(3)$ in Section 5.1, which applies just as well to tensor representations of $SO(N)$. The idea was to construct representations of $SO(N)$ by constructing objects that transform according to these representations. These objects were called "tensors", and denoted by a letter with many indices, for instance: T_{ijk}. The number of indices (called the "rank" of the tensor) tells us how these objects transform under a rotation $R \in SO(3)$:

$$T'_{abc} = \sum_{i,j,k=1}^{3} R_{ai} R_{bj} R_{ck} T_{ijk}. \tag{5.5.1}$$

If we were to put all independent components of T_{ijk} in a column vector, the transformation rule above would define a matrix representation of $SO(3)$, which in this case would be the 27-dimensional tensor product representation $\mathbf{27} = \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$.

We also noticed that symmetry properties of the objects T_{ijk} under permutations of indices are preserved by the transformation rules (for instance, symmetric tensors are mapped to symmetric tensors). Thus restricting to tensors with specific symmetry properties defines proper invariant subspaces, or subrepresentations. In this way we can build all kinds of tensor representations of $SO(3)$. We claimed in Section 5.1 that for $SO(3)$, all irreducible representations correspond to symmetric traceless tensors, which are indexed by a non-negative integer j (the rank of the tensor, or the number of indices), and have dimension $2j+1$.

The case of $SO(N)$ is very similar. Tensors are still defined by how they transform under n-dimensional rotations, as in (5.5.1), but with $R \in SO(N)$. For instance, a tensor T_{ijk}, as in (5.5.1), transforms in the N^3-dimensional tensor product representation $\mathbf{N}^3 = \mathbf{N} \otimes \mathbf{N} \otimes \mathbf{N}$. As for $SO(3)$, tensors with particular symmetry properties define subrepresentations of these tensor product representations. However, for $SO(N)$ irreducible representations are not simply given by symmetric traceless tensors anymore. The study of irreducible representations of $SO(N)$ in terms of symmetry properties of tensors uses **Young tableaux**, which we will not discuss in this class (but feel free to research them if you are interested!). The idea is to look at the general tensor product representation $\mathbf{N} \otimes \cdots \otimes \mathbf{N}$, which is generally reducible, and look at how it decomposes as a direct sum of irreducible representations. This involves the study of **Clebsch-Gordan coefficients,** which is another interesting topic that we will not study in this class.

Let us also remark that we can identify the special representation known as the adjoint representation (which is the exponentiated version of the adjoint representation for the Lie algebra) in terms of tensors for $SO(N)$ in general. Rank two anti-symmetric tensors $A_{ij} = -A_{ji}$ transform under the adjoint representation. Indeed, a rank two anti-symmetric tensor has $\frac{1}{2}N(N-1)$ independent components, which is the dimension of the adjoint representation of $SO(N)$ (also the dimension of the Lie group).

5.5.2 Tensor representations of $SU(N)$

Let us now move on to the study of tensor representations of $SU(N)$. We now want to contruct tensors that transform in certain ways under special unitary transformations $U \in SU(N)$. The key difference with $SO(N)$ is that the matrices U are unitary, and, in particular, complex-valued. So the transformations U and $U^\dagger$ are different transformations. So we should really construct objects that transform according to products of Us and $U^\dagger$s. More precisely, what we are saying here is that the defining $\mathbf{N}$ representation of $SU(N)$ is complex-valued, so we can construct tensor products of $\mathbf{N}$ with itself, but also with its complex conjugate representation $\overline{\mathbf{N}}$.

There is a neat way of keeping track of objects that transform according to U and $U^\dagger$. We will use upper and lower indices. An object with upper indices will transform according to products of Us, while an object with lower indices will transform according to products of $U^\dagger$s. We can of course also have objects with both lower and upper indices, which will transform accordingly. Concretely, what we are doing is identifying the operation of complex conjugation as raising or lowering an index:

$$\psi_i := (\psi^*)^i.$$

This is of course just a notational convention, but it is very useful to keep track of the transformation properties of tensors.

We then denote the components of a unitary matrix U by $U^i{}_a$. Its transpose matrix $U^\dagger$ then have components $(U^\dagger)^a{}_i$. With this convention the upshot is that to construct objects with appropriate transformation properties under unitary transformations, we should always sum only over one index that is upper and one that is lower. This follows because of the unitary property that $U^\dagger U = I$, or, in index notation,

$$\sum_{j=1}^{N} (U^\dagger)^b{}_j U^j{}_a = \delta^b_a.$$

As an example, a tensor T^{ij}_k transforms as

$$(T')^{ij}_k = \sum_{a,b,c=1}^{N} U^i{}_a U^j{}_b (U^\dagger)^c{}_k T^{ab}_c.$$

It thus defines the tensor product representation $\mathbf{N} \otimes \mathbf{N} \otimes \overline{\mathbf{N}}$.

The task now is to identify proper invariant subspaces, or subrepresentations, of these tensor representations. Using the exact same argument as for $SO(N)$, it is straightforward to show that symmetry properties of tensor according to permutations of upper indices and lower indices separately are preserved by the transformation rule. For instance, a tensor that is fully symmetric under permutations of its upper indices is mapped to a tensor that is also fully symmetric under permutations of its upper indices. However, we cannot exchange upper and lower indices; symmetries under such permutations would not be preserved by the transformation rule.

Is there an analog of the trace that was preserved by special orthogonal transformations? Yes, but it must now involve summing over an upper and lower index. Indeed, one sees that, for instance,

$$\begin{aligned}
\sum_{i=1}^{N} (T')^{ji}_i &= \sum_{i=1}^{N} \sum_{a,b,c=1}^{N} U^j{}_a U^i{}_b (U^\dagger)^c{}_i T^{ab}_c \\
&= \sum_{a,b,c=1}^{N} U^j{}_a \left(\sum_{i=1}^{N} U^i{}_b (U^\dagger)^c{}_i \right) T^{ab}_c \\
&= \sum_{a,b,c=1}^{N} U^j{}_a \, (\delta^c_b) \, T^{ab}_c \\
&= \sum_{a=1}^{N} U^j{}_a \left(\sum_{b=1}^{N} T^{ab}_b \right).
\end{aligned}$$

Therefore, the object $\sum_{i=1}^{N} T^{ji}_i$ transforms like a tensor with one single upper index (i.e. in the $\mathbf{N}$ defining representation of $SU(N)$). So summing over one upper and one lower index for any tensor creates a subrepresentation. We call this operation taking the trace of a tensor.

Since taking the trace always defines a subrepresentation, to understand irreducible representations of $SU(N)$ we know that we should look for traceless tensors that have appropriate symmetry properties upstairs and downstairs. In general, the study of irreducible representations from this point of view involves Young tableaux, as already mentioned for $SO(N)$. The idea is similar, where we now look at general tensor products $\mathbf{N} \otimes \cdots \mathbf{N} \otimes \overline{\mathbf{N}} \otimes \cdots \otimes \overline{\mathbf{N}}$, and study their decompositions as direct sums of irreducible representations using Clebsch-Gordan coefficients.

We will not discuss Young tableaux in this course, but let us at least enumerate the first few non-trivial tensor representations of $SU(N)$:

Table 5.5.1 The first few tensor representations of $SU(N)$

Tensor	Symmetry property	Dimension of the representation
T^i	-	N
$T^{ij} = -T^{ji}$	Anti-symmetric	$\frac{1}{2}N(N-1)$
$T^{ij} = T^{ji}$	Symmetric	$\frac{1}{2}N(N+1)$
T^i_j	Traceless ($\sum_{i=1}^N T^i_i = 0$)	$N^2 - 1$

We note that the N-dimensional, $\frac{1}{2}N(N-1)$-dimensional and $\frac{1}{2}N(N+1)$-dimensional representations also have complex conjugate representations, corresponding to similar tensors but with lower indices. We also note that the representation given by the traceless tensor T^i_j is the adjoint representation of $SU(N)$, which is equivalent to its complex conjugate.

5.5.3 Tensor representations of $SU(2)$

We already know the irreducible representations of $SU(2)$, since they are in one-to-one correspondence, through exponentiation, with the irreducible representations of its Lie algebra $\mathfrak{su}(2)$. We already constructed those using the highest weight construction, and obtained an infinite family indexed by a non-negative half-integer j, with dimensions $2j+1$. Let us see how those arise from the point of view of tensor representations.

Just as for $SO(3)$, it turns that $SU(2)$ is very special. In the case of $SO(3)$, what was particular about it is that irreducible representations were all given by symmetric traceless tensors. In the case of $SU(2)$, it turns out that *all irreducible representations are constructed from symmetric tensors with only upper indices*. In other words, for $SU(2)$, we do not need tensors with lower indices. We will see in a second why.

So all irreducible representations of $SU(2)$ correspond to fully symmetric tensors $T^{i_1 \cdots i_n}$. So we obtain an infinite family of irreducible representations, indexed by a non-negative integer n. What is the dimension of such a representation? Recall that the indices i_1 to i_n can only take values 1 or 2, since we are considering $SU(2)$, and its defining representation is two-dimensional. Thus the independent components of a fully symmetric tensor $T^{i_1 \cdots i_n}$ are:

$$T^{11\cdots 1}, \qquad T^{21\cdots 1}, \qquad T^{221\cdots 1}, \qquad \ldots \qquad T^{22\cdots 2}.$$

There are exactly $n+1$ of them. So we obtain an infinite family of irreducible representations indexed by a non-negative integer n with dimensions $n+1$. This is exactly what we obtained before with the highest weight construction, if we let $n = 2j$ with j a half-integer. Great!

But why is it that we only have to consider tensors with upper indices for $SU(2)$? Concretely, this means that we are only considering tensor products of the defining representation **2** with itself: we do not consider tensor products with its complex conjugate $\overline{\mathbf{2}}$. Why?

Well, let us look at the defining representation **2** more closely. Recall that it is given by the Pauli matrices (4.4.1):

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The question is: is the group representation obtained by exponentiating the Pauli matrices complex, real, or pseudo-real (see Section 2.12)? If it was real or pseudo-real, this would mean that it is equivalent to its complex conjugate representation, which would justify only looking at tensor products of **2** with itself (since $\overline{\mathbf{2}}$ would be equivalent to **2**).

Lemma 5.5.2 The defining representation of $SU(2)$ is pseudo-real. *The defining, two-dimensional, representation of $SU(2)$ is pseudo-real.*

Proof. Recall that a representation is real or pseudo-real if it is equivalent to its complex conjugate representation. That is, there exists an invertible matrix S such that

$$T(g)^* = ST(g)S^{-1}$$

for all $g \in SU(2)$. It is real if S is symmetric, and pseudo-real if S is anti-symmetric.

The defining representation of $SU(2)$ is obtained by exponentiating a general real linear combination of Pauli matrices:

$$T(g) = \exp(i \sum_{k=1}^{3} \theta_k T_k).$$

Then its complex conjugate is:

$$T(g)^* = \exp(-i \sum_{k=1}^{3} \theta_k T_k^*).$$

Thus,

$$T(g)^* = ST(g)S^{-1}$$

for some S if and only if

$$T_k^* = -ST_k S^{-1}$$

for the Pauli matrices T_k, $k = 1, 2, 3$.

One can check using matrix multiplication that Pauli matrices satisfy the property that

$$T_i T_j = -T_j T_i \qquad \text{for } i \neq j, \qquad T_i^2 = I.$$

Since T_1 and T_3 are real-valued, it thus follows that

$$T_2 T_1^* T_2 = T_2 T_1 T_2 = -T_1, \qquad T_2 T_3^* T_2 = T_2 T_3 T_2 = -T_3.$$

As for T_2, we see that $T_2^* = -T_2$, and thus

$$T_2 T_2^* T_2 = -T_2 T_2 T_2 = -T_2.$$

Thus, if we set $S = T_2$, and hence $S^{-1} = T_2$, we get that $T_k^* = -ST_k S^{-1}$, as required. Finally, since T_2 is anti-symmetric, we conclude that the defining representation of $SU(2)$ is pseudo-real. ■

This is the key property that makes $SU(2)$ special. Since its defining representation is pseudo-real, that is, $\mathbf{2} \cong \overline{\mathbf{2}}$, we do not need to consider tensor products involving both $\mathbf{2}$ and $\overline{\mathbf{2}}$, since

$$\mathbf{2} \otimes \cdots \otimes \mathbf{2} \otimes \overline{\mathbf{2}} \otimes \cdots \otimes \overline{\mathbf{2}} \cong \mathbf{2} \otimes \cdots \otimes \mathbf{2}.$$

Concretely, what this means is that we only need to consider tensors with upper indices for $SU(2)$.

Note that this is certainly not true for $SU(N)$ in general though. The defining representation is pseudo-real only for $SU(2)$. For $N > 2$, the defining representation is complex.

5.5.4 Tensor representations of $SU(3)$

Let us now look at $SU(3)$. In this case, the defining representation $\mathbf{3}$ is complex, and we need to consider general tensor products $\mathbf{3} \otimes \cdots \mathbf{3} \otimes \overline{\mathbf{3}} \otimes \cdots \otimes \overline{\mathbf{3}}$.

For $SU(3)$, the statement is that *all irreducible representations are given by traceless tensors that are fully symmetric upstairs an downstairs individually.* That is, tensors of the form $T^{i_1\cdots i_p}_{k_1\cdots k_q}$ that are traceless and fully symmetric under permutations of the i_a and the k_b separately.

Thus we conclude that irreducible representations of $SU(3)$ are indexed by two non-negative integers (p, q), corresponding to the number of upper and lower indices respectively. What is the dimension of such a representation?

Lemma 5.5.3 Dimension of irreducible representations of $SU(3)$. *A traceless tensor $T^{i_1\cdots i_p}_{k_1\cdots k_q}$ that is fully symmetric under permutations of the i_a and the k_b separately has*

$$\frac{1}{2}(p+1)(q+1)(p+q+2)$$

independent components. Thus, the dimension of the (p, q) irreducible representation of $SU(3)$ is $\frac{1}{2}(p+1)(q+1)(p+q+2)$.

Proof. A tensor $T^{i_1\cdots i_p}_{k_1\cdots k_q}$ that is fully symmetric under permutations of the i_a and the k_b separately has

$$\left(\frac{1}{2}(p+1)(p+2)\right)\left(\frac{1}{2}(q+1)(q+2)\right) \tag{5.5.2}$$

independent components. We then need to impose the tracelessness condition, which will impose a number of constraints. Let us count the number of constraints, which are independent, and then substract this number from the number of above to get the number of independent components of a traceless tensor that is fully symmetric upstairs and downstairs.

To impose that a tensor is traceless, we need to sum over one upper and one lower index. We can choose any of those, since the tensor is fully symmetric upstairs and downstairs. Thus we need to impose the conditions

$$\sum_{a=1}^{3} T^{ai_2\cdots i_p}_{ak_2\cdots k_p} = 0.$$

How many constraints does this give? Well, the indices $i_2, \ldots, i_p$ and $k_2, \ldots, k_p$ can take any values here. So the number of constraints is the number of independent components of a tensor with $p-1$ upper indices and $q-1$ lower indices that is fully symmetric upstairs and downstairs. By (5.5.2), this is

$$\left(\frac{1}{2}(p)(p+1)\right)\left(\frac{1}{2}(q)(q+1)\right). \tag{5.5.3}$$

Therefore, the number of independent components of a traceless tensor $T^{i_1\cdots i_p}_{k_1\cdots k_q}$ that is fully symmetric under permutations of the i_a and the k_b separately is (5.5.2) minus (5.5.3), that is,

$$\frac{1}{4}(p+1)(p+2)(q+1)(q+2) - \frac{1}{4}p(p+1)q(q+1) = \frac{1}{2}(p+1)(q+1)(p+q+2).$$

■

To end this section, let us list the first few irreducible representations of $SU(3)$ in terms of the integers (p, q). We write the representation in terms of dimension, and give its name when appropriate.

Table 5.5.4 The first few irreducible representations of $SU(3)$

(p, q)	Representation (name)
$(0, 0)$	$\mathbf{1}$ (trivial)
$(1, 0)$	$\mathbf{3}$ (fundamental)
$(0, 1)$	$\overline{\mathbf{3}}$ (anti-fundamental)
$(2, 0)$	$\mathbf{6}$
$(0, 2)$	$\overline{\mathbf{6}}$
$(1, 1)$	$\mathbf{8}$ (adjoint)
$(3, 0)$	$\mathbf{10}$
$(0, 3)$	$\overline{\mathbf{10}}$

In fact, there is an interesting story here. In the early 1960s, nine short-lived baryonic particles had been observed. They all shared similar properties and masses. So it was believes that those should live some irreducible representation of the symmetry group of the theory. Using the fact that 8 spin 0 mesons and 8 spin 1/2 baryons, Gell-Mann proposed that the symmetry group of the strong force should be $SU(3)$, and that those came in the adjoint representation of $SU(3)$. Then, since there is no 9-dimensional irreducible representation for $SU(3)$, he made the striking prediction that there should be a 10th short-lived baryonic particle with similar properties and masses as the nine other ones, so that they transformed in the 10-dimensional irreducible representations of $SU(3)$. The missing particle was soon found. Gell-Mann's prediction is a remarkable achivement of representation theory in particle physics!

5.6 The Standard Model of particle physics and GUTs

Objectives

You should be able to:

- Recognize that particles of the Standard Model can be labelled as representations of the gauge group $SU(3) \times SU(2) \times U(1)$.
- State the general idea, from the perspective of representation theory, of Grand Unified Theories (GUTs).

We are now ready to do some cool physics and see how particles in the Standard Model can be labeled by representations, and how the idea of Grand Unified Theories (GUTs) then naturally arises.

5.6.1 The Standard Model of particle physics

In this section we summarize some representation-theoretic aspects of the Standard Model in particle physics. For more on the mathematical formulation of the Standard Model, see for instance this wikipedia page.

It has been emphasized a number of times in this course that states of a quantum mechanical systems can be labeled using irreducible representations of the group of symmetries. In quantum field theory, these states correspond to particles. So in the particle physics, we can label particles in terms of how they transform under the group of symmetries, i.e., in terms of the corresponding representations.

There are two sources of symmetries in the Standard Model of particle physics. One is the Lorentz group (or Poincare group) corresponding to symmetries of spacetime. So we can label particles according to how they transform under the Lorentz group. We have not seen yet the representations of the Lorentz group, but, roughly speaking, they are similar to the representations of $SO(3)$. In particular, this source of symmetry gives rise to the spin of a particle, which tells us in which representation of the Lorentz group it transforms.

Another source of symmetries is the gauge group of the Standard Model, which is a group of transformations of the theory that leaves the physics invariant (such as the $U(1)$ transformations of the electromagnetic potential in electromagnetism). It turns out that the gauge group of the Standard Model is $SU(3)\times SU(2)\times U(1)$. The $SU(2)\times U(1)$ factor comes from the electroweak force, while the $SU(3)$ factor corresponds to the strong force. Therefore, particles are labeled by how they transform under this gauge group: in other words, they are labeled by representations of $SU(3)$, $SU(2)$ and $U(1)$.

Recall the irreducible representations of $U(1)$ are all one-dimensional (since $U(1)$ is abelian), and take the form $e^{ik\theta}$ with $\theta \in [0, 2\pi)$ and k a non-negative integer. So they are indexed by an integer, which we call the charge. It is customary in the Standard Model to specify this charge as being the so-called "weak hypercharge", which is not quite an integer, but rather a rational number.

So we usually specify a representation of $SU(3) \times SU(2) \times U(1)$ as $(\mathbf{p}, \mathbf{q})_n$, where $\mathbf{p}$ is the p-dimensional representation of $SU(3)$, $\mathbf{q}$ is the q-dimensional representation of $SU(2)$, and n is the weak hypercharge. With this notation, we can list the particles of the Standard Model, with their spin (representation of the Lorentz group) and their representation under $SU(3) \times SU(2) \times U(1)$. The following standard table is pretty much taken from wikipedia.

Table 5.6.1 The Standard Model of particle physics

Particle	Name	Representation
Spin 1		
B	Z boson	$(\mathbf{1},\mathbf{1})_0$
W	W boson	$(\mathbf{1},\mathbf{3})_0$
G	gluon	$(\mathbf{8},\mathbf{1})_0$
Spin 1/2		
q_L	left-handed quark	$(\mathbf{3},\mathbf{2})_{1/3}$
u_L^c	left-handed antiquark (up)	$(\overline{\mathbf{3}},\mathbf{1})_{-4/3}$
d_L^c	left-handed antiquark (down)	$(\overline{\mathbf{3}},\mathbf{1})_{2/3}$
ℓ_L	left-handed lepton	$(\mathbf{1},\mathbf{2})_1$
ℓ_L^c	left-handed antilepton	$(\mathbf{1},\mathbf{1})_2$
Spin 0		
H	Higgs boson	$(\mathbf{1},\mathbf{2})_1$

What is pretty cool is that you should now understand what the notation stands for, and be able to read this table!

A few things to note:

- The spin 1 particles, which correspond to the gauge bosons (that carry forces), all come in the adjoint representations of $SU(3)$, $SU(2)$ and $U(1)$. This is a general feature, gauge bosons generally transform according to the adjoint representation.
- The matter particles, i.e. those that compose matter, such as electrons,

muons, quarks, etc. all correspond to spin 1/2 particles (i.e. they are spinors).

- The Higgs boson is the only spin 0 particle.
- However, with the recent discovery of neutrino masses and neutrino oscillation, it is generally believed that one more particle should be added to the Standard Model, which would responsible for giving masses to the neutrinos. This particle is the "right-handed neutrino", which would transform according to the trivial representation $(\mathbf{1},\mathbf{1})_0$. The existence of the right-handed neutrino has not been confirmed experimentally however.

This is all very cool. But, from a theoretical physics viewpoint, the representations that appear in the Standard Model appear rather random. Why are there particles transforming in the $(\mathbf{3},\mathbf{2})_{1/3}$ representation, but not in, say, the $(\overline{\mathbf{3}},\mathbf{2})_{1/3}$ representation? Who chose this seemingly random list of representations?

One answer to this question could be that there is no answer, and that this is just what Nature says. That is a fine answer. But often, when there is something that seems rather random or mysterious, this means that we are missing something. And going a little deeper into representation theory, one sees that indeed, these representations are not random. This is the fundamental idea behind Grand Unified Theories (GUTs).

5.6.2 $SU(5)$ GUT

The idea of GUTs is that perhaps the gauge group of the Standard Model is in fact larger than $SU(3)\times SU(2)\times U(1)$, but is somehow broken to $SU(3)\times SU(2)\times U(1)$ at some higher energy scale. Why would that be a good idea? You will see!

Perhaps the simplest embedding of $SU(3)\times SU(2)\times U(1)$ as a subgroup of a larger group is

$$SU(3)\times SU(2)\times U(1)\subset SU(5).$$

This embedding can be seen in terms of the defining representations. Roughly speaking, one constructs a subgroup of 5×5 special unitary matrices by looking at those that are block diagonal, with a 3×3 block with determinant one ($SU(3)$) and a 2×2 block with determinant one ($SU(2)$). You also include diagonal matrices of the form $\text{diag}(a,a,b,b,b)$ with $a^2b^3=1$, which generates a copy of $U(1)$.

Now given irreducible representations of $SU(5)$, one can restrict to the subgroup $SU(3)\times SU(2)\times U(1)\subset SU(5)$. Those will generally now be reducible representations of the subgroup. One can work out how those decompose as direct sums of irreducible representations of the subgroups. Such restriction and decomposition are called branching rules in physics, and some textbooks include tons of table of such branching rules (see for instance Slansky's book Group Theory for Unified Model Building.)

Let us then look at the simplest irreducible representations of $SU(5)$. We know what those are: they are the tensor representations that we constructed earlier. Here are the first few, with how they decompose as direct sums of irreducible representations of the subgroup $SU(3)\times SU(2)\times U(1)\subset SU(5)$:

$$\begin{aligned}\mathbf{1}\to&(\mathbf{1},\mathbf{1})_0,\\ \mathbf{5}\to&(\mathbf{3},\mathbf{1})_{-1/3}\oplus(\mathbf{1},\mathbf{2})_{1/2},\end{aligned}$$

$$\begin{aligned}
\overline{\mathbf{5}} \to &(\overline{\mathbf{3}}, \mathbf{1})_{1/3} \oplus (\mathbf{1}, \mathbf{2})_{-1/2},\\
\mathbf{10} \to &(\mathbf{3}, \mathbf{2})_{1/6} \oplus (\overline{\mathbf{3}}, \mathbf{1})_{-2/3} \oplus (\mathbf{1}, \mathbf{1})_{1},\\
\mathbf{24} \to &(\mathbf{8}, \mathbf{1})_0 \oplus (\mathbf{1}, \mathbf{3})_0 \oplus (\mathbf{1}, \mathbf{1}_0 \oplus (\mathbf{3}, \mathbf{2})_{-5/6} \oplus (\overline{\mathbf{3}}, \mathbf{2})_{5/6}.
\end{aligned}$$

The last representation is the adjoint representation, whose decomposition always includes the adjoint representation of the subgroup, and extra stuff.

Now the really cool thing here is that all the seemingly random representations corresponding to the matter content of the Standard Model (see Table 5.6.1) are in fact the irreducible representations that appear in the decomposition of the representation $\overline{\mathbf{5}} \oplus \mathbf{10}$ of $SU(5)$! (Note that the weak hypercharge in the decomposition above has been rescaled by 2 in comparison with Table 5.6.1, which is just a choice of convention.) Thus the representations do not appear so random anymore. This is so nice that it cannot simply be a coincidence.

As for the other particles of the Standard Model, the gauge bosons arise from the decomposition of the adjoint **24** of $SU(5)$, and the Higgs arises from the **5** and $\overline{\mathbf{5}}$. The right-handed neutrino, if observed, would need to be added as coming from the trivial representation **1** of $SU(5)$.

This is what motivated Georgi and Glashow to try to come up with a gauge theory based on $SU(5)$ that would underly the Standard Model. However, the simplest such models run into trouble; for instance, they predict proton decay with a lifetime that is too short to be consistent with experiments (proton decay has never been observed). But this just means that the simple Georgi-Glashow model cannot be correct. The beauty of the representation-theoretic unification may appear in a different context. It is just too nice to be coincidental.

5.6.3 $SO(10)$ GUT

In fact, one can keep going and look at further embeddings into larger gauge groups. The next level of embedding consists in looking at the subgroup

$$SU(5) \times U(1) \subset SO(10).$$

As you will see, things become even nicer!

As before, we look at irreducible representations of $SO(10)$, and see how their restriction to the subgroup decomposes as a direct sum of irreducible representations of the subgroup. The first few non-trivial representations (we include the spin representation **16**) decompose as follows.

$$\begin{aligned}
\mathbf{10} \to &\mathbf{5}_{-2} \oplus \overline{\mathbf{5}}_2,\\
\mathbf{16} \to &\mathbf{10}_1 \oplus \overline{\mathbf{5}}_{-3} \oplus \mathbf{1}_5,\\
\mathbf{45} \to &\mathbf{24}_0 \oplus \mathbf{10}_{-4} \oplus \overline{\mathbf{10}}_4 \oplus \mathbf{1}_0.
\end{aligned}$$

Amazingly, all the matter content, *including the right-handed neutrino*, arises from the decomposition of a single representation, the spin representation **16** of $SO(10)$! Isn't that amazing? This cannot be a coincidence.

As for the rest of the particles, the gauge bosons as always come from the decomposition of the adjoint **45** of $SO(10)$, while the Higgs comes from the **10** of $SO(10)$.

The $SO(10)$ GUTs do not suffer from the proton decay problem of $SU(5)$. However, they have other issues, such as the so-called "doublet-triplet splitting problem".

Nevertheless, the point here is that the representation-theoretic "unification" of the particles of the Standard Model is really quite elegant. What it points to is that the larger groups $SU(5)$ and/or $SO(10)$ may perhaps play a role in whatever theory of physics goes beyond the Standard Model. If this is the case, it would be a tremendous achievement of representation theory in particle physics.

5.6.4 Beyond $SO(10)$

Why stop at $SO(10)$? In fact, there is a natural chain of embeddings:

$$SU(5) \subset SO(10) \subset E_6 \subset E_7 \subset E_8,$$

where E_6, E_7, E_8 are exceptional Lie groups. The fundamental representation of E_6 is 27-dimensional, and it turns out that for the subgroup $SO(10) \times U(1) \subset E_6$ it decomposes as

$$\mathbf{27} \to \mathbf{16}_1 \oplus \mathbf{10}_{-2} \oplus \mathbf{1}_4.$$

In particular, it now gives rise to not only the matter content of the Standard Model ($\mathbf{16}_1$), but also the Higgs ($\mathbf{10}_{-2}$). All those unify into a single representation of E_6! Let me state that one more time: all the matter content of the Standard Model, including the Higgs and the right-handed neutrino, all combine into the fundamental representation of E_6. Wow!

As for the gauge bosons, as always, they come from the decomposition of the adjoint of E_6.

Why then stop at E_6? Well, now there is a good reason to stop. E_6 is the only exceptional group that has complex representations, and those are needed to give rise to chiral fermions as in the Standard Model through standard symmetry breaking by a mechanism of Higgs-type. So standard GUT theories must somehow stop at E_6. It is no good to try to construct standard GUT theories with gauge group larger than E_6. In any case, all the matter content has already unified into a single representation, so there is no reason really to look at larger gauge groups.

However, in string theory we can go further, and in fact we often must. But in string theory we can break symmetries using other mechanisms, such as string mechanisms, so we are allowed to go further. In fact, some flavours of string theory (such as heterotic string theory) naturally come with an E_8 gauge group. So it seems like this representation-theoretic unification does take place in string theory, at least in some string models. Of course, we do not know yet whether string theory is a valid description of Nature. But it is quite nice that it naturally gives rise to a representation-theoretic unification of the particles of the Standard Model.

In any case, I hope that I have convinced you in this section that representation theory is essential to understand the Standard Model of particle physics. And that it points towards some sort of unification of the matter content into representations of a larger gauge group, which, if physically correct, would be a striking prediction of representation theory.

5.7 Representations of the Lorentz group

Objectives

You should be able to:

- Construct the ordinary and spin representations of $SO(4)$ by looking at its spin group.
- Recognize the representations of the Lorentz group, and determine which

ones are spin representations.

So far we studied representations of $SU(N)$, and ordinary and spin representations for $SO(3)$. But the symmetry group of spacetime in special relativity is the Lorentz group, which is composed of Lorentz transformations. Thus in quantum field theory, which is the unique formalism that naturally combines quantum mechanics and special relativity, we want to think of our fields in terms of how they transform under the Lorentz group. That is, we want to think of them in terms of representations of the Lorentz group. This is what we study in this section.

5.7.1 Representations of $SO(4)$

Before we look at the Lorentz group it is instructive to look at the group of rotations in four dimensions, $SO(4)$. This is because the Lorentz group $SO(3,1)$ is also a group of rotations in four dimensions, but with respect to the Minkowski metric instead of the Euclidean metric. As a result, many of the properties of representations of $SO(4)$ carry through to the Lorentz group $SO(3,1)$.

Remark 5.7.1 More precisely, in this section we focus on the component of $SO(3,1)$ that is connected to the identity, which is often denoted by $SO(3,1)^+$. This is the group that is physically represented in terms of Lorentz transformations of spacetime. To avoid cluttering notation in the following we will denote this component by $SO(3,1)$, but keep in mind that we always only consider its component connected to the identity.

5.7.1.1 The spin group $Spin(4)$

The double cover of $SO(4)$ is the spin group $Spin(4)$. We know that the representations of $Spin(4)$ all descend to either ordinary or spin representations of $SO(4)$, depending on whether they assign the same matrix to the elements of the kernel of the map $Spin(4) \to SO(4)$ (if they do, they descend to ordinary representations; if they do not, they descend to spin representations). Thus, to understand the representations of $SO(4)$, we want to study representations of $Spin(4)$.

In the case of $SO(3)$, we found that $Spin(3) \cong SU(2)$, which was nice and easy. It turns out that $Spin(4)$ is also isomorphic to nice and easy group: $Spin(4) \cong SU(2) \times SU(2)$.

Remark 5.7.2 Note that the spin groups $Spin(N)$ in general are not necessarily isomorphic to such nice and easy groups. But there are such isomorphisms for small enough N. For instance, $Spin(6) \cong SU(4)$.

How can we see that $Spin(4) \cong SU(2) \times SU(2)$? Intuitively, one can think of it as follows. The group $SU(2)$ is in fact the unit three-sphere S^3 as a manifold, as we saw in Example 4.1.10. In other words, an element of $SU(2)$ corresponds to a point on a S^3, that is, a point in $\mathbb{R}^4$ that satisfies the equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1.$$

Thus, specifying an element of the product group $SU(2) \times SU(2)$ is equivalent to specifying two elements of $SU(2)$, or, in other words, two points on the three-sphere S^3. But the group of symmetries of a three-sphere is the group of rotations $SO(4)$ in four dimensions. Thus, these two points will be related by a rotation. In summary, to an element of $SU(2) \times SU(2)$, we can assign an element of $SO(4)$, corresponding to the rotation that brings the first point on S^3 to the second. Finally, one sees that this mapping is not one-to-one, since the two points $U, V \in S^3$ and the opposite points $-U, -V \in S^3$ will be mapped to the same rotation in $SO(4)$.

From the point of view of Lie algebras, the statement that $SU(2) \times SU(2)$ is a double cover of $SO(4)$ implies that they must share the same Lie algebra $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ (the Lie algebra of a product group is the direct sum of the Lie algebras of the individual factors). Let us see this explicitly.

Lemma 5.7.3 The isomorphism between $\mathfrak{so}(4)$ and $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$. *The Lie algebra $\mathfrak{so}(4)$ is isomorphic to the Lie algebra $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$.*

Proof. Following along the same lines as we did for $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$, we can find an explicit description of the Lie algebra $\mathfrak{so}(4)$. It is six-dimensional, and there is a natural choice of basis, which we denote by J_i, K_i with $i = 1, 2, 3$, with the bracket:

$$[J_i, J_j] = i \sum_{k=1}^{3} \epsilon_{ijk} J_k, \tag{5.7.1}$$

$$[J_i, K_j] = i \sum_{k=1}^{3} \epsilon_{ijk} K_k, \tag{5.7.2}$$

$$[K_i, K_j] = i \sum_{k=1}^{3} \epsilon_{ijk} J_k. \tag{5.7.3}$$

To see the isomorphism with the Lie algebra $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, we do a change of basis. We define a new basis

$$J_{\pm,i} = \frac{1}{2}(J_i \pm K_i), \qquad i = 1, 2, 3.$$

It is straightforward to check that the bracket becomes, in this new basis,

$$[J_{+,i}, J_{-,j}] = 0, \qquad \text{for all } i, j = 1, 2, 3,$$

$$[J_{+,i}, J_{+,j} = i \sum_{k=1}^{3} \epsilon_{ijk} J_{+,k},$$

$$[J_{-,i}, J_{-,j} = i \sum_{k=1}^{3} \epsilon_{ijk} J_{-,k}.$$

We recognize that the $J_{+,i}$ and the $J_{-,i}$ generate two copies of the $\mathfrak{su}(2)$ Lie algebra. Moreover, those commute, and hence the Lie algebra is a direct sum $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$. We conclude that $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$. ∎

5.7.1.2 Representations of $SO(4)$

Now that we have established that $Spin(4) \cong SU(2) \times SU(2)$, we can find the irreducible representations of $Spin(4)$. The irreducible representations of $SU(2) \times SU(2)$ are given by tensor products $T \otimes S$ of irreducible representations S and T of $SU(2)$. But from Section 5.3, we already know the irreducible representations of $SU(2)$. There is an infinite tower of finite-dimensional irreducible representations, indexed by a non-negative half-integer j, of dimension $2j + 1$. Therefore, the irreducible representations of $SU(2) \times SU(2)$ are indexed by two non-negative half-integers j_1 and j_2, and have dimensions $(2j_1 + 1)(2j_2 + 1)$.

We know that all these representations descend to either ordinary irreducible representations of $SO(4)$, or spin representations. We also know that all irreducible ordinary and spin representations of $SO(4)$ arise from irreducible representations of $Spin(4)$. So to conclude our study of representation theory of $SO(4)$, all we need to do is determine which of the representations above descend to ordinary representations of $SO(4)$, and which descend to spin representations.

Theorem 5.7.4 Representations of $SO(4)$**.** *The finite-dimensional irreducible projective representations of $SO(4)$ are indexed by two non-negative half-integers j_1 and j_2 and have dimensions $(2j_1+1)(2j_2+1)$. They are ordinary representations if $j_1+j_2 \in \mathbb{Z}$, and spin representations otherwise.*

Proof. We will simply give a rough argument here. To determine whether a representation of $SU(2)\times SU(2)$ descends to ordinary representation of $SO(4)$, we look at the kernel of the double covering $SU(2) \times SU(2) \to SO(4)$. If the given representation assigns the same matrix to all elements of the kernel, then it descends to an ordinary representation of $SO(4)$. If it doesn't, it descends to a spin representation of $SO(4)$.

The kernel of the double covering $SU(2) \times SU(2) \to SO(4)$ is $\mathbb{Z}_2$. Its non-trivial element is given by the tensor product of the two non-trivial elements in the kernel of the mapping $SU(2) \to SO(3)$ that we studied before. In terms of the fundamental representation of $SU(2)$ given by 2×2 special unitary matrices, the kernel of $SU(2) \to SO(3)$ was given by the two matrices $\{I, -I\}$, where I is the 2×2 identity matrix. For the mapping $SU(2)\times SU(2) \to SO(4)$, the kernel is then given by the tensor products $\{U \otimes I, (-I) \otimes (-I)\}$.

Let us continue with this particular representation. It corresponds to $j_1 = j_2 = 1/2$. It is a four-dimensional representation of $SU(2) \times SU(2)$. Does it descend to an ordinary representation of $SO(4)$? To the identity element it assigns the matrix $I \otimes I$, which is of course the 4×4 identity matrix. To the non-trivial element in the kernel, it assigns the 4×4 matrix given by the tensor product

$$(-I) \otimes (-I) = I \otimes I,$$

which is again the 4×4 identity matrix. Therefore, it descends to an ordinary representation of $SO(4)$: it is in fact the fundamental representation of $SO(4)$ (which is interestingly built by taking the tensor product of two copies of the spin-1/2 representation of $SU(2)$).

In general, the non-trivial element of the kernel is given by the matrix obtained as the tensor product $e^{2\pi i T_3^{(1)}} \otimes e^{2\pi i T_3^{(2)}}$, where we use the notation of Section 5.4. Recall that for a given irreducible representation of $SU(2)$ indexed by a half-integer j, the matrix T_3 is a diagonal matrix with entries that are half-integers (but not integers). Thus $e^{2\pi i T_3}$ becomes minus the identity matrix. If j is an integer, then T_3 is a diagonal matrix with integer entries, and hence $e^{2\pi i T_3}$ is the identity matrix. Therefore, since minus signs cancel in the tensor product, we conclude that a representation of $SU(2) \times SU(2)$ descends to an ordinary representation if and only if $j_1 + j_2$ is an integer. It is spin otherwise. ∎

In the case of $SO(3)$, we showed in Lemma 5.5.2 that the 2-dimensional spin representation of $SO(3)$, which descends from the fundamental representation of $Spin(3) \cong SU(2)$, is pseudo-real. For $SO(4)$, we have found two lowest-dimensional spin representations, the $(1/2, 0)$ and $(0, 1/2)$, which also have dimension 2. One can ask: are those complex conjugate of each other, or are they both independently real or pseudo-real?

We will not prove this here, but it turns out that for $SO(4)$, the two 2-dimensional spin representations are pseudo-real. In fact, there is a nice classification result about spin representations of $SO(N)$ in general. One needs to distinguish between $SO(2N)$ and $SO(2N+1)$. For $SO(2N)$, there are always two lowest-dimensional spin representations, and they have dimension 2^{N-1}. For $SO(2N+1)$, there is only one lowest-dimensional spin representation, and it has dimension 2^{N-1}. The type of these spin representations (complex, real or pseudo-real) is summarized in the table below. It turns out that the pattern is periodic, with periodicity eight. So, if we write $SO(8k+m)$ for some integers $k \in \mathbb{Z}$ and $0 \leq m \leq 7$, then the type of spin representations depends only on m.

Table 5.7.5 Spin representations for $SO(8k+m)$

m	0	1	2	3	4	5	6	7
Type	Real	Real	Complex	Pseudo-real	Pseudo-real	Pseudo-real	Complex	Real

To end this section, let us check that the table is consistent with what we have found so far. $SO(3)$ corresponds to $k=0$ and $m=3$, which says that the one 2-dimensional spin reprepresentation is pseudo-real, which is what we have found. For $SO(4)$, we have $k=0$ and $m=4$, and hence from the table we conclude that the two 2-dimensional spin representations are pseudo-real, which is indeed what we stated above. Great!

5.7.2 Representation theory of the Lorentz group

This is all very nice, but what does it have to do with the Lorentz group $SO(3,1)$? The Lorentz group is certainly more complicated. For instance, it is non-compact, which implies that it has no non-trivial finite-dimensional unitary representations. Fortunately, it is however still semi-simple, so all representations are equivalent to direct sums of irreducible representations. So we only need to care about irreducible representations. In any case, at the level of Lie algebras, the irreducible representations of $SO(3,1)$are indexed in a way very similar to those of $SO(4)$. Let us see why.

The Lie algebra $\mathfrak{so}(4)$ was presented in (5.7.3). It turns out that the Lie algebra $\mathfrak{so}(3,1)$ of the Lorentz group is very similar. It is also six-dimensional, and can be written in terms of generators J_i, K_i with $i=1,2,3$ as

$$[J_i, J_j] = i\sum_{k=1}^{3} \epsilon_{ijk} J_k, \tag{5.7.4}$$

$$[J_i, K_j] = i\sum_{k=1}^{3} \epsilon_{ijk} K_k, \tag{5.7.5}$$

$$[K_i, K_j] = -i\sum_{k=1}^{3} \epsilon_{ijk} J_k. \tag{5.7.6}$$

The only difference with (5.7.3) is the sign in the bracket $[K_i, K_j]$. This sign is however quite important.

To determine the representations of $SO(3,1)$, we want to construct its double cover, as we did for $SO(4)$. Let us first do it at the level of Lie algebras. As in (5.7.3), we want to find a change of basis to construct an isomorphism of Lie algebras.

The key is to realize that if we do a change of basis $K_j \mapsto iK_j$, the brackets in (5.7.6) become exactly equal to those in (5.7.3). But we are only allowed to do such a change of basis if we consider the "complexified Lie algebra", that is, we consider the complex vector space with basis given by the J_i, K_i. We denote this complex Lie algebra as $\mathfrak{so}(3,1)_{\mathbb{C}}$. Thus what we have found is that the complexified Lie algebra $\mathfrak{so}(3,1)_{\mathbb{C}}$ and $\mathfrak{so}(4)_{\mathbb{C}}$ are isomorphic. In particular, doing the change of basis as for (5.7.3), we conclude that

$$\mathfrak{so}(3,1)_{\mathbb{C}} \cong \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}}.$$

But in the end we want to work with the real Lie algebra $\mathfrak{so}(3,1)$. So we want to turn the right-hand-side into the complexification of a single Lie algebra (instead of direct sum of complex Lie algebras).

To do so we proceed through a sequence of isomorphisms of Lie algebra. The key is the isomorphism $\mathfrak{su}(2)_{\mathbb{C}} \cong \mathfrak{sl}_2(\mathbb{C})$, which can be checked from the commutation relations of those. Then we get:

$$\begin{aligned} \mathfrak{so}(3,1)_{\mathbb{C}} \cong& \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}} && (5.7.7)\\ \cong& \mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C}) && (5.7.8)\\ \cong& \mathfrak{sl}_2(\mathbb{C}) \oplus i\mathfrak{sl}_2(\mathbb{C}) && (5.7.9)\\ \cong& \mathfrak{sl}_2(\mathbb{C})_{\mathbb{C}}, && (5.7.10) \end{aligned}$$

where in the last line we mean the complexification of the real Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ associated to the Lie group $SL(2,\mathbb{C})$.

As a result, we get $\mathfrak{so}(3,1)_{\mathbb{C}} \cong \mathfrak{sl}_2(\mathbb{C})_{\mathbb{C}}$, which is an isomorphism between the complexifications of two Lie algebras. Upon restriction to the real Lie algebras this becomes the isomorphism

$$\mathfrak{so}(3,1) \cong \mathfrak{sl}_2(\mathbb{C}).$$

Thus the Lie groups $SO(3,1)$ and $SL(2,\mathbb{C})$ share the same Lie algebra.

It turns out that $SL(2,\mathbb{C})$ is simply connected, and is in fact the universal cover of $SO(3,1)$. Moreover, the map $SL(2,\mathbb{C}) \to SO(3,1)$ is a double covering, and hence $Spin(3,1) \cong SL(2,\mathbb{C})$. We have found the spin group $Spin(3,1)$! Thus all finite-dimensional irreducible projective representations of $SO(3,1)$ descend from irreducible finite-dimensional representations of $SL(2,\mathbb{C})$.

What are the irreducible representations of $SL(2,\mathbb{C})$? Well, it is easier to work at the level of Lie algebras. I will be rather sketchy here for brevity. The isomorphisms (5.7.10) are very useful. First, it turns out that the complex representations of the complexification $\mathfrak{sl}_2(\mathbb{C})_{\mathbb{C}}$ are in one-to-one correspondence with the real representations of $\mathfrak{sl}_2(\mathbb{C})$. Thus, the latter are in one-to-one correspondence with the complex representations of $\mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}}$. But those are equivalent to the representations we constructed earlier when we studied $SO(4)$. They are indexed by two half-integers j_1 and j_2 and have dimensions $(2j_1+1)(2j_2+1)$. Therefore, the same is true of the real irreducible representations of $SL(2,\mathbb{C})$.

It turns out that as for $SO(4)$, the representations with $j_1 + j_2 \in \mathbb{Z}$ descend to ordinary irreducible representations of $SO(3,1)$, while the other ones are spin. The $(1/2,0)$ and $(0,1/2)$ are the two 2-dimensional spin representations; the $(1/2,1/2)$ is the four-dimensional representation (which corresponds to how a four-vector in spacetime transforms).

As for $SO(4)$, we can ask whether the spin representations are real, pseudo-real, or complex. It turns out that for $SO(3,1)$, the two 2-dimensional spin representations are in fact complex conjugate of each other. This is one key difference with $SO(4)$, in which case they were pseudo-real.

In fact, the table Table 5.7.5 can be generalized for arbitrary special orthogonal groups $SO(p,q)$. The type of spin representations is again periodic with periodic eight, but it now depends on $p-q$ mod 8. We get the following generalized table, which reduces to Table 5.7.5 for $q=0$.

Table 5.7.6 Spin representations for $SO(p,q)$

$p-q$ mod 8	0	1	2	3	4	5	6	7
Type	Real	Real	Complex	Pseudo-real	Pseudo-real	Pseudo-real	Complex	Real

5.8 Unitary representations of the Poincare group

Objectives

You should be able to:

- Sketch how unitary irreducible representations of the Poincare group are constructed using little groups.
- State Wigner's classification of non-negative energy unitary irreducible representations of the Poincare group.

In the previous section we studied finite-dimensional representations of the Lorentz group. As we saw, because the Lorentz group is non-compact, all those representations are not unitary, since there is not non-trivial unitary finite-dimensional representation. But in quantum mechanics and in particle physics in general, we are interested in unitary representations. The reason is that the representations act on the Hilbert space of states, i.e. they tell us how wave-functions transform. Since the norm square of the wavefunction gives the probability that the system is observed in a given state, we want the representations to preserve the norm square, or, in other words, to preserve the inner product. That is, we want unitary representations.

Moreover, the symmetry group of spacetime is larger than the Lorentz group: it also includes translations. This is the so-called Poincare group (also called "inhomogeneous Lorentz group" in older references). Putting this together, in particle physics we can in fact identify particles by how they transform under the group of symmetries. So we think of particles as irreducible unitary representations of the Poincare group. More precisely, since wavefunctions are only defined up to scale, we allow irreducible unitary projective representations. This is what we study in this section.

5.8.1 The Poincare group

Let us start by recalling the main features of the Poincare group. The Poincare group acts on a four-dimensional spacetime with coordinates x^μ as

$$x^\mu \mapsto (x')^\mu = \sum_{\nu=1}^{4} L^\mu_\nu x^\nu + a^\mu,$$

where L^μ_ν are the components of a Lorentz transformation. Thus the Poincare group includes both Lorentz transformations and translations.

We can extract the Lie algebra of the Poincare group as usual. It is 10-dimensional. A convenient choice of basis uses the generators $J_{\mu\nu} = -J_{\nu\mu}$ for the Lorentz transformations, and four generators P_μ for the translations. The commutation relations can be written as

$$\begin{aligned}
[P_\mu, P_\nu] =& 0,\\
[J_{\mu\nu}, P_\rho] =& -i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu),\\
[J_{\mu\nu}, J_{\rho\sigma}] =& -i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}),
\end{aligned}$$

where $\eta_{\mu\nu}$ are the components of the Minkowski metric (i.e. the diagonal matrix diag$(-1, 1, 1, 1)$).

Remark 5.8.1 The generators $J_{\mu\nu}$ for Lorentz transformations are connected to the generators of rotation J_i and boosts K_i introduced in (5.7.6) by

$$J_i = \frac{1}{2} \sum_{m,n=1}^{3} \epsilon_{imn} J^{mn}, \qquad K_i = J_{i0},$$

where we raised indices using the Minkowksi metric as usual.

5.8.2 Wigner's classification

The goal of this section is to classify unitary representations of the Poincare group. First, since the group is non-compact, we know that the only finite-dimensional unitary representation is the trivial representation. So all non-trivial unitary representations are infinite-dimensional.

The classification of unitary representations of the Poincare group is known as **Wigner's classification**, since it was originally achieved in Wigner's famous paper "On unitary representations of the inhomogeneous Lorentz group". This is a very nice paper in fact, have a look!

More precisely, Wigner was interested in representations that are physical, and may correspond to physical particles. Thus, what he classified are irreducible unitary representations that have non-negative, real, mass ($m \geq 0$) - negative mass representations are not expected to be physical. We will see in a second how mass arises from the point of view of representations of the Poincare group.

The method that he used is the method of induced representations, also known as the method of **little groups** in physics. In this approach, one constructs the representations of a group starting from the representations of a subgroup. We will not explain how the method of induced representations works in these notes, but will simply go through the main steps in the context of the Poincare group. We will be very hand-wavy, but I suppose that's ok. For more information one could for instance refer to these lecture notes.

The starting point is to consider the subgroup of the Poincare group consisting of translations. From the point of view of Lie algebras, we consider the subspace generated by the translation generators P_μ. Those generators commute, and hence can be simultaneously diagonalized. We pick an eigenvector $|p\rangle$, with (real) eigenvalues p_μ:

$$P_\mu |p\rangle = p_\mu |p\rangle.$$

Remark 5.8.2 We note in passing that the Poincare algebra has two Casimir invariants. One of them is $P^2 = \sum_{\mu=1}^{4} P_\mu P^\mu$, and the other is constructed as $W^2 = \sum_{\mu=1}^{4} W_\mu W^\mu$ with

$$W_\sigma = -\frac{1}{2} \sum_{\mu,\nu,\sigma=1}^{4} \epsilon_{\mu\nu\rho\sigma} J^{\mu\nu} P^\sigma.$$

W_σ is known as the **Pauli-Lubanski pseudovector**. In any case, since those are Casimir invariants, we can label the irreducible representations by the eigenvalues of these operators. In particular, the eigenvalue of P^2 on the state $|p\rangle$ is $-m^2 := \sum_{\mu=1}^{4} p_\mu p^\mu$. What this means is that the "mass squared" m^2 is one of the label for representations of the Poincare group. Thus mass appears naturally in this context; it is one of the label that specifies the representations of the Poincare group.

In any case, going back to the classification, we fix an eigenvector $|p\rangle$, with eigenvalues p_μ. The idea of the method of little groups is to now consider the subgroup of the Poincare group that leaves p_μ invariant: this is called the **little group** (in mathematics, it is called the "stabilizer subgroup"). The idea of the method is to induce unitary irreducible representations of the whole group from the unitary irreducible representations of the little group. In other words, we classify the states with a fixed momentum p_μ in terms of the unitary irreducible representations of the little group.

Concretely, we need to separate the classification into three cases, depending on the choice of starting eigenvector $|p\rangle$:

1. Positive mass $m > 0$, in which case we choose $|p\rangle$ to have eigenvalues $p_\mu = (m, 0, 0, 0)$ (that is, we use Lorentz transformations to bring ourselves to the rest frame of the particle);

2. Zero mass $m = 0$ but with non-zero p_μ, in which case we choose $|p\rangle$ to have eigenvalues $p_\mu = (p, 0, 0, p)$, $p > 0$;

3. Zero mass $m = 0$ and zero eigenvalues $p_\mu = (0, 0, 0, 0)$.

The next step is to determine the little group for each of these cases, and classify the irreducible representations of the little groups.

5.8.2.1 Zero mass and zero momentum

Let us start with the case $m = 0$ and $p_\mu = (0, 0, 0, 0)$. In this case, the little group that leaves p_μ fixed is the whole Lorentz group. But we know that its only finite-dimensional unitary representation is the trivial representation. Thus what this case corresponds to is the trivial representation for the Poincare group. We call a state that transforms according to that representation the **vacuum**, since it is invariant under all symmetries of the Poincare group.

5.8.2.2 Positive mass

Next, let us consider the case where $m > 0$ and we are in the rest frame of the particle, that is, $p_\mu = (m, 0, 0, 0)$. In this case, the little group is the subgroup of the Poincare group that leaves $(m, 0, 0, 0)$ invariant. It turns out that this is given by the group of rotations in three-dimensional space, that is, $SO(3)$.

Nice! Because we already know the irreducible unitary projective representations of $SO(3)$. We know that there is an infinite family of them, indexed by a non-negative half-integer j called the spin, and with dimensions $2j + 1$. We know that when $j \in \mathbb{Z}$ we get an ordinary representation, while when $j \notin \mathbb{Z}$ we get a spin representation.

Thus, massive particles are classified by an irreducible representation of $SO(3)$, identified by its spin j, which is half-integer.

5.8.2.3 Zero mass with non-zero p_μ

Let us now do the case of massless particles, but with non-zero p_μ. We choose a frame such that $p_\mu = (p, 0, 0, p)$.

The little group here that fixes $(p, 0, 0, p)$ is not so obvious to see. But it turns out to be given by the special Euclidean group $SE(2)$, which consists of rotations and translations in two dimensions (with Euclidean signature). Thus, to complete Wigner's classification, we need to understand the unitary representations of $SE(2)$.

The Lie algebra $\mathfrak{se}(2)$ associated to the special Euclidean group $SE(2)$ is three-dimensional. One generator J is the generator of infinitesimal rotations, with the two other generators P_1 and P_2 being generators of infinitesimal translations. The commutation relations are:

$$[J, P_1] = iP_2, \qquad [J, P_2] = -iP_1, \qquad [P_1, P_2] = 0.$$

We are interested in irreducible unitary representations of $SE(2)$.

Since $SE(2)$ is non-compact, all its non-trivial unitary representations are infinite-dimensional. But they come in two types: so-called "finite spin representations" and "continuous spin representations".

The easiest way to construct representations of $SE(2)$ is to in fact use the method of induced representations again. In this case, we consider the subspace of the Lie algebra generated by translations, and pick an eigenstate $|k\rangle$ with eigenvalues k_i under translations:

$$P_i|k\rangle = k_i|k\rangle, \qquad i = 1, 2.$$

We then consider the two cases with either $k_i = 0$ or $k_i \neq 0$.

In the $k_i = 0$ case, the little group is the group of rotations $SO(2)$. Thus irreducible unitary representations of $SE(2)$ in this case are induced from the unitary irreducible representations of $SO(2)$, which are all one-dimensional, and indexed by an integer h which we call the **helicity**. In fact, since we also allow spin representations (more precisely, the little group should have been the double cover of $SE(2)$), we allow the helicity h to be a half-integer.

In the case where $k_i \neq 0$, we choose a frame such that $k_i = (k, 0)$. The little group is $SO(1)$, which is trivial. In fact, this case gives rise to so-called "continuous spin" (or "infinite spin") representations. There are two of them here. We will not comment further on this, since those do not appear to be physical.

5.8.2.4 The end result of Wigner's classification

The end result of the classification is rather remarkable. What it says is that unitary irreducible representations of the Poincare group, which we call "particles" in physics, are indexed by two parameters: a continuous parameter m, that we call the "mass of the particle", and a discrete parameter $j \in \frac{1}{2}\mathbb{Z}$ or $h \in \frac{1}{2}\mathbb{Z}$, which we call the "spin" or "helicity" of a particle. More precisely, focusing on the case of non-negative mass, particles come in two kinds (we omit the vacuum case here and the continuous spin representations):

- Massive particles, which are indexed by a positive real number $m > 0$ (the mass), and a half-integer j (the spin).
- Massless particles, which are indexed by a zero mass ($m = 0$), and a half-integer h (the helicity).

Isn't that great? This is why it makes sense to talk about the mass and spin (or helicity) of particles: that's because those index the unitary irreducible representations of the Poincare group! Moreover, the distinction between massive and massless particles which is crucial in physics appear naturally here in terms of representations. Beautiful!

5.9 Classification of simple Lie algebras

Objectives

You should be able to:

- Sketch the main ideas behind the classification of simple Lie algebras.
- State the result of the classification of simple Lie algebras.

In this section we will sketch the main ideas behind one of the most fundamental result in Lie theory, the classification of simple Lie algebras.

5.9.1 Simple Lie algebras

Recall the definition of simple groups (see Definition 1.9.5): a group is simple if it has no non-trivial (or proper) normal subgroups. It turns out that there is a similar concept of a simple Lie algebra, which we now define.

Definition 5.9.1 Lie subalgebra. Let $\mathfrak{g}$ be a Lie algebra. A **Lie subalgebra** is a subspace $\mathfrak{h} \subset \mathfrak{g}$ that is closed under the bracket. That is, for all $X, Y \in \mathfrak{h}$, $[X, Y] \in \mathfrak{h}$. ♢

Definition 5.9.2 Invariant Lie subalgebra. A Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is **invariant** if for all $X \in \mathfrak{h}$ and all $Y \in \mathfrak{g}$, $[X, Y] \in \mathfrak{h}$. In mathematical terminology, this is also called an **ideal** in the Lie algebra $\mathfrak{g}$. ♢

The notion of invariant subalgebras is closely connected to the notion of normal subgroups. Indeed, the definition is pretty much the same, with the operation of conjugation for groups replaced by the bracket operation for algebras. Concretely, one can show that invariant Lie subalgebras give rise to normal subgroups of Lie groups by exponentiation.

Definition 5.9.3 Simple Lie algebra. A Lie algebra is **simple** if it has no non-trivial invariant Lie subalgebras. It is **semisimple** if it is a direct sum of simple Lie algebras. ♢

This is the algebra counterpart to the notion of simple groups. It turns out that simple Lie algebras are very nice, and in fact can be fully classified. This is what we now turn to.

5.9.2 Construction of the classification

To classify simple Lie algebras one needs to introduce a number of concepts that we have not discussed yet. We will be very brief here and only sketch the construction. The starting point is a generalization of the highest weight construction of irreducible representations of $\mathfrak{su}(2)$ to general Lie algebras. It turns out that the adjoint representation plays a special role in this construction, and in fact knowing the adjoint representation is sufficient to recover the whole Lie algebra (as expected, since the adjoint is determined by the structure constants). In the end, we use an explicit description of the adjoint representation in terms of weights to classify all simple Lie algebras.

Let us start by introducing a few concepts. Recall that the starting point of the highest weight construction for $\mathfrak{su}(2)$ was to diagonalize one of the generators of the Lie algebra. In general, we want to simultaneously diagonalize as many generators as possible. For this to be possible, these generators must commute. We are led to the following definition:

Definition 5.9.4 Cartan subalgebra. Let $\mathfrak{g}$ be a Lie algebra. The **Cartan subalgebra** $\mathfrak{h} \subset \mathfrak{g}$ is the Abelian subalgebra spanned by the largest set of commuting Hermitian generators H_i, $i = 1, \dots, m$. We call m the **rank** of the Lie algebra $\mathfrak{g}$. ♢

Note that by definition, the generators H_i of the Cartan subalgebra satisfy

$$H_i = H_i^\dagger, \qquad [H_i, H_j] = 0.$$

Remark 5.9.5 One should not confuse the dimension of a Lie algebra $\mathfrak{g}$ with its rank. The dimension of $\mathfrak{g}$ is the dimension of the vector space, while the rank of $\mathfrak{g}$ is the dimension of its Cartan subalgebra. For instance, $\mathfrak{su}(2)$ has dimension 3, while its rank is 1, since it has only one Cartan generator (see Section 5.3).

Example 5.9.6 Cartan subalgebra of $\mathfrak{su}(3)$. Let us first extract the Lie algebra $\mathfrak{su}(3)$. In general, it is straightforward to show that the Lie algebra $\mathfrak{su}(N)$ is the algebra of $N \times N$ traceless Hermitian matrices, with bracket given by the standard matrix commutator (more precisely, this defines the fundamental representation of $\mathfrak{su}(N)$).

In the case of $\mathfrak{su}(3)$, the dimension of the space of traceless 3×3 Hermitian matrices is $3^2 - 1 = 8$. We can write down an explicit basis for 3×3 traceless Hermitian matrices. This is given by the so-called **Gell-Mann matrices**:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{5.9.1}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \tag{5.9.2}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{5.9.3}$$

It is conventional to define the generators of $\mathfrak{su}(3)$ in terms of the Gell-Mann matrices as $L_i = \frac{1}{2}\lambda_i$. The commutation relations then give

$$[L_i, L_j] = i \sum_{k=1}^{8} c_{ijk} L_k,$$

with the non-vanishing structure constants given by (with other non-vanishing ones related by symmetry):

$$\begin{aligned} c_{123} =& 1, \\ c_{147} =& -c_{156} = c_{246} = c_{257} = c_{345} = -c_{367} = \frac{1}{2}, \\ c_{458} =& c_{678} = \frac{\sqrt{3}}{2}. \end{aligned}$$

Looking at the Gell-Mann matrices, we see that there are two diagonal generators with real eigenvalues: L_3 and L_8. We conclude that these generate the Cartan subalgebra of $\mathfrak{su}(3)$. Thus, while the dimension of $\mathfrak{su}(3)$ is 8, its rank is 2. In general, the dimension of $\mathfrak{su}(N)$ is $N^2 - 1$ while its rank is $N - 1$. □

We want to continue mimicking the highest weight construction for $\mathfrak{su}(2)$. We construct representations of $\mathfrak{g}$ by constructing the vector space on which they act. We take a basis for that vector space consisting of eigenvectors of the simultaneously diagonalizable Cartan generators. To each such basis vector $|\alpha\rangle$ is associated a vector $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ of eigenvalues of this basis vector under the action of the Cartan generators $H_1, \dots, H_m$. That is,

$$H_i|\alpha\rangle = \alpha_i|\alpha\rangle.$$

Definition 5.9.7 Weight vector. The **weight vectors** of a representation Γ of a Lie algebra $\mathfrak{g}$ are the vectors of eigenvalues of the eigenvectors of the Cartan generators. These vectors live in $\mathbb{R}^m$, where m is the rank of $\mathfrak{g}$. We call $\mathbb{R}^m$ the **weight space** of $\mathfrak{g}$. ♢

Remark 5.9.8 Note that the weight space is the same for all representations of a Lie algebra $\mathfrak{g}$, but the weight vectors change. In fact, one can think of a representation as being determined by a set of weights in weight space.

This is indeed the natural generalization of the highest weight construction for $\mathfrak{su}(2)$. In this case, there was only one Cartan generator, so the weight space was one-dimensional, i.e. it was just $\mathbb{R}$. The weight vectors of a representation were given by the half-integers $-j, -j+1, \dots, j-1, j \in \mathbb{R}$, for some non-negative half-integer j.

Example 5.9.9 The weights of the fundamental representation of $\mathfrak{su}(3)$. Going back to $\mathfrak{su}(3)$, with generators given by the Gell-Mann matrices (see (5.9.3), we can extract the weights of the representation by calculating the eigenvalues of the Cartan generators for the basis vectors $e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$, $e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ on which the representation acts. We get

$$\begin{aligned}
T_3 e_1 =& \frac{1}{2} e_1, \qquad T_8 e_1 = \frac{1}{2\sqrt{3}} e_1, \\
T_3 e_2 =& -\frac{1}{2} e_2, \qquad T_8 e_2 = \frac{1}{2\sqrt{3}} e_2, \\
T_3 e_3 =& 0, \qquad T_8 e_3 = -\frac{1}{\sqrt{3}}.
\end{aligned}$$

So the weight vectors of the fundamental representation are $(\frac{1}{2}, \frac{1}{2\sqrt{3}})$, $(-\frac{1}{2}, \frac{1}{2\sqrt{3}})$ and $(0, -\frac{1}{\sqrt{3}})$, which live in the weight space $\mathbb{R}^2$. □

It turns out that the adjoint representation plays a very special role.

Definition 5.9.10 Root. The **roots** of a Lie algebra $\mathfrak{g}$ are the non-zero weights of its adjoint representation. ♢

The roots play an important role in the highest weight construction of representations. First, one can show that if α is a root, then $-\alpha$ is also a root. So roots come in pairs $\pm\alpha$. It turns out that the generators of the Lie algebra that are not generators of the Cartan subalgebra can also be grouped in pairs $E_{\pm\alpha}$, which play the role of raising and lowering operators. Indeed, one can show that if $|\mu\rangle$ is an eigenvector of the Cartan generators H_i, $i = 1, \dots, m$ with eigenvalue $\mu = (\mu_1, \dots, \mu_m)$, then

$$H_i E_{\pm\alpha}|mu\rangle = (\mu_i \pm \alpha_i) E_{\pm\alpha}|\mu\rangle.$$

In other words, $E_{\pm\alpha}|\mu\rangle$ is also an eigenvector of the Cartan generators, but with eigenvalues raised or lowered by the value of the corresponding root α.

In this way, we can define the notion of highest weight vector, and construct representations using the raising and lowering operators associated to the roots of the Lie algebra.

In fact we can also use the result above to determine the roots. If we can rearrange the remaining generators of the Lie algebra into pairs of raising and lowering operators $E_{\pm\alpha}$, when we can determine the roots by evaluating $[H_i, E_{\pm\alpha}] = \pm\alpha_i E_{\pm\alpha}$. The α_i give the components of the root vector associated to this pair of lowering and raising operators. Note that to calculate these commutators, we can use whatever representation we want. So in this way, if we can identity the pairs of raising and lowering operators, we can obtain the roots without ever writing down the adjoint representation, for instance by calculating commutators using the simpler fundamental representation.

Example 5.9.11 Roots of $\mathfrak{su}(3)$. One way to find the roots of $\mathfrak{su}(3)$ is to write down the 8×8 matrices corresponding to the Cartan generators T_3 and T_8 in the adjoint representation, and extract their eigenvalues. Or, we can use our knowledge of the Gell-Mann matrices (5.9.3) and deduce the three pairs of raising and lowering operators. Indeed, (T_1, T_2), (T_4, T_5) and (T_6, T_7) form pair of Pauli matrices embedded in 3×3 matrices. So the pairs of raising and lowering operators for $\mathfrak{su}(3)$ should be:

$$\begin{aligned} E_{\pm\alpha^{(1)}} &= \frac{1}{\sqrt{2}}(T_1 \pm iT_2), \\ E_{\pm\alpha^{(2)}} &= \frac{1}{\sqrt{2}}(T_4 \pm iT_5), \\ E_{\pm\alpha^{(3)}} &= \frac{1}{\sqrt{2}}(T_6 \pm iT_7). \end{aligned}$$

Then we can calculate the commutators to determine the roots $\alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}$.

First, for $\alpha^{(1)}$, the commutator with T_3 is just the same as for $\mathfrak{su}(2)$, so we get:

$$[T_3, E_{\pm\alpha^{(1)}}] = \pm E_{\pm\alpha^{(1)}},$$

which means that $\alpha_1^{(1)} = 1$. Furthermore, T_8 commutes with T_1 and T_2, and hence

$$[T_8, E_{\pm\alpha^{(1)}}] = 0,$$

that is $\alpha_2^{(1)} = 0$. Thus the first two roots are $\pm(1, 0)$.

For $\alpha^{(2)}$ and $\alpha^{(3)}$ we can calculate the commutator explicitly using the fundamental representation (Gell-Mann matrices). We get:

$$\begin{aligned} [T_3, E_{\pm\alpha^{(2)}}] &= \pm\frac{1}{2}E_{\pm\alpha^{(2)}}, & [T_8, E_{\pm\alpha^{(2)}}] &= \pm\frac{\sqrt{3}}{2}E_{\pm\alpha^{(2)}}, \\ [T_3, E_{\pm\alpha^{(3)}}] &= \pm\frac{1}{2}E_{\pm\alpha^{(3)}}, & [T_8, E_{\pm\alpha^{(3)}}] &= \mp\frac{\sqrt{3}}{2}E_{\pm\alpha^{(3)}}. \end{aligned}$$

Thus the remaining four roots are $\pm(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $\pm(\frac{1}{2}, -\frac{\sqrt{3}}{2})$. The resulting root system is shown in Figure 5.9.12.

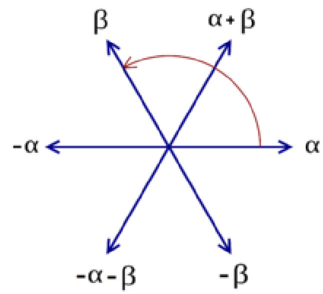

Figure 5.9.12 The root system for $\mathfrak{su}(3)$ (taken from wikipedia).

□

But our goal here is not to construct representations, but to classify simple Lie algebras. It turns out that the roots are key for this as well. Indeed, the roots tell us not only something about the adjoint representation, but in fact they can be used to reconstruct the whole algebra itself. This is perhaps expected, since the adjoint representation is defined in terms of the structure constants, and hence should package the information content of the Lie algebra itself.

To see how this goes we need a few more definitions.

Definition 5.9.13 Positive and simple root. A **positive weight** is a weight of a representation such that the first non-zero component is positive. A **positive root** is a positive weight for the adjoint representation. A **simple root** is a positive root that cannot be written as a linear combination of other positive roots with all coefficients being positive. ◊

One can show that knowing the simple roots is in fact sufficient to determine all roots of a Lie algebra, because of the symmetries of a root system. The simple roots in fact form a basis for $\mathbb{R}^m$, so, in particular, there are m simple roots, where m is the rank of the Lie algebra.

Example 5.9.14 Simple roots for $\mathfrak{su}(3)$. In Example 5.9.11 we determined that the roots of $\mathfrak{su}(3)$ are $\pm(1,0)$, $\pm(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $\pm(\frac{1}{2}, -\frac{\sqrt{3}}{2})$. The positive roots are then $(1,0)$, $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$. Since

$$(1,0) = (\frac{1}{2}, \frac{\sqrt{3}}{2}) + (\frac{1}{2}, -\frac{\sqrt{3}}{2}),$$

we see that $(1,0)$ is not a simple root. The other two positive roots, $(\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$, are the simple roots of $\mathfrak{su}(3)$. □

5.9.3 Root systems

The discussion of the preceding section shows that to a Lie algebra $\mathfrak{g}$, we can associate a bunch of roots living in weight space $\mathbb{R}^m$ by looking at the eigenvalues of the adjoint representation. These roots form what is called a **root system**. It turns out that root systems are very constrained. Random bunch of vectors in $\mathbb{R}^m$ certainly do not form root systems for some Lie algebra. It turns out that root systems can be defined abstractly, independently of Lie algebras, which we now do. In fact, we note here that root systems appear in many different contexts in mathematics, not just in the context of Lie theory.

Definition 5.9.15 Root system. A **root system** Φ in $\mathbb{R}^m$ is a finite number of vectors (roots) in $\mathbb{R}^m$ such that:

- The roots span $\mathbb{R}^m$.
- For any root $\alpha \in \Phi$, the only scalar multiple of α that is also in Φ is $-\alpha$.
- For every $\alpha \in \Phi$, Φ is closed under reflections through the hyperplane perpendicular to α. This can be rephrased as the condition that for any two $\alpha, \beta \in \Phi$, we must have that

$$\sigma_\alpha(\beta) = \beta - 2\frac{\alpha \cdot \beta}{\alpha \cdot \alpha} \in \Phi.$$

- For any two $\alpha, \beta \in \Phi$,

$$2\frac{\alpha \cdot \beta}{\alpha \cdot \alpha} \in \mathbb{Z}.$$

♢

Note that one can construct root systems by "putting together" smaller root systems. Two roots systems can be combined into a bigger one by considering the spaces $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$ as mutually orthogonal subspaces of a bigger space $\mathbb{R}^{m_1+m_2}$. Then the two roots systems form a bigger root system in $\mathbb{R}^{m_1+m_2}$. We say that root systems that cannot be obtained in this way are **irreducible**. In other words, a root system Φ is irreducible if it cannot be partitioned into two roots systems $\Phi = \Phi_1 \cup \Phi_2$ such that $\alpha \cdot \beta = 0$ for all $\alpha \in \Phi_1$ and $\beta \in \Phi_2$.

The crucial result in our context is that *there is a one-to-one correspondence between simple Lie algebras and irreducible roots systems (up to isomorphisms)*. This is absolutely key. We have already seen how, given a Lie algebra $\mathfrak{g}$, we can calculate its roots, which are the weights of its adjoint representation. Then one can show that for any simple Lie algebra, these roots form an irreducible root system, according to the definition above. So we know that we can assign a unique irreducible root system to a simple Lie algebra.

What needs to be shown is the reverse direction. First, we must show that starting with an irreducible root system, we can assign at most one associated simple Lie algebra. In other words, two non-isomorphic simple Lie algebras cannot share the same irreducible root system. This is not so trivial, but intuitively, it is expected, since two non-isomorphic simple Lie algebras cannot have the same structure constants, and hence cannot have the same roots. Second, we must show that to any irreducible root system one can assign at least one simple Lie algebra. In other words, that any irreducible root system arises as the root system of a simple Lie algebra. As we will see, this is clear for the four infinite families of irreducible root systems, since those arise as root systems of basic matrix Lie groups. But it is not so obvious to prove for the five exceptional root systems.

5.9.4 Classification of simple Lie algebras

The fact that simple Lie algebras are in one-to-one correspondence with irreducible root systems (up to isomorphisms) imply that to classify simple Lie algebras, we only need to classify irreducible root systems, which is a nice combinatorial problem. Root systems are in fact nicely encapsulated into so-called **Dynkin diagrams**.

We construct the Dynkin diagram associated to a root system Φ as follows. We assign a node to each of the simple roots. Then we draw edges between nodes according to the angle between the roots. The definition of a root system implies that the only possible angles between root vectors are $\pi/2$, $2\pi/3$, $3\pi/4$ and $5\pi/6$. We draw and edge between nodes as follows:

- No edge if the root vectors are perpendicular ($\pi/2$);
- A single edge if the angle between the root vectors is $2\pi/3$;
- A double edge if the angle between the root vectors is $3\pi/4$;
- A triple edge if the angle between the root vectors is $5\pi/6$.

Moreover, in the last two cases, one can show that the two root vectors cannot have equal length. Thus we draw an arrow pointing towards the shorter vector.

Example 5.9.16 Dynkin diagram for $\mathfrak{su}(3)$. Going back to $\mathfrak{su}(3)$, we found that the two simple roots are given by $(\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$. So the Dynkin diagram will have two nodes. The angle between the two simple roots is $2\pi/3$, and hence the two nodes are connected by a single edge. The Dynkin diagram of $\mathfrak{su}(3)$ is shown in Figure 5.9.17.

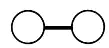

Figure 5.9.17 The Dynkin diagram for $\mathfrak{su}(3)$.

□

Using Dynkin diagrams, one can classify all possible irreducible root systems, and hence all simple Lie algebra (up to isomorphisms). The classification is shown in Figure 5.9.18. It turns out that there are four infinite families, usually denoted by A_n, B_n, C_n and D_n, and five exceptional cases, denoted by G_2, F_4, E_6, E_7 and E_8.

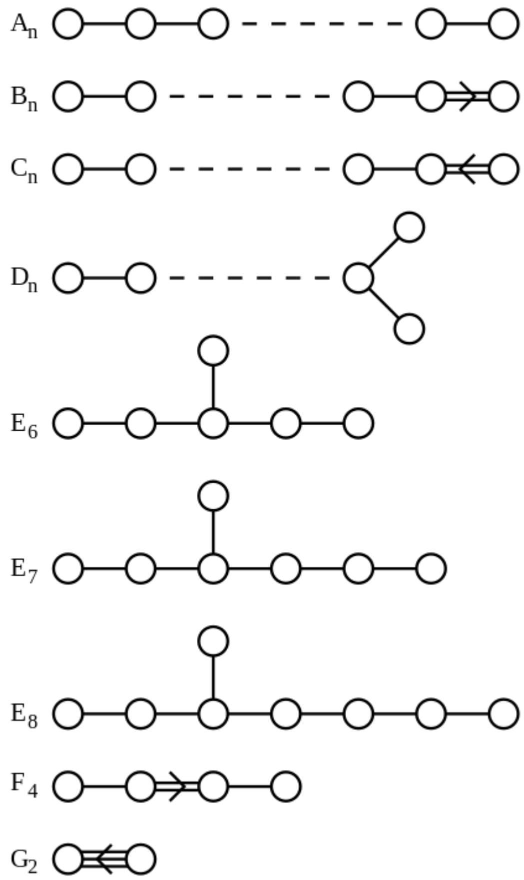

Figure 5.9.18 Classification of irreducible root systems in terms of Dynkin diagrams (taken from wikipedia).

The infinite families correspond to the root systems of well-known simple Lie algebras. We have that $A_n \cong \mathfrak{su}(n+1)$, $B_n \cong \mathfrak{so}(2n+1)$, $C_n \cong \mathfrak{sp}(2n)$, and $D_n \cong \mathfrak{so}(2n)$. The only family that we have not studied in this class is $C_n \cong \mathfrak{sp}(2n)$, which corresponds to the Lie algebra associated to $2n \times 2n$ symplectic matrices.

The five exceptional root systems also correspond to simple Lie algebras, but those are not realized in terms of matrices with simple properties (such as orthogonal, unitary, symplectic, etc.) But these exceptional Lie algebras are common in physics: for instance, in one flavour of string theory (so-called "heterotic string theory"), the starting gauge group of the theory is $E_8 \times E_8$, where E_8 is the Lie group whose Lie algebra has root system given by the E_8 exceptional case.

Remark 5.9.19 Looking at Figure 5.9.18, we see that the infinite families A_n and D_n, as well as the exceptional cases E_6, E_7, E_8 are special because they only have single edges. We call such diagrams **simply laced**. Those Dynkin diagrams arise often in many different contexts in mathematics. Whenever these diagrams arise, we refer to their classification in mathematics as an **ADE classification**.