Representation Theory [KU block 1A, 2015-16]
Lecture notes

Gorjan Alagic

August 29, 2016

Contents

Introduction

This document contains a (rough and preliminary, work-in-progress) set of lecture notes for the 2015-2016 block 1A course on Representation Theory at the University of Copenhagen, taught by the author. The course is divided into two parts. In Part 1, we discuss the basic theory of representations and characters of finite groups. One could view this part of the course as having two goals: (i.) to understand just enough theory in order to construct some of the most common examples oneself, and (ii.) to gain a basic understanding of the group algebra and the Fourier transform. In Part 2, we will move on to compact groups, once again with two central goals: (i.) to understand the Peter-Weyl theorem and the crucial differences from the finite case, and (ii.) to understand just enough Lie theory to see an example of the "theorem of the highest weight" in action. Our part I will be based on Part I of Serre's well-known text [4]; part II will be assembled from parts I and II of Brian Hall's rather accessible text [1]. In some sections, I will follow the exposition of these texts quite closely. It's important to remark that, throughout the course, we will only consider representations over $\mathbb{C}$.

Given the huge number of easily accessible representation theory textbooks and lecture notes, one might question the need for these notes. I do not claim to have any special insight or method. The aim of making these notes is simply to prepare myself for lecturing; it also helps to have a single source from which I can tell a unified story to the students.

Instead of having a lengthy preliminary section at the beginning, I chose to spread the background material throughout the notes. I will try to introduce (or at least mention) background material as close as possible to when it is actually required. One upside is that we can start thinking about representations early on. A downside is that the notes might not flow as nicely.

Finally, let me describe my basic expectations regarding exercises and exams. First, there are no exams, and your grade will be based entirely on (nearly) weekly homework sets. One could argue at length about the advantages or drawbacks of this approach. My hope is that this structure will result in a better absorption of the material, through practice that is both frequent and significant in quantity. I encourage collaboration, so long as each student writes up their own solutions entirely by themselves. All answers must be accompanied by complete and clearly explained proofs. *To be clear: a correct answer with no explanation will receive a zero score.* Make sure to indicate the points in the proof where you applied a theorem from the lectures. Keep in mind that it is difficult to explain things without using at least some words. As in all writing, these words should almost always come in complete sentences. For open-ended (e.g., yes/no) questions, think carefully about what constitutes a "proof" that your response is correct. For example, in some cases the proof will consist of a single counterexample; in that case, the counterexample should be explicit (i.e., specify the group in question, specify the matrices or operators defining the representation in question, and so on.)

Chapter 1

Finite groups

1.1 Basic definitions

This section covers the following concepts, in the setting of finite groups and finite-dimensional spaces over $\mathbb{C}$. We start with a quick reminder of some needed concepts from finite groups and vector spaces (subsection 1.1.1). We then define the following: representations (subsection 1.1.2), subrepresentations, direct sums, and irreducibility (subsection 1.1.3), and tensor products (subsection 1.1.4). Some very basic examples are sprinkled throughout. The exercises for this Section appear in subsection 1.1.5.

1.1.1 Preliminaries

Here, we briefly recall some very basic notions from group theory and vector spaces. Ideally, students have been exposed to this material previously, and this will be a (very brief) refresher. If you are already comfortable with this material, then you should skip ahead and refer back to this section later if needed—which is likely, since we will now also fix some basic notation and conventions.

Groups. For now, all groups will be finite. Recall that a *finite group* is a finite set G together with an associative operation (a map $G \times G \to G$, written $(x, y) \mapsto xy$), a distinguished identity element 1 (satisfying $1x = x1 = x$ for all x), and inverses (each $x \in G$ has a unique $x^{-1} \in G$ satisfying $xx^{-1} = x^{-1}x = 1$). The *order* $|G|$ of the group G is simply the number of elements. If $xy = yx$ for all $x, y \in G$, we say that G is abelian; otherwise it is called nonabelian. A subset H of a group G which is closed under the group operation is called a *subgroup*. A function $f : G_1 \to G_2$ from a group G_1 to another group G_2 is called a *homomorphism* if it preserves operations, i.e., if $f(x)f(y) = f(xy)$ for all $x, y \in G_1$. A homomorphism which is bijective is called an *isomorphism*; two groups are said to be isomorphic if there exists an isomorphism from one to the other.

Describing groups can be done using so-called *presentations*. A presentation lists a set of *generators*, together with a complete set of *relations*. The group is then formed by taking the set of all possible words in the generators (just like words in an alphabet), and modding out by the equivalence relation defined by the relations. For example, the cyclic group (of order n) is defined by the presentation $\mathbb{Z}/n\mathbb{Z} = \langle x : x^n = 1\rangle$. The words are made just by stringing together some finite number of xs; two words are equivalent if the number of xs in them is equivalent modulo n. It should be clear now that $\mathbb{Z}/n\mathbb{Z}$ (as defined) has order n, and that it is isomorphic to the integers $\{0, 1, \dots, n-1\}$ with the operation of addition modulo n. For a more complicated (and non-abelian) example, consider the group S_n consisting of all permutations of a set of n objects, say the set $[n] = \{1, 2, \ldots, n\}$. It is convenient to write elements of S_n in *cycle notation*; for example, the element that sends 1 to 2, 2 to 3, and 3 back to 1 (while fixing all others) is succinctly written as (123). The element that transposes (swaps) 1 and 3 as well as 5 and 7 would be written (13)(57). It's not hard to see that S_n is generated by the $n-1$ adjacent transpositions $\sigma_j := (j\, j+1)$. Each transposition is its own inverse, and transpositions that affect disjoint pairs commute with each other. As it turns out, adding one more relation is sufficient to fully characterize S_n:

$$S_n = \left\langle \sigma_1, \sigma_2, \ldots, \sigma_{n-1} \;:\; \sigma_i^2 = 1;\, \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1};\, \sigma_i\sigma_j = \sigma_j\sigma_i \text{ if } |j-i| > 1 \right\rangle .$$

Vector spaces. For now, all vector spaces will be finite-dimensional, with scalars the complex numbers $\mathbb{C}$. If V is such a space, then it is isomorphic to $\mathbb{C}^d$ for some finite d, called the *dimension* of V (written $\dim V$). The set V is an abelian group under the operation of addition of vectors (with identity element 0), and is closed under scaling by elements of $\mathbb{C}$, which is distributive with respect to vector addition. A *subspace* of V is a subset of V which is itself a vector space under the same operations as V. A *proper subspace* is one which is equal to neither V nor the trivial subspace (i.e., the set $\{0\}$.)

Given two vector spaces V and W, a map $f : V \to W$ is called a *linear operator* or *linear map* if it satisfies $f(ax + y) = af(x) + f(y)$ for all $a \in \mathbb{C}$ and $x, y \in V$. A bijective linear operator is called a *linear isomorphism*, or just isomorphism when the context is clear. The set of linear maps from V to itself is denoted $\mathrm{End}(V)$; the subset of $\mathrm{End}(V)$ consisting of isomorphisms (or equivalently, the invertible maps) is denoted $\mathrm{GL}(V)$, and is a group under the composition operation.

Although a pure mathematician might be hesitant to do so, we will frequently and happily fix bases for our vector spaces. Recall that a basis for V is a finite set of $\dim V$-many vectors in V which span V (i.e., each $v \in V$ is a linear combination of them) and are linearly independent (i.e., none of them is a linear combination of the others). There are many consequences of choosing a particular basis for V, but the two most crucial ones are: *(i.)* each vector $v \in V$ is associated with a tuple of $\dim V$ complex numbers (namely the coefficients of v in the basis expansion), and *(ii.)* each linear map $L \in \mathrm{End}(V)$ is associated to a $\dim V \times \dim V$ *matrix* (L_{ij}) of complex numbers. Each column $(L_{i\cdot})$ of this matrix is a vector; it is the image of the ith basis element under the linear map L. If $L, K \in \mathrm{End}(V)$, then the matrix of their product LK is simply the matrix product of the matrix of L with the matrix of K. If $L \in \mathrm{GL}(V)$, then the matrix of L is invertible, and its inverse is again an element of $\mathrm{GL}(V)$; a basis choice for V thus makes $\mathrm{GL}(V)$ into a *matrix group*, i.e. a group where the elements are matrices and the operation is matrix product. Of course, this is an infinite group.

If we fix a particular basis B and then wish to change to another basis B', this can be done via a basis change operation. Assembling the vectors of B' (written in the basis B) as columns results in an invertible matrix M. To rewrite a vector v from B to B', simply compute $M^{-1}v$. To rewrite a matrix L from B to B', compute $M^{-1}LM$.

Given two vector spaces V and W, their *direct sum* $V \oplus W$ is another vector space, whose elements are pairs (v, w) (written $v \oplus w$) with $v \in V$ and $w \in W$ and whose operations are coordinatewise, i.e., $v \oplus w + v' \oplus w' = (v + v') \oplus (w + w')$ and $a(v \oplus w) = av \oplus aw$. Dimension is additive with respect to direct sum. Indeed, the direct sum of $\mathbb{C}^n$ and $\mathbb{C}^m$ is isomorphic to $\mathbb{C}^{n+m}$. The direct sum of vector spaces induces a direct sum operation on linear maps. For example, for $A \in \mathrm{End}(V)$ and $B \in \mathrm{End}(W)$, the map $A \oplus B : (v \oplus w) \mapsto (Av \oplus Bw)$ is an element of $\mathrm{End}(V \oplus W)$. A closely associated notion to direct sum is *projection*; a *projection operator* is a map $L \in \mathrm{End}(V)$ which satisfies $L^2 = L$. One can easily show that each such L decomposes V into a direct sum, namely $V = \mathrm{im}\, L \oplus \ker L$.

Another operation on two vector spaces V and W is *tensor product*. The result is again a vector space, denoted $V \otimes W$, whose elements are *tensors* $v \otimes w$, and where the operation is *bilinear*. In other words, the operation is addition modulo the relations $a(v \otimes w) = (av \otimes w) = (v \otimes aw)$ and $(v \otimes w) + (v \otimes w') = (v \otimes (w + w'))$ and $(v \otimes w) + (v' \otimes w) = ((v + v') \otimes w)$. Dimension is multiplicative with respect to tensor product. In particular, $\mathbb{C}^n \otimes \mathbb{C}^m$ is isomorphic to $\mathbb{C}^{nm}$. Just as with direct sum, the tensor product of vector spaces induces a tensor product operation on linear maps. For example, for $A \in \mathrm{End}(V)$ and $B \in \mathrm{End}(W)$, the map $A \otimes B : (v \otimes w) \mapsto (Av \otimes Bw)$ is an element of $\mathrm{End}(V \otimes W)$.

If you lack intuition about direct sums and tensor products, I strongly encourage you to do the following simple exercises. Don't be afraid to work out very simple examples (e.g., in dimensions 1 and 2) to make sure you understand what's happening. These concepts will be central to the entire course, starting from the very first lecture.

- what is the direct sum decomposition of the plane induced by projection to the x-axis?
- pick an explicit basis for $\mathbb{C}^n$ and $\mathbb{C}^m$; compute corresponding bases for $\mathbb{C}^n \oplus \mathbb{C}^m$ and $\mathbb{C}^n \otimes \mathbb{C}^m$.
- continuing with the previous exercise, given basis expansions of vectors $v \in \mathbb{C}^n$ and $w \in \mathbb{C}^n$, how do you compute the basis expansions of $v \oplus w$ and $v \otimes w$?
- continuing with the previous exercise, given an $n \times n$ matrix A and an $m \times m$ matrix B, how do you compute the matrices $A \oplus B$ and $A \otimes B$?

1.1.2 Representations

Definition 1.1.1. *A **representation** of a finite group G is a homomorphism $\rho : G \to \mathrm{GL}(V)$, for some finite-dimensional complex vector space V.*

Written explicitly, the definition means that $\rho(xy) = \rho(x)\rho(y)$ for every x and y in G; we may say that ρ maps the group product (on the left-hand side) to composition of linear operators (on the right-hand side). For the sake of compactness, it is common to say "(ρ, V) is a representation of G" when defining representations. Once the choice of G, V, and ρ has been fixed, it is standard to refer to both the space V and the map ρ as "representations." In keeping with this custom, we define the dimension (or *degree*) of ρ to be $\dim V$.

Fixing a particular basis of V, we can write each operator $\rho(x)$ explicitly as a matrix, with matrix entries $\rho(x)_{ij}$. The representation ρ thus associates to each group element an explicit matrix $(\rho(x)_{ij})$ of complex numbers, in a way that maps the group product into the matrix product. This is sometime called a *matrix representation* of G. For a matrix representation, the product can be written out explicitly in the usual way:

$$\rho(xy)_{ij} = \sum_k \rho(x)_{ik}\rho(y)_{kj} .$$

From now on, we will move the subscripts closer to ρ, writing $\rho_{ij}(x)$ in place of $\rho(x)_{ij}$. This notation makes sense, since the object ρ_{ij} is a sensible thing: it is a map from G to $\mathbb{C}$. We will refer to this map as a *matrix entry* of the representation ρ.

As with all things in mathematics, we are interested in classification. This is only feasible with a natural notion of equivalence.

Definition 1.1.2. *Two representations $\rho : G \to \mathrm{GL}(V)$ and $\sigma : G \to \mathrm{GL}(W)$ are **isomorphic** (written $\rho \cong \sigma$) if there exists a linear isomorphism $T : V \to W$ such that $\sigma(g) = T\rho(g)T^{-1}$.*

Note that the existence of T already implies that V is isomorphic to W. In particular, if we chose bases for V and W and view ρ and σ as matrix representations, then they are isomorphic precisely when their matrices differ only by a change of basis.

Examples.

1. Following Definition 1.1.1, we see that a one-dimensional representation of a finite group G (sometimes called a *character* of G) is a homomorphism from G to $\mathbb{C}^\times$, the "circle group" of unit-length complex numbers under multiplication. For any G, the map that assigns every $g \in G$ to the number 1 is a one-dimensional representation called the *trivial representation.* Another example is the sign representation of $\mathbb{Z}/2\mathbb{Z} = \langle x : x^2 = 1 \rangle$, defined by setting $1 \mapsto 1$ and $x \mapsto -1$.

2. Fix any group G and a vector space V with $\dim V = |G|$. We can then identify some basis $\{e_x : x \in G\}$ of V with the group elements of G. Define $\rho : G \to \mathrm{GL}(V)$ by setting $\rho(y) : e_x \mapsto e_{yx}$. It's easy to check that ρ is a representation; it's called *the (left) regular representation* of G, and we will denote it by Reg_G. Looking again at $\mathbb{Z}/2\mathbb{Z}$, we easily compute the regular representation:
$$1 \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad x \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

3. In general, any action of G on some finite set X also defines a representation ρ of G. The vector space V has a basis $\{e_x\}_{x\in X}$ identified with X, and the representation is defined by the rule $g : e_x \mapsto e_{gx}$. Such representations are called *permutation representations* or *combinatorial representations.* The left regular representation is the special case where G acts on itself by left-multiplication.

1.1.3 Subrepresentations, direct sums, and irreducible representations

The ability to create new representations from existing ones is crucial in the task of characterizing all inequivalent representations. Our first tool for doing this is straightforward, and is based on the vector space direct sum.

Definition 1.1.3. *Let (ρ, V) and (σ, W) be representations of a group G. The representation $\rho \oplus \sigma : G \to \mathrm{GL}(V \oplus W)$ defined by $[\rho \oplus \sigma](g) := \rho(g) \oplus \sigma(g)$ is called the* ***direct sum*** *of ρ and σ.*

It should be clear that, once we fix bases for V and W, the matrices $[\rho\oplus\sigma](g)$ in the resulting basis of $V \oplus W$ are (matrix) direct sums of the matrices $\rho(g)$ and $\sigma(g)$.

It will also be useful to discover "smaller" representations sitting inside larger ones. To that end, let (ρ, V) be a representation of a group G, and suppose that there exists a proper ρ-invariant subspace W of V; that is to say, suppose that $\rho(g)w \in W$ for all $g \in G$ and all $w \in W$. It is straightforward to check that the map $\rho^W : G \to GL(V)$ defined by $\rho^W(g) := \rho(g)|_W$ defines a representation of G; we say that ρ^W is a *subrepresentation* of ρ. We thus make the following definitions.

Definition 1.1.4. *Let (ρ, V) and (σ, W) be representations of G. Then (σ, W) is a* ***subrepresentation*** *of (ρ, V) (sometimes written $\sigma \prec \rho$) if there exists a ρ-invariant subspace W' of V such that the restriction of ρ to W' is isomorphic to σ.*

Definition 1.1.5. *A representation (ρ, V) is* ***irreducible*** *if there are no proper ρ-invariant subspaces of V.*

The fact that the concept of irreducibility is of great value is apparent from the following.

Theorem 1.1.1. *Every representation is a direct sum of irreducible representations.*

Proof. The proof proceeds in two steps. First we show that subrepresentations have a "complement" representation; this fact first appeared in a paper by Maschke in 1989 [2] and is thus sometimes called Maschke's Theorem. Then, we (actually, you) will use this fact recursively in order to break down any given representation into irreducible summands.

Let's prove Maschke's Theorem. Let (ρ, V) be a representation of G and $(\rho|_W, W)$ a subrepresentation. Pick a projection $\Pi_W : V \to W$ and note that $V = W \oplus \ker \Pi_W$. We would like to show that $\ker \Pi_W$ is ρ-invariant, i.e., that $\Pi_W \rho(x) w' = 0$ for all $x \in G$ and all $w' \in \ker \Pi_W$. This would be easy if Π_W and $\rho(x)$ commuted, but in general they do not. We can fix this problem by *symmetrizing* Π_W as follows. Set

$$\Pi_W^\rho = \frac{1}{|G|} \sum_{g \in G} \rho(g) \Pi_W \rho(g)^{-1} .$$

One easily checks (see homework) that Π_W^ρ is a projection operator onto W, and that it commutes with $\rho(x)$ for all $x \in G$. Setting $W' = \ker \Pi_W^\rho$, we have that

$$\ker \Pi_W^\rho \rho(x) w' = \rho(x) \ker \Pi_W^\rho w' = \rho(x) 0 = 0$$

i.e., W' is ρ-invariant. It follows that $\rho = \rho|_W \oplus \rho|'_W$.

It remains to show that, given an arbitrary representation ρ, one can recursively apply the above until ρ is expressed as a direct sum of irreducible representations (see homework). □

Examples.

1. Consider again the left regular representation $\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}$ of $\mathbb{Z}/2\mathbb{Z} = \langle x : x^2 = 1 \rangle$, defined by

$$\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

It's easy to spot an invariant vector, which defines a copy of the trivial representation:

$$\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(x) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Since $\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}$ is two-dimensional, we can pick any vector which is linearly independent from the above; it spans another one-dimensional representation of $\mathbb{Z}/2\mathbb{Z}$ which we have already seen:

$$\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(1) \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{and} \qquad \mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}(x) \begin{bmatrix} 1 \\ -1 \end{bmatrix} = - \begin{bmatrix} 1 \\ -1 \end{bmatrix} .$$

We can thus write $\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}} = \chi_1 \oplus \chi_{-1}$ where $\chi_1 : x \mapsto 1$ is the trivial representation and $\chi_{-1} : x \mapsto -1$ is the nontrivial one.

1.1.4 Tensor products

In addition to taking direct sums, we can also form new representations from old ones by taking tensor products. This proceeds in the expected way.

Definition 1.1.6. *Let (ρ, V) and (σ, W) be representations of a group G. The representation $\rho \otimes \sigma : G \to \mathrm{GL}(V \otimes W)$ defined by $[\rho \otimes \sigma](g) := \rho(g) \otimes \sigma(g)$ is called the **tensor product** of ρ and σ.*

Recalling that $(A \otimes B)(C \otimes D) = (AC \otimes BD)$ for linear operators A, B, C, D, we can easily check that the above really does define a representation.

$$\begin{aligned}[\rho \otimes \sigma](x)[\rho \otimes \sigma](y) &= (\rho(x) \otimes \sigma(x))(\rho(y) \otimes \sigma(y)) = (\rho(x)\rho(y) \otimes \sigma(x)\sigma(y)) \\ &= \rho(xy) \otimes \sigma(xy) = [\rho \otimes \sigma](xy)\,.\end{aligned}$$

Computing the matrix entries of a tensor product representation is fairly straightforward, but does involve some bookkeeping of indices. The point is that, if we fix bases for V and W (and hence also of $V \otimes W$), then the matrices $[\rho \otimes \sigma](x)$ are matrix tensor products of the matrices $\rho(x)$ and $\sigma(x)$. If the indices i and j index basis elements of V, and the indices k and l index basis elements of W, then the matrix entries will satisfy

$$[\rho \otimes \sigma]_{ik,jl} = \rho_{ij}\sigma_{kl}\,.$$

1.1.5 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. If ρ is a representation of a finite group, what is $\rho(1)$?

2. Is the image $\mathrm{im}(\rho) = \{\rho(x) : x \in G\}$ of a representation always a group?

3. Recall that the matrix entries $\rho_{ij} : G \to \mathbb{C}$ of a representation ρ (in some basis) are defined by $\rho_{ij}(g) = \rho(g)_{ij}$. Are the ρ_{ij} homomorphisms?

4. Write out the regular representation of S_3 in explicit matrix form.

5. Let G be a group. Is the regular representation of G reducible or irreducible?

6. Let (ρ, V) be a representation. Suppose that there exists $v \in V$ such that $\{\rho(g)v : g \in G\}$ forms a basis of V. Prove that ρ is isomorphic to the (left) regular representation of G.

7. We defined the *left* regular representation in lecture. How do you think the *right* regular representation is defined? Be sure to check that what you wrote down is truly a representation. In fact, it will be a combinatorial representation. What is the group action that defines it?

8. Prove that the right regular representation is isomorphic to the left regular representation.

9. Write a complete proof of Theorem 1.1.1 (Every representation is a direct sum of irreducible representations.) In addition to the steps proven in lecture, you will need to (a.) prove that the symmetrized operator Π_W^ρ is a projection onto W which commutes with $\rho(x)$ for all $x \in G$, and (b.) use Maschke's Theorem recursively to finish the proof.

10. Write out the matrices of the representation $\mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}} \otimes \mathrm{Reg}_{\mathbb{Z}/2\mathbb{Z}}$ explicitly. Find (and explicitly describe) its irreducible decomposition.

11. Please list any typos and mistakes that you found in the lecture notes and exercises so far. Thanks!

1.2 Schur's Lemma and Character Theory

This section contains the first basic result in the course: Schur's Lemma. It also sets down some basic facts about character theory, and consequences thereof (canonical decompositions, decomposition of the regular representation.)

1.2.1 Preliminaries

As in the previous section, we now rapidly review a few basic concepts from outside representation theory that will be needed for this section. First, recall that *conjugation* by a group element g of some group G is the operation on G defined by

$$h \mapsto ghg^{-1}.$$

It's easy to show that this is an equivalence relation on G, and thus partitions G into equivalence classes, which are called *conjugacy classes*. In the cyclic groups $\mathbb{Z}/n\mathbb{Z}$ (and indeed all abelian groups), the conjugacy classes are uninteresting: each group element forms a class, so that there are $|G|$ conjugacy classes. In the symmetric group S_n, each conjugacy class consists of all permutations with the same cycle structure. For example, the transpositions $\{(a\ b) : 1 \leq a < b \leq n\}$ form one conjugacy class, and so do all permutations with cycle structure $(a\ b)(c\ d\ e)$, and so on. In general, cycle structures (and hence also conjugacy classes of S_n) are in bijective correspondence with integer partitions of n.

Recall that an *eigenvalue* of a linear operator A is a number λ such that $Av = \lambda v$ for some nonzero vector v. The vector v is called the *eigenvector* of A corresponding to the eigenvalue λ. Since we are dealing with the complex field, our linear operators will always have at least one eigenvalue (this follows from the fundamental theorem of algebra, applied to the characteristic polynomial of the operator.) The sum of the eigenvalues of A is called the *trace* of A, and is denoted $\mathrm{Tr}[A]$. Given a basis, the trace is simply the sum of the diagonal elements of A, i.e. $\mathrm{Tr}[A] = \sum_i A_{ii}$. The definition based on eigenvalues is clearly basis-independent; it remains to convince yourself that the two definitions are the same.

Given a vector space V, an *inner product* on V is a map $\langle \cdot | \cdot \rangle : V \times V \to \mathbb{C}$ which is conjugate-symmetric: $\langle x|y\rangle = \overline{\langle y|x\rangle}$, linear in the first argument: $\langle ax + y|z\rangle = a\langle x|z\rangle + \langle y|z\rangle$, and positive-definite: $\langle x|x\rangle = 0$ for $x = 0$ and $\langle x|x\rangle > 0$ otherwise. Since all of our vector spaces are finite dimensional over $\mathbb{C}$, we can always fix an isomorphism between our space and $\mathbb{C}^d$, and then use the inner product

$$\langle x|y\rangle = \sum_{j=1}^{d} x_j y_j^*$$

for $\mathbb{C}^d$.

Inner products are very useful; they allow us to think about geometrically intuitive notions like lengths of vectors and angles between pairs of vectors, even in "unfamiliar" abstract vector spaces. Once we have fixed an inner product, we have a notion of *norm* (or length) $\|v\| = \sqrt{|\langle v|v\rangle|}$ for all vectors. We also have a notion of *orthogonality*: we say that two vectors v, w are orthogonal if $\langle v|w\rangle = 0$. This is identical to the notion of perpendicular lines in Euclidean space.

The choice of inner product also provides us with a useful operation on linear operators, called the *adjoint*. Given a linear operator A, its adjoint is the unique operator $A^\dagger$ which satisfies $\langle Av|w\rangle = \langle v|A^\dagger w\rangle$ for all v, w. You should check that if $A \in \mathrm{End}(\mathbb{C}^d)$ is given by a matrix, then the matrix of $A^\dagger$ is the *conjugate transpose* of A (i.e., the matrix you get by transposing A and then conjugating every matrix entry). A special class of operators in $\mathrm{GL}(V)$ are the so-called *unitary operators*. These are invertible operators which preserve the norm and inner product, i.e. $\|Uv\| = \|v\|$ and $\langle Uv|Uw\rangle = \langle v|w\rangle$ for all $v, w \in V$. An equivalent definition is to say that $U \in \mathrm{End}(V)$ is a unitary operator iff $U^\dagger U = \mathbb{1}_V$.

Inner products and unitary operators will come up very frequently in the course, so get used to working with them! Here's a few simple exercises to try for practice:

- check that all the transpositions in S_n are conjugates of each other;
- prove that the set of group elements $\{e_x : x \in G\}$ in the regular representation is an orthonormal basis.
- prove that the trace is basis-independent;
- prove that the two definitions of "unitary" are equivalent;
- prove that the set of unitary operators on V forms a group, the so-called unitary group $\mathrm{U}(V)$;

1.2.2 Schur's Lemma and orthogonality relations

We now come to the first basic result in the course: Schur's Lemma, proved by Issai Schur in 1905 [3]. Schur used it to prove orthogonality of characters; we will do this in the next subsection.

Lemma 1.2.1. [Schur's Lemma] *Let (ρ, V_ρ) and (σ, V_σ) be irreducible representations of a finite group G, and let f be a linear map from V_ρ to V_σ such that*

$$f \circ \rho(g) = \sigma(g) \circ f \tag{1.1}$$

for all $g \in G$. If $\rho \cong \sigma$, then $f = \lambda \mathbb{1}_{V_\rho}$ for some $\lambda \in \mathbb{C}$; otherwise $f = 0$.

Before we get to the proof, here's a brief remark. Another way of stating the condition (1.1) on f is that the following diagram commutes for every $g \in G$:

$$\begin{array}{ccc} V_\rho & \xrightarrow{\rho(g)} & V_\rho \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ V_\sigma & \xrightarrow[\sigma(g)]{} & V_\sigma \end{array}$$

A map satisfying this condition is sometimes called an *intertwiner* of ρ and σ, or a *G-equivariant map*, where it is understood that it is equivariant with respect to the G-actions on V_ρ and V_σ defined by ρ and σ. Intertwiners (and more generally, equivariant maps) crop up very frequently in representation theory (and more generally, mathematics). Schur's Lemma is thus a crucial result that you should get very familiar with.

Proof. First, an easy case: if $f = 0$, then we are done. So suppose $f \neq 0$, and consider its kernel $\ker(f)$. Pick $v \in \ker(f)$, and note that

$$(f\rho(g)) \cdot v = (\sigma(g)f) \cdot v = \sigma(g) \cdot 0 = 0.$$

This implies that $\rho(g) \cdot v \in \ker(f)$, i.e., that $\ker(f)$ is a subrepresentation of ρ. Since ρ is irreducible, there are two possibilities: $\ker(f) = 0$ or $\ker(f) = V_\rho$. We already assumed $f \neq 0$, so it must be that $\ker f = 0$. Next, consider $w \in \mathrm{im}(f) \subset V_\sigma$. Let $v \in V_\rho$ such that $f \cdot v = w$, and check that

$$(f\rho(g)) \cdot v = (\sigma(g)f) \cdot v = \sigma(g) \cdot w,$$

i.e., $\sigma(g) \cdot w$ is also in $\mathrm{im}(f)$. Hence $\mathrm{im}(f)$ is a subrepresentation of σ, and by irreducibility of σ either $\mathrm{im}(f) = 0$ or $\mathrm{im}(f) = V_\sigma$. The former case is eliminated by the assumption $f \neq 0$.

So we now have that $\ker f = 0$ and $\mathrm{im}(f) = V_\sigma$. It follows that f is an isomorphism of vector spaces, and that $V_\rho \cong V_\sigma$. The condition (1.1) now says that $\rho \cong \sigma$. Since $f \neq 0$, we can pick a nonzero eigenvalue λ of f, with eigenvector v_λ. Define $f' = f - \lambda \mathbb{1}_{V_\rho}$; it's easy to check that f' is again an intertwiner of ρ with itself. Now the argument from the previous paragraph (applied to f') tells us that either $f' = 0$ or $\ker(f') = 0$. The latter cannot be the case, since $f' \cdot v_\lambda = 0$, so we have $f' = 0$; this, in turn, implies that $f = \lambda \mathbb{1}_{V_\rho}$. □

As a first consequence of Schur's Lemma, we will now show that the matrix entries of irreducible representations are orthogonal. To this end, it will be convenient to think about *unitary representations*; these are representations (ρ, V_ρ) where every $\rho(x)$ is a unitary operator, i.e. $\rho(x) \in U(V_\rho) \subset \mathrm{GL}(V_\rho)$ for all $x \in G$. In fact, any representation can be turned into a unitary representation, as follows. Let $\langle \cdot | \cdot \rangle_0$ be an inner product on V_ρ, and define

$$\langle v|w\rangle := \sum_{g\in G} \langle \rho(g)v|\rho(g)w\rangle_0 \,. \tag{1.2}$$

One easily checks (see exercises) that this defines another inner product, and that each $\rho(x)$ is a unitary operator with respect to this new inner product. It follows that any representation can be "unitarized", and thus we do not lose any generality by developing all of our theory using unitary representations only.

Another inner product space which we will discuss frequently is the *group algebra* $\mathbb{C}G = \{f : G \to \mathbb{C}\}$. This is the space of all complex-valued functions on the group. It has dimension $|G|$, and a natural inner product defined by

$$\langle f|f'\rangle_G = \frac{1}{|G|} \sum_{g\in G} f(g) f'(g)^*. \tag{1.3}$$

We have already encountered important elements of $\mathbb{C}G$: given a choice of basis for the space V_ρ of any representation ρ, we defined the "matrix entries of ρ" by

$$\begin{aligned} \rho_{ij} : G &\longrightarrow \mathbb{C} \\ g &\longmapsto \rho(g)_{ij} \,. \end{aligned}$$

When taking inner products of these functions, it is important to note that $\rho_{ij}(g)^* = \rho_{ji}(g^{-1})$; this is an easy consequence of the fact that $\rho(g)^\dagger = \rho(g^{-1})$, which follows from the unitarity of ρ.

As it turns out, the matrix entries of the irreducible representations of G form an *orthonormal family*, in the following sense.

Proposition 1.2.1. *Let (ρ, V_ρ) and (σ, V_σ) be irreducible representations of a finite group G, together with a choice of basis for both V_ρ and V_σ. Then*

$$\langle \rho_{ij} | \sigma_{kl} \rangle_G = \begin{cases} 1/d_\rho & \text{if } \rho \cong \sigma \text{ and } i = k, j = l\,; \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $M : V_\sigma \to V_\rho$, and check that

$$M' := \frac{1}{|G|} \sum_{g \in G} \rho(g) M \sigma(g^{-1})$$

is an intertwiner, i.e., that $\rho(x)M' = M'\sigma(x)$. By Schur's Lemma, if $\rho \not\cong \sigma$ then $M' = 0$. In terms of matrix entries, this means that

$$\frac{1}{|G|} \sum_{g \in G} \sum_{a,b} \rho_{ia}(g) M_{ab} \sigma_{bk}(g^{-1}) = 0$$

for all i, k. Here, i and a index basis elements of V_ρ while b and k index basis elements of V_σ. Rearranging slightly, applying $\sigma_{bk}(g^{-1}) = \sigma_{kb}(g)^*$, and recalling the definition of the inner product in (1.3), the above becomes

$$\sum_{a,b} M_{ab} \langle \rho_{ia} | \sigma_{kb} \rangle_G = 0\,,$$

again for all i, k. Recalling that M was arbitrary, we can fix some particular indices j and l and pick $M_{ab} = \delta_{aj}\delta_{bl}$ (here δ_{ts} denotes the Kronecker delta, defined by $\delta_{ts} = 1$ if $t = s$ and $\delta_{ts} = 0$ if $t \neq s$; in particular, M has a 1 in the j, l spot and zeroes elsewhere.) The left-hand side thus becomes the inner product $\langle \rho_{ij} | \sigma_{kl} \rangle_G$ we are interested in, and the right-hand side is always zero. This establishes the claim of the Proposition in the case $\rho \not\cong \sigma$.

Now consider the case $\rho \cong \sigma$, where (again by Schur's Lemma) $M' = \lambda \mathbb{1}_{V_\rho}$. To compute λ, we compute the trace of both sides, which yields

$$d_\rho \lambda = \mathrm{Tr}[M'] = \frac{1}{|G|} \sum_{g \in G} \mathrm{Tr}\left[\rho(g) M \rho(g^{-1})\right] = \mathrm{Tr}[M]\,.$$

We now proceed exactly as before. We first write $M' = \lambda \mathbb{1}_{V_\rho}$ in terms of matrix entries.

$$\frac{1}{|G|} \sum_{g \in G} \sum_{a,b} \rho_{ia}(g) M_{ab} \rho_{bk}(g^{-1}) = \lambda \delta_{ak} = \frac{\mathrm{Tr}[M]}{n} \delta_{ak}\,.$$

Then we rewrite this slightly to get

$$\sum_{a,b} M_{ab} \langle \rho_{ia} | \rho_{kb} \rangle_G = \frac{\mathrm{Tr}[M]}{n} \delta_{ik}\,,$$

which again holds for any i, k. Finally, we pick $M_{ab} = \delta_{aj}\delta_{bl}$ for an arbitrary choice of j and l, so that the left-hand side becomes $\langle \rho_{ij} | \sigma_{kl} \rangle_G$. The right-hand side is only nonzero when $i = k$ and $j = l$ (the latter implies $\mathrm{Tr}[M] = 1$). This establishes the claim in the case $\rho \cong \sigma$. □

1.2.3 Character theory

Working with representations directly is too cumbersome for simple tasks like computing decompositions and testing equivalence. Thankfully, there is an alternative: character theory. We begin by defining characters, and stating a few basic properties.

Definition 1.2.1. *Let (ρ, V) be a representation of a finite group G. The* ***character*** *of ρ is the map $\chi_\rho : G \to \mathbb{C}$ defined by $\chi_\rho(g) = \mathrm{Tr}[\rho(g)]$.*

Proposition 1.2.2. *Let ρ and σ be representations of a finite group G. Then $\chi_{\rho\oplus\sigma}(g) = \chi_\rho(g) + \chi_\sigma(g)$ and $\chi_{\rho\otimes\sigma}(g) = \chi_\rho(g)\chi_\sigma(g)$ for all $g \in G$.*

Proof. See exercises. □

We established orthonormality for matrix entries of irreducible representations in the previous subsection. This may have seemed a bit dry, but now there's a payoff: we can show that characters admit similar orthonormality relations; this, in turn, dramatically simplifies the task of characterizing representations.

Proposition 1.2.3. *Let ρ and σ be irreducible representations of a finite group G. If $\rho \cong \sigma$, then $\langle\chi_\rho|\chi_\sigma\rangle_G = 1$; otherwise $\langle\chi_\rho|\chi_\sigma\rangle_G = 0$.*

Proof. See exercises. □

Theorem 1.2.1. *Let ρ be a representation of a finite group G, with irreducible direct sum decomposition $\rho \cong \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$. Given an irreducible representation σ of G, the number of ρ_j isomorphic to σ is equal to $\langle\chi_\rho|\chi_\sigma\rangle_G$.*

Proof. By Proposition 1.2.2,

$$\langle\chi_\rho|\chi_\sigma\rangle_G = \langle\chi_{\rho_1}|\chi_\sigma\rangle_G + \langle\chi_{\rho_2}|\chi_\sigma\rangle_G + \cdots + \langle\chi_{\rho_k}|\chi_\sigma\rangle_G \,.$$

By Proposition 1.2.3, each term on the right-hand side above is either one or zero, depending on whether ρ_j is isomorphic to σ or not. The claim follows. □

Noting that rearrangement of direct summands is an isomorphism of representations yields the following important corollary.

Corollary 1.2.1. *Two representations are isomorphic if and only if they have the same character.*

Recall that the set of group elements of a finite group can be partitioned into *conjugacy classes*; each conjugacy class is a subset $X \subset G$ such that $x, y \in X$ implies $y = gxg^{-1}$ for some $g \in G$. A function $f : G \to \mathbb{C}$ is called a *class function* if it is constant on conjugacy classes, i.e., if $f(yxy^{-1}) = f(x)$ for all $x, y \in G$. We will denote the space of class functions on G by $\mathbb{C}_{\mathrm{cl}}G$; it is clearly a subspace of the space $\mathbb{C}G$ of *all* functions on G. Moreover, its dimension is equal to the number of conjugacy classes of G.

Theorem 1.2.2. *Let G be a finite group. Then the set $\{\chi_\rho : \rho \in \hat{G}\}$ is an orthonormal basis for the space $\mathbb{C}_{\mathrm{cl}}G$ of class functions on G.*

Proof. That each χ_ρ is actually a class function follows from the cyclic property of the trace:

$$\chi_\rho(yxy^{-1}) = \mathrm{Tr}[\rho(yxy^{-1})] = \mathrm{Tr}[\rho(y^{-1})\rho(y)\rho(x)] = \chi_\rho(x)\,.$$

By Proposition 1.2.3, the irreducible characters also form an orthonormal system inside $\mathbb{C}_{\mathrm{cl}}G$. It remains to check that this system is complete; we will do this by showing that any class function orthogonal to all of the irreducible characters is zero. Let f be such a function, and ρ an irreducible representation. Consider the "inner product" of f with ρ:

$$\rho_f := \sum_g f(g)^* \rho(g)\,. \tag{1.4}$$

Note that, since f is a class function, we have

$$\rho(x)\rho_f\rho(x)^{-1} = \sum_g f(g)^*\rho(xgx^{-1}) = \sum_h f(x^{-1}hx)^*\rho(h) = \sum_h f(h)^*\rho(h) = \rho_f\,.$$

By Schur's Lemma, ρ_f is a scalar multiple of the identity. To find the scalar, we compute the trace.

$$\mathrm{Tr}[\rho_f] = \sum_g f(g)^*\,\mathrm{Tr}[\rho(g)] = \sum_g f(g)^*\chi_\rho(g) = \langle f|\chi_\rho\rangle = 0\,.$$

By Proposition 1.2.2, ρ_f is actually zero even when ρ is not irreducible. In particular, if ρ is the regular representation of G and e_g denotes the basis vector corresponding to the group element g, we have that

$$\rho_f \cdot e_1 = \sum_g f(g)^*\rho(g)\cdot e_1 = \sum_g f(g)^* e_g$$

is zero. Since the e_g are orthogonal, it follows that $f^* = f = 0$. □

The following important fact is an immediate consequence of the above and Proposition 1.2.3.

Corollary 1.2.2. *For finite groups, the number of irreducible representations is equal to the number of conjugacy classes.*

We end this section with two standard applications of character theory to the problem of computing irreducible decompositions. Given a group G, we denote the set of all irreducible representations of G (modulo equivalence) by $\hat{G}$. The following crucial fact tells us that the regular representation Reg_G of G contains all the elements of $\hat{G}$ as subrepresentations. In a concrete sense, this means that Reg_G encapsulates all of the representation-theoretic data about G.

Theorem 1.2.3. *Let G be a finite group. Then every $\rho \in \hat{G}$ appears in the direct sum decomposition of Reg_G with multiplicity equal to its dimension d_ρ. In particular, $|G| = \sum_{\rho\in\hat{G}} d_\rho^2$.*

Proof. See exercises. □

Looking back at Theorem 1.1.1 (every representation is a direct sum of irreducibles), note that we made no mention of *uniqueness* of decompositions. Indeed, the choice of projection operator (or inner product) may affect the decomposition. Nonetheless, there is still a sense in which decompositions are unique. Determining this unique form and explicitly computing the subspaces is yet another place where characters are quite useful.

Theorem 1.2.4. *Let (ρ, V_ρ) be a representation of a finite group G, and let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be the irreducible subrepresentations of ρ. For each j, let W_j denote the direct sum of all the irreducible subrepresentations of ρ isomorphic to σ_j. Then the decomposition*

$$V_\rho = W_1 \oplus W_2 \oplus \cdots W_k$$

is unique, i.e., it does not depend on the choice of full irreducible decomposition of ρ.

Moreover, the operator

$$\Pi_{\sigma_j} := \frac{d_{\sigma_j}}{|G|} \sum_{x \in G} \chi_{\sigma_j}(x)^* \rho(x)$$

is a projection operator on V_ρ with image W_j.

Proof. We will prove the second claim, regarding the operators Π_{σ_j}; the first claim will then follow by the uniqueness of the Π_{σ_j}. First, consider the restriction $\Pi_{\sigma_j}|_\tau$ of Π_j to some irreducible subrepresentation τ of ρ. Since the character χ_{σ_j} is a class function, we already saw in the proof of Theorem 1.2.2 that $\Pi_{\sigma_j}|_\tau$ is a scalar multiple of the identity. To compute the scalar, we take the trace as usual.

$$\mathrm{Tr}[\Pi_{\sigma_j}|_\tau] = \frac{d_{\sigma_j}}{|G|} \sum_{x \in G} \chi_{\sigma_j}(x)^* \chi_\tau(x) = d_{\sigma_j} \langle \chi_{\sigma_j} | \chi_\tau \rangle \,.$$

The scalar is thus 1 if σ_j is isomorphic to τ, and 0 otherwise. Note that $\Pi_{\sigma_j}|_{W_i}$ is a direct sum of some number of copies of $\Pi_{\sigma_j}|_{\sigma_i}$, and that

$$\Pi_{\sigma_j} = \bigoplus_{i=1}^{k} \Pi_{\sigma_j}|_{W_i} \,.$$

Each of the above summands is either the identity operator (if $\sigma_j \cong \sigma_i$) or the zero operator (if $\sigma_j \ncong \sigma_i$). The result follows. □

We will sometimes refer to the above as the *canonical decomposition theorem.*

1.2.4 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Check that the "ρ-averaged inner product" in (1.2) is truly an inner product, and that each $\rho(x)$ is a unitary operator with respect to this inner product.

2. Let ρ be a representation. What is $M := \sum_{g \in G} \rho(g)$? What is the image of M?

3. Use Proposition 1.2.1 (orthogonality of matrix entries) and the definition of characters to prove Proposition 1.2.2 (character of direct sum representations) and Proposition 1.2.3 (orthogonality of irreducible characters).

4. Use character theory to prove Theorem 1.2.3 (decomposition of the regular representation).

5. Let (ρ, V_ρ) be a representation of G and let V_ρ^* be the *dual space* of V_ρ, consisting of all linear functionals $\varphi : V_\rho \to \mathbb{C}$ on V_ρ. Show that there exists a representation (ρ^*, V_ρ^*) which satisfies

   $$(\rho^*(g)\varphi) : \rho(g)v \mapsto \varphi(v)$$

   for all $g \in G$, $v \in V_\rho$ and $\varphi \in V_\rho^*$. This representation is called the *dual* or *contragredient* representation of ρ. What is the character of ρ^*, in terms of χ_ρ?

6. Prove that all irreducible representations of abelian groups are one-dimensional. Give three different proofs: one by decomposing the regular representation, another using Corollary 1.2.2, and another one using Schur's Lemma directly.

7. Let G be a finite abelian group, and $\hat{G}$ its set of irreducible representations (i.e., characters). Prove that $\hat{G}$ is itself a group (called the *dual group* of G), and that it is isomorphic to G.

8. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

1.3 Constructions and Examples

1.3.1 Preliminaries

We now refresh two basic notions from group theory that we will need: product groups, and coset decompositions. Let G_1 and G_2 be two finite groups. Then we can form their *direct sum* $G_1 \times G_2$ to be the set $\{(g_1, g_2) : g_1 \in G, g_2 \in G\}$ of all pairs of group elements (one from G_1 and one from G_2), together with coordinatewise composition:

$$(g_1, g_2)(h_1, h_2) = (g_1 h_1, g_2 h_2)\,.$$

Direct products are quite useful; for example, we need them to classify all abelian groups (they are all direct products of cyclic groups).

Now let G be a finite group, and H be a subgroup (i.e., a subset of G which is closed under the group operation.) We will frequently write $H \leq G$ for short. Given $g \in G$, the set $gH := \{gh : h \in H\}$ is called a *left coset* of H in G. If two elements $x, y \in G$ belong to the same coset of H (equivalently, if $x^{-1}y \in H$), then we say that they are *congruent modulo* H, and write $x \equiv y \bmod H$. It's easy to check that the cosets of H form a partition G; there are thus $|G|/|H|$ distinct cosets. The number $|G|/|H|$ is called the *index* of H in G. If we pick one element from each coset to form a set R, then we say that R is a *system of representatives* for the cosets of H in G.

To check your understanding, try the following simple exercise. Let G_1 and G_2 be finite groups, and let $G := G_1 \times G_2$ be their direct product. Let $H \leq G$ be the subgroup defined by $H := \{(g_1, 1) : g_1 \in G_1\}$. Compute a coset decomposition and a set of coset representatives for H in G. What is the index?

In addition to the cyclic groups $\mathbb{Z}/n\mathbb{Z}$ and the symmetric groups S_n we introduced before, in this Section we will also talk about the dihedral groups D_n. We will define them below, in the Examples section.

1.3.2 Products

We have previously seen how to take representations of a particular group G, and use them to build other representations of G (e.g., via subrepresentations, direct sums, and tensor products.) Now we will consider how to start from representations of G and build representations of *other* groups related to G. A straightforward example is for product groups.

Proposition 1.3.1. *Let G_1 and G_2 be finite groups, and $G_1 \times G_2$ their product. Given irreducible representations (ρ_1, V_{ρ_1}) of G_1 and (ρ_2, V_{ρ_2}) of G_2, the map*

$$\begin{aligned} \rho_1 \otimes \rho_2 : G_1 \times G_2 &\longrightarrow GL(V_{\rho_1} \otimes V_{\rho_2}) \\ (g_1, g_2) &\longmapsto \rho_1(g_1) \otimes \rho_2(g_2) \end{aligned}$$

defines an irreducible representation of $G_1 \times G_2$*. Moreover, all irreducible representations of* $G_1 \times G_2$ *arise in this way.*

Proof. The fact that $\rho_1 \otimes \rho_2$ is a representation follows directly from the rule for composition of tensor products of operators:

$$\begin{aligned}[\rho_1 \otimes \rho_2](a,b) \cdot [\rho_1 \otimes \rho_2](c,d) &= (\rho_1(a) \otimes \rho_2(b)) \cdot (\rho_1(c) \otimes \rho_2(d)) \\ &= \rho_1(ac) \otimes \rho_2(bd) = [\rho_1 \otimes \rho_2](ac,bd)\,.\end{aligned}$$

Irreducibility follows from character theory. First, note that the character of $\rho_1 \otimes \rho_2$ is

$$\chi_{\rho_1 \otimes \rho_2}(x,y) = \mathrm{Tr}[[\rho_1 \otimes \rho_2](x,y)] = \mathrm{Tr}[\rho_1(x) \otimes \rho_2(y)] = \mathrm{Tr}[\rho_1(x)]\,\mathrm{Tr}[\rho_2(y)] = \chi_{\rho_1}(x)\chi_{\rho_2}(y)\,.$$

Now one should exercise care in keeping track of which inner product we are using. In the following, I have distinguished them by subscripts. We have

$$\langle \chi_{\rho_1 \otimes \rho_2} | \chi_{\rho_1 \otimes \rho_2} \rangle_{G_1 \times G_2} = \langle \chi_{\rho_1} | \chi_{\rho_1} \rangle_{G_1} \langle \chi_{\rho_2} | \chi_{\rho_2} \rangle_{G_2} = 1\,,$$

where the last step follows from the irreducibility of ρ_1 and ρ_2. We conclude that $\rho_1 \otimes \rho_2$ is irreducible.

To see that all of the irreducible representations of $G_1 \times G_2$ are of this type, we just apply Theorem 1.2.3 and do some simple dimension counting.

$$\sum_{\rho_1 \in \hat{G}_1} \sum_{\rho_2 \in \hat{G}_2} d^2_{\rho_1 \otimes \rho_2} = \sum_{\rho_1 \in \hat{G}_1} \sum_{\rho_2 \in \hat{G}_2} d^2_{\rho_1} d^2_{\rho_2} = \sum_{\rho_1 \in \hat{G}_1} d^2_{\rho_1} \sum_{\rho_2 \in \hat{G}_2} d^2_{\rho_2} = |G_1||G_2| = |G_1 \times G_2|\,.$$

The extreme left-hand side is the total squared dimension of all representations constructed as above. The extreme right-hand side is (by Theorem 1.2.3) the total squared dimension of all representations of $G_1 \times G_2$. So, by the above equality, we have them all. □

It is important to note the difference between the above tensor product, and the tensor product of representations defined in the previous section. The crucial difference is that, in this case, we start with representations of G_1 and G_2 and end up with a representation of $G_1 \times G_2$. In the tensor product from Definition 1.1.6, we start with two representations of G, and end up again with a representation of G. This can become even more confusing when considering the representations of $G \times G$; then we can take the tensor product of an irreducible representation ρ with itself in two different ways:

$$[\rho \otimes \rho](g) = \rho(g) \otimes \rho(g) \qquad \text{vs.} \qquad [\rho \otimes \rho](x,y) = \rho(x) \otimes \rho(y)\,.$$

The former is a representation of G, while the latter is a representation of $G \times G$. When thinking about tensor products, you should always be clear about which tensor product you mean!

1.3.3 Restriction and Induction

We already saw that restricting a representation to a proper invariant subspace yields another representation (of the same group.) We can also restrict representations to subgroups, as follows. Let (ρ, V_ρ) be a representation of a finite group G and H a subgroup of G. Define

$$\begin{aligned}\mathrm{Res}^G_H[\rho] : H &\longrightarrow \mathrm{GL}(V_\rho) \\ h &\longmapsto \rho(h)\,.\end{aligned}$$

It's straightforward to check that the $(\mathrm{Res}^G_H[\rho], V_\rho)$ is a representation of H.

It turns out that we can also "lift" representations of subgroups to representations of a larger group. This is called *induction*, and is a bit more complicated than restriction. We start with two examples. For both examples, we let $H \leq G$ and R a set of representatives for the cosets of H in G.

First, take χ_H to be the trivial representation of H. Let V be the vector space spanned by $\{e_r : r \in R\}$. Then G acts on the basis of V by $g \cdot e_r \mapsto e_{g\cdot r}$. This is the same as the action of G on the cosets of H in G. In particular, notice the slight trick in notation: $g \cdot r$ is not a product of group elements g and r; instead, it is the image of the representative r under the action of g. This, in turn, is again an element of R. To compute it explicitly, take the group product gr, figure out which coset it belongs to, and take its representative. In any case, as we saw earlier, this action of G yields a permutation representation (ρ, V) of G. We say that this representation is *induced by* χ_H, and write

$$\rho = \mathrm{Ind}^G_H\left[\chi_H\right] .$$

For the second example, take the regular representation (Reg_H, V_H) of H. Define a vector space W consisting of $|R| = |G|/|H|$ copies of V_H, i.e.,

$$W = \bigoplus_{r\in R} V_{(r,H)}$$

where each $V_{(r,H)}$ is isomorphic to V_H. We will define a representation of G on W. This representation will permute around the spaces $V_{(r,H)}$ in some way, as well as acting inside each $V_{(r,H)}$. Concretely, the coset decomposition

$$G = \bigcup_{r\in R} \{rh : h \in H\}$$

allows us to identify the elements G with the basis elements $v_{r,h}$ of W. Here, $\{v_{r,h} : h \in H\}$ spans $V_{(r,H)}$. It's then straightforward to check that the action of any particular $g \in G$ on $v_{r,h}$ is given by left multiplication: g sends rh to grh, which can then be decomposed again as $grh = r'h'$ for some $r' \in R$ and $h' \in H$. We then define $g \cdot v_{r,h} = v_{r',h'}$. If you're not there yet, you should now convince yourself that the representation on W thus defined is isomorphic to the regular representation Reg_G of G.

We now define the notion of induced representation, generalizing both examples above.

Definition 1.3.1. *Let $H \leq G$ with a set R of coset representatives, and let (σ, V) be a representation of H. Then the induced representation $\mathrm{Ind}^G_H[\rho]$ is defined on $\bigoplus_{r\in R} V_r$ (i.e., $|R|$ copies of V) by setting*

$$\mathrm{Ind}^G_H[\rho](g) \cdot \sum_r v_r = \rho(h) v_{g\cdot r} ,$$

where h is the unique element of H defined by $gr = (g \cdot r)h$.

The last statement may require some clarification: it says that the product of the group elements gr is an element of some left coset of H; the representative of that coset is $g \cdot r$ (here $\cdot$ is the action of G on the set of left cosets), which precisely means that we can write $gr = (g \cdot r)h$ for some $h \in H$.

1.3.4 Examples : cyclic and dihedral groups

In this section, we will construct all the irreducible representations of some common families of groups.

Cyclic groups. Recall that the cyclic groups are defined by $\mathbb{Z}/n\mathbb{Z} = \langle x | x^n = 1\rangle$. Explicitly, we can think of this as the set $\{0, 1, \ldots, n-1\}$ of integers with addition modulo n. Since it is an abelian group, all of the representations will be one-dimensional; we may as well consider only the unitary ones. So we are looking for homomorphisms

$$\chi : \mathbb{Z}/n\mathbb{Z} \longrightarrow \left\{e^{i\theta} : \theta \in [0, 2\pi)\right\}$$

to the group of invertible complex numbers of modulus 1. From the presentation of $\mathbb{Z}/n\mathbb{Z}$ it is clear that we need only decide where to map x; to preserve homomorphism, the image should have order n. There are n choices of how to do this:

$$\chi_k : x \mapsto e^{2\pi i x k/n} ,$$

one for each $k \in \{0, 1, \ldots, n-1\}$. To show that these are non-isomorphic, we verify that the inner products satisfy

$$\langle \chi_j | \chi_k \rangle = \delta_{jk} .$$

By the decomposition of the regular representation, we know that this is all of them.

Dihedral groups. The dihedral group D_n is the set of all reflections and rotations of a regular n-gon. For concreteness, we can imagine the n-gon centered at the origin in the plane, and rotated and rescaled so that one of the vertices is placed at $(1, 0)$. If r is the counterclockwise rotation by $2\pi/n$, then r has order n, and hence $\langle r\rangle$ is a cyclic subgroup of D_n isomorphic to $\mathbb{Z}/n\mathbb{Z}$. Let s be any reflection, and check that $s^2 = 1$ and $srs^{-1} = r^{-1}$. This is a complete set of relations, so we can write

$$D_n = \left\langle r, s \,|\, r^n = 1, s^2 = 1, srs^{-1} = r^{-1} \right\rangle .$$

Let's find the irreducible representations of D_n. We consider two separate cases: n even, and n odd. First, for n even, both r and s have even order. We can thus assign each of them to 1 or -1 to get a homomorphism into the unit circle. There are four choices for how to do this, which yields four nonisomorphic irreducible representations. To compute the remaining representations, we induce up from the characters of the subgroup $H \cong \mathbb{Z}/n\mathbb{Z}$ of pure rotations. Let $\rho_k = \mathrm{Ind}_H^{D_n}[\chi_k]$. There are two cosets: H and sH. It follows that ρ_k will be two-dimensional; the action on the cosets is trivial for the identity, and a swap for s. The action on the non-identity coset is also "twisted" (by a minus sign in the exponent.) We thus have

$$\rho_k(r^j) = \begin{pmatrix} \omega^{jk} & 0 \\ 0 & \omega^{-jk} \end{pmatrix} \qquad \rho_k(sr^j) = \begin{pmatrix} 0 & \omega^{-jk} \\ \omega^{jk} & 0 \end{pmatrix}$$

where $\omega_n = e^{2\pi i/n}$. It's straightforward to check that $\rho_0 \cong \chi_1 \oplus \chi_2$ where χ_1 is the trivial representation and χ_2 is the character defined by $r \mapsto 1$ and $s \mapsto -1$. One also sees that $\rho_k \cong \rho_{n-k}$ (via the isomorphism that swaps the two standard basis elements.) For the remaining irreps, a direct calculation confirms that

$$\rho_k(r) \cdot v = \alpha v$$

only if v spans one of the two coordinate axes; these axes, in turn, are swapped by $\rho_k(s)$. We thus have that the $\{\rho_k : 0 < k < n/2\}$ are irreducible. Summing the squared dimensions of the irreducible representations identified so far, we get

$$1 \cdot 4 + 2^2 \cdot (n/2 - 1) = 2n = |D_n| ,$$

so we have all of them.

For odd n, the order of r is no longer even, so we only get two one-dimensional representations by setting $s \mapsto 1$ or $s \mapsto -1$. We can construct two-dimensional representations ρ_k exactly as above, and the same arguments still hold. Since n is odd, the number of pairwise non-isomorphic ρ_k is now $(n-1)/2$. Counting dimensions, we see again that this is all there is: $1 \cdot 2 + 2^2 \cdot (n-1)/2 = 2n$.

1.3.5 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Suppose ρ is an irreducible representation of a finite group G, and H a subgroup of G. Is $\mathrm{Res}_H^G[\rho]$ irreducible?

2. Suppose ρ is an irreducible representation of a group H, which is a subgroup of another group G. Is $\mathrm{Ind}_H^G[\rho]$ irreducible?

3. Compute the character table of $\mathbb{Z}/4\mathbb{Z}$. Be sure to include a proof that it is correct and complete.

4. Compute the character table of D_4. Be sure to include a proof that it is correct and complete.

5. Let $G = H \times K$, and σ an irreducible representation of K. Show that $\mathrm{Ind}_K^G[\sigma]$ is isomorphic to $\mathrm{Reg}_H \otimes \sigma$.

6. Let $G = (D_k)^n$ be the product $D_k \times \cdots \times D_k$ of n copies of the dihedral group D_k, for some $n, k > 1$. Completely characterize all of the irreducible representations of G.

7. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

1.4 The group algebra and the Fourier transform

1.4.1 The group algebra

Recall that an *associative algebra* is a vector space V equipped with an associative bilinear product $V \times V \to V$. The product of two vectors $v, w \in V$ will be denoted by $v * w$, and must satisfy associativity, distributivity and bilinearity conditions:

- $(u * v) * w = u * (v * w)$;
- $(v + w) * u = v * u + w * u$;
- $u * (v + w) = u * v + u * w$;
- $(av) * (bw) = ab(v * w)$.

We will only deal with unital algebras, i.e. those that possess an identity element 1 satisfying $1 * v = v$ for all $v \in V$. A typical example is the set $\mathrm{End}(W)$ of linear operators on a finite dimensional complex vector space W. More concretely, the space $M_n(\mathbb{C})$ of all $n \times n$ complex matrices is an algebra, where $*$ is the matrix product. This is sometimes called the *full matrix algebra* over $\mathbb{C}^n$. If you have not encountered algebras (or matrix algebras) before, you should verify explicitly that this is indeed an algebra. In particular, you should understand the structure of the underlying vector space, as well as why the $*$ product satisfies all of the conditions above.

For us, the most important algebra (besides the full matrix algebras) is the *group algebra* of a finite group G. We defined this before, but now we will take it apart and study it very carefully. Recall that it is defined as the space of functions from G to the complex numbers, i.e.,

$$\mathbb{C}G = \{f : G \to \mathbb{C}\} .$$

The vector space structure is given by addition and scaling of functions in the expected way: for $f, g \in \mathbb{C}G$ and $a \in \mathbb{C}$, it's immediate that $f + g$ and af are also elements of $\mathbb{C}G$, and that this operation makes $\mathbb{C}G$ into a vector space. To make $\mathbb{C}G$ into an algebra, we need to specify a product. This is the so-called *convolution product,* and is defined as follows.

$$[f * g](x) = \frac{1}{|G|} \sum_{y \in G} f(y) g(y^{-1}x) .$$

It's straightforward to check that this is associative, distributive, and bilinear, and thus indeed an algebra product. The identity element is δ_e, i.e., the function which takes the value 1 on the identity element $e \in G$, and 0 elsewhere.

As we discussed before, $\mathbb{C}G$ also comes equipped with a natural inner product

$$\langle f|g\rangle_G = \frac{1}{|G|} \sum_{x \in G} f(x) g(x)^* .$$

This inner product also gives us a notion of length (or norm) for functions in $\mathbb{C}G$, by setting

$$\|f\|_2^2 = |\langle f|f\rangle| = \frac{1}{|G|} \sum_{x \in G} |f(x)|^2 .$$

We have already seen two bases of $\mathbb{C}G$ which are orthonormal with respect to the above inner product. The first is the group basis $\{\delta_g : g \in G\}$ where $\delta_g(h) = \delta_{gh}$. You can also think about this as the basis of "point functions." The other basis consists of the matrix entries of irreducible representations of G. To be more precise, this "other" basis is in fact a family of bases; fixing one element of this family requires choosing a basis for each $\rho \in \hat{G}$. Once this is done, then Proposition 1.2.1 says that the functions $\rho'_{ij} := \sqrt{d_\rho}\rho_{ij}$ are orthonormal, i.e.

$$\langle \rho'_{ij}|\sigma'_{kl}\rangle = \begin{cases} 1 & \text{if } \rho \cong \sigma \text{ and } i = k, j = l\,, \\ 0 & \text{otherwise.} \end{cases}$$

This basis has some interesting properties related to the convolution product. To see this, let's compute the convolution of two matrix entries.

$$\begin{aligned}
[\rho_{ij} * \sigma_{kl}](x) &= \frac{1}{|G|} \sum_{y \in G} \rho_{ij}(y) \sigma_{kl}(y^{-1}x) \\
&= \frac{1}{|G|} \sum_{y \in G} \rho_{ij}(y) \sum_{t=1}^{d_\sigma} \sigma_{kt}(y^{-1}) \sigma_{tl}(x) \\
&= \sum_t \sigma_{tl}(x) \frac{1}{|G|} \sum_{y \in G} \rho_{ij}(y) \sigma_{tk}(y)^* \\
&= \sum_t \sigma_{tl}(x) \langle \rho_{ij}|\sigma_{tk}\rangle .
\end{aligned}$$

This is zero unless $\rho \cong \sigma$ and $i = t$ and $j = k$; in that case it equals $\rho_{il}(x)$. Summarizing, we have

$$\rho_{ij} * \sigma_{kl} = \begin{cases} \rho_{il} & \text{if } \rho \cong \sigma \text{ and } i = k\,, \\ 0 & \text{otherwise.} \end{cases}$$

This looks interesting. What does it remind you of?

Let's look at this from another point of view. Let L_x be the "left multiplication by x" operator on $\mathbb{C}G$. We previously wrote this as $\mathrm{Reg}_G(x)$. Note that $[L_x f](y) = f(x^{-1}y)$ for any $f \in \mathbb{C}G$. Let's see how this operator interacts with convolution.

$$\begin{aligned}[L_x[f * g]](y) = [f * g](x^{-1}y) &= \sum_z f(z)g(z^{-1}x^{-1}y) \\ &= \sum_t f(x^{-1}t)g(t^{-1}y) = [[L_x f] * g](y)\,,\end{aligned}$$

where we applied the substitution $t = xz$ in the third step. The above means that the operator "convolve with g" (let's call it C_g) commutes with L_x for every $x \in G$. Recalling the decomposition

$$\mathrm{Reg}_G \cong \bigoplus_{\rho \in \hat{G}} \rho^{\oplus d_\rho}$$

of the regular representation of G and applying Schur's Lemma to C_g, we see that C_g is block diagonal (in the first direct sum above), and is described by a $d_\rho \times d_\rho$ complex matrix inside each space $\rho^{\oplus d_\rho}$. To see this, it is instructive to draw the $d_{|G|} \times d_{|G|}$ matrix for C_g, separate it into $|\hat{G}| \times |\hat{G}|$ blocks, and fill in each block according to Schur's Lemma. But what are these $d_\rho \times d_\rho$ matrices? As it turns out, understanding this question (as we will do in the next subsection) sheds significant light on the structure of $\mathbb{C}G$.

Before we go on, let's compute one more thing about $\mathbb{C}G$: it's center. Letting δ_x be the point function at $x \in G$, we check that

$$\begin{aligned}[f * \delta_x](y) &= \sum_z f(z)\delta_x(z^{-1}y) = f(yx^{-1}) \\ [\delta_x * f](y) &= \sum_z \delta_x(z)f(z^{-1}y) = f(x^{-1}y)\end{aligned}$$

Note that f is in the center of $\mathbb{C}G$ if and only if $f * \delta_x = \delta_x * f$ for all $x \in G$. The above tells us that this is true if and only if $f(xy) = f(yx)$ for all $x, y \in G$, i.e., f is a class function.

1.4.2 The Fourier transform on finite groups

The perhaps somewhat mysterious results of the previous section can be understood more clearly via the *Fourier transform* on G. Simply put, the Fourier transform is a change of basis from the group basis to the orthonormal basis of matrix entries ρ'_{ij} of irreducible representations. For this reason, the latter is sometimes called "the Fourier basis" (although by now you know that this is not unique.) Recall that expanding a vector v in an orthonormal basis B is written

$$v = \sum_{b \in B} \langle v|b\rangle b\,.$$

In our case, expanding functions $f \in \mathbb{C}G$ in terms of the Fourier basis functions is as follows.

$$f = \sum_{\rho \in \hat{G}} \sum_{i,j=1}^{d_\rho} \langle f | \rho'_{ij} \rangle \rho'_{ij} \,. \tag{1.5}$$

The terms $\langle f | \rho'_{ij} \rangle$ are called the *Fourier coefficients of* f, and are defined by

$$\hat{f}(\rho_{ij}) = \langle f | \rho'_{ij} \rangle = \frac{\sqrt{d_\rho}}{|G|} \sum_{x \in G} f(x) \rho_{ij}(x)^* \,.$$

It is convenient to gather up all of the coefficients involving ρ into a matrix, which then allows us to forget the choice of basis for ρ. We thus define

$$\hat{f}(\rho) = \frac{\sqrt{d_\rho}}{|G|} \sum_{x \in G} f(x) \rho(x)^\dagger \,.$$

The *Fourier transform* is defined to be the map

$$\begin{aligned} \mathcal{F} : \mathbb{C}G &\longrightarrow \bigoplus_{\rho \in \hat{G}} M_{d_\rho}(\mathbb{C}) \\ f &\longmapsto \bigoplus_{\rho \in \hat{G}} \hat{f}(\rho) \,. \end{aligned} \tag{1.6}$$

We will shortly see that it is in fact an *algebra isomorphism*. First, we can rewrite Equation 1.5 to get the *Fourier inversion formula*:

$$\begin{aligned} f(g) &= \sum_{\rho \in \hat{G}} \sum_{i,j=1}^{d_\rho} \langle f | \rho'_{ij} \rangle \rho'_{ij}(g) \\ &= \sum_{\rho \in \hat{G}} \sqrt{d_\rho} \sum_{i,j=1}^{d_\rho} \hat{f}(\rho)_{ji} \rho_{ij}(g) \\ &= \sum_{\rho \in \hat{G}} \sqrt{d_\rho} \, \mathrm{Tr}\left[\hat{f}(\rho) \rho(g) \right] \,. \end{aligned} \tag{1.7}$$

Now we check that the algebra product on $\mathbb{C}G$ (i.e., convolution) is mapped to the algebra product on $\bigoplus_{\rho \in \hat{G}} M_{d_\rho}(\mathbb{C})$ (i.e., matrix product).

$$\begin{aligned} \widehat{f * g}(\rho) &= \frac{\sqrt{d_\rho}}{|G|} \sum_{x \in G} [f * g](x) \rho(x)^\dagger \\ &= \frac{\sqrt{d_\rho}}{|G|^2} \sum_{x \in G} \sum_{y \in G} f(y) g(y^{-1}x) \rho(x^{-1}) \\ &= \frac{\sqrt{d_\rho}}{|G|^2} \sum_{x \in G} \sum_{y \in G} f(y) g(y^{-1}x) \rho(x^{-1}yy^{-1}) \\ &= \frac{\sqrt{d_\rho}}{|G|^2} \sum_{z \in G} g(z) \rho(z^{-1}) \sum_{y \in G} f(y) \rho(y^{-1}) \\ &= \frac{1}{\sqrt{d_\rho}} \hat{g}(\rho) \cdot \hat{f}(\rho) \,, \end{aligned} \tag{1.8}$$

where we used the substitution $y^{-1}x \mapsto z$ in the second-to-last step. The issue of the constant $d_\rho^{-1/2}$ is a minor one: we could have easily made the constant 1 by rescaling the definition of the Fourier transform; the advantage of the present scaling is that it makes the Fourier transform unitary—as you will verify in the homework.

We finish with an important example: the *discrete Fourier transform*, or DFT. This transform is just a special case of $\mathcal{F}$, for the cyclic group $\mathbb{Z}/n\mathbb{Z}$. Its applications are legion. We start with a function

$$f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C};$$

we can imagine that the function values represent n equispaced samples from some periodic signal. The standard approach in signal analysis (or any of the other untold applications of the DFT) is to decompose our signal into distinct frequencies. These frequencies are precisely the characters of $\mathbb{Z}/n\mathbb{Z}$. The frequency components of f are

$$\hat{f}(k) = \sum_{j=0}^{n-1} f(j)\chi_k(j)^* = \sum_{j=0}^{n-1} f(j)e^{-2\pi ijk/n} .$$

and the decomposition of the function into these components is

$$f(x) = \sum_{k=0}^{n-1} \hat{f}(x)\chi_k(x)^* = \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} f(j)\chi_k(jx^{-1}) = \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} f(j)e^{2\pi ik(j-x)/n} .$$

One of the many advantages of this decomposition is that convolution (which is useful in signal analysis, but costly to compute directly) can be computed just by multiplying together n complex numbers.

1.4.3 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Let G be a finite group. Compute the following Fourier transforms:
   - $\hat{\delta_e}$, where δ_e is the function that takes the value 1 on the identity and 0 elsewhere.
   - $\hat{\chi_1}$, where χ_1 is the trivial representation.

2. Let G be a finite group, and ρ an irreducible representation of G. Compute the following Fourier transforms:
   - $\hat{\rho_{ij}}$ for any $1 \le i, j \le d_\rho$;
   - $\hat{\chi_\rho}$.

3. Let G be a finite group, and ρ, σ two (possibly inequivalent) irreducible representations of G. Compute $\chi_\rho * \chi_\sigma$.

4. Let G be a finite abelian group. Define an inner product on $\mathbb{C}\hat{G}$ such that the Fourier transform is unitary, i.e., $\|f\|_2 = \|\hat{f}\|_2$ for every $f \in \mathbb{C}G$.

5. Recall that the space $M_n(\mathbb{C})$ of $n \times n$ matrices is an inner product space, with inner product

   $$\langle A, B\rangle = \mathrm{Tr}[A^\dagger B] .$$

   Now extend your solution in Problem 4 to cover non-abelian groups.

6. Let G be a finite abelian group, and $f \in \mathbb{C}G$. Define

$$\mathbf{supp}\, f = \{x \in G : f(x) \neq 0\} \qquad \text{and} \qquad \mathbf{supp}\, \hat{f} = \{\chi \in \hat{G} : \hat{f}(\chi) \neq 0\}\,.$$

Recall that $\|f\|_\infty^2 = \max\{|f(x)| : x \in G\}$. Prove that

$$\|f\|_\infty^2 \leq \sum_{\chi \in \hat{G}} |\hat{f}(\chi)|^2 \sum_{\chi \in \mathbf{supp}\, \hat{f}} 1\,.$$

7. Continuing from Problem 5, use your solution to Problem 4 to prove the following uncertainty principle for finite abelian groups:

$$|\mathbf{supp}\, f||\mathbf{supp}\, \hat{f}| \geq |G| \qquad \text{for all } f \in \mathbb{C}G\,.$$

Give an example (of a particular f, for any G) where you get equality above.

8. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

Chapter 2

Compact groups

2.1 Basic facts: passing from finite to compact

2.1.1 Topology, measure, and integration

We begin with a quick refresh of basic ideas from point-set topology. If you have not seen formal topology before, it may seem dry at first. Just keep in mind that virtually everything you want to do with infinite sets (take limits, discuss continuity, integrate functions, etc.) requires topology. Even if you have not seen the formal definitions before, you have certainly been implicitly using them! As you read the next few paragraphs, it might be useful to occasionally flip down and look at some of the examples before continuing.

Topological spaces, connectedness and compactness. Recall that a **topology** on a set X is a family $\mathcal{T}$ of subsets such that (i.) the empty set and X are in $\mathcal{T}$, (ii.) any union of elements of $\mathcal{T}$ is also an element of $\mathcal{T}$, and (iii.) any finite intersection of elements of $\mathcal{T}$ is also an element of $\mathcal{T}$. A set X equipped with a topology is called a **topological space,** or simply space. The elements of $\mathcal{T}$ are called **open** sets, and their complements are **closed** sets. An arbitrary subset of X can be closed, open, both, or neither. A **basis** for a topology is a family of sets which generate $\mathcal{T}$ under the operation of arbitrary unions; in other words, any open set can be written as a union of basis sets. Starting from a topological space X, any subset $Y \subseteq X$ can also be given the structure of a topological space, by setting the open sets of Y to be the intersections of Y with the open sets of X, i.e. $\mathcal{T}_Y := \{Y \cap O : O \in \mathcal{T}\}$. The set Y equipped with the topology $\mathcal{T}_Y$ is then called a **subspace** of X. The product $X \times Y$ of two topological spaces X and Y is again a topological space, with open sets the arbitrary unions of products $U \times V$ where U is open in X and V is open in Y. This extends in the obvious way to any finite number of products.

A topological space which is equal to a disjoint union of two nonempty open sets is said to be disconnected. If a topological space cannot be written as a disjoint union of nonempty open sets, then it is said to be **connected.** Any space X can be partitioned into a collection of connected subspaces; the elements of the finest such partition are called the **connected components** of that space.

Unlike that of connectedness, the formal definition of compactness is slightly unwieldy; thinking about examples will help clarify what it means. Formally, a topological space X is **compact** if any open cover contains a finite subcover. This means that for every (possibly infinite) collection $\{U_j\}$ of open sets such that $X \subseteq \cup_j U_j$, there is a finite set J such that $X \subseteq \cup_{j\in J} U_j$. An equivalent definition[1] is in terms of limit points. A **limit point** of an infinite set S is a point x such that every open set containing x also contains some point $s \in S$ not equal to x. A space X is then compact if every infinite subset S of X has a limit point in X.

Continuity, measure and integration. A topology allows us to define a notion of continuous function. Roughly speaking, these are functions from one topological space to another, which map nearby points to nearby points. This is defined in terms of neighborhoods; a neighborhood of a point x is simply an open set containing x. Formally, a function $f : X \to Y$ is **continuous** at a point x if for every neighborhood V of $f(x)$, there exists a neighborhood U of x such that $f(U) \subseteq V$. If f is continuous at every point of X, then we simply say that f is continuous.

We will require the notion of measure. A measure assigns a notion of "mass" to the subsets of a topological space, and is a crucial step towards integration. Measure theory can be a somewhat complicated subject to treat in complete rigor, so we will settle for a semi-formal treatment. For us, a measure μ on a compact set X will be a map which assigns to each subset[2] $M \subseteq X$ a real number $\mu(M)$, such that (i.) $\mu(\emptyset) = 0$ and $\mu(X) = 1$, (ii.) $\mu(M) \geq 0$ for all $M \subseteq X$, and (iii.) $\mu(\cup_j E_j) = \sum_j \mu(E_j)$.

Once we have fixed a measure for our space X, then (up to a lot of details we will not discuss) we also get integrals. This means that, for any[3] function $f : X \to \mathbb{C}$ on a compact space X with measure μ, we can define its **integral**

$$\int_X f(x)\, d\mu\,.$$

The integral captures the familiar notion of "area under the curve." It can be defined completely by first setting its value on characteristic functions of subsets of X, as follows.

$$\int_X \chi_M(x)\, d\mu = \mu(M) \qquad \text{where} \qquad \chi_M(x) = \begin{cases} 1 & \text{if } x \in M \\ 0 & \text{if } x \notin M\,. \end{cases}$$

We then extend the integral to functions which are linear combinations of such characteristic functions (on disjoint subset) by requiring that the integral should be linear, i.e.

$$\int_X [\alpha f + g](x)\, d\mu = \alpha \int_X f(x)\, d\mu + \int_X g(x)\, d\mu \qquad \text{for all } f, g : X \to \mathbb{C} \text{ and } \alpha \in \mathbb{C}\,.$$

The remaining step is to incorporate limits. Roughly speaking, we require that the integral of a limit of functions should be equal to the limit of the integrals.

Examples. A finite set X is typically given the so-called discrete topology, where the basis consists of singleton sets $\{\{x\} : x \in X\}$. The topology then consists of all subsets of X. Every set of X is both open and closed. It is a disconnected space, with connected components being the singleton sets. It is compact, since any cover of X by open sets will necessarily contain the finite subcover $\{\{x\}x \in X\}$. By following the definition, it's easy to see that any function from a finite set to another set is continuous (basically this is because $\{x\}$ is a neighborhood of x.) We

[1] Actually, these definitions are not equivalent in general. However, they are equivalent on metric spaces, and this will be enough for us.

[2] Technically, this will not hold for all subsets; we will not be concerned with this, as all the sets we will encounter will be "measurable."

[3] Here again, this will hold for all functions we will encounter, but is not true in general.

can define a measure on X by setting $\mu(\{x\}) = 1/|X|$ for each $x \in X$. The integral on X is then defined by

$$\int_X f(x)\, d\mu := \frac{1}{|X|} \sum_{x \in X} f(x)\,.$$

The real line $\mathbb{R}$ is a topological space with basis consisting of all the open intervals $(a, b) = \{x : a < x < b\}$, for all $a < b$. The notion of limit points, openness, and closedness is the natural one you already know. For example, the limit points of the set (a, b) are a and b, and the closure of (a, b) is the set $[a, b] = \{x : a \leq x \leq b\}$, which is a closed set and contains all of its limit points. The space $\mathbb{R}$ is connected: if $\mathbb{R}$ is a disjoint union of X and Y, then we can pick an interval $[a, b]$ with one endpoint in X, the other endpoint in Y, and show that $[a, b]$ contains some points neither in X nor in Y, a contradiction. On the other hand, $\mathbb{R}$ is not compact: the open cover $\{(x, x+2) : x \in \mathbb{Z}\}$ contains no finite subcover; alternatively, the infinite subset $\{x : x \in \mathbb{Z}\} \subset \mathbb{R}$ has no limit point in $\mathbb{R}$. Since we did not worry about measure and integration for non-compact sets, we will not discuss it for $\mathbb{R}$. However, the closed interval $[0, 1]$ is compact, and admits the usual notion of measure and integration. This is defined by assigning measure to intervals in the obvious way: $\mu((a, b)) = b - a$. The resulting integral is the familiar Riemann (or Lebesgue) integral, and allows us to write things like

$$\int_{x=0}^{1} 2x + ix\, d\mu = 2 \int_{x=0}^{1} x\, d\mu + i \int_{x=0}^{1} x\, d\mu = 1 + i/2\,.$$

One similarly defines a topology on $\mathbb{R}^n$. In all of these spaces, one can show that compactness is equivalent to being closed and bounded (bounded here means that it lies inside a ball of some finite radius). Measure and integration can be defined appropriately for compact subsets of $\mathbb{R}^n$. By viewing $\mathbb{C}$ as $\mathbb{R}^2$, all of this extends to finite-dimensional complex vector spaces as well. In all of these spaces, continuity of functions is precisely the notion of continuity you are familiar with from calculus.

The circle $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \cong \{e^{i\theta} : 0 \leq \theta < 2\pi\}$ is also a topological space, with basis the open intervals $(e^{i\theta_1}, e^{i\theta_2})$ for $\theta_2 > \theta_1$. Each such interval is assigned measure equal to the (normalized) difference between the initial and the final angle. So, for $0 < \theta_1 < \theta_2 < 2\pi$, we have $\mu((e^{i\theta_1}, e^{i\theta_2})) = (\theta_2 - \theta_1)/2\pi$. In other words, we are viewing S^1 as the interval $[0, 2\pi]$ of the real line, with the two endpoints identified, and the measure rescaled by $1/2\pi$. So, for instance,

$$\int_{\theta=-\pi/4}^{\pi/4} 1\, d\mu = \mu((e^{-i\pi/4}, e^{i\pi/4}) = \frac{1}{2}\,.$$

The sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ is also a compact, connected topological space. We can re-parametrize the sphere in terms of spherical polar coordinates (θ, ϕ) with $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. This means that any $\vec{x} \in S^2 \subset \mathbb{R}^3$ can be written as $\vec{x}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ for some (θ, ϕ) in the allowed range. The integral of a function $f : S^2 \to \mathbb{C}$ (expressed in polar coordinates) is given by

$$\int_{S^2} f(\vec{x})\, d\mu := \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} f(\vec{x}(\theta, \phi)) \sin\theta\, d\theta\, d\phi\,.$$

The $1/4\pi$ term ensures that the integral over the entire sphere is 1; the $\sin\theta$ term comes from changing the variables of integration; roughly speaking, it is the "distortion term" we must introduce to fit the integral over the rectangle $[0, 2\pi] \times [0, \pi]$ onto the sphere.

2.1.2 Topological groups and the Haar integral

A **topological group** is a group G equipped with a topology such that both the group product (viewed as a map from $G \times G$ to G) and group inversion (viewed as a map from G to G) are continuous functions. The real line $\mathbb{R}$ with addition, the unit circle S^1 under complex multiplication, and all finite groups (with the discrete topology) are all topological groups. While the first is not compact, the latter two are; we call such groups **compact groups.** It is a theorem that compact groups admit a special kind of measure, called **Haar measure.** Haar measure is a measure μ (as defined above) with an additional properties: bi-invariance under the group product. This means that for any $X \subset G$ and any $y \in G$, the left translation $y \cdot X = \{yx : x \in X\}$ and the right translation $X \cdot y = \{xy : x \in X\}$ have the same measure as X. In other words,

$$\mu(X) = \mu(y \cdot X) = \mu(X \cdot y)\,.$$

In fact, one can show that Haar measure is the unique bi-invariant measure.

With Haar measure established, we can now follow the above prescription for constructing an integral. The result is the so-called Haar integral. It inherits the bi-invariance property from Haar measure. In other words, we have

$$\int_{x \in G} f(yx)\, d\mu = \int_{x \in G} f(xy)\, d\mu = \int_{x \in G} f(x)\, d\mu$$

for all $y \in Y$ and all complex-valued functions f on G. The Haar integral will be crucial in understanding the representation theory of compact groups, since it provides a generalized notion of "averaging" over the group.

Examples. Recall that finite sets (with the discrete topology) are compact, and that all functions on finite sets are continuous. It follows that any finite group G is a compact topological group, with Haar measure defined by

$$\mu(X) = \frac{|X|}{|G|} \qquad \text{for any } X \subseteq G\,,$$

and Haar integral defined by

$$\int_G f(x)\, d\mu = \frac{1}{|G|} \sum_{x \in G} f(x)\,.$$

It's straightforward to check that both are invariant under left-multiplication and right-multiplication by elements of G.

The set $S^1 = \{e^{i\theta} : 0 \leq \theta < 2\pi\}$ with product $e^{i\phi}e^{i\theta} = e^{i(\phi+\theta)}$ is also a compact topological group. Again, the Haar measure and Haar integral are precisely equal to the "standard" measure and integral for the space S^1, as discussed above. Bi-invariance is straightforward to check. It should be noted that S^1 is isomorphic to $SO(2)$, the group of rotations of the plane. Elements of this group are 2×2 orthogonal matrices with determinant one. In other words,

$$SO(2) = \{M \in M_2(\mathbb{R}) : M^{\text{tr}} M = \mathbb{1}_2 \text{ and } \det(M) = 1\}\,.$$

Recall that a general matrix of this form can be written as

$$M = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\,.$$

Note that, from this point of view, functions $f : SO(2) \to \mathbb{C}$ map matrices to complex numbers, and we now know how to integrate such functions.

The set of origin-preserving rotations of three-dimensional Euclidean space is also a compact group. This is the group

$$SO(3) = \{M \in M_3(\mathbb{R}) : M^{\mathrm{tr}} M = \mathbb{1}_2 \text{ and } \det(M) = 1\} .$$

Defining the Haar integral on SO(3) is slightly complicated, and proceeds roughly as follows:

1. Prove that the set H of z-axis preserving rotations is a subgroup isomorphic to $SO(2)$; this enables us to integrate any function supported only on H by using the integral on $SO(2)$. Moreover, we can use the invariance property (under multiplication by x) of the integral to integrate any function supported on an arbitrary coset xH.

2. Prove that the coset space $SO(3)/H$ is in bijective correspondence with the sphere S^2. Hence any function can be written as a sum of functions, each one defined only on one coset.

3. Now write the integral on $SO(3)$ as an integral over H, followed by an integral over the coset space:
$$\int_{SO(3)} f(x)\, dx = \int_{y \in S^2} \int_{h \in SO(2)} f(yh)\, dy\, dh ,$$
where the two integrals are defined as above.

2.1.3 Representation theory of compact groups: general facts

We now essentially review most of the central results of Chapter I, but in the setting of compact groups. In many cases, the definitions, theorems, and proofs are identical (or almost identical.) When there are significant differences, we will sketch them out.

We begin by defining representations, which are now required to be continuous.

Definition 2.1.1. *A* ***representation*** *of a compact group G is a continuous homomorphism $\rho : G \to \mathrm{GL}(V)$ for some finite-dimensional complex vector space V.*

Just as before, if we fix a basis for V, then ρ becomes a matrix representation, and we can talk about its matrix entries. The notion of isomorphism is also the same, and still amounts to a simultaneous basis change. The notions of subrepresentation, direct sum, tensor product, and irreducible representation are identical to the case of finite groups.

The first difference arises in trying to prove Maschke's theorem; recall that this was a step in the proof that every representation is a direct sum of irreducible representations. The proof for compact groups is different in only one step: the construction of the symmetrized projection operator. Before this was done by averaging over the finite group; we can now replace this with a Haar integral. Letting (ρ, V_ρ) be a representation, W an invariant subspace, and Π_W a projection operator onto W, we set

$$\Pi_W^\rho := \int_G \rho(x) \Pi_W \rho(x)^{-1}\, d\mu(x) , \tag{2.1}$$

where the integral is the Haar integral. Note that the above is in fact an integral of an operator-valued function. One way to define this is via a basis, and then to show basis independence by using the linearity of the integral. The integral of a matrix-valued function $F : G \to M_n(\mathbb{C})$ is then defined to be the unique operator

$$M = \int_G F(x)\, d\mu(x) \in M_n(\mathbb{C}) \qquad \text{satisfying} \qquad M_{ij} = \int_G F(x)_{ij}\, d\mu(x) \text{ for all } 1 \le i, j \le n .$$

With this definition, we can now return to (2.1), and prove that it is a projection operator which commutes with $\rho(x)$ for all $x \in G$. The remainder of the proof is as before.

Schur's Lemma is stated and proved precisely as in the finite group case. The next step is to show orthogonality of matrix entries of representations. In order for this to make sense, we need an inner product for functions on G. This is defined as follows.

$$\langle f|g\rangle_G = \int_G f(x)g(x)^* \, d\mu(x) \, .$$

This also yields a notion of norm for such functions, by setting

$$\|f\|_2 = \sqrt{|\langle f|f\rangle|} \, .$$

Since these norms may not always be finite, we will henceforth restrict ourselves to working only with functions of finite norm. We thus define, for any compact group G, the space

$$L^2(G) = \{f : G \to \mathbb{C} : \|f\|_2 < \infty\} \, .$$

In words, this is the space of square-integrable complex-valued functions on the group. We can now state the orthogonality condition for matrix entries of representations. It states that

$$\langle \rho_{ij}|\sigma_{kl}\rangle = \begin{cases} 1/d_\rho & \text{if } \rho \cong \sigma \text{ and } i = k, j = l\,; \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Recall that the proof involved choosing a matrix $M_{ab} : V_\sigma \to V_\rho$ with a 1 in the a, b position and zeroes elsewhere, and then showing that a certain averaged version of M_{ab} is an intertwiner. We do this again, now setting

$$M'_{ab} = \int_G \rho(x) M_{ab} \sigma(x^{-1}) \, d\mu(x) \, .$$

We then check that this is an intertwiner and apply Schur's Lemma to conclude $M'_{ab} = \lambda_{ab}\mathbb{1}$, with $\lambda_{ab} = 0$ when $\rho \ncong \sigma$. The precise orthogonality condition now follows by computing traces.

Now that we have the orthogonality of matrix entries, we immediately also get orthogonality for irreducible characters. We thus have

$$\langle \chi_\rho|\chi_\sigma\rangle = \begin{cases} 1 & \text{if } \rho \cong \sigma \\ 0 & \text{otherwise.} \end{cases}$$

One topic that we have not addressed so far is what happens to the regular representation. Recall that the regular representation of G had a basis identified with the group elements of G. This makes sense when G is finite, but in general it would give a vector space of uncountable dimension when G is compact. Instead, let's develop the analogy from a different point of view. We can view the regular representation of the finite group as the space $\mathbb{C}G = \{f : G \to \mathbb{C}\}$ with the group action $y \cdot f(x) = f(y^{-1}x)$. This definition translates more easily to the compact setting, except the space $\mathbb{C}G$ for compact G is again too big. Instead, we choose the space $L^2(G)$, which is also a representation of G (with, in general, countably infinite dimension) under the action $y \cdot f(x) = f(y^{-1}x)$. You might wonder if we discarded too much by passing from arbitrary functions on G to square-integrable functions on G. The following theorem says that this is not the case.

Theorem 2.1.1. Peter-Weyl. *Let G be a compact topological group. The set of matrix entries of irreducible representations of G are dense in the space $L^2(G)$. Moreover, we have the following decomposition of representations:*

$$L^2(G) \cong \bigoplus_{\rho \in \hat{G}} \rho^{\oplus d_\rho} .$$

We emphasize that $\hat{G}$ may now be *countably infinite*, unlike in the case of finite groups. Notice that we have already shown one direction of the first statement, i.e., that the matrix entries are contained in $L^2(G)$. The rest of the proof is somewhat involved and requires tools from functional analysis. In the interest of getting to examples (i.e., Lie groups), we will not cover it here.

Before doing that, let's discuss the Fourier transform in the setting of compact groups. Recall that for finite groups, the Fourier transform was a basis-change transformation on $\mathbb{C}G$, taking us from the basis of group elements to the Fourier basis of matrix entries of irreducible representations of G. While there is no obvious "group basis" in the general compact case, there is still the Fourier basis (as given by the Peter-Weyl theorem.) Starting with a function $f \in L^2(G)$, we expand f in the orthonormal basis specified by the theorem and Equation 2.2. The coefficients of the expansion form the Fourier transform of f, as follows.

$$\langle f | \rho_{ij} \rangle = \int_G f(x) \rho_{ij}(x) \, d\mu(x) .$$

We then collect together the coefficients corresponding to a particular irreducible representation, and define

$$\hat{f}(\rho) = \sqrt{d_\rho} \int_G f(x) \rho(x)^\dagger .$$

We can then define the Fourier transform of f as before:

$$\begin{aligned} \mathcal{F} : L^2(G) &\longrightarrow \bigoplus_{\rho \in \hat{G}} M_{d_\rho}(\mathbb{C}) \\ f &\longmapsto \bigoplus_{\rho \in \hat{G}} \hat{f}(\rho) . \end{aligned} \tag{2.3}$$

The Fourier inversion formula is again computed by expressing the "basis change" in a way that does not require choosing bases for the irreducible representations. It is as follows.

$$\begin{aligned} f(g) &= \sum_{\rho \in \hat{G}} \sum_{i,j=1}^{d_\rho} \langle f | \rho'_{ij} \rangle \rho'_{ij}(g) \\ &= \sum_{\rho \in \hat{G}} \sqrt{d_\rho} \sum_{i,j=1}^{d_\rho} \hat{f}(\rho)_{ji} \rho_{ij}(g) \\ &= \sum_{\rho \in \hat{G}} \sqrt{d_\rho} \, \mathrm{Tr} \left[\hat{f}(\rho) \rho(g) \right] . \end{aligned}$$

A little bit of caution is called for here: the above expansion is actually a series, and not a sum (since, in general, $\hat{G}$ is countably infinite.) However, by the Peter-Weyl theorem, if f is in $L^2(G)$, the series will necessarily converge.

Recall that, for finite groups, the Fourier transform was actually an isomorphism of algebras. Is that still true here? Let's define convolution and see. For $f, g \in L^2(G)$, define

$$[f * g](x) = \int_G f(y) g(y^{-1}x)\, d\mu(y)\,.$$

Now let's compute the Fourier transform of a convolution.

$$\begin{aligned}
\widehat{f * g}(\rho) &= \sqrt{d_\rho} \int_G [f * g](x) \rho(x)^\dagger \, d\mu(x) \\
&= \sqrt{d_\rho} \int_G \int_G f(y) g(y^{-1}x) \rho(x^{-1}) \, d\mu(x)\, d\mu(y) \\
&= \sqrt{d_\rho} \int_G \int_G f(y) g(y^{-1}x) \rho(x^{-1} y y^{-1}) \, d\mu(x)\, d\mu(y) \\
&= \sqrt{d_\rho} \int_G g(z) \rho(z^{-1}) \, d\mu(z) \int_G f(y) \rho(y^{-1}) \, d\mu(y) \\
&= \frac{1}{\sqrt{d_\rho}} \hat{g}(\rho) \cdot \hat{f}(\rho)\,,
\end{aligned}$$

We conclude that the Peter-Weyl isomorphism is in fact also an isomorphism of algebras, mapping convolution to matrix product.

2.1.4 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Prove that $\mathrm{SO}(2) = \{M \in M_2(\mathbb{R}) : M^{\mathrm{tr}} M = \mathbb{1}, \det M = 1\}$ is isomorphic to the circle group $S^1 = \{e^{i\theta} : 0 \leq \theta < 2\pi\}$.

2. Find all the irreducible representations of $SO(2)$. Write down the Fourier transform and the Fourier inversion formula for this group.

3. Write down generators and relations for the group D_∞ consisting of all origin-preserving rotations and reflections of the plane. Describe the generators explicitly in geometric terms: what is their action on the plane? Find all the irreducible representations of D_∞.

4. Prove Maschke's Theorem for compact groups: given a representation ρ of a compact group G and an invariant subspace W, there exists another invariant subspace W' such that $\rho = \rho|_W \oplus \rho|_{W'}$.

5. Let (ρ, V_ρ) be a representation of a compact group G. Prove that there is an inner product on V_ρ such that ρ is a unitary representation.

6. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

2.2 Matrix Lie groups

For the rest of the course, we will study the (arguably) most important examples of compact groups, and their representation theory. These are the compact Lie groups. In order to minimize the course requirements, I have decided to study these groups from the *matrix group* point of view. We will be following the excellent text of B. Hall [1]. Unfortunately, due to the compressed schedule of the course, we will have to skip many of the parts of this text in order to reach the interesting representation-theoretic constructions.

2.2.1 Definition and basic examples

We begin with a few basic definitions. We let $M_n(\mathbb{C})$ denote the space of all $n \times n$ complex matrices, which we will sometimes identify with the usual complex Euclidean space $\mathbb{C}^{n^2}$. This gives $M_n(\mathbb{C})$ a topology, and in particular a notion of open and closed sets. We let $\mathrm{GL}_n(\mathbb{C})$ denote the subset of all $n \times n$ invertible matrices, and we let $\mathrm{GL}_n(\mathbb{R})$ denote the subset of those which only have real entries. We give both of these spaces the subspace topology inherited from $M_n(\mathbb{C})$.

Definition 2.2.1. *A* ***matrix Lie group*** *is a closed subgroup G of* $\mathrm{GL}_n(\mathbb{C})$.

In place of saying that the subgroup should be closed, we can also ask that it contains its limit points, in the following sense. If A_n is a sequence of matrices in $M_n(\mathbb{C})$, we say that A_n converges to a matrix A as $n \to \infty$ if it does so entrywise, i.e., if the j, k entry of A_n converges to $A_{j,k}$ for all j, k. That G is a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ is then equivalent to saying that the limit of any sequence of matrices in G is either in G, or is not invertible. For an example of a subgroup of $\mathrm{GL}_n(\mathbb{C})$ which is not closed, take the subgroup of matrices with rational entries. It's then clear that there are limits of such matrices which are invertible, but are not in the subgroup.

Next, we will familiarize ourselves with the most important examples, beginning with the general and special linear groups. Recalling that the entire space itself is always closed in its topology, we see that $\mathrm{GL}_n(\mathbb{C})$ satisfies the definition of a matrix Lie group. We call it the **general linear group**. Moreover, since the limit of matrices with real entries has real entries, $\mathrm{GL}_n(\mathbb{R})$ is also a matrix Lie group. You have already encountered several examples. The group R^* of nonzero real numbers (under multiplication) is isomorphic to $\mathrm{GL}_1(\mathbb{R})$. Likewise, the group C^* of nonzero complex numbers (under multiplication) is isomorphic to $\mathrm{GL}_1(\mathbb{C})$. The vector space $\mathbb{R}^n$ (as a group under vector addition) can be viewed as a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ via the map

$$(x_1, x_2, \ldots, x_n) \longmapsto \begin{pmatrix} e^{x_1} & & 0 \\ & \ddots & \\ 0 & & e^{x_n} \end{pmatrix} .$$

Recall that the determinant of an $n \times n$ matrix A is defined by

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{j=1}^{n} A_{i,\sigma(i)} ,$$

Here the sum runs over all permutations of n, and $\mathrm{sgn}(\sigma)$ denotes the signature of the permutation σ. If σ is a product of an even number of transpositions, then $\mathrm{sgn}(\sigma) = 1$; otherwise it is -1. An equivalent definition is that $\det(A)$ is the product of the eigenvalues of A (with multiplicity.) The latter definition makes it clear that the determinant is a continuous function of A; in particular, limits of matrices with determinant one will also have determinant one. We can thus define the **special linear groups** $\mathrm{SL}_n(\mathbb{C})$ and $\mathrm{SL}_n(\mathbb{R})$, consisting of invertible matrices with determinant one, with matrix entries from the appropriate field.

We have already encountered unitary matrices. It turns out that they also form a Lie group. Before, defined the unitary matrices to be the matrices $A \in M_n(\mathbb{C})$ which satisfy $A^\dagger A = \mathbb{1}$. Here

$A^\dagger$ denotes the adjoint of A, i.e., the conjugate-transpose of A, with matrix entries

$$A^\dagger_{ij} = \overline{A_{ji}}\,.$$

Recall that this is equivalent to requiring that A preserves the standard inner product

$$\langle v|w\rangle = \sum_j v_j w_j^*$$

on $\mathbb{C}^n$. After all,

$$\langle Av|Aw\rangle = \langle A^\dagger Av|w\rangle = \langle v|w\rangle\,.$$

Note that, since $A^\dagger A = \mathbb{1}$, the matrix A has an inverse, namely $A^\dagger$. Hence the unitary matrices form a subset of $\mathrm{GL}_n(\mathbb{C})$. To see that it is a subgroup, we check that

$$(AB)^\dagger(AB) = B^\dagger A^\dagger AB = \mathbb{1} \qquad \text{and} \qquad (A^{-1})^\dagger A^{-1} = AA^\dagger = \mathbb{1}\,.$$

We can thus define the **unitary group** $\mathrm{U}(n)$ to be the subgroup of $\mathrm{GL}_n(\mathbb{C})$ consisting of unitary matrices. We also define the **special unitary group** $\mathrm{SU}(n)$ to be the subgroup of unitary matrices with determinant one.

A matrix $A \in M_n(\mathbb{R})$ is said to be **orthogonal** if $A^{\mathrm{tr}}A = \mathbb{1}$, where A^{tr} denotes the transpose of A, with matrix entries

$$A^{\mathrm{tr}}_{ij} = A_{ji}\,.$$

We see that orthogonal matrices have inverses, since $A^{\mathrm{tr}}A = \mathbb{1}$ implies $A^{-1} = A^{\mathrm{tr}}$. We can thus define the **orthogonal group** $O(n)$ to be the subgroup of $\mathrm{GL}_n(\mathbb{R})$ consisting of all orthogonal matrices. We also define the **special orthogonal group** $\mathrm{SO}(n)$ to be the subgroup of orthogonal matrices with determinant one. We have already encountered $\mathrm{SO}(2)$ and $\mathrm{SO}(3)$ in the previous section, so we know that they consist of all rigid origin-preserving rotations of the corresponding Euclidean spaces.

There are many more examples of Lie groups, with plentiful applications. We may encounter some other examples later in the course.

2.2.2 Topological properties

For now, we will discuss only three simple topological properties which a matrix Lie group may have. They are compactness, connectedness, and simple-connectedness. The first two you have encountered already.

Definition 2.2.2. *A matrix Lie group G is said to be **compact** if it is a compact set as a subset of $M_n(\mathbb{C})$.*

Recall that since $M_n(\mathbb{C})$ is topologically equivalent to the Euclidean space $\mathbb{C}^{n^2}$, a subset of $M_n(\mathbb{C})$ is compact if and only if it is closed and bounded. Note that closure is now required inside $M_n(\mathbb{C})$; the property of being a matrix Lie group only required closure inside $\mathrm{GL}_n(\mathbb{C})$. Taken together, this means that compactness is equivalent to two conditions: there must be a universal constant which bounds the matrix entries of all elements in G, and limits of sequences of matrices in G must lie in G.

In the exercises, you will show that unitary and orthogonal matrices have bounded matrix entries. This will establish boundedness. To see closedness, one can express the defining conditions (e.g., $A^\dagger A = \mathbb{1}$) in terms of matrix entries, and show that this property is closed under taking limits (in $M_n(\mathbb{C})$, not just $\mathrm{GL}_n(\mathbb{C})$.) To see that several of our examples above are not compact, consider the following sequence of matrices in $\mathrm{SL}_n(\mathbb{R})$:

$$A_n := \begin{pmatrix} m & & & & \\ & 1/m & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

For connectedness, we will use a slightly different definition from what you might expect. It is called path-connectedness, and it turns out to be equivalent to connectedness (for matrix Lie groups, but not in general.)

Definition 2.2.3. *A matrix Lie group G is said to be* ***connected*** *if for all $A, B \in G$ there exists a continuous path $A(t)$ in G such that $A(0) = A$ and $A(1) = B$.*

Given an element $A \in G$, we say that the **connected component** of A is the set of all $B \in G$ such that a path $B(t)$ exists which satisfies $B(0) = A$ and $B(1) = B$. Note that the property of being connected by a path is transitive. To prove connectedness for a matrix Lie group G, it is thus sufficient to show that every $A \in G$ is connected to the identity $\mathbb{1}$ via a continuous path.

We first show that $\mathrm{GL}_n(\mathbb{C})$ is connected. Fix $A \in \mathrm{GL}_n(\mathbb{C})$ and recall that every matrix is similar to an upper triangular matrix with the eigenvalues on the diagonal. Since A is invertible, it follows that there exists invertible C such that $A = CBC^{-1}$ with

$$B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

with each λ_j is nonzero. Define a path $A(t) = CB(t)C^{-1}$ where $B(t)_{ii} = B_{ii}$ and $B(t)_{ij} = (1-t)B_{ij}$ for $i \neq j$. We now have that $A(1) = CDC^{-1}$ where $D = B(1)$ is a diagonal matrix with entries the λ_j. Since $\mathbb{C}$ is path-connected, we can define for each j a path from λ_j to 1. This gives a path from D to the identity, and thus also a path from A to the identity. To adapt to the case of $\mathrm{SL}_n(\mathbb{C})$, we need to ensure that the second part of the path above maintains the property that $\prod_j \lambda_j = 1$. This is easily done by choosing arbitrary paths (to 1) $\lambda_j(t)$ for $1 \leq j \leq n-1$ and setting

$$\lambda_{n-1}(t) = \frac{1}{\prod_j \lambda_j(t)}. \tag{2.4}$$

For the unitary groups, we use the fact that any unitary matrix has an orthonormal basis of eigenvectors, with eigenvalues having absolute value one. Given $A \in \mathrm{U}(n)$, let $A = UDU^{-1}$ be the corresponding decomposition into a diagonal matrix D (with entries of the form $\lambda_j = e^{i\theta_j}$) conjugated by some other unitary U. We then define a path $A(t) = UD(t)U^{-1}$ which will send each diagonal element of D to 1 via the path $\lambda_j(t) = e^{i(1-t)\theta_j}$. To adapt to the case of $\mathrm{SU}(n)$, we again choose arbitrary paths for the first $n-1$ eigenvalues and choose the nth path as in (2.4). You will show that the group $\mathrm{SO}(n)$ is connected in the exercises.

We end with the notion of simple-connectedness.

Definition 2.2.4. *A matrix Lie group G is said to be* ***simply connected*** *if it is connected, and every loop in G can be continuously deformed to a point in G.*

By a loop, we mean a path $A(t)$ from a group element A to itself, i.e. $A(0) = A(1) = A$. One can continuously deform paths to one another by adding an additional parameter s which also varies from 0 to 1. In our case, we are only intereseted in deforming loops to a point. In that case, this means that $A(t, s)$ is a continuous family of loops (i.e., continuous in both t and s) such that $A(0, s) = A(1, s) = A$ for all s, and $A(t, 0) = A(t)$ and $A(t, 1) = A$ all t. In the exercises you will show that SU(2) is topologically equivalent to the complex two-sphere, which is simply connected.

2.2.3 Lie group homomorphisms

The notion of homomorphism and isomorphism is the same as for general topological groups.

Definition 2.2.5. *Let G and H be matrix Lie groups. A map $\Phi : G \to H$ is called a **Lie group homomorphism** if it is a continuous group homomorphism. If it is also bijective with continuous inverse, then it is called a **Lie group isomorphism**.*

If G and H are matrix Lie groups and there exists a Lie group isomorphism $\Phi : G \to H$, then we will say that G and H are isomorphic. We've already seen two simple examples of Lie group homomorphisms, namely the determinant

$$\det : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}^*$$

and the map from $\mathbb{R}$ to SO(2) defined by

$$r \mapsto \begin{pmatrix} \cos r & -\sin r \\ \sin r & \cos r \end{pmatrix}.$$

These are both easily checked to be continuous homomorphisms.

Let's now consider a more complicated and interesting example, concerning SU(2) and SO(3). Recall that the group SO(3) consists of all rigid rotations of $\mathbb{R}^3$; equivalently, it is the group of inner-product-preserving transformations of $\mathbb{R}^3$ which have determinant one. For this example, we will think about $\mathbb{R}^3$ in a slightly different way. Specifically, we will identify $\mathbb{R}^3$ with the *real* vector space

$$V := \{X \in M_2(\mathbb{C}) : X^\dagger = X \text{ and } \mathrm{Tr}[X] = 0\}.$$

The identification is given by the map

$$(x_1, x_2, x_3) \longmapsto \begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix}.$$

You should check that the matrices of the form given on the right hand side of this identification are precisely all of the matrices in V. It's also straightforward to check that this is an inner-product-preserving identification, where we take the usual inner product on $\mathbb{R}^3$ and the inner product

$$\langle X_1 | X_2 \rangle = \frac{1}{2}\,\mathrm{Tr}[X_1 X_2]$$

on V. Now we define an action of SU(2) on V by defining, for each $U \in \mathrm{SU}(2)$,

$$\begin{aligned} \Phi_U &: V \to V \\ &X \mapsto UXU^{-1}. \end{aligned}$$

One easily checks that UXU^{-1} is again in V whenever $U \in \mathrm{SU}(2)$, and that $\Phi_{U_1U_2} = \Phi_{U_1}\Phi_{U_2}$, so that this is indeed a valid action of SU(2) on V. We also check that

$$\langle UX_1U^{-1}|UX_2U^{-1}\rangle = \langle X_1|X_2\rangle$$

for every $X_1, X_2 \in V$, i.e., that the map Φ_U is inner-product preserving for each $U \in \mathrm{SU}(2)$. Taken together, this means that Φ is a homomorphism from SU(2) to O(3). It's also continuous, and thus a Lie group homomorphism. We checked before that SU(2) is connected; in the homework, you will show that O(3) is not connected and that SO(3) is the connected component of the identity. Recalling that continuous functions map connected spaces to connected spaces, we can write

$$\begin{aligned} \Phi : \mathrm{SU}(2) &\to \mathrm{SO}(3) \\ U &\mapsto \Phi_U \end{aligned}$$

as a Lie group homomorphism.

Proposition 2.2.1. *The map $U \mapsto \Phi_U$ is a two-to-one and onto map of* SU(2) *to* SO(3)*.*

Proof. To show that the map is two-to-one, we compute the kernel. We check that, if $UXU^{-1} = X$ for all X, then U commutes with all elements of V; in fact, it must commute with all of $M_2(\mathbb{C})$. This is equivalent to being a scalar multiple of the identity, and since we're in SU(2) it is equivalent to being $\mathbb{1}$ or $-\mathbb{1}$. Hence Φ is a two-to-one map.

To see that it is onto, first observe that any element of SO(3) can be expressed as a planar rotation by angle θ about some eigenvector X (which we will think of as lying in V). If we diagonalize X, we can write it as

$$X = V \begin{pmatrix} x_1 & 0 \\ 0 & -x_1 \end{pmatrix} V^{-1}$$

for some $V \in \mathrm{U}(2)$. The plane orthogonal to X (again viewed in V) is the space of matrices of the form

$$Y = V \begin{pmatrix} 0 & x_2 + ix_3 \\ x_2 - ix_3 & 0 \end{pmatrix} V^{-1}.$$

Now set

$$U = V \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} V^{-1}$$

and observe that $U \in \mathrm{SU}(2)$, and that $UXU^{-1} = X$. Let's see how U acts on the plane orthogonal to X; this amounts to computing UYU^{-1}, which yields

$$UYU^{-1} = V \begin{pmatrix} 0 & e^{i\theta}(x_2 + ix_3) \\ e^{-i\theta}(x_2 - ix_3) & 0 \end{pmatrix} V^{-1}.$$

Expanding $e^{i\theta} = \cos\theta + i\sin\theta$, it's straightforward to check that the above action corresponds to rotating the (x_2, x_3) plane by angle θ. □

2.2.4 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Prove that a complex matrix is unitary if and only if its column vectors are orthonormal under the standard inner product on $\mathbb{C}^n$. Conclude that the matrix entries of a unitary matrix A satisfy $|A_{jk}| \leq 1$. Prove that if A is a unitary matrix, then $|\det A| = 1$.

2. Prove that a real matrix is orthogonal if and only if its column vectors are orthonormal under the standard inner product on $\mathbb{R}^n$. Conclude that the matrix entries of an orthogonal matrix A satisfy $|A_{jk}| \leq 1$. Prove that if A is orthogonal, then $\det(A) = \pm 1$.

3. Prove that the groups $\mathrm{U}(n)$ and $\mathrm{O}(n)$ are closed as subsets of $\mathrm{GL}_n(\mathbb{C})$, and are thus matrix Lie groups.

4. Show that $\mathrm{SO}(n)$ is path-connected.

5. Show that $\mathrm{O}(n)$ is not connected. How many connected components does it have, and what are they?

6. Show that there is a bijection between $\mathrm{SU}(2)$ and the complex two-sphere, i.e., $\{\alpha, \beta \in \mathbb{C} : |\alpha|^2 + |\beta|^2 = 1\}$, via the map
$$(\alpha, \beta) \longmapsto \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}.$$
Conclude that $\mathrm{SU}(2)$ is simply-connected.

7. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

2.3 Matrix exponential and matrix logarithm

The exponential mapping plays a crucial role in the point of view we will take on the representation theory of Lie groups. In fact, it's an integral part of Lie theory in general. We first define the exponential of a matrix, simply by generalizing the usual power series defining the exponential function.

Definition 2.3.1. *The **exponential** of a matrix $X \in M_n(\mathbb{C})$ is the matrix*
$$\exp(X) = \sum_{m=0}^{\infty} \frac{X^m}{m!}.$$

We need to check that this series actually converges. This can be done using the matrix norm on $M_n(\mathbb{C})$ which you have already encountered:
$$\|X\|_2 = \mathrm{Tr}[X^\dagger X]^{1/2}.$$
This is sometimes called the Hilbert-Schmidt norm. By Cauchy-Schwarz, it satisfies $\|XY\| \leq \|X\|\|Y\|$ for all $X, Y \in M_n(\mathbb{C})$. It's easy to check that a sequence X_m of matrices converges to a matrix X if and only if $\|X_m - X\|$ converges to zero. Note that
$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \leq \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} < \infty$$

by the convergence of the usual exponential function. To show continuity of exp, observe that the partial sums of the series are continuous. The Weierstrass M-test then tells us that the entire series converges uniformly on bounded sets; by the uniform limit theorem, exp is continuous on all bounded sets, and hence on all of $M_n(\mathbb{C})$.

We now list several basic properties of the exponential map, which you will prove in the homework. We will sometimes write $\exp(X)$ as e^X instead; they mean the same thing.

1. $\exp(0) = \mathbb{1}$.
2. $(\exp(X))^\dagger = \exp(X^\dagger)$.
3. $\exp(X)$ is invertible, with inverse $\exp(-X)$.
4. $\exp(a + bX) = \exp(aX)\exp(bX)$ for all $a, b \in \mathbb{C}$.
5. If $XY = YX$ then $e^{X+Y} = e^X e^Y = e^Y e^X$.
6. If C is invertible, then $e^{CXC^{-1}} = Ce^X C^{-1}$.
7. If X is diagonalizable, then so is e^X, and $\det[e^X] = e^{\mathrm{Tr}[X]}$.

We will also need the following property, which is arrived at by differentiating the power series for e^{tX} term-by-term, and entry-by-entry.

Proposition 2.3.1. *Let $X \in M_n(\mathbb{C})$. Then e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X$$

An immediate consequence of the Proposition is that

$$\left.\frac{d}{dt}e^{tX}\right|_{t=0} = X\,.$$

Definition 2.3.2. *For a matrix $A \in M_n(\mathbb{C})$, define* $\log A$ *to be*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A - \mathbb{1})^m}{m}$$

whenever the series converges.

Theorem 2.3.1. *The function* $\log(\cdot)$ *is defined and continuous on*

$$B_n(\mathbb{1}; 1) := \{A \in M_n(\mathbb{C}) : \|A - \mathbb{1}\|_2 < 1\}.$$

For $A \in B_n(\mathbb{1}; 1)$ we have $e^{\log A} = A$. For X satisfying $\|X\| < \log 2$, $\|e^X - \mathbb{1}\| < 1$ and $\log e^X = X$.

Proof. For the first part, we proceed just as with the exponential mapping. Since $\|(A - \mathbb{1})^m\|_2 \leq \|A - \mathbb{1}\|_2^m$, absolute convergence of the series defining the matrix logarithm (for $\|A - \mathbb{1}\| < 1$) follows from absolute convergence of the series defining the usual logarithm (for $|z - 1| < 1$). Convergence and continuity now follows by the same arguments as before.

Now let $A \in B_n(\mathbb{1};1)$, and suppose for the moment that it is diagonalizable, with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then there is an invertible C such that

$$(A - \mathbb{1})^m = C \begin{pmatrix} (\lambda_1 - 1)^m & & 0 \\ & \ddots & \\ 0 & & (\lambda_n - 1)^m \end{pmatrix} C^{-1}.$$

Since $A \in B_n(\mathbb{1};1)$, the eigenvalues λ_j satisfy $|\lambda_j - 1| < 1$ (check this by showing that the operator norm is bounded by the Hilbert-Schmidt norm), and so we can write

$$\log A = C \begin{pmatrix} \log \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \log \lambda_n \end{pmatrix} C^{-1}.$$

By the properties of the exponential given above, we conclude that $e^{\log A} = A$. Finally, we can approximate non-diagonalizable A by a sequence of diagonalizable matrices, and apply the continuity of exp and log.

A very similar argument to the above proves the remaining part of the theorem, this time appealing to the fact that the standard complex logarithm satisfies $\log(\exp v) = v$ for $|v| < \log 2$. □

Here's an application of the matrix logarithm which will be quite useful. As it turns out, there is still much we can say about e^{X+Y} even in the case where X and Y do not commute. One important result in this direction is the Lie Product Formula (also known as the Trotter product formula.) We state it here; for a proof, see [1].

Theorem 2.3.2. *For all $X, Y \in M_n(\mathbb{C})$, we have*

$$e^{X+Y} = \lim_{m\to\infty} \left(e^{X/m} e^{Y/m}\right)^m .$$

2.3.1 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Prove properties (1)-(7) of the exponential mapping.

2. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

2.4 Lie Algebras

2.5 Basic definitions

We begin with an abstract definition of Lie algebras.

Definition 2.5.1. *A finite-dimensional **Lie algebra** is a finite-dimensional vector space $\mathfrak{g}$ together with a map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which:*

1. *is bilinear: $[aX, Y] = [X, aY] = a[X, Y]$ for all scalars a and $X, Y \in \mathfrak{g}$;*

2. *is skew-symmetric:* $[X, Y] = -[Y, X]$ *for all* $X, Y \in \mathfrak{g}$*;*

3. *satisfies the Jacobi identity:*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

for all $X, Y, Z \in \mathfrak{g}$.

If the vector space is over $\mathbb{R}$, then we say that $\mathfrak{g}$ is a real Lie algebra; if it is over $\mathbb{C}$, then we say that $\mathfrak{g}$ is a complex Lie algebra. The operation $[\cdot, \cdot]$ is called the **bracket** of $\mathfrak{g}$. If X and Y satisfy $[X, Y] = 0$, then we say that they commute. The **center** of a Lie algebra $\mathfrak{g}$ is the set of all $X \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $Y \in \mathfrak{g}$. If the center is equal to the entire Lie algebra, then we say that the Lie algebra is **commutative**.

One way to construct Lie algebras is to start with an associative algebra. Recall that an associative algebra (for us) is simply a finite-dimensional vector space with an associative product. Let $\mathcal{A}$ be an associative algebra and let $\mathfrak{g}$ be a subspace of $\mathcal{A}$ such that $XY - YX \in \mathfrak{g}$ for all $X, Y \in \mathfrak{g}$. Then it's easy to check that $\mathfrak{g}$ is a Lie algebra with the bracket

$$[X, Y] = XY - YX\,.$$

Verifying the Jacobi identity is a little bit of work, but is useful to do, because it illustrates that associativity of the algebra is crucial.

Definition 2.5.2. *A* ***Lie subalgebra*** *of a Lie algebra* $\mathfrak{g}$ *is a subspace* $\mathfrak{h}$ *of* $\mathfrak{g}$ *which is closed under the bracket operation.*

We remark that there is an ambiguity in this definition: it is possible to consider a *real* Lie subalgebra of a *complex* Lie algebra. We will emphasize this with more specific language when necessary.

For an example of Lie subalgebras, consider the Lie algebra $M_n(\mathbb{C})$ of all complex matrices with the bracket operation $[X, Y] = XY - YX$, and take the subspace $\mathrm{sl}_n(\mathbb{C})$ of matrices with zero trace. Since the bracket of any pair X, Y satisfies $\mathrm{Tr}[[X, Y]] = \mathrm{Tr}[XY - YX] = \mathrm{Tr}[XY] - \mathrm{Tr}[YX] = 0$, we see that $\mathrm{sl}_n(\mathbb{C})$ is a Lie subalgebra of $M_n(\mathbb{C})$.

One can also build many examples by taking direct sums of Lie algebras.

Definition 2.5.3. *If* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$ *are Lie algebras, the direct sum of* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$ *is the Lie algebra defined on the vector space* $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ *by the bracket*

$$[X_1 \oplus X_2, Y_1 \oplus Y_2] = [X_1, X_2] \oplus [Y_1, Y_2]\,.$$

We may also ask when a Lie algebra $\mathfrak{g}$ is a direct sum of two Lie subalgebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ of $\mathfrak{g}$. This is the case precisely when $\mathfrak{g}$ is equal to their direct sum as a vector space, and $[X_1, X_2] = 0$ for all $X_1 \in \mathfrak{h}_1$ and $X_2 \in \mathfrak{h}_2$.

If we choose a basis $X_1, \ldots, X_N$ for a Lie algebra $\mathfrak{g}$, then the bracket operation can be described by a set of constants c_{jkl}, where

$$[X_j, X_k] = \sum_{l=1}^{N} c_{jkl} X_l\,.$$

The scalars c_{jkl} are called the **structure constants** of $\mathfrak{g}$. They of course depend crucially on the choice of basis.

2.6 Lie algebra homomorphisms

As you might expect, Lie algebra homomorphisms are maps which preserve the bracket operation.

Definition 2.6.1. *Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras. A linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is called a* ***Lie algebra homomorphism*** *if*

$$\phi([X,Y]) = [\phi(X), \phi(Y)]$$

for all $X, Y \in \mathfrak{g}$. If ϕ is bijective, then it is called a ***Lie algebra isomorphism****.*

An important generic example is the **adjoint map**. It is a map which associates to each element X of a Lie algebra $\mathfrak{g}$ the linear map

$$\begin{aligned} \mathrm{ad}_X : \mathfrak{g} &\longrightarrow \mathfrak{g} \\ Y &\longmapsto [X,Y]\,. \end{aligned}$$

This is analogous to how we can let each element of a finite group act on the group by (say) left multiplication. The adjoint map itself is $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, since bracketing by a fixed element is a linear operation. To show that it is a Lie algebra homomorphism, first note that $\mathrm{End}(\mathfrak{g})$ is itself a Lie algebra under the usual bracket. We then need to establish that

$$\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$$

where the left-hand bracket takes place in $\mathfrak{g}$ and the right-hand bracket takes place in $\mathrm{End}(\mathfrak{g})$. This amounts to showing that

$$\mathrm{ad}_{[X,Y]}(Z) = [[X,Y],Z] = -[Z,[X,Y]]$$

is equal to

$$[\mathrm{ad}_X, \mathrm{ad}_Y](Z) = (\mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X)(Z) = [X,[Y,Z]] - [Y,[X,Z]] = [X,[Y,Z]] + [Y,[Z,X]]\,.$$

But this is precisely the Jacobi identity.

2.7 The Lie Algebra of a Lie group

We now connect Lie algebras to Lie groups. This will later enable us to study the representations of both simultaneously.

Definition 2.7.1. *Let G be a matrix Lie group. The* ***Lie algebra of*** *G, denoted $\mathfrak{g}$, is the set of all matrices X such that e^{tX} is in G for all real numbers t, with the bracket operation $[X,Y] = XY - YX$.*

We must now check that the set defined above is actually a Lie algebra under the given bracket operation. We will prove this in a few steps. Throughout, it is helpful to recall that e^{tX} defined a smooth curve in $M_n(\mathbb{C})$, whose derivative at $t = 0$ is X.

Theorem 2.7.1. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$, and $X, Y \in \mathfrak{g}$. Then*

1. *$AXA^{-1} \in \mathfrak{g}$ for all $A \in G$.*
2. *$sX \in \mathfrak{g}$ for all real s.*
3. *$X + Y \in \mathfrak{g}$.*

4. $XY - YX \in \mathfrak{g}$.

Proof. For the first claim, we have $e^{tAXA^{-1}} = Ae^{tX}A^{-1} \in G$ for all real t, by the properties of the exponential map we proved previously. The second claim is clear from the definition. For the third claim, recall the Lie product formula

$$e^{t(X+Y)} = \lim_{m\to\infty} \left(e^{tX/m}e^{tY/m}\right)^m .$$

Each term of the sequence on the right is clearly in G, and by closure of G, so is the limit. For the fourth claim, we use the product rule for smooth matrix-valued functions:

$$\frac{d}{dt}[A(t)B(t)] = \frac{dA}{dt}B(t) + A(t)\frac{dB}{dt} .$$

We compute the following.

$$\frac{d}{dt}\left(e^{tX}Ye^{-tX}\right)\Big|_{t=0} = Xe^{tX}Ye^{-tX} - e^{tX}YXe^{-tX}\big|_{t=0} = XY - YX .$$

By the above and the first claim, $XY - YX$ is the derivative of a smooth curve in $\mathfrak{g}$. Since $\mathfrak{g}$ is a subspace of $M_n(\mathbb{C})$, it is topologically closed. By the limit definition of the derivative, it follows that $XY - YX$ is itself in $\mathfrak{g}$. □

One nice consequence of the last calculation in the proof is the following. We have that, for $X, Y \in M_n(\mathbb{C})$, the commutator is given by

$$[X, Y] = \frac{d}{dt}\left(\frac{d}{ds}e^{tX}e^{sY}e^{-tX}\right)\Big|_{s,t=0}$$

If G is commutative, then the term inside the derivatives simplifies to e^{sY}, and the entire right-hand side then evaluates to zero. Hence, if G is a commutative matrix Lie group, then its Lie algebra $\mathfrak{g}$ is a commutative Lie algebra.

2.7.1 Examples

We now compute the Lie algebras of the basic matrix Lie groups we've been studying so far.

Proposition 2.7.1. *The Lie algebra of* $\mathrm{GL}_n(\mathbb{C})$ *is* $M_n(\mathbb{C})$*, and the Lie algebra of* $\mathrm{GL}_n(\mathbb{R})$ *is* $M_n(\mathbb{R})$*. The Lie algebra of* $\mathrm{SL}_n(\mathbb{C})$ *is the set of traceless* $n \times n$ *complex matrices, and the Lie algebra of* $\mathrm{SL}_n(\mathbb{R})$ *is the set of traceless* $n \times n$ *real matrices.*

We will denote the Lie algebra of $\mathrm{GL}_n(\mathbb{C})$ by $\mathrm{gl}_n(\mathbb{C})$, and the Lie algebra of $\mathrm{SL}_n(\mathbb{C})$ by $\mathrm{sl}_n(\mathbb{C})$, and likewise for the real case.

Proof. First, note that e^{tX} is invertible (with inverse e^{-tX}), for any matrix X. Moreover, if X is real then so is e^{tX}. On the other hand, if X is a complex matrix and e^{tX} is real for every t, then

$$X = \frac{d}{dt}e^{tX}\Big|_{t=0}$$

must also be real.

For the second part, note that $\det(e^{tX}) = e^{t\,\mathrm{Tr}[X]}$ for all X. This is obviously true for diagonalizable X, and can be extended to arbitrary X by using the fact that diagonalizable matrices are dense. It follows that, if X has trace zero, then e^{tX} has determinant one for all t, i.e., $X \in \mathrm{sl}_n(\mathbb{C})$. On the other hand, if X satisfies $\det(e^{tX}) = 1$ for all real t, then

$$\mathrm{Tr}[X] = \frac{d}{dt} e^{t\,\mathrm{Tr}[X]}\Big|_{t=0} = \frac{d}{dt}\det(e^{tX})\Big|_{t=0} = 0\,.$$

The same argument works for the case of $\mathrm{SL}_n(\mathbb{R})$. □

Recall that a complex matrix X is called Hermitian if $X^\dagger = X$. We say that a complex matrix X is *skew-Hermitian* if it satisfies $X^\dagger = -X$.

Proposition 2.7.2. *The Lie algebra of* $\mathrm{U}(n)$ *consists of all skew-Hermitian complex matrices, and the Lie algebra of* $\mathrm{SU}(n)$ *consists of all skew-Hermitian matrices with trace zero.*

We will denote the Lie algebra of $\mathrm{U}(n)$ by $\mathrm{u}(n)$, and the Lie algebra of $\mathrm{SU}(n)$ by $\mathrm{su}(n)$.

Proof. For a given real t, the matrix e^{tX} is unitary if and only if

$$(e^{tX})^{-1} = (e^{tX})^\dagger\,.$$

The left-hand side is e^{-tX}, and it was checked in the homework that the right-hand side is $e^{tX^\dagger}$. By taking the derivative of both sides, we see that that e^{tX} is unitary (for all real t) if and only if X is skew-Hermitian. Just as before, requiring that the group elements have determinant one is equivalent to requiring that the Lie algebra elements have trace zero. □

Proposition 2.7.3. *The Lie algebra of* $\mathrm{O}(n)$ *consists of all skew-Hermitian real matrices, and the Lie algebra of* $\mathrm{SO}(n)$ *is equal to the Lie algebra of* $\mathrm{O}(n)$

We will denote the Lie algebra of $\mathrm{O}(n)$ by $\mathrm{o}(n)$, and the Lie algebra of $\mathrm{SO}(n)$ by $\mathrm{so}(n)$ (and, as the proposition says, $\mathrm{o}(n) = \mathrm{so}(n)$).

Proof. The proof is essentially the same as in the unitary case. A real matrix e^{tX} is orthogonal if and only if e^{-tX} is equal to its transpose, which in turn is true if and only if $X^{\mathrm{tr}} = -X$. Within the space of real skew-Hermitian matrices, $X_{ii} = -(X^{\mathrm{tr}})_{ii} = -X_{ii}$, which implies that all the diagonal entries are zero; in particular, $\mathrm{Tr}[X] = 0$. The additional condition of determinant one in $\mathrm{SO}(n)$ thus does not change the Lie algebra. □

It will be convenient to think about low-dimensional examples, and for those examples a (convenient) basis can be illuminating. We will use the following basis for su(2):

$$E_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \qquad E_2 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad E_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\,.$$

This basis satisfies commutation relations $[E_1, E_2] = E_3$, $[E_2, E_3] = E_1$, $[E_3, E_1] = E_2$.

We will also use the following basis for so(3):

$$F_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad F_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad F_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,.$$

This basis satisfies commutation relations $[F_1, F_2] = F_3$, $[F_2, F_3] = F_1$, and $[F_3, F_1] = F_2$.

2.7.2 Relationship between Lie group and Lie algebra homomorphisms

We now see how to construct a Lie algebra homomorphism from a homomorphism of the corresponding Lie groups.

Theorem 2.7.2. *Let G and H be matrix Lie groups, and suppose $\Phi : G \to H$ is a Lie group homomorphism. Then there exists a unique linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ satisfying $\Phi(e^X) = e^{\phi(x)}$ for all $X \in \mathfrak{g}$. It also satisfies the following properties.*

1. *$\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$ for all $X \in \mathfrak{g}$ and all $A \in G$;*
2. *$\phi([X,Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$;*
3. *$\phi(X) = \frac{d}{dt}\Phi(e^{tX})|_{t=0}$ for all $X \in \mathfrak{g}$.*

Proof. Let $X \in \mathfrak{g}$. The set $\{e^{tX} : t \in \mathbb{R}\}$ is clearly a subgroup of G; for obvious reasons, such subgroups are called *one-parameter subgroups.* We claim that the image of this subgroup of G under the map Φ is another one-parameter subgroup; more precisely, we claim that there exists Z such that $\Phi(e^{tX}) = e^{tZ}$. This may seem obvious, but it does in fact require proof, which makes use of the matrix logarithm (see Theorem 2.14 in [1].) One consequence is that we can compute Z by

$$Z = \frac{d}{dt}\Phi(e^{tX})\bigg|_{t=0} .$$

We set $\phi(X) = Z$, which satisfies the third property. By setting $t = 1$ we also see that $\Phi(e^X) = e^Z$ for all $X \in \mathfrak{g}$. We now need to check properties one and two, and verify linearity and uniqueness.

For linearity, note that $\Phi(e^{t(sX)}) = e^{t(sZ}$ which implies $\phi(sX) = sZ = s\phi(X)$. Given X and Y in $\mathfrak{g}$, we use the Lie product formula (twice) and continuity of Φ to compute

$$\begin{aligned} e^{t\phi(X+Y)} &= \Phi(e^{t(X+Y)}) = \Phi\left(\lim_{m\to\infty}(e^{tX/m}e^{tY/m})^m\right) \\ &= \lim_{m\to\infty}\left(\Phi(e^{tX/m})\Phi(e^{tY/m})\right)^m = \lim_{m\to\infty}\left(e^{t\phi(X)/m}e^{t\phi(Y)/m}\right)^m \\ &= e^{t(\phi(X)+\phi(Y))} \end{aligned}$$

Differentiating both sides at $t = 0$ yields $\phi(X+Y) = \phi(X) + \phi(Y)$, which establishes linearity of ϕ.

To show uniqueness, suppose that there were another linear map $\phi' : \mathfrak{g} \to \mathfrak{h}$ with the property that $\Phi(e^X) = e^{\phi'(X)}$ is true for all $X \in \mathfrak{g}$. By linearity, it would follow that $e^{t\phi(X)} = e^{t\phi'(X)}$ for all t, and differentiating both sides at $t = 0$ yields $\phi(X) = \phi'(X)$.

To verify property 1, note that, for any $A \in G$,

$$e^{t\phi(AXA^{-1})} = \Phi(e^{tAXA^{-1}}) = \Phi(A)\Phi(e^{tX})\Phi(A)^{-1} = \Phi(A)e^{t\phi(X)}\Phi(A)^{-1} .$$

Now again we differentiate both sides at $t = 0$.

To verify property 2, we pick $X, Y \in \mathfrak{g}$, and recall that the commutator can be defined via the differential, so that

$$\begin{aligned} \phi([X,Y]) &= \phi\left(\frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0}\right) = \frac{d}{dt}\phi(e^{tX}Ye^{-tX})\bigg|_{t=0} = \frac{d}{dt}\Phi(e^{tX})\phi(Y)\Phi(e^{-tX})\bigg|_{t=0} \\ &= \frac{d}{dt}e^{\phi(tX)}\phi(Y)e^{-\phi(tX)}\bigg|_{t=0} = [\phi(X), \phi(Y)] , \end{aligned}$$

where we used property 1 in the third step. □

For an example of this theorem, recall the "double-cover" Lie group homomorphism $\Phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$. To compute the corresponding Lie algebra homomorphism $\phi : \mathrm{su}(2) \to \mathrm{so}(3)$, we use the third property. Let $X \in \mathrm{su}(2)$ and $Y \in V$. Recalling the action of SU(2) on V, we have

$$\frac{d}{dt}\Phi(e^{tX})Y\bigg|_{t=0} = \frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0} = [X, Y]\,.$$

So the corresponding action of su(2) on V is the bracket action: X acts on V by $Y \mapsto [X, Y]$. You will calculate what this action is explicitly in the homework.

2.7.3 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Let $\Phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ be the "double-cover" matrix Lie group homomorphism discussed in the lecture. Calculate the images of the basis elements E_1, E_2, E_3 of su(2) under the induced Lie algebra homomorphism $\phi : \mathrm{su}(2) \to \mathrm{so}(3)$.

2. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

2.8 Representations of Lie Groups and Lie Algebras

We now define notions of representations for both Lie groups and Lie algebras, and discuss how the two are connected. First, recall that if V is a finite-dimensional real or complex vector space, then (given a basis choice for V), the group $\mathrm{GL}(V)$ is a matrix Lie group. A basis choice for V also gives the algebra $\mathrm{gl}(V) := \mathrm{End}(V)$ of all linear operators on V the structure of a Lie algebra, with the standard bracket $[X, Y] = XY - YX$. It's not hard to check that this basis choice is largely irrelevant, since different basis choices yield isomorphic Lie groups and isomorphic Lie algebras. This allows us to make the following definitions.

Definition 2.8.1. *A representation of a matrix Lie group G is a Lie group homomorphism*

$$\Pi : G \to \mathrm{GL}(V)$$

for some finite-dimensional real or complex vector space V. A representation of a real or complex Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism

$$\pi : \mathfrak{g} \to \mathrm{gl}(V)$$

for some finite-dimensional real or complex vector space V.

We say that the representation in question is real (resp. complex) if the relevant vector space V is defined over the real (resp., complex) numbers. Just as before, we will frequently think of each representation as defining an action of the Lie group (or of the Lie algebra) on the vector space V, and will use the corresponding terminology to talk about group elements (or algebra elements) as *acting* on vectors in V. Of course, in one case the action will be the left-action, while in the other case the action will be the bracket action.

Many representation-theoretic notions we encountered in the setting of finite groups and general compact groups carry over naturally and immediately to the setting of Lie groups and Lie algebras. Some of these (using notation as in Definition 2.8.1) are:

- **faithful**: a Lie group (or Lie algebra) representation is called faithful if it is a one-to-one (i.e., injective) homomorphism.

- **matrix representation**: a choice of basis for V turns any Lie group (or Lie algebra) representation into a matrix representation, whereby each (group or algebra) element is mapped to a matrix.

- **invariant, nontrivial, irreducible**: given a choice of Lie representation, a subspace W of V is called invariant if it is preserved under the action of the representation. It is called nontrivial if it is neither equal to the 0 subspace nor to V itself. The representation on V is said to be irreducible if it contains no nontrivial invariant subspaces.

- **intertwiners, isomorphism**: Given two representations (Π, V) and (Σ, W) of a matrix Lie group G, a linear map $f : V \to W$ satisfying $f\Pi(g) = \Sigma(g)f$ for all $g \in G$ is called an intertwiner. If f is invertible, then we say that Π and Σ are isomorphic as representations. The notions of intertwiner and representation isomorphism are defined analogously for Lie algebra representations.

- **direct sums**: Given representations (Π, V) and (Σ, W) of a matrix Lie group G, we define their direct sum $(\Pi \oplus \Sigma, V \oplus W)$ by setting $[\Pi \oplus \Sigma](g) = \Pi(g) \oplus \Sigma(g)$. Direct sums for Lie algebra representations are defined similarly.

Next, we record a straightforward consequence of Theorem 2.7.2 to representations. This fact will be crucial in developing a connection between the representation theory of a matrix Lie group G and the representation theory of its Lie algebra $\mathfrak{g}$.

Corollary 2.8.1. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and Π a representation of G. Then there is a unique representation π of $\mathfrak{g}$ such that $\Pi(e^X) = e^{\pi(X)}$ for all $X \in \mathfrak{g}$. It is computed by*

$$\pi(X) = \frac{d}{dt}\Pi(e^{tX})\bigg|_{t=0}$$

and satisfies $\pi(AXA^{-1}) = \Pi(A)\pi(X)\Pi(A)^{-1}$ for all $X \in \mathfrak{g}$ and all $A \in G$.

Returning to previously-encountered concepts from representation theory, we now consider the **tensor product**. Given a representation (Π, V) of a matrix Lie group G, and a representation (Σ, W) of a matrix Lie group H, we define a representation $(\Pi \otimes \Sigma, V \otimes W)$ of the matrix Lie group $G \times H$ by setting $[\Pi \otimes \Sigma](g, h) = \Pi(g) \otimes \Sigma(h)$. By Corollary 2.8.1, the representation $\Pi \otimes \Sigma$ of $G \times H$ corresponds to a unique representation (which we will denote by $\pi \otimes \sigma$) of the Lie algebra of $G \times H$, which is $\mathfrak{g} \oplus \mathfrak{h}$ (you will prove this in the Exercises). To compute this, first check that the derivative of the tensor product satisfies the following product rule:

$$\frac{d}{dt}(u(t) \otimes v(t)) = \frac{du}{dt} \otimes v(t) + u(t) \otimes \frac{dv}{dt}\,.$$

We then compute

$$\begin{aligned}
[\pi \otimes \sigma](X \oplus Y) &= \frac{d}{dt}[\Pi \otimes \Sigma]\big(e^{t(X \oplus Y)}\big)\Big|_{t=0} \\
&= \frac{d}{dt}[\Pi \otimes \Sigma]\big(e^{tX}, e^{tY}\big)\Big|_{t=0} \\
&= \frac{d}{dt}\Pi\big(e^{tX}\big) \otimes \Sigma\big(e^{tY}\big)\Big|_{t=0} \\
&= \left(\frac{d}{dt}\Pi\big(e^{tX}\big)\right) \otimes \Sigma\big(e^{tY}\big) + \Pi\big(e^{tX}\big) \otimes \left(\frac{d}{dt}\Sigma\big(e^{tY}\big)\right)\Bigg|_{t=0} \\
&= \pi(X) \otimes \mathbb{1} + \mathbb{1} \otimes \sigma(Y)\,.
\end{aligned}$$

Motivated by this calculation, we define, in general, the tensor product $\pi \otimes \sigma$ of two Lie algebra representations π of $\mathfrak{g}$ and σ of $\mathfrak{h}$ to be the representation of $\mathfrak{g} \oplus \mathfrak{h}$ defined by

$$[\pi \otimes \sigma](X \oplus Y) = \pi(X) \otimes \mathbb{1} + \mathbb{1} \otimes \sigma(Y)\,.$$

Why could we not define it to be simply $\pi(X) \otimes \sigma(Y)$?

Following the above, we also define the tensor product of representations of the same group (or algebra). Specifically, given representations Π and Σ of a Lie group G, we define the tensor product representation $\Pi \otimes \Sigma$ of G by $[\Pi \otimes \Sigma](A) = \Pi(A) \otimes \Sigma(A)$. Likewise, given representations π and σ of a Lie algebra $\mathfrak{g}$, we define the tensor product representation $\pi \otimes \sigma$ of $\mathfrak{g}$ by $[\pi \otimes \sigma](X) = \pi(X) \otimes \mathbb{1} + \mathbb{1} \otimes \sigma(X)$.

Before we continue, we'll need the following fact. We only sketch the proof here; for a complete proof, see Corollary 3.47 (and the results referenced therein) of [1].

Theorem 2.8.1. *Let G be a connected matrix Lie group, and $A \in G$. Then $A = e^{X_1} e^{X_2} \cdots e^{X_m}$ for some $X_1, X_2, \ldots, X_m \in \mathfrak{g}$.*

Proof. (Sketch) The proof proceeds as follows.

1. using the properties of the matrix exponential and the matrix logarithm established in previous sections, prove that there exists $0 < \epsilon < \log 2$ such that the following holds for all $B \in V_\epsilon := \exp(\{X \in M_n(\mathbb{C}) : \|X\| < \epsilon\})$: $B \in G$ if and only if $\log B \in \mathfrak{g}$.

2. using the connectedness of G, choose a continuous path $A(t)$ such that $A(0) = \mathbb{1}$ and $A(1) = A$. Using the previous fact, show that there exists $\delta > 0$ such that $|s - t| < \delta$ implies that $A(s)^{-1}A(t) \in V_\epsilon$ and hence there exists $X \in \mathfrak{g}$ such that $A(s)^{-1}A(t) = \exp(X)$.

3. divide $[0, 1]$ into m segments of length δ and expand

$$A = A(0)^{-1}A(1) = A(0)^{-1}A(1/m)A(1/m)^{-1}A(2/m) \cdots A(1)\,.$$

Choosing the appropriate X_j so that $\exp(X_j) = A(j/m)^{-1}A((j+1)/m)$ yields the final result.

□

The above theorem is necessary in the following.

Proposition 2.8.1. *A representation Π of a connected matrix Lie group G is irreducible if and only if the corresponding representation π of the Lie algebra $\mathfrak{g}$ is irreducible.*

Proof. We first assume that Π is irreducible, and let W be a subspace which is invariant under the action of π. By Theorem 2.8.1, any $A \in G$ can be written as $A = e^{X_1} e^{X_2} \cdots e^{X_m}$ for some $X_j \in \mathfrak{g}$. Since W is invariant under $\pi(X_j)$, it will also be invariant under $\exp(\pi(X_j))$. By Corollary 2.8.1, W is also invariant under

$$\Pi(A) = \Pi(e^{X_1} e^{X_2} \cdots e^{X_m}) = \Pi(e^{X_1})\Pi(e^{X_2}) \cdots \Pi(e^{X_m}) = e^{\pi(X_1)} e^{\pi(X_2)} \cdots e^{\pi(X_m)} .$$

By the irreducibility of Π, we conclude that W is trivial, and hence that π is also irreducible.

For the other direction, we begin with the assumption that π is irreducible, and we choose a subspace W which is invariant under the action of Π. For every $X \in \mathfrak{g}$, W is invariant under $\Pi(e^{tX})$, and hence also under

$$\pi(X) = \frac{d}{dt} \Pi(e^{tX}) \bigg|_{t=0} .$$

It follows that Π is irreducible. □

Proposition 2.8.2. *Two representations Π_1, Π_2 of a connected matrix Lie group G are isomorphic if and only if the corresponding representations π_1, π_2 of the Lie algebra $\mathfrak{g}$ are isomorphic.*

Proof. See Exercises. □

Examples.

- **trivial representation.** Every matrix Lie group G has a trivial representation $G \to \mathrm{GL}_1(\mathbb{C})$ defined by $A \mapsto 1$ for all $A \in G$. For Lie algebras, the trivial representation $\mathfrak{g} \to \mathrm{gl}_1(\mathbb{C})$ is defined by $X \mapsto 0$ for all $X \in G$.

- **standard representation.** Every matrix Lie group G has a standard representation, where each element is represented by itself. This representation is simply the inclusion map $G \hookrightarrow \mathrm{GL}_n(\mathbb{C})$. If the Lie group is real, then the standard representation can also be viewed as the inclusion $G \hookrightarrow \mathrm{GL}_n(\mathbb{R})$, yielding a real representation. The standard representation of Lie algebras is defined in the same way.

- **adjoint representation.** Recall the following example of a Lie algebra homomorphism: the adjoint map. It is defined to be the homomorphism $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}) = \mathrm{gl}(\mathfrak{g})$ defined by setting, for each $X \in \mathfrak{g}$,

  $$\begin{aligned} \mathrm{ad}_X : \mathfrak{g} &\to \mathfrak{g} \\ Y &\mapsto [X, Y] . \end{aligned}$$

  The map $\mathrm{ad} : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$ is clearly a Lie algebra representation.

  We can define an analogous representation for matrix Lie groups. If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then for each $A \in G$ we define $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ by setting $\mathrm{Ad}_A(X) = AXA^{-1}$. The map $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ defined by $A \mapsto \mathrm{Ad}_A$ is then a Lie group representation.

  What is the relationship between these two representations? This is easy to compute:

  $$\mathrm{ad}_X(Y) = \frac{d}{dt} \mathrm{Ad}_{e^{tX}}(Y) \bigg|_{t=0} = \frac{d}{dt} e^{tX} Y e^{-tX} \bigg|_{t=0} = [X, Y] .$$

  In other words, ad is precisely the unique Lie algebra representation corresponding to the Lie group representation Ad.

The representations of SU(2) **and** $\mathrm{sl}_2(\mathbb{C})$**.** We now consider a more detailed example. Consider the matrix Lie group SU(2), and let V_m be the vector space of homogeneous polynomials of degree m in two complex variables. Each $U \in \mathrm{SU}(2)$ defines a representation Π_m on V_m by

$$[\Pi_m(U) \cdot f](z) = f(U^{-1}z)$$

where $z = (x, y)$ is the vector consisting of formal (complex) variables. Let's check this directly. Each element $f \in V_m$ has the form

$$f(x, y) = a_0 x^m + a_1 x^{m-1} y + \cdots + a_m y^m ,$$

where the a_j are arbitrary complex coefficients. The corresponding representation of su(2) is given by

$$(\pi_m(X)f)(z) = \frac{d}{dt} f(e^{-tX}z)\bigg|_{t=0} .$$

Letting $z(t) = (x(t), y(t))$ be the curve defined by $z(t) = e^{-tX}z$, we have

$$\pi_M(X)f = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\bigg|_{t=0} .$$

Since $dz/dt|_{t=0} = -Xz$, we get

$$\pi_m(X)f = -\frac{\partial f}{\partial x}(X_{11}x + X_{12}y) - \frac{\partial f}{\partial y}(X_{21}x + X_{22}y) .$$

It is convenient to think about π_m as a representation of $\mathrm{sl}_2(\mathbb{C})$, which is isomorphic to the Lie algebra su(2) with complex coefficients (this is called *complexification*). To check this, note that every traceless complex matrix Z can be written as the sum of two traceless anti-Hermitian matrices, namely $(Z - Z^\dagger)/2$ and $i(Z + Z^\dagger)/2$. This change has essentially no effect on the representation theory. On the basis

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} ,$$

the action of π is now given by

$$\pi_m(H) = -x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} , \qquad \pi_m(X) = -y\frac{\partial}{\partial x} , \qquad \pi_m(Y) = -x\frac{\partial}{\partial y} .$$

The action of these operators on the above basis of V_m is

$$\begin{aligned} \pi_m(H)(x^{m-k}y^k) &= (-m + 2k)x^{m-k}y^k \\ \pi_m(X)(x^{m-k}y^k) &= (-m - k)x^{m-k-1}y^{k+1} \\ \pi_m(Y)(x^{m-k}y^k) &= -kx^{m-k+1}y^{k-1} . \end{aligned}$$

Note that each basis vector is an eigenvector of H, and that X (resp., Y) shifts the vectors in such a way that the eigenvalue is raised (resp., lowered) by 2. Note that the commutation relations for this basis are given by

$$[H, X] = 2X \qquad [H, Y] = -2Y \qquad [X, Y] = H .$$

Theorem 2.8.1. *The representations π_m of $\mathrm{sl}_2(\mathbb{C})$ defined above are irreducible. Moreover, if π is an irreducible complex representation of $\mathrm{sl}_2(\mathbb{C})$ with dimension $m+1$, then it is isomorphic to π_m.*

Since we are running short on time in the course, we leave the proof to [1].

2.8.1 Exercises

Please see the final paragraph of the Introduction for my expectations regarding homework exercises.

1. Let G and H be matrix Lie groups. Prove that the Lie algebra of $G \times H$ is $\mathfrak{g} \oplus \mathfrak{h}$.

2. Prove Proposition 2.8.2: Two representations Π_1, Π_2 of a connected matrix Lie group G are isomorphic if and only if the corresponding representations π_1, π_2 of the Lie algebra $\mathfrak{g}$ are isomorphic.

3. Show that the adjoint and standard representations of the Lie group SO(3) are isomorphic.

4. Consider the usual action of SO(2) on $\mathbb{R}^2$. Prove that this makes $\mathbb{R}^2$ a real, irreducible representation of SO(2). Prove that Schur's Lemma fails in this case, by demonstrating an intertwiner of this representation with itself which is not a scalar multiple of the identity.

5. Please list all of the typos and mistakes that you found in the lecture notes and exercises in this Section. Thanks!

2.9 Representation theory of $\mathrm{sl}_3(\mathbb{C})$

Building on the example of $\mathrm{sl}_2(\mathbb{C})$ from the last section, we will now describe the representations of $\mathrm{sl}_3(\mathbb{C})$. As before, this is equivalent to characterizing the representations of su(3), and hence also the representations of SU(3). While these might seem like a few isolated examples, their representation theory in fact contains all of the basic features of the completely general case, and is thus a very useful model to understand.

We begin by selecting a basis for $\mathrm{sl}_3(\mathbb{C})$:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad Y_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Note that the span of H_1, X_1, Y_1 and the span of H_2, X_2, Y_2 are subalgebras, both isomorphic to $\mathrm{sl}_2(\mathbb{C})$. We thus already know the bracket relations inside those algebras: they are the same as in $\mathrm{sl}_2(\mathbb{C})$. Note also that $[H_1, H_2] = 0$, i.e., H_1 and H_2 commute. The simultaneous eigenvectors (and the corresponding eigenvalue pairs) of H_1 and H_2 play a special role in the representation theory of $\mathrm{sl}_3(\mathbb{C})$.

Definition 2.9.1. *Let (π, V) be a representation of $\mathrm{sl}_3(\mathbb{C})$. An ordered pair $\mu = (m_1, m_2) \in \mathbb{C}^2$ is called a* ***weight*** *for π if there exists nonzero $v \in V$ such that $\pi(H_1)v = m_1 v$ and $\pi(H_2)v =$*

m_2v. Any such vector v is called a ***weight vector****, and the space of all such vectors is called the* ***weight space*** *corresponding to the weight μ. The* ***multiplicity*** *of a weight is the dimension of the corresponding weight space.*

Since π is a complex representation, $\pi(H_1)$ has at least one eigenvalue λ_1. Since π is a homomorphism, $\pi(H_1)$ commutes with $\pi(H_2)$, which means that $\pi(H_2)$ preserves the λ_1-eigenspace of $\pi(H_1)$. The restriction of $\pi(H_2)$ to that subspace has an eigenvalue as well, say λ_2. It follows that every representation has at least one weight (m_1, m_2). As we saw in the previous section, the eigenvalues of H in $\mathrm{sl}_2(\mathbb{C})$ are always integers. Since we can view $\langle H_j, X_j, Y_j\rangle$ as a copy of $\mathrm{sl}_2(\mathbb{C})$, this implies that m_1 and m_2 (being eigenvalues of H_1 and H_2) are integers as well.

Definition 2.9.2. *A nonzero ordered pair $\alpha = (a_1, a_2) \in \mathbb{C}^2$ is called a* ***root*** *if it is a nonzero weight of the adjoint representation.*

Recall that the definition of the adjoint representation states that $\pi(H_j) = ad_{H_j}$. It follows that (a_1, a_2) is a root if and only if there exists a nonzero $Z \in \mathrm{sl}_3(\mathbb{C})$ such that $[H_1, Z] = a_1 Z$ and $[H_2, Z] = a_2 Z$; such a Z is called a **root vector** corresponding to α. By computing all of the commutation relations of the above basis of $\mathrm{sl}_3(\mathbb{C})$, one can show that there are six roots for $\mathrm{sl}_3(\mathbb{C})$, written below as (root, eigenvector) pairs.

$$\begin{array}{ll} ((2,-1), X_1) & ((-2,1), Y_1) \\ ((-1,2), X_2) & ((1,-2), Y_2) \\ ((1,1), X_3) & ((-1,-1), Y_3) \end{array}$$

Lemma 2.9.1. *Let $\alpha = (a_1, a_2)$ be a root and Z_α the corresponding root vector. Let $\mu = (m_1, m_2)$ be a weight for a representation π, with weight vector v. Then*

$$\pi(H_1)\pi(Z_\alpha)v = (m_1 + a_1)\pi(Z_\alpha)v\,,$$
$$\pi(H_2)\pi(Z_\alpha)v = (m_2 + a_2)\pi(Z_\alpha)v\,.$$

Proof. First, we have that $[\pi(H_1), \pi(Z_\alpha)] = \pi([H_1, Z_\alpha]) = a_1\pi(Z_\alpha)$, which implies that

$$\pi(H_1)\pi(Z_\alpha) = \pi(Z_\alpha)\pi(H_1) + a_1\pi(Z_\alpha)\,.$$

Similarly, we also have

$$\pi(H_2)\pi(Z_\alpha) = \pi(Z_\alpha)\pi(H_2) + a_2\pi(Z_\alpha)\,.$$

The above two equations, when applied to v, yield the claim. □

The above should be compared with the structure of $\mathrm{sl}_2(\mathbb{C})$ from the previous section.

Definition 2.9.3. *Let $\alpha_1 = (2, -1)$ and $\alpha_2 = (-1, 2)$, and μ_1, μ_2 two weights of $\mathrm{sl}_3(\mathbb{C})$. Then μ_1 is* ***higher*** *than μ_2 (written $\mu_1 \prec \mu_2$) if*

$$\mu_1 - \mu_2 = a\alpha_1 + b\alpha_2$$

with a, b both nonnegative. A weight μ_0 of a representation π of $\mathrm{sl}_3(\mathbb{C})$ which is higher than all the other weights of π is called ***the highest weight*** *of π.*

We are now ready to state the theorem of the highest weight for $\mathrm{sl}_3(\mathbb{C})$.

Theorem 2.9.1.

1. *Every irreducible representation of* $\mathrm{sl}_3(\mathbb{C})$ *is the direct sum of its weight spaces.*
2. *Every irreducible representation of* $\mathrm{sl}_3(\mathbb{C})$ *has a unique highest weight* μ*. The highest weight is of the form* $\mu = (m_1, m_2)$ *where* m_1 *and* m_2 *are integers.*
3. *Two irreducible representations of* $\mathrm{sl}_3(\mathbb{C})$ *with the same highest weight are isomorphic.*
4. *For every pair* (m_1, m_2) *of non-negative integers, there exists an irreducible representation of* $\mathrm{sl}_3(\mathbb{C})$ *with highest weight* (m_1, m_2)*.*

Since there is only one lecture remaining, we will work out an example instead of talking about the proof (which would likely take up multiple lectures anyway.)

Example: highest weight $(1,1)$**.** Let's see how one can construct the irreducible representation of $\mathrm{sl}_3(\mathbb{C})$ with highest weight $(1,1)$ (which exists, by the theorem.) We begin by identifying the highest weights of two representations we have already seen: the standard representation, and its dual representation.

The standard representation is of course just the identity map $X \mapsto X$. Its weight vectors are thus the simultaneous eigenvectors of H_1 and H_2. These are just the standard basis vectors e_1, e_2, e_3. The corresponding weights are $(1,0)$, $(-1,1)$ and $(0,-1)$. Recalling the definition of highest weight for $\mathrm{sl}_3(\mathbb{C})$, we see that $(1,0)-(-1,1) = (2,-1) = \alpha_1$ and $(1,0)-(0,-1) = (1,1) = \alpha_1 + \alpha_2$. We conclude that the standard representation has highest weight $(1,0)$, and will thus denote it by $\pi_{1,0}$.

The dual representation is the map $X \mapsto -X^{\mathrm{tr}}$. You can check this by starting from the dual representation $X \mapsto (X^{-1})^{\mathrm{tr}}$ we defined for groups and computing the corresponding Lie algebra representation. We can again take the standard basis vectors, and immediately see that they are also the simultaneous eigenvectors of $-H_1^{\mathrm{tr}}$ and $-H_2^{\mathrm{tr}}$, but now with different eigenvalues. The weights are now $(-1,0)$, $(1,-1)$, $(0,1)$. As before, we check that $(0,1)-(-1,0) = (1,1) = \alpha_1+\alpha_2$ and $(0,1)-(1,-1) = (-1,2) = \alpha_2$, meaning that $(0,1)$ is the highest weight of this representation. We will thus denote it by $\pi_{0,1}$.

Now let's take the tensor product of the standard and the dual representation. Why should we do this? Because the weights of tensor products behave in a very nice and intuitive way. Let's check this directly. Define $\pi = \pi_{1,0} \otimes \pi_{0,1}$ and recall that

$$\pi(Z) = \pi_{1,0}(Z) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{0,1}(Z)$$

for any $Z \in \mathrm{sl}_3(\mathbb{C})$. Let $v = e_1 \otimes e_3$ and check that

$$\begin{aligned}
\pi(H_1)v &= [\pi_{1,0}(H_1) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{0,1}(H_1)](e_1 \otimes e_3) = e_1 \otimes e_3 + 0 = v \\
\pi(H_2)v &= [\pi_{1,0}(H_2) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{0,1}(H_2)](e_1 \otimes e_3) = 0 + e_1 \otimes e_3 = v\,.
\end{aligned}$$

It follows that v is a weight vector of π with weight $(1,1)$.

How does this help us build $\pi_{1,1}$, the irreducible representation of $\mathrm{sl}_3(\mathbb{C})$ with highest weight $(1,1)$? Notice what Lemma 2.9.1 and Theorem 2.9.1 say, when put together: once we have a weight vector (such as v above), we can "discover" all the other weight vectors by acting on v by the images of the X_1, X_2, X_3 and Y_1, Y_2, Y_3. As a matter of fact, since the X_j only "raise", and we're starting from the highest weight, we need only consider the action of the Y_j. Since $Y_3 = -[Y_1, Y_2]$, considering the action of Y_1 and Y_2 is sufficient. These actions, in the standard representation, are

$$\begin{aligned}
&Y_1 e_1 = e_2; \qquad && Y_1 e_2 = 0; \qquad && Y_1 e_3 = 0 \\
&Y_1 e_1 = 0; \qquad && Y_1 e_2 = e_3; \qquad && Y_2 e_3 = 0\,.
\end{aligned}$$

For the dual representation, it's convenient to use a slightly different basis: $f_1 = e_3$, $f_2 = -e_2$, $f_3 = e_1$. The action is then given by

$$\begin{aligned} &\pi_{0,1}(Y_1)f_1 = 0; && \pi_{0,1}(Y_1)f_2 = f_3; && \pi_{0,1}(Y_1)f_3 = 0 \\ &\pi_{0,1}(Y_2)f_1 = f_2; && \pi_{0,1}(Y_2)f_2 = 0; && \pi_{0,1}(Y_2)f_3 = 0\,. \end{aligned}$$

To generate the other weight vectors, starting with $v = e_1 \otimes f_1$, we will need to repeatedly apply the operators

$$\pi_{1,1}(Y_1) = Y_1 \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{0,1}(Y_1) \qquad \text{and} \qquad \pi_{1,1}(Y_2) = Y_2 \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{0,1}(Y_2)\,.$$

From this we can generate the entire space, and check that it is spanned by the following vectors: $e_1 \otimes f_1, e_2 \otimes f_1, e_1 \otimes f_2, e_3 \otimes f_1 + e_2 \otimes f_2, e_2 \otimes f_2 + e_1 \otimes f_3, e_2 \otimes f_3, e_3 \otimes f_2$, and $e_3 \otimes f_3$. The dimension of $\pi_{1,1}$ is thus eight.

Bibliography

[1] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer, 2003.

[2] Heinrich Maschke. Ueber den arithmetischen Charakter der Coefficienten der Substitutionen endlicher linearer Substitutionsgruppen. *Math. Ann.*, 50(4):492–498, 1898.

[3] Issai Schur. *Neue Begrndung der Theorie der Gruppencharaktere" (New foundation for the theory of group characters), Sitzungsberichte der Kniglich Preuischen Akademie der Wissenschaften zu Berlin*, pages 406–432. 1905.

[4] J.P. Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer New York, 1996.