MAT 449: Representation theory

These lecture notes are in progress, and will change without notice. Please send your comments/corrections to `jl0793@princeton.edu`. Last edited: April 26, 2023

Contents

1 Introduction

Lecture 1 starts here

When we first meet groups, we typically think of them as acting on sets. We recall that a (left) action of a group G on a set X is by definition a map

$$G \times X \to X \tag{1.0.1}$$
$$(g, x) \mapsto g \cdot x \tag{1.0.2}$$

such that $g \cdot (h \cdot x) = (gh) \cdot x$ for all $g, h \in G$ and $x \in X$ and $e \cdot x$ for all $x \in X$, where $e \in G$ is the identity element. Equivalently, writing $\mathrm{Sym}(X)$ for the group of all bijections $X \to X$, an action is defined by a homomorphism $G \to \mathrm{Sym}(X)$. We say X is a G-set. Some highlights of group actions include the following:

- Every group admits a faithful action on some set. (Cayley's theorem)
- Every G-set is a disjoint union of *transitive* G-sets.
- Every transitive G-set is isomorphic to G/H, the set of left cosets under some subgroup $H \leq G$ with G action given by left multiplication. There is an isomorphism of G-sets $G/H \simeq G/H'$ if and only if $H' = gHg^{-1}$ for some $g \in G$.

It is easy to motivate why G-sets are reasonable objects to study: after all, groups were invented by Galois precisely to study how Galois groups act on roots of polynomial. In this course we will *linearize* the situation, and consider actions of groups on vector spaces by linear transformations; this is called *representation theory*. Why is this a good thing to do?

- Linear algebra is the only thing in mathematics we truly understand, so linearization is always a good idea.

- This gives rise to a very complete theory for finite groups, which generalizes well to compact Lie groups, algebras and other algebraic structures.
- This has intrinsic applications to group theory, whose statement does not involve representation theory. For example, we will prove later in the course Burnside's theorem that every finite group of order $p^a q^b$ with p, q prime is solvable using techniques from representation theory.
- There are plenty of applications to and connections with number theory, algebraic geometry, combinatorics, physics, chemistry, ...
- Any time a group G acts on an object X (like a manifold, variety, other group, ...), we can try to linearize the situation (taking tangent spaces, cohomology, ...) and study this linear problem instead. It turns out that this perspective is extremely profitable, see e.g. the study of Galois representations.

It is also just very fun!

2 Basic definitions of representation theory

2.1 Linear algebra preliminaries

Convention 2.1. *All vector spaces considered in this course will be finite-dimensional unless explicitly stated otherwise.*

Let F be a field and V be a vector space over F. Define

$$\mathrm{End}(V) \coloneqq \mathrm{Hom}(V,V) = \{F\text{-linear maps } V \to V\}, \tag{2.1.1}$$
$$\mathrm{GL}(V) \coloneqq \{F\text{-linear bijections } V \to V\}. \tag{2.1.2}$$

After choosing bases (equivalently, choosing an isomorphism $V \simeq F^n$), we can represent linear maps using matrices. For every integer $n \geq 1$ define

$$\mathrm{GL}_n(F) \coloneqq \{\text{invertible } n \times n\text{-matrices over } F\} \tag{2.1.3}$$
$$= \{A \in \mathrm{Mat}_n(F) \mid \det A \neq 0\}. \tag{2.1.4}$$

Let $\{e_1, \dots, e_n\}$ be a basis of V. Given $\phi \in \mathrm{End}(V)$, define the matrix $A_\phi = (a_{ij})_{1 \leq i,j \leq n} \in \mathrm{Mat}_n(F)$ via $\phi(e_j) = \sum_i a_{ij} e_i$.

Proposition 2.2.

1. *The map $\phi \mapsto A_\phi$ induces an isomorphism $\mathrm{GL}(V) \simeq \mathrm{GL}_n(F)$.*
2. *If $\phi \mapsto B_\phi$ is the isomorphism $\mathrm{GL}(V) \simeq \mathrm{GL}_n(F)$ defined using a different basis of V, then there exists a matrix $X \in \mathrm{GL}_n(F)$ such that $B_\phi = X A_\phi X^{-1}$ for all $\phi \in \mathrm{GL}(V)$.*

Proof. Linear algebra. (Exercise) □

Conclusion: every linear map $V \to V$ gives rise to a matrix once a basis is chosen, and changing the basis replaces the matrix by a conjugate.

2.2 Three equivalent definitions

We will now define a representation, in three equivalent ways. This parallels the three different ways of defining a G-action on a (finite) set X: as a map $G \times X \to X$ satisfying some properties, as a homomorphism $G \to \mathrm{Sym}(X)$, or as a homomorphism $G \to S_n$ (once a bijection $X \simeq \{1, 2, \ldots, n\}$ is chosen).

Definition 2.3. *A linear action of a group G on an F-vector space V is an action*

$$G \times V \to V \tag{2.2.1}$$
$$(g, v) \mapsto g \cdot v \tag{2.2.2}$$

such that for every $g \in G$ the induced map $g \cdot (-) \colon V \to V$ is F-linear. In other words, we require that $g \cdot (\lambda v + \mu w) = \lambda(g \cdot v) + \mu(g \cdot w)$ for all $g \in G$, $v, w \in V$ and $\lambda, \mu \in F$. We also say that G acts linearly on V, or that V is a G-module.

Definition 2.4. *A representation of a group G on a vector space V is a homomorphism*

$$\rho \colon G \to \mathrm{GL}(V). \tag{2.2.3}$$

We also say that V is a G-representation or G-rep for short, with the homomorphism ρ being implicitly understood.

Definition 2.5. *A matrix representation of a group G is a homomorphism*

$$R \colon G \to \mathrm{GL}_n(F). \tag{2.2.4}$$

We now relate these three definitions. Given a linear action of G on V, the assignment $\rho(g)(v) := g \cdot v$ defines a homomorphism $\rho \colon G \to \mathrm{GL}(V)$ and hence a representation of G on V. Conversely, a representation $\rho \colon G \to \mathrm{GL}(V)$ defines a linear G-action on V via $g \cdot v := \rho(g)(v)$. Given a representation $\rho \colon G \to \mathrm{GL}(V)$ and a basis of V, we obtain a matrix representation $R \colon G \to \mathrm{GL}_n(F)$ using Proposition 2.2. Conversely, a matrix representation gives rise to a representation of G on the vector space $V := F^n$, using the 'natural' isomorphism $\mathrm{GL}(F^n) \simeq \mathrm{GL}_n(F)$.

Conclusion: the following notions are the same

$$\{\text{Linear actions}\} = \{\text{Representations}\} \tag{2.2.5}$$
$$\{\text{Matrix representations}\} = \{\text{Representations } + \text{ choice of basis of } V\} \tag{2.2.6}$$

We will freely switch between linear actions and representations without mention, and view representations as matrix representations once we have chosen a basis of V.

2.3 Morphisms of representations

Definition 2.6. *A G-homomorphism (or simply morphism) between G-representations V and W is a linear map $f \colon V \to W$ such that $\phi(g \cdot v) = g \cdot \phi(v)$ for all $g \in G$ and $v \in V$. This is also called a G-linear map or G-equivariant map. Write*

$$\mathrm{Hom}_G(V, W) \subset \mathrm{Hom}(V, W) \tag{2.3.1}$$

for the subset of all G-homomorphisms between V and W. This is a linear subspace of $\mathrm{Hom}(V, W)$.

Definition 2.7. *A G-homomorphism ϕ between two representations is called a G-isomorphism (or simply an isomorphism) if ϕ is bijective. We say two representations are isomorphic or equivalent if there exists an isomorphism between them.*

Proposition 2.8. *Let $R, R' \colon G \to \mathrm{GL}_n(F)$ be two matrix representations. Then R and R' are equivalent as representations if and only if there exists a matrix $X \in \mathrm{GL}_n(F)$ such that $\rho'(g) = X\rho(g)X^{-1}$ for all $g \in G$.*

Proof. Exercise. (Problem set 1) ☐

2.4 Properties of representations

Definition 2.9. *Let $\rho\colon G \to \mathrm{GL}(V)$ be a representation.*

- *The dimension of V is called the degree or dimension of the representation ρ.*
- *We say ρ is faithful if ρ is injective, equivalently the action of G on V is faithful.*
- *We say $W \subset V$ is a subrepresentation of V if W is stable under the G-action, i.e. $g \cdot w \in W$ for all $w \in W$. In that case W is itself a representation of G.*
- *We say V is irreducible if $V \neq \{0\}$ and the only subrepresentations V are $\{0\}$ and V itself.*

3 Examples

3.1 First examples

Definition 3.1. *We call $V = \{0\}$ (with its unique G-action) the zero representation. We call $V = F$ (with the trivial G-action) the trivial representation.*

Examples 3.2. *One-dimensional representations are the same as homomorphisms $G \to F^\times$. They are automatically irreducible. Two one-dimensional representations are isomorphic if and only if they are* equal*. For example, if $G = C_n = \{1, g, \ldots, g^{n-1}\}$ is the cyclic group of order n and $F = \mathbb{C}^\times$, then homomorphisms $G \to \mathbb{C}^\times$ are all of the form $g \mapsto \zeta^k$ for some $k \in \{0, \ldots, n-1\}$, where $\zeta \in \mathbb{C}^\times$ is a primitive nth root of unity. This gives n irreducible representations of G over $\mathbb{C}$; we will soon show (Theorem 5.6) that these are all the irreducible representations of G up to equivalence.*

Example 3.3. *Let $G = C_2 = \{1, g\}$. Then giving a homomorphism $G \to \mathrm{GL}(V)$ is the same as giving an element $\phi \in \mathrm{GL}(V)$ satisfying $\phi^2 = \mathrm{Id}$. Every eigenvalue of ϕ is ± 1 and if $2 \neq 0$ in F, ϕ is diagonalizable:*

$$v = \frac{1}{2}(v + \phi(v)) + \frac{1}{2}(v - \phi(v)). \tag{3.1.1}$$

It follows that every representation of C_2 is equivalent to one where the action of ϕ is diagonal and the entries on the diagonal are ± 1.

Example 3.4. *Let $G = C_4 = \{1, g, g^2, g^3\}$ be the cyclic group of order 4. Since the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ has $A^4 = I$, the map $R\colon G \to \mathrm{GL}_2(\mathbb{R}), g^i \mapsto A^i$ defines a matrix representation of G over $\mathbb{R}$. This representation is faithful. We claim it is also irreducible. Indeed, let $W \subset \mathrm{GL}_2(\mathbb{R})$ be a one-dimensional G-stable subspace. Then W is a representation of G, defined by a homomorphism $G \to \mathrm{GL}(W) = \mathbb{R}^\times$. But every element $x \in \mathbb{R}^\times$ with $x^4 = 1$ equals 1 or -1. It follows that W is a ± 1-eigenspace for A. Since the eigenvalues of A are $\pm i$, we obtain a contradiction. However, we can also consider R as a representation over $\mathbb{C}$, by composing with the inclusion $\mathrm{GL}_2(\mathbb{R}) \to \mathrm{GL}_2(\mathbb{C})$. Then A becomes conjugate to $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. You will show on the problem set that every element of finite order in $\mathrm{GL}_n(\mathbb{C})$ is diagonalizable. This is the advantage of working with $\mathbb{C}$.*

Example 3.5. *Let $G = D_n = \langle r, s \mid r^n = s^2 = 1, srs = r^{-1} \rangle$ be the dihedral group of order $2n$. By definition, G is a subgroup of $\mathrm{GL}_2(\mathbb{R})$ preserving a regular n-gon; this defines a 2-dimensional representation $G \to \mathrm{GL}_2(\mathbb{R})$. Composing with the inclusion $\mathrm{GL}_2(\mathbb{R}) \to \mathrm{GL}_2(\mathbb{C})$, this also defines a representation of G over $\mathbb{C}$. If $n \geq 3$ then this representation is irreducible.*

Lecture 2 starts here

Example 3.6. *Let $G = Q_8 = \langle i, j \mid i^4 = 1, i^2 = j^2, jij^{-1} = i^{-1} \rangle$ be the quaternion group of order 8. Writing $-1 = i^2 = j^2$ and $k = ij$, its elements are $\{\pm 1, \pm i, \pm j, \pm k\}$. It can be realized as a subgroup of $\mathrm{GL}_2(\mathbb{C})$ via*

$$i \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \; j \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.1.2}$$

This defines the standard representation of Q_8. In the problem set, you will show that every irreducible representation V of Q_8 over $\mathbb{C}$ with $\dim V \geq 2$ is isomorphic to the standard representation. Moreover, you will show that this representation cannot be realized over $\mathbb{R}$: that is, it cannot be conjugated inside $\mathrm{GL}_2(\mathbb{R})$.

3.2 Permutation representations

It turns out every group action gives rise to a representation.

Definition 3.7. *Let G be a group acting on a finite set X. Let FX be the vector space with basis $\{e_x \mid x \in X\}$ indexed by X, so an element of FX is of the form $\sum_{x \in X} c_x e_x$, where $c_x \in F$ and e_x should be interpreted as a formal symbol. Then the association $g \cdot e_x := e_{g \cdot x}$ extends to a linear G-action on FX. This is called the permutation representation associated to the G-action on X.*

Very often, we will even just write an element of FX as a sum $\sum_{x \in X} c_x x$, where we simply write x instead of e_x.

Remark 3.8. *Warning: some people call a G-action on a finite set a permutation representation. I will stick to calling them simply actions, or permutation actions, reserving the term permutation representation for the linear action on the vector space FX.*

Example 3.9. *Let $G = S_3$ be the symmetric group on three letters and let $X = \{1, 2, 3\}$ on which G acts by definition. Let $V = \mathbb{C}X$ be the associated permutation representation. In the basis $\{e_1, e_2, e_3\}$ of V, the action of G is given by the matrices:*

$$\mathrm{Id} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; (123) \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; (132) \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.2.1}$$

$$(12) \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; (23) \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; (13) \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.2.2}$$

These are all permutation matrices, something which is always true for permutation representations (in the natural basis of FX). Observe that $e_1 + e_2 + e_3$ is fixed by every element of G, hence the $\mathbb{C}$-span of this element is isomorphic to the trivial representation.

We know that every group acts on itself by left multiplication. This shows that every finite group has a 'god-given' representation:

Definition 3.10. *Let G be a finite group and consider the action of G on itself by left multplication. The associated permutation representation FG is called the regular representation of G.*

The regular representation is extremely important. For example, we will show pretty soon that every irreducible representation 'appears' in the regular representation when $F = \mathbb{C}$.

4 Complete reducibility and Maschke's theorem

4.1 Direct sums, indecomposability

Definition 4.1. *Given two G-representations V and W, their direct sum is the vector space of pairs $V \oplus W = \{(v, w) \mid v \in V, w \in W\}$ under componentwise addition. The association $g \cdot (v, w) := (g \cdot v, g \cdot w)$ defines a linear G-action on $V \oplus W$, called the direct sum representation.*

Remark 4.2. *Recall from linear algebra the comparison of the above 'external' direct sum to the following notion of 'internal' direct sum: if U, W are subspaces of a vector space V, we say V is the direct sum of U and W and we write $V = U \oplus W$ if $V = U + W$ and $U \cap W = \{0\}$. To compare the two notions, note that if V is an internal direct sum of U and W then the summing map $U \oplus W \to V$ from the external direct sum to V is an isomorphism.*

Definition 4.3. *A nonzero G-representation V is called decomposable if we can write $V = U \oplus W$ where U and W are nonzero subrepresentations of V, and indecomposable otherwise.*

Lemma 4.4. *Every irreducible representation is indecomposable.*

Proof. If the representation were decomposable, it would be isomorphic to $U \oplus W$ where U and W are nonzero. But then U would be a subrepresentation, contradicting irreducibility. □

Let's interpret irreducibility and indecomposability in terms of matrix representations.

Given a subrepresentation $W \leq V$, choose a basis $\{e_1, \ldots, e_m\}$ of W and extend it to a basis $\{e_1, \ldots, e_n\}$ of V. Let $R_W : G \to \mathrm{GL}_m(F)$ and $R_V : G \to \mathrm{GL}_n(F)$ be the associated matrix representations. Then $R_V(g)$ has a block upper triangular form:

$$R_V(g) = \left(\begin{array}{c|c} R_W(g) & * \\ 0 & * \end{array}\right). \tag{4.1.1}$$

Now suppose there exists another subrepresentation $U \leq V$ such that $V = W \oplus U$. Choose bases of W and U, giving a basis of V and matrix representations R_V, R_W, R_U. Then R_V has block *diagonal* form

$$R_V(g) = \left(\begin{array}{c|c} R_W(g) & 0 \\ 0 & R_U(g) \end{array}\right). \tag{4.1.2}$$

Maschke's theorem shows that we can alway transform the block upper triangular form of R_V into a block diagonal form after changing the basis (under certain conditions on G and F).

4.2 Maschke and its proof

Theorem 4.5 (Maschke)**.** *Let G be a* finite *group and suppose that the order of G is invertible in F. Let V be a G-representation and $W \leq V$ a subrepresentation. Then there exists a G-invariant complement to W: there exists a subrepresentation $U \leq V$ such that $V = W \oplus U$.*

We will give two proofs. For the first one, we will assume $F = \mathbb{C}$.

Definition 4.6. *A Hermitian inner product on a complex vector space V is a map $\langle -, - \rangle : V \times V \to \mathbb{C}$ such that for all $v, w \in V, \lambda, \mu \in \mathbb{C}$:*

- *(Hermitian)* $\langle v, w\rangle = \overline{\langle w, v\rangle}$;
- *(Sesquilinear)* $\langle \lambda v + \mu v', w\rangle = \lambda\langle v, w\rangle + \mu\langle v', w\rangle$;
- *(Positive definite)* $\langle v, v\rangle > 0$ *if* $v \neq 0$.

If V is a G-representation, we say h is G-invariant if $\langle gv, gw\rangle = \langle v, w\rangle$ for all $v, w \in V, g \in G$.

Lemma 4.7. *Let V be a G-representation over $\mathbb{C}$ with a G-invariant Hermitian inner product h and let $W \leq V$ be a subrepresentation. Then $W^\perp = \{v \in V \mid \langle w, v\rangle = 0, \forall\, w \in W\}$ is a subrepresentation of V and $V = W \oplus W^\perp$.*

Proof. The direct sum decomposition $V = W \oplus W^\perp$ holds for any subspace W, so it suffices to show that $W^\perp$ is stable under G. Let $v \in W^\perp$ and $g \in G$. Then $\langle v, w\rangle = 0$ for all $w \in W$. By G-invariance, $\langle g \cdot v, g \cdot w\rangle = 0$ for all $w \in W$. Since W is G-invariant, $g \cdot W = W$ so $\langle g \cdot v, w\rangle = 0$ for all $w \in W$. Hence $g \cdot v \in W^\perp$ and the claim is proven. $\square$

Theorem 4.8. *(Weyl's unitary trick) Suppose that G is finite and let V be a G-representation over $\mathbb{C}$. Then there exists a G-invariant Hermitian inner product on V.*

Proof. The proof is beautiful. Start with any Hermitian inner product $\langle -, -\rangle_0$ on V. Then $\langle -, -\rangle_0$ is not necessarily G-invariant, but we can make a new one that will be: set

$$\langle v, w\rangle := \frac{1}{|G|}\sum_{g\in G}\langle g \cdot v, g \cdot w\rangle_0. \tag{4.2.1}$$

Then $\langle -, -\rangle$ is again a Hermitian inner product on V, being a sum of such things. Moreover, $\langle -, -\rangle$ is G-invariant by construction. Indeed, for every $g' \in G$, we compute:

$$\langle g'v, g'w\rangle = \frac{1}{|G|}\sum_{g\in G}\langle gg'v, gg'w\rangle_0 \tag{4.2.2}$$

As g runs over G, gg' runs over G too, so the sum equals

$$\frac{1}{|G|}\sum_{g\in G}\langle gv, gw\rangle_0 = \langle v, w\rangle, \tag{4.2.3}$$

as claimed. $\square$

Since every Hermitian inner product is equivalent to the standard one on $\mathbb{C}^n$, given by $\langle (x_i), (y_i)\rangle = \sum x_i \bar{y}_i$, and since the subgroup of $\mathrm{GL}_n(\mathbb{C})$ leaving the standard form invariant is

$$\mathrm{U}_n(\mathbb{C}) := \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A\bar{A}^t = I\}, \tag{4.2.4}$$

we get as a consequence:

Corollary 4.9. *Every finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ is conjugate to a subgroup of $\mathrm{U}_n(\mathbb{C})$.*

First proof of Maschke's theorem, assuming $F = \mathbb{C}$. Let $\langle -, -\rangle$ be a G-invariant inner product (which exists by Theorem 4.8) on V and set $U = W^\perp$. By Lemma 4.7, U is G-stable and $V = W \oplus U$. $\square$

Lecture 3 starts here

We will now give a second proof, just assuming that G is finite and F is a general field whose characteristic does not divide the order of $|G|$. We will use the following lemma and a similar 'averaging' idea. If W is a subspace of V, we call a map $p\colon V \to W$ a projector if $p(w) = w$ for all $w \in W$.

Lemma 4.10. *Let V be a G-representation, $W \leq V$ a subrepresentation and $p\colon V \to W$ a G-equivariant projector. Then $\ker(p) \leq V$ is a subrepresentation and $V = W \oplus \ker(p)$.*

Proof. The map $p\colon V \to W$ is surjective, and injective when restricted to W; it follows that $W \cap \ker(p) = \{0\}$ and $V = W + \ker(p)$ by dimension considerations, hence $V = W \oplus \ker(p)$. It suffices to prove that $\ker(p)$ is a subrepresentation of V. But if $v \in \ker(p)$ and $g \in G$, then $p(gv) = gp(v) = g \cdot 0 = 0$, so $g \cdot v \in \ker(p)$. □

Second proof of Maschke's theorem. Let $p_0\colon V \to W$ be *any* projector (not necessarily G-equivariant). Define the linear map $p\colon V \to W$ via:

$$p(v) = \frac{1}{|G|} \sum_{g \in G} g \cdot p_0(g^{-1} \cdot v). \tag{4.2.5}$$

(This makes sense since we can divide by $|G|$ in F.) We claim that p is a G-equivariant projector. Indeed, if $w \in W$ then

$$p(w) = \frac{1}{|G|} \sum_{g \in G} g \cdot p(g^{-1} \cdot w) = \frac{1}{|G|} \sum_{g \in G} g \cdot (g^{-1} \cdot w) = w \tag{4.2.6}$$

So p is indeed a projector. To check that it is G-invariant, let $h \in G$ and compute:

$$h \cdot p(v) = \frac{1}{|G|} \sum_{g \in G} hg \cdot p(g^{-1} \cdot v) \tag{4.2.7}$$

$$= \frac{1}{|G|} \sum_{g \in G} hg \cdot p((hg)^{-1} h \cdot v) \tag{4.2.8}$$

Since the map $G \to G, g \mapsto hg$ is a bijection, we may replace hg by g in the above sum which equals

$$\frac{1}{|G|} \sum_{g \in G} g \cdot p(g^{-1}(h \cdot v)) = p(h \cdot v), \tag{4.2.9}$$

proving that p is G-equivariant, as claimed. Using Lemma 4.10, we see that $U = \ker(p)$ is a G-stable complement to W. □

4.3 Corollaries of Maschke's theorem

Definition 4.11. *A G-representation V is said to be completely reducible (or semisimple) if V is a direct sum of irreducible representations: there exists irreducible representations $V_1, \ldots, V_n$ such that $V \simeq V_1 \oplus \cdots \oplus V_n$.*

Theorem 4.12. *Suppose that G is finite and $|G| \neq 0$ in F. Then every G-representation is semisimple.*

Proof. This follows from induction on the dimension of V and repeatedly applying Maschke's theorem. Indeed, suppose V is a G-representation. If $\dim V = 1$, then V is irreducible. If $\dim V \geq 2$, then either V is irreducible (in which case we're done), or V has a proper subrepresentation $0 \neq W \leq V$. By Maschke's theorem, there exists a subrepresentation $U \leq V$ with $V = W \oplus U$. Since $\dim W$ and $\dim U$ are $< \dim V$, U and W are completely reducible by the induction hypothesis. It follows that V is completely reducible too. □

Remark 4.13. *In abstract terms, this shows that the category of $\mathbb{C}$-representations of a finite group is 'semisimple'. So there are no nontrivial 'extensions' between representations.*

Here is another corollary of Maschke's theorem, which we will reprove later using character theory. We did not cover this in class, and consequently it is non-examinable.

Corollary 4.14. *Suppose that G is finite and $|G| \neq 0$ in F. Let V be an irreducible representation of G. Then V is isomorphic to a subrepresentation of the regular representation FG.*

Proof. We will first exhibit V as a *quotient* of V, and then apply Maschke's theorem to turn the quotient into a subrepresentation. Let $v \in V$ be a nonzero element. Let $\phi \colon FG \to V$ be the map sending $\sum_g c_g e_g$ to $\sum_g c_g(g \cdot v)$. Then ϕ is a G-homomorphism, since this can be checked on basis elements of FG. Therefore the image $\mathrm{image}(\phi) \leq V$ is a nontrivial subrepresentation. Since V is irreducible, $\mathrm{image}(\phi) = V$ and ϕ is surjective. By Maschke's theorem, the kernel of ϕ has a G-invariant complement $W \leq FG$, and the composition $W \hookrightarrow FG \xrightarrow{\phi} V$ is an isomorphism. The claim follows. □

This already implies that there are only finitely many isomorphism classes of irreducible representations of G.

5 Schur's lemma and isotypical decomposition

5.1 Schur's lemma

Schur's lemma is simple to state and prove but has surprisingly many consequences. We continue to work with a group G and vector spaces over a field F.

Theorem 5.1 (Schur's lemma)**.** *Let V, W be irreducible G-representations.*

1. *Let $\phi \colon V \to W$ be a G-equivariant map. Then ϕ is either zero or invertible.*
2. *Suppose F is algebraically closed. Let $\phi \colon V \to V$ be a G-equivariant endomorphism. Then ϕ is a scalar multiple of the identity map.*

Proof.

1. The crucial observation to make is that $\ker(\phi)$ and $\mathrm{image}(\phi)$ are subrepresentations of V and W respectively. Indeed, if $v \in \ker(\phi)$ and $g \in G$ then $\phi(g \cdot v) = g \cdot \phi(v) = 0$, so $g \cdot v \in \ker(\phi)$. If $w = \phi(v) \in \mathrm{image}(\phi)$, then $g \cdot w = \phi(g \cdot v) \in \mathrm{image}(\phi)$. Then the proof is easy: if ϕ is nonzero, $\ker(\phi) \neq V$ and $\mathrm{image}(\phi) \neq 0$. By irreducibility of V and W, $\ker(\phi) = 0$ and $\mathrm{image}(\phi) = W$. In other words, ϕ is bijective, hence invertible.
2. Let $\phi \colon V \to V$ be a G-equivariant map. Let $\lambda \in F$ be an eigenvalue of ϕ; such a λ exists because F is algebraically closed (hence the characteristic polynomial of ϕ has a root in F). Then the linear map $\phi - \lambda \,\mathrm{Id} \colon V \to V$ is again G-equivariant. However, $\phi - \lambda\,\mathrm{Id}$ is not injective because every λ-eigenvector for ϕ lies in the kernel of this map. By Part (1), we conclude that $\phi - \lambda\,\mathrm{Id} = 0$, that is $\phi = \lambda\,\mathrm{Id}$.

□

Remark 5.2. *If F is not necessarily algebraically closed, we can deduce that for any irreducible G-representation, $\mathrm{End}_G(V) = \mathrm{Hom}_G(V, V)$ is a* division algebra *under composition, i.e. every nonzero element is invertible. Since the only division algebra over an algebraically closed field is the field itself, this explain the second part of Theorem 5.1.*

Let's discuss some consequences of Schur's lemma. We will now assume, until explicitly stated otherwise, that

Convention: G is finite and $F = \mathbb{C}$.

5.2 Abelian groups

For a group G, let $Z(G) = \{g \in G \mid gh = hg, \forall h \in G\}$ be its center.

Proposition 5.3 (Central character). *Let V be an irreducible G-representation. Let $Z(G)$ be the center of G. Then there exists a homomorphism $\lambda \colon Z(G) \to \mathbb{C}^\times$, called the central character, such that $z \cdot v = \lambda(z) \cdot v$ for all $z \in Z(G)$ and $v \in V$. In other words, the restriction of $G \to \mathrm{GL}(V)$ to $Z(G)$ lands in the subset of scalar multiples of* Id.

Proof. We claim that if $z \in Z(G)$ then the map $\phi_z := \rho(z) \colon V \to V, v \mapsto z \cdot v$ is G-equivariant. Indeed, since z commutes with every element of G, we have $\phi_z(g \cdot v) = z \cdot (g \cdot v) = g \cdot (z \cdot v) = g \cdot \phi_z(v)$. By Part 2 of Schur's lemma, the map $v \mapsto zv$ is given by multiplication by a scalar $\lambda(z) \in \mathbb{C}^\times$. The claim follows. $\square$

This has the following not-so obvious corollary.

Corollary 5.4. *Suppose that G admits a faithful irreducible representation. Then the center $Z(G)$ of G is cyclic.*

Proof. This implies that the central character $Z(G) \to \mathbb{C}^\times$ is injective. But every finite subgroup of $\mathbb{C}^\times$ is cyclic. $\square$

Remark 5.5 (Non-examinable). *Not every group with cyclic center has a faithful irreducible representation. Indeed, suppose S_3 acts on C_3 via the non-trivial action of $C_2 = S_3/C_3$ on C_3. Then $C_3 \rtimes S_3$ has trivial center but no irreducible faithful representation. (Exercise for enthousiasts.)*

Theorem 5.6. *Let G be a finite abelian group. Then every irreducible G-representation is one-dimensional. The number of irreducible representations of G (up to equivalence) is finite and equals $|G|$.*

Proof. Since G is abelian, $G = Z(G)$. By Proposition 5.3, the action of G on an irreducible representation V is given by multiplication by the central character $\lambda \colon G \to \mathbb{C}^\times$. In particular, every subspace of V is a subrepresentation. Since V is irreducible, it follows that V is one-dimensional.

To count the number of irreps of G up to isomorphism, it suffices to count the number of homomorphisms $G \to \mathbb{C}^\times$. By the structure theorem for finitely generated abelian groups, G is a direct product of cyclic groups $C_{n_1}, \dots, C_{n_k}$. Since $\mathrm{Hom}(A \times B, \mathbb{C}^\times) = \mathrm{Hom}(A, \mathbb{C}^\times) \times \mathrm{Hom}(B, \mathbb{C}^\times)$ for abelian groups A, B, we may assume that G is cyclic of order n, with generator n. If $\zeta \in \mathbb{C}^\times$ is a primitive nth root of unity, every homomorphism $G \to \mathbb{C}^\times$ is of the form $g \mapsto \zeta^k$ for some $0 \leq k \leq n-1$. $\square$

We will see later that the number of irreps up to equivalence for a (not necessarily abelian) group is equal to the number of conjugacy classes in G.

5.3 Isotypical decomposition

Lecture 4 starts here

Recall from linear algebra that if $\phi \colon V \to V$ is a diagonalizable linear map with eigenvalues $\lambda_1, \dots, \lambda_k$ then

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_k), \tag{5.3.1}$$

where $V(\lambda_i)$ denotes the λ_i-eigenspace of ϕ. This decomposition is 'canonical', i.e. it does not require us to choose any bases. Something similar happens for representations. To see this, we will need the following consequence of Schur's lemma. Recall that $\mathrm{Hom}_G(V, W)$ denotes the set of G-homomorphisms $V \to W$

Lemma 5.7. *Let V, W be irreducible G-representations. Then*

$$\dim_{\mathbb{C}} \operatorname{Hom}_G(V, W) = \begin{cases} 0 & \text{if } V \not\simeq W, \\ 1 & \text{if } V \simeq W. \end{cases} \tag{5.3.2}$$

Proof. By Schur's lemma, every element of $\operatorname{Hom}_G(V, W)$ is either an isomorphism or the zero map. If $V \not\simeq W$, this implies $\operatorname{Hom}_G(V, W) = 0$. Suppose $V \simeq W$ and let $\psi\colon V \to W$ be an isomorphism of G-representations. We will show that every G-linear map $\phi\colon V \to W$ is a scalar multiple of ψ, proving that $\operatorname{Hom}_G(V, W)$ is one-dimensional. Indeed, the composition $\psi^{-1}\phi\colon V \to V$ is a G-equivariant endomorphism of V. Again by Schur's lemma, this endomorphism is a scalar multiple of the identity map, so $\psi^{-1}\phi = \lambda \operatorname{Id}$ for some $\lambda \in \mathbb{C}$. Therefore $\phi = \lambda\psi$, proving that $\operatorname{Hom}_G(V, W) = \operatorname{span}\{\psi\}$. □

Lemma 5.8. *Let V, V_1, V_2 be G-representations. Then there are isomorphisms of vector spaces*

$$\operatorname{Hom}_G(V, V_1 \oplus V_2) \simeq \operatorname{Hom}_G(V, V_1) \oplus \operatorname{Hom}_G(V, V_2), \tag{5.3.3}$$

$$\operatorname{Hom}_G(V_1 \oplus V_2, V) \simeq \operatorname{Hom}_G(V_1, V) \oplus \operatorname{Hom}_G(V_2, V). \tag{5.3.4}$$

Proof. We will write down the 'obvious maps' and their inverses for the first isomorphism, leaving the verifications and the second isomorphism as an exercise. Given a G-homomorphism $\phi\colon V \to V_1 \oplus V_2$, let $\phi_i\colon V \to V_i$ be the postcomposition with the projection $V_1 \oplus V_2 \to V_i$. Given two G-homomorphisms $\phi_i\colon V \to V_i$ for $i = 1, 2$, let define the map $\phi\colon V \to V_1 \oplus V_2$ via $\phi(v) = (\phi_1(v), \phi_2(v))$. This establishes a bijection between linear maps $\phi\colon V \to V_1 \oplus V_2$ and pairs of linear maps $\phi_1\colon V \to V_1$ and $\phi_2\colon V \to V_2$. It is easy to check that ϕ is G-equivariant if and only if ϕ_1 and ϕ_2 are G-equivariant. □

Proposition 5.9. *Let V be a G-representation and write $V = V_1 \oplus \cdots \oplus V_k$ where each V_i is irreducible. Let S be an irreducible G-representation. Then*

$$\dim_{\mathbb{C}} \operatorname{Hom}_G(S, V) = \#\{1 \leq i \leq k \mid V_i \simeq S\}. \tag{5.3.5}$$

Proof. By Lemma 5.8, $\operatorname{Hom}_G(S, V) = \operatorname{Hom}_G(S, V_1) \oplus \cdots \oplus \operatorname{Hom}_G(S, V_k)$. By Lemma 5.7, $\operatorname{Hom}_G(S, V_i)$ is one-dimensional or zero according to whether $S \simeq V_i$ or not. Combining these two sentences proves the proposition. □

Theorem 5.10 (Isotypic decomposition). *Let V be a G-representation.*

1. *There exists mutually non-isomorphic irreducible representations $V_1, \ldots, V_k$, nonnegative integers $a_1, \ldots, a_k$ and a decomposition*

$$V \simeq V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k}. \tag{5.3.6}$$

2. *The integers a_i are uniquely determined. In other words, if $V \simeq V_1^{\oplus b_1} \oplus \cdots \oplus V_k^{\oplus b_k}$ then $a_i = b_i$ for all $1 \leq i \leq k$.*

3. *The image of $V_i^{a_i}$ in V is independent of the choice of isomorphism* (5.3.6). *This subrepresentation of V is called the V_i-isotypic component of V.*

4. *Let W be another representation and suppose that $W \simeq V_1^{\oplus b_1} \oplus \cdots \oplus V_k^{\oplus b_k}$. Then there exist isomorphisms of vector spaces*

$$\operatorname{Hom}_G(V, W) \simeq \bigoplus_{i=1}^{k} \operatorname{Hom}(V_i^{\oplus a_i}, V_i^{\oplus b_i}) \tag{5.3.7}$$

$$\simeq \bigoplus_{i=1}^{k} \operatorname{Mat}_{a_i \times b_i}(\mathbb{C}) = \{a_i \times b_i\text{-matrices}\}. \tag{5.3.8}$$

Proof. Part (1) is a restatement of Theorem 4.12. For Part (2), note that $a_i = \dim_{\mathbb{C}} \mathrm{Hom}_G(V_i, V)$ by Proposition 5.9, so the a_i do not depend on the isomorphism (5.3.6). To prove (4), it follows from Lemmas 5.8 and 5.7 that

$$\mathrm{Hom}_G(V, W) \simeq \bigoplus_{i=1}^{k} \mathrm{Hom}(V_i^{\oplus a_i}, V_i^{\oplus b_i}) \simeq \bigoplus_{i=1}^{k} \mathrm{Hom}_G(V_i, V_i)^{a_i \times b_i}. \tag{5.3.9}$$

By Schur's lemma, $\mathrm{Hom}_G(V_i, V_i) \simeq \mathbb{C}$, which concludes, so $\mathrm{Hom}_G(V_i, V_i)^{a_i \times b_i}$ can be seen as the space of $a_i \times b_i$-matrices. To prove (3), it suffices to prove that any isomorphism $V_1^{a_1} \oplus \cdots \oplus V_k^{a_k} \xrightarrow{\sim} V_1^{a_1} \oplus \cdots \oplus V_k^{a_k}$ maps each $V_i^{a_i}$ to itself. This follows immediately from (5.3.7). $\square$

Definition 5.11. *A decomposition of V of the form*

$$V = W_1 \oplus \cdots \oplus W_k \tag{5.3.10}$$

where $W_i \simeq V_i^{\oplus n_i}$, V_i is irreducible and $V_i \not\simeq V_j$ for all $i \neq j$ is called the isotypic decomposition, and W_i is called the V_i-isotypic component. The integer n_i is called the multiplicity of V_i in V.

By Theorem 5.10, this decomposition is unique hence it is really justified to speak of *the* isotypic decomposition.

Remark 5.12. *We can characterize the V_i-isotypic component of V in the following 'basis-independent' way: it is the union of $\mathrm{image}(\phi)$, where ϕ runs over all elements of $\mathrm{Hom}_G(V_i, V)$. Using tensor products (to be discussed later), it can also be characterized as the image of the natural map $V_i \otimes \mathrm{Hom}_G(V_i, V) \to V$.*

Remark 5.13. *Even though the V_i-isotypic component $W_i \leq V$ is uniquely determined, we cannot write $W_i \simeq V_i^{n_i}$ without making any choices. Going back to the analogy with eigenspaces, the eigenspaces of a diagonalizable operator are uniquely determined but choosing eigenvectors requires choosing a basis. In fact, giving an isomorphism $W_i \simeq V_i^{n_i}$ is the same as giving a basis of $\mathrm{Hom}_G(V_i, V)$.*

6 Character theory

Recall that G is assumed to be a finite group and all representations are finite-dimensional over $\mathbb{C}$.

6.1 Basic definitions

Definition 6.1. *The character of a representation $\rho\colon G \to \mathrm{GL}(V)$ is the function $\chi_V : G \to \mathbb{C}, g \mapsto \mathrm{tr}(\rho(g))$. We call any function $\chi\colon G \to \mathbb{C}$ of the form χ_V for some G-representation V a character, and in that case we say that χ is afforded by V.*

Alternatively, after choosing a basis of V the character is the function $g \mapsto \mathrm{tr}(R(g))$, where $R\colon G \to \mathrm{GL}_n(\mathbb{C})$ is the corresponding matrix representation. Since the trace of a matrix is invariant under conjugation, this definition is well defined and does not depend on a choice of basis. Note that the character χ_V is merely a function $G \to \mathbb{C}$, not necessarily a group homomorphism. We can therefore assign a single complex number to an element of $g \in G$. At this point it might be hard to imagine how this function $G \to \mathbb{C}$ can contain much information about the representation, which in a basis assigns a whole matrix of numbers to every $g \in G$. Amazingly, it will turn out that the character of V will determine V!

Definition 6.2. *Let $\chi = \chi_V$ be a character. We say χ is faithful, irreducible, trivial if V is. We call $\dim V$ the degree of χ. We say χ is linear if it is of degree 1. In that case χ is a group homomorphism $G \to \mathbb{C}^\times$.*

We start with some easy but very useful properties of characters.

Proposition 6.3. *Let V be a G-representation with character χ_V.*

1. $\chi_V(1) = \dim V$.
2. $\chi_V : G \to \mathbb{C}$ *is a class function: for all $g, h \in G$ we have*
$$\chi_V(ghg^{-1}) = \chi_V(h). \tag{6.1.1}$$
3. $\chi_V(g^{-1}) = \overline{\chi_V(g)}$ *for all $g \in G$.*
4. *If V, W are two representations then $\chi_{V \oplus W} = \chi_V + \chi_W$.*

Proof.

1. $\chi_V(1) = \operatorname{tr}(\mathrm{Id}_V) = \dim V$.
2. Choose a basis of V and let $R \colon G \to \mathrm{GL}_n(\mathbb{C})$ be the associated matrix representation. Then $\chi_V(ghg^{-1}) = \operatorname{tr}(R(ghg^{-1})) = \operatorname{tr}(R(g)R(h)R(g)^{-1}) = \operatorname{tr}(R(h)) = \chi_V(h)$, where the third equality is justified by the fact that the trace is conjugation invariant.
3. You have shown in Problem Set 1 (Q4) that $\rho(g) \in \mathrm{GL}(V)$ is diagonalizable; suppose that it has eigenvalues $\lambda_1, \dots, \lambda_n$, counted with multiplicity. Then $\rho(g^{-1})$ has eigenvalues $\lambda_1^{-1}, \dots, \lambda_n^{-1}$. Since g is of finite order, each λ_i is a root of unity and $\lambda_i^{-1} = \bar{\lambda}_i$. It follows that $\chi_V(g^{-1}) = \lambda_1^{-1} + \cdots + \lambda_n^{-1} = \bar{\lambda}_1 + \cdots + \bar{\lambda}_n = \overline{\chi_V(g)}$.
4. Choose bases of V and W. This determines a basis of $V \oplus W$ in which the action of G is given by block diagonal form (4.1.2). Since the trace of a block diagonal matrix is the sum of the traces of the blocks, the identity follows.

$\square$

Example 6.4. *Let $G = S_3$ and let $V = \mathbb{C}\{1, 2, 3\}$ be the associated permutation representation. Then $V \simeq (\text{trivial}) \oplus W$, where W is the subset of V whose coordinates sum to zero. Using the fact that $\chi_V = \chi_W + \chi_{triv}$ and the matrices from* (3.2.1)*, we can quickly compute the character of χ_W. Namely $\chi_W(1) = 2$, $\chi_W((12)) = 0$ and $\chi_W((123)) = -1$.*

Definition 6.5. *A class function is a function $f \colon G \to \mathbb{C}$ that is constant on conjugacy classes of G, in other words that satisfies $f(hgh^{-1}) = f(g)$ for all $g, h \in G$. Write $\mathcal{C}(G)$ for the vector space of all class functions $G \to \mathbb{C}$ under pointwise addition.*

Let $\mathcal{C}_1, \dots, \mathcal{C}_k$ be the conjugacy classes of G. Since a function $f \colon G \to \mathbb{C}$ is a class function if and only if it is constant on conjugacy classes, an explicit basis $\{\delta_i \mid 1 \le i \le k\}$ of $\mathcal{C}(G)$ is given by the characteristic functions of the conjugacy classes $\mathcal{C}_i$:

$$\delta_i(g) = \begin{cases} 1 & \text{if } g \in \mathcal{C}_i, \\ 0 & \text{otherwise.} \end{cases} \tag{6.1.2}$$

It follows that $\dim \mathcal{C}(G) = k = \#\{\text{number of conjugacy classes of } G\}$.

Definition 6.6 (Inner product of class functions)**.** *Given two class functions $f, f' \colon G \to \mathbb{C}$, we define*

$$\langle f, f' \rangle := \frac{1}{|G|} \sum_{g \in G} f(g)\overline{f'(g)}. \tag{6.1.3}$$

This is a Hermitian inner product on $\mathcal{C}(G)$.

Choose representatives $g_i \in \mathcal{C}_i$. Then we may also write

$$\langle f, f' \rangle = \frac{1}{|G|} \sum_{i=1}^{k} |\mathcal{C}_i| f(g_i) \overline{f'(g_i)} = \sum_{i=1}^{k} \frac{1}{|C_G(g_i)|} f(g_i) \overline{f'(g_i)}. \tag{6.1.4}$$

Here $C_G(g_i) = \{g \in G \mid gg_i = g_i g\}$ is the centralizer of g_i and the second equality follows from the orbit-stabilizer formula.

6.2 Completeness of characters + consequences

The next theorem is one of the cornerstones of this course and of representation theory of finite groups.

Theorem 6.7 (Completeness of characters)**.** *The irreducible characters of G form an orthonormal basis of $\mathcal{C}(G)$. In other words:*

1. *Let V, V' be two irreducible G-representations with characters χ, χ'. Then*

$$\langle \chi, \chi' \rangle = \begin{cases} 1 & \text{if } V \simeq V', \\ 0 & \text{if } V \not\simeq V'. \end{cases} \tag{6.2.1}$$

2. *Every class function $f \in \mathcal{C}(G)$ is a $\mathbb{C}$-linear combination of irreducible characters.*

The proof will be postponed to next lecture. We first look at some staggering consequences.

Corollary 6.8. *The number of irreducible representations of G up to equivalence equals the number of conjugacy classes of G.*

Proof. Both irreducible characters and characteristic functions of conjugacy classes form bases of $\mathcal{C}(G)$. □

Corollary 6.9. *Two representations V and V' are isomorphic if and only if they have the same character.*

Proof. Let V be a G-representation. We know that $V \simeq V_1^{n_1} \oplus \cdots \oplus V_k^{n_k}$ where each V_i is irreducible. Moreover two representations are isomorphic if and only all the multiplicities n_i agree. But if V has character χ and V_i has character χ_i, we have $\chi = n_1\chi_1 + \cdots + n_k\chi_k$ and hence by orthonormality

$$\langle \chi, \chi_i \rangle = n_i. \tag{6.2.2}$$

It follows that the character χ determines the multiplicities n_i, hence the isomorphism class of V. □

Corollary 6.9 might fail for infinite groups. For example, if $G = \mathbb{Z}$ then the representations

$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; 1 \mapsto \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{6.2.3}$$

are not isomorphic but have the same character.

Lecture 5 starts here

Corollary 6.10. *A representation V is irreducible if and only if $\langle \chi_V, \chi_V \rangle = 1$.*

Proof. Write $V = V_1^{n_1} \oplus \cdots \oplus V_k^{n_k}$ where V_i are irreducible and non-isomorphic. Then

$$\langle \chi_V, \chi_V \rangle = n_1^2 + \cdots + n_k^2. \tag{6.2.4}$$

V is irreducible if and only if exactly one of the n_i's is 1 and all the others are zero. But this is true if and only if $n_1^2 + \cdots + n_k^2 = 1$. □

For the next result, we first determine the character of our favourite representation, the regular representation.

Lemma 6.11. *Let χ_{reg} be the character of the regular representation. Then*

$$\chi_{reg}(g) = \begin{cases} |G| & \text{if } g = 1, \\ 0 & \text{else.} \end{cases} \tag{6.2.5}$$

Proof. In the standard basis of the regular representation, the action of g is given by permutation matrices, as is the case for any permutation representation. Since the trace of a permutation matrix is equal to the number of basis elements fixed by the matrix (the number of elements on the diagonal), we see that $\chi_{reg}(g)$ equals the number of fixed points of the multiplication-by-g map $G \to G, h \mapsto gh$. This map has no fixed points if $g \neq 1$ and has $|G|$ fixed points if $g = 1$. □

Theorem 6.12. *Let $V_1, \ldots, V_k$ be a set of representatives for the isomorphism classes of irreducible representations of G. Let $\chi_1, \ldots, \chi_k$ be their irreducible characters of G and let $d_1, \ldots, d_k$ be their degrees. Then*

$$d_1^2 + \cdots + d_k^2 = |G|. \tag{6.2.6}$$

Proof. Let χ_{reg} be the character of the regular representation $\mathbb{C}G$. The proof will follow from evaluating $\langle \chi_{reg}, \chi_{reg} \rangle$ in two different ways. First of all, by Lemma 6.11,

$$\langle \chi_{reg}, \chi_{reg} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{reg}(g)\overline{\chi_{reg}(g)} = \frac{1}{|G|} |G| \cdot \overline{|G|} = |G|. \tag{6.2.7}$$

Since the regular representation is completely reducible, we can write $\chi_{reg} = m_1\chi_1 + \cdots + m_k\chi_k$, where m_i is the multiplicity of V_i in the regular representation. Again by Lemma 6.11,

$$m_i = \langle \chi_{reg}, \chi_i \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{reg}(g)\overline{\chi_i(g)} = \frac{1}{|G|} \chi_{reg}(1)\chi_i(1) = \chi_i(1) = \dim V_i. \tag{6.2.8}$$

In other words, we have $m_i = d_i$. By orthonormality, we have

$$\langle \chi_{reg}, \chi_{reg} \rangle = \sum_{1 \leq i,j \leq k} d_i d_j \langle \chi_i, \chi_j \rangle = d_1^2 + \cdots + d_k^2. \tag{6.2.9}$$

Combining (6.2.7) and (6.2.9) proves the theorem. □

The proof of Theorem 6.12 shows:

Corollary 6.13. *Let V_i be an irreducible G-representation. Then the multiplicity of V_i in the regular representation equals* $\dim V_i$.

Theorem 6.12 also shows that G is abelian if and only if $d_i = 1$ for all i. Indeed, this is both equivalent to $k = |G|$.

Corollary 6.14. *Two elements $g, g' \in G$ are conjugate if and only if $\chi(g) = \chi(g')$ for all irreducible characters χ.*

Proof. The elements g, g' are conjugate if and only if $f(g) = f(g')$ for every class function $f \in \mathcal{C}(G)$. Since the irreducible characters of G span $\mathcal{C}(G)$ as a $\mathbb{C}$-vector space, this is equivalent to $\chi(g) = \chi(g')$ for all irreducible characters χ. $\square$

6.3 Character table

Let $\mathcal{C}_1, \dots, \mathcal{C}_k$ be the conjugacy classes of G and choose representatives $g_i \in \mathcal{C}_i$ for each i.

Definition 6.15. *The character table of G is the complex $k \times k$-matrix $(\chi_i(g_j))_{1 \leq i,j \leq k}$ where χ_i runs over the irreducible characters of G.*

Note that this is a square matrix by Theorem 6.7. Usually we think of the character table not just as a matrix, but as a table where we also record the sizes of the conjugacy classes and the irreducible characters. The character table is a concise way of packaging all the information of all the irreducible representations of G.

Slogan: the character table is the '$\mathbb{C}$-linear shadow' of G.

Example 6.16. *Let $G = S_3$. Then G has three conjugacy classes, with representatives $\mathrm{Id}, (12), (123)$ and of size $1, 3, 2$ respectively. The group G has two one-dimensional representations, the trivial one and the sign character sgn. Since there are three conjugacy classes, there is exactly one other irreducible representation. Let χ be the character of the 'standard' representation of S_3, the sum-zero subspace of $\mathbb{C}^3$ where S_3 permutes the coordinates. In Example 6.4, we calculated χ and since $\langle \chi, \chi \rangle = \frac{1}{6}(1 \cdot 2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2) = 1$ we know that χ is irreducible. So the character table of S_3 is:*

	Id	(12)	(123)
1	1	1	1
sgn	1	-1	1
χ	2	0	-1

Example 6.17. *Suppose that we are given a nonabelian group G of order 6 and let us pretend for a moment that we don't know it is isomorphic to S_3. What can we say about its character table purely from first princples? It turns out everything! Indeed, let k be the numer of conjugacy classes of G and $d_1, \dots, d_k$ the dimensions of the irreducible characters. By Theorem 6.12, $d_1^2 + \cdots + d_k^2 = 6$. Since G is nonabelian, $k < 6$. The only way to write 6 as a sum of five or less squares is as $1^2 + 1^2 + 2^2$, so G has two characters of degree 1 and one character of degree 2. Let ϵ be the nontrivial linear character and χ the degree 2 irreducible character. Let g be an element of order 2 and h be an element of order 3 in G. The $1, g, h$ are representatives for the conjugacy classes of G with size $1, 3, 2$ and the character table of G looks like:*

	Id	g	h
1	1	1	1
ϵ	1	?	?
χ	1	?	?

Since h^{-1} has order 3, h^{-1} is conjugate to h. Therefore $\epsilon(h) \in \mathbb{C}^\times$ satisfies $\overline{\epsilon(h)} = \epsilon(h^{-1}) = \epsilon(h)$ and $\epsilon(h)^3 = 1$. This implies that $\epsilon(h) = 1$. Since ϵ is nontrivial, it must be $\neq 1$ on some conjugacy class, so $\epsilon(g) \neq 1$ and $\epsilon(g)^2 = 1$, hence $\epsilon(g) = -1$. So our character table looks like:

	Id	g	h
1	1	1	1
ϵ	1	-1	1
χ	1	$?$	$?$

The last row can be determined using the relations $\langle \chi, 1\rangle = 0$ and $\langle \chi, \epsilon\rangle = 0$. (Or using column orthogonality, see below.)

On Problem Set 2 you will see an example of two non-isomorphic groups with the same character table.

Remark 6.18. *Usually constructing the character table of a group is much easier than actually constructing the corresponding irreducible representations! But just knowing the character table is often enough to deduce interesting results about the group in question.*

Let us state some properties of character tables that are very useful when computing them. The relation $\langle \chi_i, \chi_j\rangle = \delta_{ij}$ for irreducible characters χ_i, χ_j is called row orthogonality, because it states the the rows of the character table are orthogonal, with the caveat that one has to weigh the entries by the size of the associated conjugacy class. We know from linear algebra that a matrix with orthonormal rows also has orthonormal columns, similarly here, we have:

Proposition 6.19 (Column orthogonality). *Let $\mathcal{C}, \mathcal{C}'$ be conjugacy classes of G with representatives $g \in \mathcal{C}$, $g' \in \mathcal{C}'$. Then*

$$\sum_{g\in G} \chi(g)\overline{\chi(g')} = \begin{cases} |C_G(g)| & \text{if } \mathcal{C} = \mathcal{C}', \\ 0 & \text{else }. \end{cases} \tag{6.3.1}$$

Proof. This essentially follows from the fact $A\bar{A}^t = I \Rightarrow \bar{A}^t A = I$, but we have to be slightly careful to remember weighting by the size of the conjugacy class. Let $A = (a_{ij})_{1\le i,j\le k} = (\chi_i(g_j))$ be the character table of G and let $D = \mathrm{diag}(|C_G(g_1)|, \dots, |C_G(g_k)|)$. Then row orthogonality translates to $AD^{-1}\bar{A}^t = I$. Since left inverses and right inverses of matrices are the same, we also have $D^{-1}\bar{A}^t A = I$, in other words $\bar{A}^t A = D$. This implies the proposition. ☐

The following lemma allows us to generate new characters from old ones, which is often useful when constructing character tables.

Lemma 6.20.

1. *If χ is an (irreducible) character, then so is $\bar{\chi}$.*
2. *If χ is an (irreducible) character and ϵ, then so is $\epsilon\chi$ for every linear (one-dimensional character) ϵ.*

Proof.

1. If χ is afforded by the matrix representation $R\colon G \to \mathrm{GL}_n(\mathbb{C})$, then $\bar{\chi}$ is afforderd by $g \mapsto \overline{R(g)}$. Since complex conjugation induces a bijection between subrepresentations of R and subrepresentations of $\bar{R}$, $\bar{\chi}$ is irreducible if χ is.
2. If χ is afforded by $R\colon G \to \mathrm{GL}_n(\mathbb{C})$, then $\epsilon\chi$ is afforded by $g \mapsto \epsilon(g)R(g)$. The remaining properties follow from (Q3) on Problem Set 1.

☐

On the problem set you will get a lot of hands on experience with computing character tables of finite groups.

6.4 First projection formula

Lecture 6 starts here

We will now work our way towards proving Theorem 6.7. We will start by proving orthonormality. This will follow from the very similarly looking Lemma 5.7. As a warm-up, we will start with the first projection formula, which implies orthogonality with the first row of the character table.

Definition 6.21. *If V is a G-representation, the subspace*

$$V^G \coloneqq \{v \in V \mid g \cdot v = v, \forall g \in G\} \tag{6.4.1}$$

is called the G-invariant subspace, or the subspace of G-fixed points. It is a subrepresentation of V.

Proposition 6.22 (First projection formula). *Let V be a representation with character χ. If $1\colon G \to \mathbb{C}$ is the trivial character then*

$$\langle \chi, 1\rangle = \langle 1, \chi\rangle = \dim V^G. \tag{6.4.2}$$

Proof. The map

$$\pi\colon V \to V,\ v \mapsto \frac{1}{|G|}\sum_{g\in G} g \cdot v \tag{6.4.3}$$

lands in V^G and is a projector, that is $\pi(v) = v$ for all $v \in V^G$. It follows that the trace of π equals $\dim V^G$. (This is a general fact about projectors.) But the definition of π shows that the trace of π is exactly $\langle \chi, 1\rangle$. Since $\dim V^G$ is a real number, $\langle \chi, 1\rangle = \langle 1, \chi\rangle$. ☐

6.5 Orthonormality of characters

To deduce orthogonality between χ_V and χ_W for irreducible G-representations, the trick is to apply the first projection formula to a third representation, namely $\mathrm{Hom}(V, W)$, defined as follows:

Definition 6.23. *Let V, W be G-representations. Then the assignment $(g \cdot \phi)(v) \coloneqq g \cdot \phi(g^{-1} \cdot v)$ defines a linear G-action on $\mathrm{Hom}(V, W)$, so $\mathrm{Hom}(V, W)$ is again a G-representation.*

You should check as an exercise why this is indeed a G-action, and why the inverse is present! In particular, if W is the trivial representation then the dual vector space $\mathrm{Hom}(V, \mathbb{C}) = V^*$ is again a G-representation via $(g \cdot \phi)(v) = \phi(g^{-1} \cdot v)$. The following lemma is immediate from the definition.

Lemma 6.24. *Let V, W be G-representations. Then*

$$\mathrm{Hom}(V, W)^G = \mathrm{Hom}_G(V, W). \tag{6.5.1}$$

The following lemma is a calculation:

Lemma 6.25. *Let V, W be G-representations. Then $\chi_{\mathrm{Hom}(V,W)}(g) = \overline{\chi_V(g)}\chi_W(g)$ for all $g \in G$.*

Proof. Let $g \in G$. Recall that g acts as a diagonalizable linear map on V and W; let $\{v_1, \dots, v_m\}$ and $\{w_1, \dots, w_n\}$ be a basis of eigenvectors of V and W respectively. Then $g \cdot v_i = \lambda_i$ and $g \cdot w_j = \mu_j$, where $\lambda_i, \mu_j \in \mathbb{C}$ are roots of unity. So $g^{-1} \cdot v_i = \bar{\lambda}_i v_i$ for all i.

For $1 \leq i \leq m$ and $1 \leq j \leq n$, let $\epsilon_{ij} \in \mathrm{Hom}(V, W)$ be the linear map defined by sending v_i to w_j, and v_k to 0 for all $k \neq i$. Then $\{\epsilon_{ij}\}$ is a basis of $\mathrm{Hom}(V, W)$. (These are like the 'elementary matrices'.) By definition of the G-action on $\mathrm{Hom}(V, W)$, $(g \cdot \epsilon_{ij})(v) = g \cdot \epsilon_{ij}(g^{-1} \cdot v)$ for all $v \in V$. In particular, $(g \cdot \epsilon_{ij})(v_k)$ equals 0 if $k \neq i$, and equals $\mu_j \bar{\lambda}_i w_j$ if $k = i$. It follows that $g \cdot \epsilon_{ij} = \bar{\lambda}_i \mu_j \epsilon_{ij}$. We conclude that

$$\chi_{\mathrm{Hom}(V,W)}(g) = \sum_{i=1}^{m} \sum_{j=1}^{n} \bar{\lambda}_i \mu_j = \left(\sum_{i=1}^{m} \bar{\lambda}_i \right) \left(\sum_{j=1}^{n} \mu_i \right) = \overline{\chi_V(g)} \chi_W(g). \tag{6.5.2}$$

□

Remark 6.26. *Later we will define the tensor product $V \otimes W$ of two representations and this will satisfy $\chi_{V \otimes W} = \chi_V \chi_W$. The proof of the previous lemma then follows from the isomorphism of G-representations $\mathrm{Hom}(V, W) \simeq V^* \otimes W$.*

Proposition 6.27. *Let V, W be G-representations. Then*

$$\langle \chi_V, \chi_W \rangle = \langle \chi_W, \chi_V \rangle = \dim \mathrm{Hom}_G(V, W). \tag{6.5.3}$$

Proof. Let $U = \mathrm{Hom}(V, W)$, seen as a G-representation. Applying the first projection formula to U shows that $\langle \chi_U, 1 \rangle = \dim U^G$. By Lemma 6.24, $U^G = \mathrm{Hom}_G(V, W)$. By Lemma 6.25, $\langle \chi_V, \chi_W \rangle = \langle \underline{1, \chi_U} \rangle$. This already shows that $\langle \chi_V, \chi_W \rangle = \dim \mathrm{Hom}_G(V, W)$. Since $\langle \chi_V, \chi_W \rangle$ is an integer, $\langle \chi_W, \chi_V \rangle = \overline{\langle \chi_V, \chi_W \rangle} = \langle \chi_V, \chi_W \rangle$. □

Proposition 6.27, which is very interesting in its own right, allows us to prove orthogonality of irreducible characters.

Proof of Part 1 of Theorem 6.7. Follows from Proposition 6.27 combined with Schur's lemma (more precisely, Lemma 5.7). □

6.6 Proof that characters form a basis

To prove Theorem 6.7, it remains to prove that every class function is a $\mathbb{C}$-linear combination of characters. When we will discuss representation theory of semisimple algebras, there will be a very natural proof of this fact. Here we content ourselves with the following slightly magical argument.

Lemma 6.28. *Let $f \colon G \to \mathbb{C}$ be a class function and $\rho \colon G \to \mathrm{GL}(V)$ a representation. Then*

$$\sum_{g \in G} f(g) \rho(g) \in \mathrm{End}(V) \tag{6.6.1}$$

is a G-linear endomorphism, in other words lies in $\mathrm{End}_G(V)$.

Proof. We can just check this directly. Indeed, call this homomorphism ϕ. If $h \in G$ and $v \in V$, then

$$\phi(h \cdot v) = \sum_{g \in G} f(g) g \cdot (hv) = \sum_{g \in G} f(g) gh \cdot v. \tag{6.6.2}$$

Replacing g by hgh^{-1}, this equals

$$\sum_{g \in G} f(hgh^{-1}) hg \cdot v = \sum_{g \in G} f(g) hg \cdot v = h \cdot \phi(v). \tag{6.6.3}$$

□

Proposition 6.29. *Every class function $f \in \mathcal{C}(G)$ is a $\mathbb{C}$-linear combination of irreducible characters.*

Proof. Since $\langle -, - \rangle$ is a Hermitian inner product on $\mathcal{C}(G)$, it suffices to prove that the orthogonal complement of the span of the irreducible characters is trivial. So let $f \in \mathcal{C}(G)$ be a class function with $\langle \chi, f \rangle = 0$ for every irreducible character χ. We need to show that f is zero. Given a representation $\rho\colon G \to \mathrm{GL}(V)$, define

$$\phi_V \coloneqq \sum_{g \in G} \overline{f(g)} \rho(g) \in \mathrm{End}(V), \tag{6.6.4}$$

which is an element of $\mathrm{End}_G(V)$ by Lemma 6.28. We claim that ϕ_V is the zero endomorphism for all V. Indeed, by complete reducibility it suffices to prove this for irreducible V. In that case $\phi_V \in \mathrm{End}_G(V) = \mathbb{C} \cdot \mathrm{Id}$ is a scalar by Schur's lemma. So write $\phi_V = \lambda \,\mathrm{Id}$ for some $\lambda \in \mathbb{C}$. How do we know that this scalar is zero? Take the trace! Indeed, the trace of ϕ_V equals $|G| \langle \chi_V, f \rangle$, which is zero by assumption. But the trace of $\lambda \,\mathrm{Id}$ is $\lambda \dim V$. This implies that $\lambda = 0$ and so $\phi_V = 0$ for all representations V. Here comes the punchline: take $V = \mathbb{C}G$ to be the regular representation. Then the endomorphism ϕ_V sends the basis element e_1 to $\sum_{g \in G} \overline{f(g)} e_g$. Since $\phi_V = 0$ and $\{e_g\}$ forms a basis of $\mathbb{C}G$, $f(g) = 0$ for all $g \in G$. □

7 Some multilinear algebra

Lecture 7 starts here

Before doing more representation theory, we have to first introduce (or recall) some concepts from multilinear algebra. In the first three sections, we will work over $\mathbb{C}$ but everything works over a general field of characteristic zero.

7.1 Tensor products

Let U, V, W be vector spaces over a field F.

Definition 7.1. *A map $b\colon V \times W \to U$ is bilinear if for every $v \in V$ and $w \in W$, the maps $b(v, -)\colon W \to U$ and $b(-, w)\colon V \to U$ are linear.*

Informally speaking, the tensor product $V \otimes W$ of V and W is the vector space such that *linear* maps $V \otimes W \to U$ naturally correspond to *bilinear* maps $V \times W \to U$.

We may concretely construct $V \otimes W$ as follows:

Definition 7.2. *Let V and W be vector spaces with bases $\{e_1, \ldots, e_m\}$ and $\{f_1, \ldots, f_n\}$, we define the tensor product of V and W to be the vector space $V \otimes W$ with basis given by the formal symbols $\{e_i \otimes f_j \mid 1 \le i \le m, 1 \le j \le n\}$. For $v = \sum_{i=1}^m x_i e_i$ and $w = \sum_{j=1}^n y_j f_j$ we define $v \otimes w$ by extending $(-) \otimes (-)$ bilinearly:*

$$v \otimes w = \sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} x_i y_j (e_i \otimes f_j) \in V \otimes W. \tag{7.1.1}$$

We will usually leave the summation indices and ranges explicit in what follows. By definition, $\dim(V \otimes W) = (\dim V) \cdot (\dim W)$.

Example 7.3. *If $V = \mathbb{C}^2$ and $W = \mathbb{C}^3$, then every element of $V \otimes W$ is of the form $x_{11} e_1 \otimes f_1 + x_{12} e_1 \otimes f_2 + x_{13} e_1 \otimes f_3 + x_{21} e_2 \otimes f_1 + x_{22} e_2 \otimes f_2 + x_{23} e_1 \otimes f_3$. We may view $V \otimes W$ as the space of 2×3-matrices,* *mapping such an element to the matrix* $\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}$*. What do the elements of the form* $v \otimes w$ *look like? Writing* $v = y_1 e_1 + y_2 e_2$ *and* $w = z_1 f_1 + z_2 f_2 + z_3 f_3$*, we compute that* $v \otimes w$ *corresponds to the matrix* $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \cdot \begin{pmatrix} z_1 & z_2 & z_3 \end{pmatrix}$*. This matrix has proportional rows, so is of rank* ≤ 1*. It follows that matrices of rank* 2 *cannot be written in this form, e.g.* $e_1 \otimes f_1 + e_2 \otimes f_2$ *is not of the form* $v \otimes w$*.*

Elements in $V \otimes W$ of the form $v \otimes w$ are called pure tensors. The above example highlights an important feature: not every element of $V \otimes W$ is a pure tensor! Of course every element is a linear combination of pure tensors.

The next proposition states the most important properties about $V \otimes W$.

Proposition 7.4. *1. The map* $V \times W \to V \otimes W, (v, w) \mapsto v \otimes w$ *is bilinear.*

2. If $\{v_1, \ldots, v_m\}$ *and* $\{w_1, \ldots, w_n\}$ *are bases of* V *and* W*, then* $\{v_i \otimes w_j\}$ *is a basis for* $V \otimes W$*.*

Proof. 1. This is very formal; we will check that $(v + v') \otimes w = v \otimes w + v' \otimes w$, leaving the other parts as an exercise. Write $v = \sum x_i e_i$, $v' = \sum x'_i e_i$ and $w = \sum y_j f_j$. Then

$$\begin{aligned} (v + v') \otimes w &= \left(\sum (x_i + x'_i) e_i\right) \otimes \left(\sum y_j f_j\right) &(7.1.2)\\ &= \sum_{i,j} (x_i + x'_i) y_j (e_i \otimes f_j) &(7.1.3)\\ &= \sum_{i,j} x_i y_j (e_i \otimes f_j) + \sum_{i,j} x'_i y_j (e_i \otimes f_j) &(7.1.4)\\ &= v \otimes w + v' \otimes w. &(7.1.5) \end{aligned}$$

2. Write $e_k = \sum_i a_{ik} v_i$ and $f_l = \sum_j b_{jl} w_j$. By Part 1, $e_k \otimes f_l = \sum_{i,j} a_{ik} b_{jl} (v_i \otimes w_j)$. It follows that the span of the set $\{v_i \otimes w_j\}$ contains all elements $e_k \otimes f_l$, so this set spans $V \otimes W$. Since $\{v_i \otimes w_j\}$ has size $mn = \dim(V \otimes W)$, this is indeed a basis.

□

Remark 7.5 (Can be safely ignored)**.** *We can use Proposition 7.4 to define the tensor product of* V *and* W *without choosing bases, as follows. Let* $F(V \times W)$ *be the (infinite-dimensional) vector space with basis given by the formal symbols* $\{e_{v,w} \mid v \in V, w \in W\}$*. Let* Z *be the subspace of* $F(V \times W)$ *spanned by the following relations for all* $v, v' \in V, w, w' \in W, \lambda, \mu \in F$*:*

$$\begin{aligned} &e_{\lambda v + \mu v', w} - \lambda e_{v,w} - \mu e_{v',w}, &(7.1.6)\\ &e_{v, \lambda w + \mu w'} - \lambda e_{v,w} - \mu e_{v,w'}. &(7.1.7) \end{aligned}$$

Then we may define the tensor product of V *and* W *as the quotient space* $F(V \times W)/Z$*, and for every* $v \in V$*,* $w \in W$ *we let* $v \otimes w$ *to be the image of* $e_{v,w} \in F(V \times W)$ *in this quotient space. It can be checked that this produces a vector space isomorphic to our original definition. In fact, the map* $F(V \times W) \to V \otimes W$ *defined by sending* $e_{v,w}$ *to* $v \otimes w$ *has the subspace* Z *in its kernel (by Part 1 of Proposition 7.4), hence induces a map* $F(V \times W)/Z \to V \otimes W$*. Since* $V \otimes W$ *is spanned by pure tensors, this map is surjective. Moreover using the relations of* Z *the set* $\{e_i \otimes f_j\}$ *spans* $F(V \times W)/Z$*. By dimension considerations, the map must be an isomorphism.*

7.2 Basis properties of tensor products

The next proposition makes the relationship between tensor products and bilinear maps precise. It gives an easy way to construct maps *out of* tensor products. It is called the universal property of tensor products.

Proposition 7.6. *If V, W, U are vector spaces, the map*

$$\mathrm{Hom}(V \otimes W, U) \to \{\textit{Bilinear maps } V \times W \to U\} \tag{7.2.1}$$

$$(\phi\colon V \otimes W \to U) \mapsto b_\phi(v, w) := \phi(v \otimes w) \tag{7.2.2}$$

is an isomorphism of vector spaces. (The right hand side is a vector space in the obvious way.)

Proof. Since $(v, w) \mapsto v \otimes w$ is bilinear and ϕ is linear, b_ϕ is bilinear. Therefore the association $\phi \mapsto b_\phi$ is well-defined. Recall that $V \otimes W$ has basis $\{e_i \otimes f_j\}$. Moreover giving a bilinear form $b\colon V \times W \to U$ is the same as specifying the elements $b(e_i, f_j) \in U$, and any such specification gives rise to a bilinear form. The last two sentences quickly imply the proposition. To be completely explicit (so that we are all happy), if $\phi\colon V \otimes W \to U$ is such that $b_\phi = 0$, then $b_\phi(e_i, f_j) = \phi(e_i \otimes f_j) = 0$ for all i, j. Since $\{e_i \otimes f_j\}$ forms a basis of $V \otimes W$, this implies that $\phi = 0$, so the map $\phi \mapsto b_\phi$ is injective. To prove that it is surjective, let $b\colon V \times W \to U$ be an arbitrary bilinear map. Let $\phi\colon V \otimes W \to U$ be the unique linear map satisfying $\phi(e_i \otimes f_j) = b(e_i, f_j)$ for all i, j. Then $b_\phi = b$, since both bilinear forms agree on the pairs (e_i, f_j). Therefore $\phi \mapsto b_\phi$ is surjective, proving the proposition. □

We conclude that writing down a linear map $V \otimes W \to U$ is the same as writing down a bilinear map $V \times W \to U$. Let's see this concretely in action.

Given two linear maps $\alpha\colon V \to V'$ and $\beta\colon W \to W'$, the map $V \times W \to V' \otimes W', (v, w) \mapsto \alpha(v) \otimes \beta(w)$ is bilinear, so by Proposition 7.6 corresponds to a linear map $\alpha \otimes \beta\colon V \otimes W \to V' \otimes W'$.

Definition 7.7. *The linear map $\alpha \otimes \beta\colon V \otimes W \to V' \otimes W'$ constructed above is called the tensor product of α and β.*

This has all the nice properties you would expect, like $(\alpha' \otimes \beta') \circ (\alpha \otimes \beta) = (\alpha' \circ \alpha) \otimes (\beta' \circ \beta)$. In category theory lingo, we say that the tensor product is *functorial*.

Proposition 7.8. *There is a 'natural' isomorphism $V^* \otimes W \simeq \mathrm{Hom}(V, W)$.*

Proof. We will just define the linear map $V^* \otimes W \to \mathrm{Hom}(V, W)$, leaving the verification that it is an isomorphism as an exercise. By Proposition 7.6, we just need to write down a bilinear map $V^* \times W \to \mathrm{Hom}(V, W)$. Here is one: given $(\alpha, w) \in V^* \times W$, let $f_{\alpha,w}\colon V \to W$ be the linear map defined by $f_{\alpha,w}(v) = \alpha(v)w$. To verify that it is an isomorphism, choose bases and think about elementary matrices! □

We briefly indicate how to take tensor products of n vector spaces $V_1, \ldots, V_n$. If $B_1, \ldots, B_n$ are bases of $V_1, \ldots, V_n$ then $V_1 \otimes \cdots \otimes V_n$ is the vector space with basis $\{b_1 \otimes \cdots \otimes b_n \mid b_i \in B_i\}$. We may similarly define $v_1 \otimes \cdots \otimes v_n \in V_1 \otimes \cdots \otimes V_n$ if $v_i \in V_i$ by writing each v_i as a linear combination of elements in B_i and expanding linearly in each variable. We then analogously have a bijective correspondence:

$$\left\{\begin{matrix}\text{Linear maps} \\ V_1 \otimes \cdots \otimes V_n \to U\end{matrix}\right\} \leftrightarrow \left\{\begin{matrix}\text{Multilinear maps} \\ V_1 \times \cdots \times V_n \to U\end{matrix}\right\} \tag{7.2.3}$$

(A multilinear map is a map that this linear in every variable when all the others are fixed.) We also have natural isomorphisms $(V_1 \otimes V_2) \otimes V_3 \simeq V_1 \otimes (V_2 \otimes V_3) \simeq V_1 \otimes V_2 \otimes V_3$, so we will not write the brackets anymore.

7.3 Symmetric and exterior powers

It is especially interesting to take the tensor product of a vector space with itself.

Definition 7.9. *For $n \geq 1$ the nth tensor power of V is*

$$V^{\otimes n} := \underbrace{V \otimes \cdots \otimes V}_{n \text{ times}}. \tag{7.3.1}$$

By convention, we set $V^{\otimes 0} = \mathbb{C}$. We can break up this space into smaller pieces. To see the general picture, let us first treat the simplest interesting case $n = 2$. The assignment $v \otimes w \mapsto w \otimes v$ uniquely extends to an involution $\tau\colon V^{\otimes 2} \to V^{\otimes 2}$. As with every involution on a vector space, we can consider its ± 1-eigenspaces:

Definition 7.10. *The symmetric and exterior square of V are respectively:*

$$\operatorname{Sym}^2 V := \{x \in V^{\otimes 2} \mid \tau(x) = x\}, \tag{7.3.2}$$

$$\bigwedge^2 V := \{x \in V^{\otimes 2} \mid \tau(x) = -x\}. \tag{7.3.3}$$

We have

$$V^{\otimes 2} = \operatorname{Sym}^2 V \oplus \bigwedge^2 V. \tag{7.3.4}$$

This parallels the decomposition of a square matrix as a sum of a symmetric and anti-symmetric matrix!

We can write down bases for $\operatorname{Sym}^2 V$ and $\bigwedge^2 V$. If $e_1, \ldots, e_m$ is a basis of V, then $\{e_i \otimes e_j \mid 1 \leq i, j \leq m\}$ is a basis of $V^{\otimes 2}$, and $\tau(e_i \otimes e_j) = e_j \otimes e_i$. It follows that $\operatorname{Sym}^2 V$ has basis

$$\{e_i \otimes e_j + e_j \otimes e_i \mid 1 \leq i \leq j \leq m\} \tag{7.3.5}$$

and $\bigwedge^2 V$ has basis

$$\{e_i \otimes e_j - e_j \otimes e_i \mid 1 \leq i < j \leq m\}. \tag{7.3.6}$$

We now move on to general n. Every $\sigma \in S_n$ induces a linear map $\sigma\colon V^{\otimes n} \to V^{\otimes n}$ via

$$v_1 \otimes \cdots \otimes v_n \mapsto v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)} \tag{7.3.7}$$

and extending linearly.

Lemma 7.11. *This defines a (left) linear action of S_n on $V^{\otimes n}$.*

Proof. The inverse in the formula ensures that this is a left action instead of a right action. Note that for each $\sigma \in S_n$ the association (7.3.7) is multilinear, so indeed σ defines a linear map. For $\sigma, \tau \in S_n$, let $w_i = v_{\tau^{-1}(i)}$. Then by definition

$$\begin{aligned}
\sigma \cdot (\tau \cdot (v_1 \otimes \cdots \otimes v_n)) &= \sigma \cdot (w_1 \otimes \cdots \otimes w_n) && (7.3.8)\\
&= w_{\sigma^{-1}(1)} \otimes \cdots \otimes w_{\sigma^{-1}(n)} && (7.3.9)\\
&= v_{\tau^{-1}(\sigma^{-1}(1))} \otimes \cdots \otimes v_{\tau^{-1}(\sigma^{-1}(n))} && (7.3.10)\\
&= v_{(\sigma\tau)^{-1}(1)} \otimes \cdots \otimes v_{(\sigma\tau)^{-1}(n)} && (7.3.11)\\
&= (\sigma\tau)(v_1 \otimes \cdots \otimes v_n). && (7.3.12)
\end{aligned}$$

So $\sigma \cdot (\tau \cdot x) = (\sigma\tau) \cdot x$ if $x \in V^{\otimes n}$ is a pure tensor. By linearity, this identity holds for all $x \in V^{\otimes n}$. □

Conclusion: $V^{\otimes n}$ defines a representation of S_n! We know from §5.3 that we can decompose any representation into its isotypic components indexed by the irreducible representations of the group. Doing this for the trivial representation and sign representation $\mathrm{sgn}\colon S_n \to \{\pm 1\}$ of S_n, we get:

Definition 7.12. *The symmetric and exterior nth power of V are defined as*

$$\operatorname{Sym}^n V = \{x \in V^{\otimes n} \mid \sigma(x) = x, \forall \sigma \in S_n\}, \tag{7.3.13}$$

$$\bigwedge^n V = \{x \in V^{\otimes n} \mid \sigma(x) = sgn(\sigma)x, \forall \sigma \in S_n\}. \tag{7.3.14}$$

If $n \geq 3$, then in general the inclusion $\operatorname{Sym}^n V \oplus \bigwedge^n V \subset V^{\otimes n}$ is strict, so the situation is more complicated than (7.3.4). In fact, armed with our knowledge of representation theory, we know exactly why the situation is more complicated: S_n has more representations than 1 and sgn if $n \geq 2$! This is the starting point of the fascinating theory of Schur functors, something we might cover when talking about representation theory of S_n.

Sometimes $\operatorname{Sym}^n V$ and $\bigwedge^n V$ are defined as quotients of $V^{\otimes n}$, instead of subspaces. In fact, the quotient definition is the correct definition when working over a field of positive characteristic. Here we can compare the two as follows. If $v_1, \ldots, v_n \in V$, define

$$v_1 \bullet \cdots \bullet v_n \coloneqq \frac{1}{n!} \sum_{\sigma \in S_n} \sigma \cdot (v_1 \otimes \cdots \otimes v_n), \tag{7.3.15}$$

$$v_1 \wedge \cdots \wedge v_n \coloneqq \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)(\sigma \cdot (v_1 \otimes \cdots \otimes v_n)). \tag{7.3.16}$$

We can directly check that $v_1 \bullet \cdots \bullet v_n \in \operatorname{Sym}^n V$ and $v_1 \wedge \cdots \wedge v_n \in \bigwedge^n V$. Since these expressions are multilinear in the variables $v_1, \ldots, v_n$, they induce linear maps

$$\pi_S \colon V^{\otimes n} \to \operatorname{Sym}^n V, \tag{7.3.17}$$

$$\pi_A \colon V^{\otimes n} \to \bigwedge^n V. \tag{7.3.18}$$

('S' for symmetric, 'A' for alternating.)

Lemma 7.13. *The maps $\pi_S \colon V^{\otimes n} \to \operatorname{Sym}^n V$ and $\pi_A \colon V^{\otimes n} \to \bigwedge^n V$ are projectors.*

Proof. Recall that this just means that $\pi_S(x) = x$ for all $x \in \operatorname{Sym}^n V$ and $\pi_A(x) = x$ for all $x \in \bigwedge^n V$. But this is obvious from the formula $\pi_S(x) = \frac{1}{n!}\sum_\sigma \sigma \cdot x$ for all $x \in V^{\otimes n}$. Similarly $\pi_A(x) = \frac{1}{n!}\sum_\sigma(\mathrm{sgn}(\sigma))\sigma \cdot x$ so $\pi_A(x) = x$ for all $x \in \bigwedge^n V$. $\square$

Note that this lemma is the reason why we include the factor $1/n!$ in (7.3.15) and (7.3.16). It follows that the maps π_S and π_A are surjective and hence realize $\operatorname{Sym}^n V$ and $\bigwedge^n V$ as quotients of $V^{\otimes n}$. These maps are also useful in writing down bases:

Proposition 7.14. *Let $\{e_1, \ldots, e_m\}$ be a basis of V. Then*

$$\{e_{i_1} \bullet \cdots \bullet e_{i_n} \mid 1 \leq i_1 \leq \cdots \leq i_n \leq m\} \tag{7.3.19}$$

is a basis of $\operatorname{Sym}^n V$, and

$$\{e_{i_1} \wedge \cdots \wedge e_{i_n} \mid 1 \leq i_1 < \cdots < i_n \leq m\} \tag{7.3.20}$$

is a basis of $\bigwedge^n V$.

Proof. By Lemma 7.13, π_S is surjective so $\operatorname{Sym}^n V$ is generated by $S = \{e_{i_1} \bullet \cdots \bullet e_{i_n} \mid 1 \leq i_1, \ldots, i_n \leq m\}$. However, not all these elements are linearly indepedendent. Indeed, note that $e_{i_1} \bullet \cdots \bullet e_{i_n} = e_{\sigma(i_1)} \bullet \cdots \bullet e_{\sigma(i_n)}$, so we only need to keep those elements of S that satisfy $i_1 \leq \cdots \leq i_n$. Using the fact that $\{e_{i_1} \otimes \cdots \otimes e_{i_n} \mid 1 \leq i_1, \ldots, i_n \leq m\}$ is a basis of $V^{\otimes n}$, it is now easy to see that any two elements in (7.3.19) are distinct and linearly independent, hence forming a basis.

The analysis for $\bigwedge^n V$ is similar: by Lemma 7.13, $\bigwedge^n V$ is generated by $S = \{e_{i_1} \wedge \cdots \wedge e_{i_n} \mid 1 \leq i_1, \ldots, i_n \leq m\}$. Using the definition (7.3.16), we can calculate (exercise!) that $\sigma(e_{i_1} \wedge \cdots \wedge e_{i_n}) = \operatorname{sgn}(\sigma) e_{i_1} \wedge \cdots \wedge e_{i_n}$. In particular, if two indices i_a and i_b are equal, then letting $\sigma = (a\, b)$ be the transposition swapping a and b shows that $e_{i_1} \wedge \cdots \wedge e_{i_n} = 0$. Therefore, we only need to keep those elements of S such that all the $i_1, \ldots, i_n$ are distinct and up to reordering (which only changes the sign) we may assume that $i_1 < \cdots < i_n$. This shows that the set of (7.3.20) spans $\bigwedge^n V$, and it is easy to check that all the elements are linearly independent. □

We remark that every map $\alpha \colon V \to V$ induces linear maps $\operatorname{Sym}^n(\alpha) \colon \operatorname{Sym}^n V \to \operatorname{Sym}^n V$ and $\bigwedge^n(\alpha) \colon \bigwedge^n V \to \bigwedge^n V$.

We also remark that Proposition 7.14 shows that $\bigwedge^m V$ (where $m = \dim V$) is one-dimensinoal. Therefore every linear map $\alpha \colon V \to V$ induces a linear map $\wedge^m(\alpha) \colon \bigwedge^m V \to \bigwedge^m V$ on a one-dimensional vector space. On Problem set 3, you will show that this map is given by multiplication by the determinant of α!

Remark 7.15. *It is very good to think of* $\operatorname{Sym}^n V$ *as homogeneous degree* n *polynomials 'in* V*'. This perspective might make the basis of* $\operatorname{Sym}^n V$ *from Proposition 7.14 more transparent.*

Remark 7.16 (Can be safely ignored)**.** *There is a universal property similar to Proposition 7.6 for* $\operatorname{Sym}^n V$ *and* $\bigwedge^n V$*. Say a multilinear map* $b \colon V \times \cdots \times V \to U$ *is* symmetric *if* $b(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = b(v_1, \ldots, v_n)$ *for all* $\sigma \in S_n$*. Then the restriction*

$$\operatorname{Hom}(\operatorname{Sym}^n V, U) \to \left\{ \begin{matrix} \textit{Symmetric multilinear maps} \\ V \times \cdots \times V \to U \end{matrix} \right\} \tag{7.3.21}$$

$$(\phi \colon \operatorname{Sym}^n V \to U) \mapsto b_\phi(v_1, \ldots, v_n) \coloneqq \phi(v_1 \bullet \cdots \bullet v_n) \tag{7.3.22}$$

is an isomorphism. Similarly a multilinear map is alternating *if* $b(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = sgn(\sigma) b(v_1, \ldots, v_n)$ *for all* $\sigma \in S_n$*. Then the restriction*

$$\operatorname{Hom}(\bigwedge^n V, U) \to \left\{ \begin{matrix} \textit{Alternating multilinear maps} \\ V \times \cdots \times V \to U \end{matrix} \right\} \tag{7.3.23}$$

$$(\phi \colon \bigwedge^n V \to U) \mapsto b_\phi(v_1, \ldots, v_n) \coloneqq \phi(v_1 \wedge \cdots \wedge v_n) \tag{7.3.24}$$

is again an isomorphism.

7.4 Tensor products of representations

Lecture 8 starts here

Let G be a finite group and $\rho \colon G \to \mathrm{GL}(V)$ and $\rho' \colon G \to \mathrm{GL}(W)$ be representations. Then we can give $V \otimes W$ the structure of a representation by mapping g to $\rho(g) \otimes \rho'(g) \in \mathrm{GL}(V \otimes W)$. In other words, the linear G-action on $V \otimes W$ is determined by specifying that

$$g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w) \tag{7.4.1}$$

for all $g \in G, v \in V, w \in W$. It can be readily checked (do this!) that this really is a linear G-action. This follows from the fact that taking tensor products of linear maps (as in Definition 7.7) behaves well with respect to composition.

Remark 7.17. *We emphasize again that (in general) not every element of $V \otimes W$ is of the form $v \otimes w$ and so Equation* (7.4.1) *should be interpreted as follows: there exists a unique linear G-action on $V \otimes W$ such that* (7.4.1) *holds; this can be justified using Proposition 7.6.*

Lemma 7.18. $\chi_{V \otimes W} = \chi_V \chi_W$.

This shows that the product of two characters is a character!

Proof. Let $g \in G$. Since g is diagonalizable, we may assume that the bases $\{e_i\}$ and $\{f_j\}$ are eigenvectors for g. If $g \cdot e_i = \lambda_i e_i$ and $g \cdot f_j = \mu_j$, then $g(e_i \cdot f_j) = \lambda_i \mu_j (e_i \otimes f_j)$. So $\chi_{V \otimes W}(g) = \sum_{i,j} \lambda_i \mu_j = \chi_V(g)\chi_W(g)$. $\square$

We have seen on Problem set 1 that the tensor product of an irreducible representation and a one-dimensional representation is still irreducible. However, for general irreducible representations V, W, the representation $V \otimes W$ is usually highly reducible. For example, if $V = W$ then $V \otimes V$ decompoes into subrepresentations $\bigwedge^2 V$ and $\mathrm{Sym}^2 V$ (see next section).

Lemma 7.19. *There exists an isomorphism of G-representations $\mathrm{Hom}(V, W) \simeq V^* \otimes W$.*

Proof. This follows from the fact that these representations have the same character: combine Lemmas 6.25 and 7.18 and the fact that $\chi_{V^*} = \overline{\chi_V}$. Alternatively, we can observe that the 'natural' isomorphism from Proposition 7.8 is G-equivariant. $\square$

7.5 Symmetric/exterior powers of representations

Let V be a G-representation. Then $V^{\otimes n}$ is again a G-representation by specifying that $g(v_1 \otimes \cdots \otimes v_n) = (gv_1) \otimes \cdots \otimes (gv_n)$. Moreover, in (7.3.7) we have seen that $V^{\otimes n}$ also has a linear S_n-action. In fact, these two actions interact well:

Lemma 7.20. *The G-action and S_n-action commute with each other: for all $x \in V^{\otimes n}$, $g \in G$ and $\sigma \in S_n$,*

$$g \cdot (\sigma \cdot x) = \sigma \cdot (g \cdot x). \tag{7.5.1}$$

Proof. Since both actions are linear, it suffices to check this identity on pure tensors $x \in V^{\otimes n}$, where it immediately follows from the definition. $\square$

This observation is the beginning of a beautiful story whose central theme is called 'Schur–Weyl duality'. In our case, we conclude that G preserves the S_n-isotypic components of $V^{\otimes n}$. In particular, the subspaces $\mathrm{Sym}^n V \subset V$ and $\bigwedge^n V \subset V^{\otimes n}$ are G-invariant and hence by restriction define G-representations.

The next proposition determines their characters if $n = 2$.

Proposition 7.21. *For every $g \in G$:*

$$\chi_{\mathrm{Sym}^2 V}(g) = \frac{1}{2}\left(\chi_V(g)^2 + \chi_V(g^2)\right), \tag{7.5.2}$$

$$\chi_{\bigwedge^2 V}(g) = \frac{1}{2}\left(\chi_V(g)^2 - \chi_V(g^2)\right). \tag{7.5.3}$$

Proof. This is an explicit eigenvalue calculation. Note that since $V^{\otimes 2} = \mathrm{Sym}^2 V \oplus \bigwedge^2 V$ we have $\chi_{V^{\otimes 2}}(g) = \chi_V(g)^2 = \chi_{\mathrm{Sym}^2 V}(g) + \chi_{\bigwedge^2 V}(g)$ so we only need to prove the formula for $\chi_{\bigwedge^2 V}(g)$. Since g is diagonalizable, V has a basis of eigenvectors $e_1, \ldots, e_m$ with $g \cdot e_i = \lambda_i e_i$. Then $\bigwedge^2 V$ has a basis $\{e_i \otimes e_j - e_j \otimes e_i \mid 1 \le i < j \le m\}$, and each $e_i \otimes e_j - e_j \otimes e_i$ is an eigenvector for g with eigenvalue $\lambda_i \lambda_j$. So

$$\chi_{\bigwedge^2 V}(g) = \sum_{i<j} \lambda_i \lambda_j. \tag{7.5.4}$$

But

$$\chi_V(g)^2 = \left(\sum_i \lambda_i\right)^2 = 2\sum_{i<j} \lambda_i \lambda_j + \sum_i \lambda_i^2, \tag{7.5.5}$$

$$\chi_V(g^2) = \sum_i \lambda_i^2. \tag{7.5.6}$$

Subtracting $\chi_V(g^2)$ from $\chi_V(g)^2$ and dividing by 2 indeed gives $\chi_{\bigwedge^2 V}(g)$, as desired. $\square$

The computation of $\chi_{\bigwedge^n V}$ will be carried out on Problem set 3.

Example 7.22. *Consider the standard representation V of $G = S_3$. Let's decompose $V^{\otimes 2}$ into irreducible components. It suffices to decompose $\mathrm{Sym}^2 V$ and $\bigwedge^2 V$ into irreducible components. To do this, we may use the characters of S_3 from Example 6.16 and orthogonality relations. We get $\chi_{V^{\otimes 2}} = 1 + sgn + \chi_V$.*

7.6 Representations of $G \times H$

Using tensor products we can give a description of all the irreducible representations of $G \times H$, given that we know all the irreducible representations of G and H already.

Given a G-representation V and a H-representation W, we denote by $V \boxtimes W$ the $G \times H$-representation whose underlying vector space is $V \otimes W$ and where the linear $G \times H$-action is defined via $(g, h) \cdot (v \otimes w) = (gv) \otimes (hw)$.

Remark 7.23. *The notation $V \boxtimes W$ might seem confusing at first. As vector spaces, there is no difference between $V \boxtimes W$ and $V \otimes W$, but we use the $\boxtimes$ to indicate that we think of $V \boxtimes W$ as a representation of the product group $G \times H$.*

Remark 7.24. *If $G = H$, then the above construction defines a linear $G \times G$-action on $V \otimes W$ for every two G-representations V, W. When restricting this G-action to the 'diagonal' $G = \{(g, g) \mid g \in G\} \subset G \times G$, we recover the usual G-action on $V \otimes W$ from §7.4.*

Theorem 7.25. *Let G and H be finite groups. If V and W are irreducible representations of G and H respectively, then $V \boxtimes W$ is an irreducible representation of $G \times H$. Conversely, every irreducible representation of $G \times H$ is of this form.*

Proof. The character of $V \boxtimes W$ at (g, h) can be computed to be $\chi_V(g)\chi_W(h)$. Therefore if V' and W' are irreps of G and H respectively then

$$\langle \chi_{V\boxtimes W}, \chi_{V'\boxtimes W'} \rangle = \frac{1}{|G \times H|} \sum_{(g,h) \in G \times H} \chi_V(g)\chi_W(h)\overline{\chi_{V'}(g)}\,\overline{\chi_{W'}(h)} \tag{7.6.1}$$

$$= \left(\frac{1}{|G|}\sum_{g \in G} \chi_V(g)\overline{\chi_{V'}(g)}\right)\left(\frac{1}{|H|}\sum_{h \in H} \chi_W(h)\overline{\chi_{W'}(h)}\right) \tag{7.6.2}$$

$$= \langle \chi_V, \chi_{V'} \rangle \langle \chi_W, \chi_{W'} \rangle. \tag{7.6.3}$$

Taking $V = V'$ and $W = W'$, we see that $V \boxtimes W$ is indeed irreducible.

To show that every irreducible representation of $G \times H$ is of this form, we use a counting argument. Recall that the number of irreps of a group equals the number of conjugacy classes. The number of conjugacy classes of $G \times H$ equals the product of the number of conjugacy classes in G and those in H. the above calculation shows that $V \boxtimes W \simeq V' \boxtimes W'$ if and only if $V \simeq V'$ and $W \simeq W'$. This shows that we have produced all irreps of $G \times H$. $\square$

8 Representation theory of algebras

We will now switch our focus for a few lectures from representations of groups to representations of algebras. This viewpoint was pioneered by Emmy Noether, who clarified many proofs in the early days of representation theory. We will connect this back to groups using the so-called group algebra of a finite group G.

Representation theory of algebras is in some sense the correct setting for many questions and problems in representation theory of finite groups, for example when studying representations defined over $\mathbb{R}$ ('Schur indicators'), $\mathbb{Q}$, or fields of positive characteristic $\mathbb{F}_p$ ('modular representation theory'). It is also very useful when studying representations of 'quivers' and Lie algebras (through their universal enveloping algebras). Here we will content ourselves with giving a new proof of Part 2 of Theorem 6.7 (class functions are linear combinations of characters) and using it in some applications of representation theory.

8.1 Basics of algebras

Because everything in this section is completely formal, we will assume that F is a general field for now. It will be useful in the future (when talking about Lie groups) to have these notions for $F = \mathbb{R}$.

Definition 8.1. *An associative unital algebra over a field F is a vector space A over F together with a bilinear map $A \times A \to A, (a, b) \mapsto a \cdot b$ such that*

1. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ *for all $a, b, c \in A$. (Associative)*
2. *There exists a (necessarily unique) $1 \in A$ such that $1 \cdot a = a \cdot 1 = a$ for all $a \in A$. (Unital)*

We will usually simply call A an algebra. A linear map between algebras $f : A \to B$ is called an algebra homomorphism if $f(1) = 1$ and $f(a_1 a_2) = f(a_1) f(a_2)$ for all $a_1, a_2 \in A$.

We do not require A to be commutative!

Example 8.2. *Let $A = \mathrm{Mat}_n(F)$ be the algebra of $n \times n$-matrices equipped with matrix multiplication. Then A is an F-algebra. To phrase this without choosing bases, if V is a vector space over F, $\mathrm{End}(V)$ is an F-algebra under composition.*

Example 8.3. *If A_1, A_2 are F-algebras, then the direct product vector space $A_1 \times A_2 = \{(a_1, a_2) \mid a_i \in A_i\}$ is an F-algebra under pointwise composition $(a_1, a_2) \cdot (b_1, b_2) := (a_1 b_1, a_2 b_2)$.*

Warning 8.4. *Later in this course we will study Lie algebras. Confusingly, they are not algebras in the sense we just defined! They are typically non-associative.*

Definition 8.5. *Let A be an algebra over a field F. A left A-module is a vector space M together with a bilinear map $A \times M \to M, (a, m) \mapsto a \cdot m$ such that $1 \cdot m = m$ and $a \cdot (b \cdot m) = (a \cdot b) \cdot m$ for all $a, b \in A$ and $m \in M$.*

Similarly, a right A-module is a vector space M with a bilinear map $M \times A \to M$ such that $m \cdot 1 = m$ and $(m \cdot a) \cdot b = m \cdot (ab)$ for all $a, b \in A$ and $m \in M$.

Lemma 8.6. *Giving a left A-module is the same as giving a homomorphism of algebras $A \to \mathrm{End}(M)$, via $a \mapsto (m \mapsto a \cdot m)$.*

Proof. Exercise. □

When we don't specify left or right, an A-module will always mean a left A-module. We also call an A-module a representation of A. We say a linear map between A-modules $f \colon M \to N$ is an A-module homomorphism if $f(a \cdot m) = a \cdot f(m)$ and we write $\mathrm{Hom}_A(M, N) \subset \mathrm{Hom}(M, N)$ for the set of all A-module homomorphisms.

Example 8.7. *$A = \mathrm{Mat}_n(F)$ has an n-dimensional left module F^n, the space of column vectors on which A acts via left multiplication. Under Lemma 8.6, this corresponds to the 'tautological' identity homomorphism $\mathrm{Mat}_n(F) \to \mathrm{Mat}_n(F) = \mathrm{End}_F(F^n)$ of F-algebras. You will show on Problem Set 3 that F^n is an irreducible $\mathrm{Mat}_n(F)$-module.*

Example 8.8. *Every algebra A has a canonical representation, where $M = A$ and A acts on M via left multiplication. (In fact, A is simultaneously a left and right module. This is called an (A, A)-bimodule, but we will not use this notion.) When $A = \mathrm{Mat}_n(F)$, then this representation is isomorphic to a direct sum of n copies of the representation F^n.*

In analogy with the case of group, we have the notions of:

- An A-module isomorphism: a bijective A-module homomorphism;
- A-submodule of A-module V: subspace $W \subset V$ such that $a \cdot w \in W$ for all $a \in A, w \in W$;
- Irreducible/simple module: A-module V with no proper nonzero submodules;
- Direct sum of A-modules: of V, W are A-modules then $V \oplus W$ is an A-module via $a \cdot (v, w) = (a \cdot v, a \cdot w)$.

8.2 Schur's lemma and central characters

Lecture 9 starts here

The representation theory of algebras is more complicated than that of finite groups, but there are some similarities. For example, Schur's lemma remains true in this context:

Lemma 8.9 (Schur's lemma for algebras)**.** *Let A be an F-algebra and Let V, W be irreducible (finite-dimensional) A-representations.*

1. *Every $\phi \in \mathrm{Hom}_A(V, W)$ is either zero or invertible.*
2. *$\mathrm{End}_A(V)$ is a division algebra, i.e. every nonzero element is invertible.*
3. *If F is algebraically closed, $\mathrm{End}_A(V) = F \cdot \mathrm{Id}_V$.*

Proof. The proof is identical to the case of groups! See Theorem 5.1. □

This has the following useful corollary, similarly to Proposition 5.3 in the case of groups. We assume again that $F = \mathbb{C}$ from now on.

Definition 8.10. *Let A be a $\mathbb{C}$-algebra. Then center of A is defined as*

$$Z(A) \coloneqq \{z \in A \mid za = az,\ \forall a \in A\}. \tag{8.2.1}$$

It is a commutative subalgebra of A.

Corollary 8.11 (Central character)**.** *Let A be a $\mathbb{C}$-algebra and let $\rho\colon A \to \operatorname{End}(V)$ be an irreducible A-module. Then $\rho(Z(A))$ is contained in the scalar matrices $\{\mathbb{C} \cdot \mathrm{Id}_V\}$ of $\operatorname{End}(V)$ and hence defines a morphism of algebras*

$$\omega_V \colon Z(A) \to \mathbb{C}, \tag{8.2.2}$$

called the central character of V.

Proof. The proof is identical to that of Proposition 5.3: if $a \in Z(A)$ then the linear map $a \cdot (-)\colon V \to V$ is an A-module homomorphism, hence by Schur's lemma (Lemma 8.9) is given by multiplication by a scalar $\omega_V(a) \in \mathbb{C}$. □

In fact, by comparing traces we see that $\omega_V(a) = \operatorname{tr}(\rho(a))/\dim V$ for all $a \in Z(A)$.

Example 8.12. *If $A = \operatorname{Mat}_n(\mathbb{C})$ and $V = \mathbb{C}^n$ then on Problem set 3 you will show that V is an irreducible A-module. It follows that the center $Z(\operatorname{Mat}_n(\mathbb{C}))$ acts on $\mathbb{C}^n$ by scalar matrices. This proves that $Z(\operatorname{Mat}_n(\mathbb{C})) = \mathbb{C} \cdot \mathrm{Id}$ in a representation theoretic way. (It can also be checked explicitly using e.g. elementary matrices.)*

8.3 The group algebra

Let G be a finite group.

Definition 8.13. *The group algebra $\mathbb{C}[G]$ is the $\mathbb{C}$-algebra with basis given by symbols $\{e_g \mid g \in G\}$ and with multiplication given on basis elements by $e_g \cdot e_{g'} = e_{gg'}$, and extending linearly.*

The group algebra is important because it connects the representation theory of finite groups with the representation theory of algebras: we have a bijection[1]

$$\{\mathbb{C}[G]\text{-modules}\} \overset{1:1}{\longleftrightarrow} \{G\text{-representations}\}. \tag{8.3.1}$$

Indeed, given a $\mathbb{C}[G]$-module V, the restriction of the map $\mathbb{C}[G] \times V \to V$ to $G \times V$ defines a linear G-action on V. Conversely, given a linear G-action on V, we may extend it to a $\mathbb{C}[G]$-module by defining $(\sum_g c_g e_g) \cdot m \coloneqq \sum_g c_g (g \cdot m)$. It can be checked that this indeed defines a left $\mathbb{C}[G]$-module structure on V.

Remark 8.14. *Note that the regular representation of G is precisely the group algebra $\mathbb{C}[G]$, seen as a left module over itself. (as in Example 8.8.)*

We will now show, perhaps surprisingly, that $\mathbb{C}[G]$ is always a product of matrix algebras. Let $V_1, \ldots, V_m$ be a set of representatives of the isomorphism classes of irreducible representations of G. Write $d_i = \dim V_i$.

Theorem 8.15. *There is an isomorphism of $\mathbb{C}$-algebras*

$$\mathbb{C}[G] \simeq \operatorname{Mat}_{d_1}(\mathbb{C}) \times \cdots \times \operatorname{Mat}_{d_m}(\mathbb{C}). \tag{8.3.2}$$

[1] for the categorically minded: an isomorphism of categories

Proof. The G-action on V_i extends to a left $\mathbb{C}[G]$-module structure on V_i. In other words, we have for each i a homomorphism of algebras $\mathbb{C}[G] \to \operatorname{End}(V_i)$. We therefore get a natural morphism of algebras $\phi\colon \mathbb{C}[G] \to \operatorname{End}(V_1) \times \cdots \times \operatorname{End}(V_m)$. We claim that ϕ is injective. Indeed, if $x \in \mathbb{C}[G]$ has $\phi(x) = 0$, then $x \cdot v_i = 0$ for all $v_i \in V_i$ and all $1 \leq i \leq m$. In particular, by complete reducibility, $x \cdot v = 0$ for all G-representations V and $v \in V$. Taking $V = \mathbb{C}[G]$ to be the regular representation and $v = e_1 = 1$, we see that $0 = x \cdot v = x \cdot e_1 = x$, so $x = 0$, showing that ϕ is indeed injective. Since ϕ is injective, it suffices to show that the dimensions of the domain and target of ϕ are equal. The dimension of $\mathbb{C}[G]$ is $|G|$. On the other hand, the dimension of $\operatorname{Mat}_{d_1}(\mathbb{C}) \times \cdots \times \operatorname{Mat}_{d_m}(\mathbb{C})$ is $d_1^2 + \cdots + d_m^2$, which also equals $|G|$ by Theorem 6.12. $\square$

Remark 8.16. *A $\mathbb{C}$-algebra is* semisimple *if every left A-module is a direct sum of simple or irreducible modules. What is really going on here is that Maschke's theorem implies that the algebra $\mathbb{C}[G]$ is* semisimple*, and that the Artin–Wedderburn theorem states that every finite-dimensional semisimple $\mathbb{C}$-algebra is a product of matrix algebras. Since we will not use these concepts in what follows, we have chosen to not introduce them in detail here, but any serious algebra student should be aware of them!*

Example 8.17. *If G is abelian, then every irreducible representation of G is one-dimensional and $\mathbb{C}[G] \simeq \mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}$. If $G = S_3$, then G has an irrep of dimension $1, 1, 2$ so $\mathbb{C}[S_3] \simeq \mathbb{C} \times \mathbb{C} \times \operatorname{Mat}_2(\mathbb{C})$.*

Remark 8.18. *Not every algebra is a group algebra. For example, $\operatorname{Mat}_2(\mathbb{C})$ is not isomorphic to $\mathbb{C}[G]$ for any group G, because a group algebra always has a one-dimensional trivial representation.*

8.4 The center of the group algebra

Let G be a finite group with conjugacy classes $\mathcal{C}_1, \ldots, \mathcal{C}_k$. Let $C_i = \sum_{g \in \mathcal{C}_i} e_g \in \mathbb{C}[G]$.

Proposition 8.19. *The center $Z(\mathbb{C}[G])$ has $\mathbb{C}$-basis given by $C_1, \ldots, C_k$.*

The center of the group algebra is not the same as the group algebra of the center!

Proof. Let $x = \sum c_g e_g \in \mathbb{C}[G]$. Since $\{e_h \mid h \in G\}$ is a basis of $\mathbb{C}[G]$, $x \in Z(\mathbb{C}[G])$ if and only if $e_h \cdot x = x \cdot e_h$ for all $h \in G$. This is true if and only if $x = e_{h^{-1}} \cdot x \cdot e_h = \sum_g c_g e_{h^{-1}gh}$ for all $h \in G$. Replacing g by hgh^{-1}, the last expression equals $\sum_g c_{hgh^{-1}} e_g$. Comparing coefficients, we see that $x \in Z(\mathbb{C}[G])$ if and only if $c_g = c_{hgh^{-1}}$ for all $g, h \in G$. In other words, x is central if and only if $g \mapsto c_g$ is a class function! Since characteristic functions of conjugacy classes form a basis of class functions, $Z(\mathbb{C}[G])$ has basis $C_1, \ldots, C_k$. $\square$

We can use our results on the group algebra to give a new proof of 'completeness of characters', i.e. the fact that every class function is a linear combination of irreducible characters. Indeed, by orthogonality of irreducible characters, it suffices to prove that the number of irreps up to isomorphism (call this number m) equals the number of conjugacy classes (call this number k) of G. But both of these numbers equals the dimension of $Z(\mathbb{C}[G])$. Indeed, Proposition 8.19 implies that $\dim Z(\mathbb{C}[G]) = k$. On the other hand, Theorem 8.15 implies that

$$Z(\mathbb{C}[G]) \simeq Z(\operatorname{Mat}_{d_1}(\mathbb{C}) \times \cdots \times \operatorname{Mat}_{d_m}(\mathbb{C})). \tag{8.4.1}$$

Using the fact that $Z(A \times B) = Z(A) \times Z(B)$ and that $Z(\operatorname{Mat}_n(\mathbb{C})) = \mathbb{C}$ (Example 8.12), we see that $\dim Z(\mathbb{C}[G]) = m$, showing that indeed $k = m$. The careful reader can check that this argument is not circular! To prove that $\mathbb{C}[G]$ is a product of matrix algebras, we have used the identity $d_1^2 + \cdots + d_m^2 = |G|$ (Theorem 6.12), but the proof of this identity does not rely on the completeness of characters.

9 Integrality in the group algebra and applications

We will now use the group algebra together with the notion of algebraic integers to prove some surprising properties about representations. Firstly, we show that the degree of a representation always divides the order of the group G. This is highly non-obvious! Secondly, we will use it to deduce that every group of order $p^a q^b$ is solvable (Burnside's theorem).

9.1 Algebraic integers

Recall that a complex number α is algebraic if $f(\alpha) = 0$ for some monic polynomial $f \in \mathbb{Q}[x]$. For example, $\sqrt{2}$ and $i/7$ are algebraic, being solutions to the polynomials $x^2 - 2$ and $x^2 + 1/49$. It turns out that numbers like e and π are not algebraic, but that requires more work.

Definition 9.1. *A complex number $\alpha \in \mathbb{C}$ is an algebraic integer if $f(\alpha) = 0$ for some monic polynomial $f \in \mathbb{Z}[x]$ with* integer *coefficients. In other words, there exists an integer $n \geq \mathbb{Z}_{\geq 1}$ and integers $c_1, \ldots, c_n$ such that*

$$\alpha^n + c_1\alpha^{n-1} + \cdots + c_n = 0. \tag{9.1.1}$$

We write $\overline{\mathbb{Z}} \subset \mathbb{C}$ for the subset of algebraic integers.

You should think of $\overline{\mathbb{Z}}$ as a generalization of the usual integers.

Example 9.2. *We have $i \in \overline{\mathbb{Z}}$ because $i^2 + 1 = 0$ and $x^2 + 1 \in \mathbb{Z}[x]$. Also $1 + \sqrt{2} \in \overline{\mathbb{Z}}$, being the solution to $x^2 - 2x - 1 \in \mathbb{Z}[x]$. However, it turns out that $i/7 \notin \overline{\mathbb{Z}}$. (But takes a little more effort to show.) The 'problem' is the 7 in the denominator!*

Here are some essential properties of algebraic integers. We did not cover the proof in class, and you don't need to understand it, however I hope you will read it and convince yourself that it is reasonably elementary.

Proposition 9.3.

1. *$\overline{\mathbb{Z}}$ is a subring of $\mathbb{C}$. In other words, if α, β are algebraic integers, then so is $\alpha + \beta$ and $\alpha\beta$.*
2. *$\overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$. In other words, if α is both rational and an algebraic integer, then α is in fact an integer.*
3. *$\alpha \in \overline{\mathbb{Z}}$ if and only if there exists a subring $R \subset \mathbb{C}$ containing α that is a finitely generated $\mathbb{Z}$-module.*

Proof. We first show Part 3. If $\alpha \in \overline{\mathbb{Z}}$, then the subring $R = \mathbb{Z}[\alpha] = \{\sum_{i=0}^m a_i\alpha^i \mid a_i \in \mathbb{Z}\} \subset \mathbb{C}$ generated by α is a finitely generated $\mathbb{Z}$-module. Indeed, by assumption $\alpha^n + c_1\alpha^{n-1} + \cdots + c_n = 0$ for some $c_i \in \mathbb{Z}$, so α^n lies in the $\mathbb{Z}$-span of $\{1, \ldots, \alpha^{n-1}\}$. We can apply the same argument to higher powers of α and it follows that $\mathbb{Z}[\alpha]$ is generated by $\{1, \alpha, \ldots, \alpha^{n-1}\}$. Conversely, let $R \subset \mathbb{C}$ be a subring that is a finitely generated $\mathbb{Z}$-module and contains α. By the classification of finitely generated abelian groups and the fact that R is torsion-free, $R \simeq \mathbb{Z}^n$ for some $n \geq 1$. The map $\phi \colon R \to R, r \mapsto \alpha r$ is $\mathbb{Z}$-linear and after choosing a basis of R it can be represented by an $n \times n$-matrix with integer coefficients. It follows that the characteristic polynomial of this matrix is a monic polynomial $f = x^n + c_1x^{n-1} + \cdots + c_n$ with integer coefficients. By Cayley–Hamilton, we have an identity of maps $\phi^n + c_1\phi^{n-1} + \cdots + c_n = 0$. Evaluating this identity at $r = 1$, we see that $\alpha^n + c_1\alpha^{n-1} + \cdots + c_n$, showing that $\alpha \in \overline{\mathbb{Z}}$.

We can use Part 3 to prove Part 1. Indeed, if $\alpha, \beta \in \overline{\mathbb{Z}}$, then $\mathbb{Z}[\alpha]$ and $\mathbb{Z}[\beta]$ are finitely generated $\mathbb{Z}$-modules. It follows that $R = \mathbb{Z}[\alpha, \beta] = \mathbb{Z}[\alpha][\beta]$ is a finitely generated $\mathbb{Z}$-module (take products of generators). Therefore $R \subset \overline{\mathbb{Z}}$ and hence $\alpha + \beta$ and $\alpha\beta$ lie in $R \subset \overline{\mathbb{Z}}$.

Part 2 follows from the fundamental theorem of arithmetic, namely that each integer can be uniquely written as a product of primes. Write $\alpha = a/b \in \overline{\mathbb{Z}} \cap \mathbb{Q}$ for integers $a, b \in \mathbb{Z}$. If α were not an integer, there exists a prime p dividing b but not a. By assumption, $\alpha \in \overline{\mathbb{Z}}$ so $\alpha^n + c_1\alpha^{n-1} + \cdots + c_{n-1}\alpha + c_n = 0$. Writing $\alpha = a/b$ and clearing out denominators, we get $a^n + c_1 a^{n-1} b + \cdots + c_n b^n = 0$. We see that every term on the left hand side is divisible by p except a^n; this is a contradiction. □

Algebraic integers will be useful because we will sometimes be able to show that certain numbers related to characters are both algebraic integers and rational, so by the above proposition they must be integers.

9.2 Characters and algebraic integers

Lecture 10 starts here

Let G be a finite group.

Lemma 9.4. *If χ is a character of G, then $\chi(g) \in \overline{\mathbb{Z}}$ for all $g \in \mathbb{Z}$.*

Proof. Roots of unity are algebraic integers, because they are roots of monic polynomials of the form $x^d - 1 \in \mathbb{Z}[x]$. Since $\chi(g)$ is a sum of roots of unity and since $\overline{\mathbb{Z}}$ is closed under addition, this proves the lemma. □

Let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ be the conjugacy classes of G and choose representatives $g_i \in \mathcal{C}_i$.

Proposition 9.5. *Let χ be an* irreducible *character of G. Then for every $1 \leq i \leq k$ we have*

$$\frac{\chi(g_i)}{\chi(1)}|\mathcal{C}_i| \in \overline{\mathbb{Z}}. \tag{9.2.1}$$

Note that this statement is not obvious because there is a $\chi(1)$ in the denominator: we cannot always 'divide' algebraic integers and stay in $\overline{\mathbb{Z}}$.

Proof. This will follow from calculations with the center of the group algebra. Indeed, recall from Proposition 8.19 that $Z(\mathbb{C}[G])$ has $\mathbb{C}$-basis given by $C_1, \ldots, C_k$ where $C_i = \sum_{g \in \mathcal{C}_i} e_g$. Let $\rho\colon G \to \mathrm{GL}(V)$ be the representation with character χ, and let $\omega = \omega_V : Z(\mathbb{C}[G]) \to \mathbb{C}$ be the associated central character (Corollary 8.11). The star of the show in this proof is the complex number $\omega(C_i)$. We first determine this number. We have

$$\mathrm{tr}\left(\sum_{g \in \mathcal{C}_i} \rho(g)\right) = \mathrm{tr}(\omega(C_i)\,\mathrm{Id}_V). \tag{9.2.2}$$

The left hand side of this equation equals $\sum_{g \in \mathcal{C}_i} \chi(g) = \chi(g_i)|\mathcal{C}_i|$. The right hand side equals $\omega(C_i) = \dim V \omega(C_i) = \chi(1)\omega(C_i)$. It follows that

$$\omega(C_i) = \frac{\chi(g_i)}{\chi(1)}|\mathcal{C}_i|. \tag{9.2.3}$$

But we now claim that $\omega(C_i)$ is an algebraic integer! Indeed, since $\{C_1, \ldots, C_k\}$ is a basis of $Z(\mathbb{C}[G])$ we have $C_p \cdot C_q = \sum_{r=1}^k a_{pqr} C_r$ for some $a_{pqr} \in \mathbb{C}$. But a_{pqr} are in fact nonnegative integers. Indeed, by writing out C_p, C_q, C_r in the basis $\{e_g \mid g \in G\}$ we see that $a_{pqr} = \#\{(x, y) \in \mathcal{C}_p \times \mathcal{C}_q \mid xy = g_r\} \in \mathbb{Z}$. Since ω is an algebra homomorphism $\omega(C_p) \cdot \omega(C_q) = \sum_{r=1}^k a_{pqr}\omega(C_r)$. It follows that $R = \mathbb{Z}\omega(C_1) + \ldots \mathbb{Z}\omega(C_k)$ is a subring of $\mathbb{C}$ that is finitely generated as a $\mathbb{Z}$-module. By Part 3 of Proposition 9.3, this implies that $R \subset \overline{\mathbb{Z}}$ so $\omega(C_i) \in \overline{\mathbb{Z}}$. □

It is important in the above proposition that χ is irreducible. Indeed, if χ is reducible then we don't have a central character $\omega\colon Z(\mathbb{C}[G]) \to \mathbb{C}$.

We now show that the degree of an irreducible character divides the order of the group. It is very unclear a priori that this must be true!

Theorem 9.6. *Let V be an irreducible representation of G. Then $\dim V$ divides $|G|$.*

Proof. Let χ be the character of V. Expanding the identity $\langle \chi, \chi \rangle = 1$ and dividing out by $\dim V = \chi(1)$, we get

$$\sum_{g \in G} \frac{\chi(g)}{\chi(1)} \overline{\chi(g)} = \frac{|G|}{\dim V}. \tag{9.2.4}$$

We need to show that the rational number $|G| / \dim V$ is an integer. By Part 2 of Proposition 9.3, it suffices to prove that it is an *algebraic* integer. For this we look at the left hand side of (9.2.4). Since χ is a class function and $\overline{\chi(g)} = \chi(g^{-1})$, we can rewrite this left hand side as

$$\sum_{i=1}^{k} \frac{\chi(g_i)}{\chi(1)} |\mathcal{C}_i| \chi(g_i^{-1}). \tag{9.2.5}$$

Now comes the *coup de grâce*. Every term $\frac{\chi(g_i)}{\chi(1)}|\mathcal{C}_i|$ is an algebraic integer by Proposition 9.5. Moreover $\chi(g_i^{-1})$ is an algebraic integer by Lemma 9.4. By Part 1 Proposition 9.3, sums and products of algebraic integers are algebraic integers, so the whole expression (9.2.5) is an algebraic integer! $\square$

That's magic.

Example 9.7. *Let G be a group of order p^n for some prime p. Then every irreducible representation has dimension p^m for some $m \leq n$. In fact, since the sum of the squares of the dimensions equals p^n and there is always the trivial representation, we must have $p^{2m} < p^n$, so $2m < n$. So every group of order p^2 has only one-dimensional irreducible representations (i.e. is abelian) and every group of order p^3 only has irreducible representations of dimension 1 and p.*

9.3 Burnside's theorem

Recall from group theory the following theorem:

Proposition 9.8. *Let $G \neq 1$ be a group of prime power order. Then $Z(G) \neq 1$.*

In particular, the only groups of prime power order that are simple are groups of the form C_p. Burnside proved the following result using character theory:

Theorem 9.9 (Burnside)**.** *Let G be a group of order $p^a q^b$, where $p \neq q$ are prime and $a + b \geq 2$. Then G is not simple.*

By induction on the order of G, it follows that every such group is solvable. Note that the result is sharp in some sense: the smallest non-abelian simple group A_5 has order $2^2 \times 3 \times 5$.

We will use the integrality results from the previous section together with the following lemma from Galois theory, whose proof we will not cover but is included for completeness:

Lemma 9.10. *Suppose that $\lambda_1, \dots, \lambda_m$ are roots of unity and that*

$$\alpha = \frac{1}{m}\sum_{j=1}^{m} \lambda_j \tag{9.3.1}$$

is an algebraic integer. Then either $\alpha = 0$ or $|\alpha| = 1$.

Proof. This lemma follows from considering the norm of α. Suppose $\lambda_m^d = 1$ for all m. Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_d)/\mathbb{Q})$. Then $\sigma(\alpha) \in \bar{\mathbb{Z}}$ for all $\sigma \in G$, since if $f(\alpha) = 0$ for some monic $f \in \mathbb{Z}[x]$, then $f(\sigma(\alpha)) = 0$ too. Let $N(\alpha) = \prod_{\sigma\in G} \sigma(\alpha)$. Them $N(\alpha) \in \bar{\mathbb{Z}}$ and by definition $\sigma(N(\alpha)) = N(\alpha)$ for all $\sigma \in G$. Therefore by Galois theory $N(\alpha) \in \mathbb{Q}$ and so $N(\alpha) \in \bar{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$. We know that $|\alpha| \leq \frac{1}{m}\sum |\lambda_j| = 1$ and similarly $|\sigma(\alpha)| \leq 1$. So $|N(\alpha)| = \prod_\sigma |\sigma(\alpha)| \leq 1$. Since $N(\alpha) \in \mathbb{Z}$, we either have $N(\alpha) = 0$ (in which case $\alpha = 0$) or $|N(\alpha)| = 1$, so all inequalities $|\sigma(\alpha)| \leq 1$ are equalities and $|\alpha| = 1$. ☐

Lemma 9.11. *Let χ be an irreducible character of G and $\mathcal{C}$ a conjugacy class such that $\chi(1)$ and $|\mathcal{C}|$ are coprime. Then for all $g \in \mathcal{C}$ we have either $\chi(g) = 0$ or $|\chi(g)| = \chi(1)$.*

Proof. We would like to apply the previous lemma to $\alpha = \chi(g)/\chi(1)$. It remains to show that α is an algebraic integer. However, we know that $\chi(g)|\mathcal{C}|/\chi(1)$ is an algebraic integer by Proposition 9.5. Since $|\mathcal{C}|$ and p are coprime, by Bezout's theorem there exist integers a, b such that $a|\mathcal{C}| + b\chi(1) = 1$. Multiplying this equation by $\chi(g)/\chi(1)$, we see that

$$a\frac{\chi(g)}{\chi(1)}|\mathcal{C}| + \chi(g)|\mathcal{C}| = \frac{\chi(g)}{\chi(1)}. \tag{9.3.2}$$

We know that the left hand side is a sum of algebraic integers, so $\chi(g)/\chi(1)$ is indeed an algebraic integer, as desired. ☐

Theorem 9.12. *Suppose that G has a conjugacy class $\mathcal{C}$ of prime power order $p^n > 1$. Then G is not nonabelian and simple.*

Proof. Assume that G is simple and nonabelian. Let $g \in \mathcal{C}$ be a representative. By column orthogonality applied to $\{1\}$ and $\mathcal{C}$,

$$1 + \sum_{\chi\neq 1} \chi(1)\chi(g) = 0, \tag{9.3.3}$$

where the sum runs over all nontrivial irreducible characters of G. We first claim that if χ is such a character then $|\chi(g)| \neq \chi(1)$. Indeed, let $\rho\colon G \to \mathrm{GL}(V)$ be the representation with character χ. Then ρ is faithful since G is simple and $\chi \neq 1$. On Problem Set 2, you have shown that $|\chi(g)| = \chi(1)$ implies that $\rho(g) = \lambda\,\mathrm{Id}$ for some $\lambda \in \mathbb{C}^\times$. Therefore $\rho(g)$ commutes with $\rho(h)$ for all $h \in G$. Since ρ is faithful, $g \in Z(G)$, contradicting the fact G is nonabelian and simple. This proves the claim. We conclude by Lemma 9.11 that for all nontrivial irreducible characters χ, either $p \mid \chi(1)$ or $\chi(g) = 0$. So the above sum simplifies to

$$1 + \sum_{\substack{\chi\neq 1\\ p\mid\chi(1)}} \chi(1)\chi(g) = 0. \tag{9.3.4}$$

Dividing by p we get

$$-\frac{1}{p} = \sum_{\substack{\chi\neq 1\\ p\mid\chi(1)}} \frac{\chi(1)}{p}\chi(g). \tag{9.3.5}$$

The right hand side is a sum of algebraic integers, hence an algebraic integer. Therefore $-1/p$ is an algebraic integer. Since $\mathbb{Q} \cap \bar{\mathbb{Z}} = \mathbb{Z}$, this is a contradiction! ☐

Proof of Theorem 9.9. If $a = 0$ or $b = 0$ then the result follows from Proposition 9.8, so assume $a \geq 1$ and $b \geq 1$ and G nonabelian simple. Let $P \leq G$ be a Sylow-p-subgroup of G. By that same proposition, $Z(P) \neq 1$; let $x \in Z(P)$ be a nontrivial element. Then the centralizer of x contains P, so $p^a = |P|$ divides $|C_G(x)|$, which in turn divides $p^a q^b$. It follows that $|C_G(x)| = p^a q^n$ for some $0 \leq n \leq b$. If $n = b$, then $x \in Z(G)$, contradicting the fact that G is nonabelian and simple. If $n < b$, then by the orbit-stabilizer formula the conjugacy class of x has size $|G|/|C_G(x)| = q^{n-b} > 1$. By Theorem 9.12, G is not nonabelian and simple, contradiction. □

Feit and Thompson proved that every nonabelian group of odd order is not simple. The proof is hundreds of pages long.

10 Induction of representations

Lecture 11 starts here

In this last section on representation theory of finite groups, we will discuss a way to make a G-representation out of an H-representation, where $H \leq G$ is a subgroup. Since we will be only spending one lecture on this topic, we will be quite brief in places.

10.1 Motivation

Let G be a finite group and $H \leq G$ a subgroup. Given a representation V of G, the restriction of V to H is the H-representation by simply restricting the G-action on V to H. It is sometimes written as $\mathrm{Res}^G_H(V)$ to emphasize that we view it as an H-representation. We therefore have a map:

$$\mathrm{Res}^G_H\colon \{G\text{-representations}\} \to \{H\text{-representations}\}. \tag{10.1.1}$$

Induction goes to other way around: given an H-representation V, we will define a G-representation $\mathrm{Ind}^G_H(V)$, so we will have a map

$$\mathrm{Ind}^G_H\colon \{H\text{-representations}\} \to \{G\text{-representations}\}. \tag{10.1.2}$$

This is not an inverse to Res^G_H, and the underlying vector spaces of V and $\mathrm{Ind}^G_H V$ will not be the same!

As a warm-up, let's try to see if we can come up with a definition of H has index two in G. If the following paragraphs don't make sense, don't worry and just read the formal definition of induction in the next section.

Since H has index 2, $G = H \sqcup tH$ for some $t \in G$. Given an H-representation V, we would like to make a G-representation $W = \mathrm{Ind}^G_H V$ somehow. The decomposition $G = H \sqcup tH$ shows that we just need to decide how H acts on W and how t acts on W, because then $th \in tH$ acts via $t \cdot (h \cdot w)$ on $w \in W$.

Our first attempt might be to just set $W = V$ as vector spaces and try to act by G on V. It's clear how H acts on V, but what is $t \cdot v$ for $v \in V$? Without extra information, there doesn't seem to be a reasonable way to assign an element of V to $t \cdot v$. To remedy this, we simply 'formally' add all vectors of the form '$t \cdot v$' to our space! In other words, we define W to be the direct sum of two copies of V, so $W = V \oplus tV$ and we write an element of W as $v_1 + tv_2$ with $v_i \in V$.

I now claim that we can define a reasonable linear G-action on W. How does H act on this space? It acts via its defining action on the first factor $V \subset W$. To describe the action on the second factor, note that if $h \in H$, we can write $ht = th'$ and so it seems intuitively clear that $h \cdot (tv)$ should be $t(h' \cdot v)$. So if $R\colon H \to \mathrm{GL}_n(\mathbb{C})$

is a matrix representation for V then h acts on W via the block $2n \times 2n$-matrix

$$\begin{pmatrix} R(h) & 0 \\ 0 & R(t^{-1}ht) \end{pmatrix}. \tag{10.1.3}$$

How do elements $g = th \in tH$ act? We have $g \cdot v = t(h \cdot v)$. Moreover $gt = h'$ for some $h' \in H$ and so we set $g \cdot (tv) = h'v$. In other words, g will act via the block matrix

$$\begin{pmatrix} 0 & R(gt) \\ R(t^{-1}g) & 0 \end{pmatrix}. \tag{10.1.4}$$

It can be checked that W is indeed a G-representation. The next section will define induction in the general case, but using the same idea.

10.2 Defining induction

We will define induction in a hands on way. Choose representatives $t_1, \ldots, t_m \in G$ of the left cosets G/H. In other words, we have $[G : H] = m$ and

$$G = t_1H \sqcup t_2H \sqcup \cdots \sqcup t_mH. \tag{10.2.1}$$

Let V be an H-representation. We define a new vector space $\operatorname{Ind}_H^G V$ as follows:

$$\operatorname{Ind}_H^G(V) := t_1V \oplus \cdots \oplus t_mV. \tag{10.2.2}$$

Here each t_iV denotes a copy of the vector space V, and we write elements of t_iV as $t_iv, v \in V$. The notation t_iV is purely formal, as it usually doesn't make sense to act by t_i on V (since the latter only has an H-action), but it will be useful to index the copies of V by these elements t_i to describe the G-action on $\operatorname{Ind}_H^G(V)$. If $g \in G$, g acts on the left cosets G/H, so there is an element $\sigma \in S_m$ and unique elements $h_1, \ldots, h_m \in H$ such that $gt_i = t_{\sigma(i)}h_i$. If $\sum t_iv_i \in \operatorname{Ind}_H^G(V)$, we define

$$g \cdot \left(\sum_{i=1}^m t_iv_i \right) := \sum_{i=1}^m t_{\sigma(i)}(h_i \cdot v_i). \tag{10.2.3}$$

It can be checked (exercise!) that this indeed defines a linear G-action on $\operatorname{Ind}_H^G V$, called the induction of V to G. We will later see as a corollary of Proposition 10.3 that (the isomorphism class of) $\operatorname{Ind}_H^G V$ does not depend on the choice of representatives $t_1, \ldots, t_m$.

10.3 Examples

Example 10.1 (Inducing the trivial representation)**.** *If $V = \mathbb{C}$ is the trivial representation of H, then* (10.2.3) *shows that $\operatorname{Ind}_H^G V$ is a vector space with basis $t_1, \ldots, t_m$ and g acts via permuting the basis elements according to how G acts on the left cosets G/H. In other words, $\operatorname{Ind}_H^G V$ is isomorphic to the permutation representation of G acting on G/H. In particular, if $H = \{1\}$ and V is trivial then $\operatorname{Ind}_H^G V$ is the regular representation of G.*

As the above example shows, the induction of a representation will typically be reducible. However, if you pick your subgroup and representation wisely then the induction can be irreducible.

Example 10.2. *Let $G = D_5 = \{1, r, \ldots, r^4, s, rs, \ldots, r^4s\}$ and $H = \langle r \rangle \leq G$. Let $\zeta = e^{2\pi i/5}$. Then $H \to \mathbb{C}^\times, r \mapsto \zeta$ defines a one-dimensional H-representation $V = \mathbb{C}e$. Let's determine $W = \mathrm{Ind}_H^G(V)$. We may take $\{1, s\}$ as our left coset representatives, so if we set $v_1 = e$ and $v_2 = se$ then $W = \mathbb{C}v_1 \oplus \mathbb{C}v_2$. Since $rs = sr^4$, r acts via ζ on v_1 and via ζ^4 on v_2. How does s act? It sends v_1 to v_2 and v_2 to v_1. It follows that in the basis $\{v_1, v_2\}$ the action of r, s is given by*

$$r \mapsto \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^4 \end{pmatrix}, \; s \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{10.3.1}$$

From Q10 on Problem Set 1, you recognize this representation as a two-dimensional irreducible representation of D_5.

10.4 The character of induction

Proposition 10.3. *Let V be an H-representation with character $\chi \colon H \to \mathbb{C}$. Let ψ be the character of $\mathrm{Ind}_H^G(V)$. Then for all $g \in G$,*

$$\psi(g) = \sum_{\substack{1 \leq i \leq m \\ t_i^{-1} g t_i \in H}} \chi(t_i^{-1} g t_i) = \frac{1}{|H|} \sum_{\substack{x \in G, \\ xgx^{-1} \in H}} \chi(x^{-1} g x). \tag{10.4.1}$$

Proof. The action of G on $\mathrm{Ind}_H^G V$ is given by 'block' permutation matrices, where the permutation is given by the action of G on the left cosets G/H. More precisely, let $R \colon H \to \mathrm{GL}_n(\mathbb{C})$ be the permutation representation associated to some basis $\{e_1, \ldots, e_n\}$ of V. Then this basis determines a basis $\{t_i e_j\}$ of $\mathrm{Ind}_H^G V$ If $g \in G$, $\sigma \in S_m$ and $h_i \in H$ are as in (10.2.3), then g acts on $\mathrm{Ind}_H^G V$ by an $mn \times mn$ block permutation matrix, where the blocks have size $n \times n$, the permutation corresponds to σ, and the ith block is $R(t_{\sigma(i)}^{-1} g t_i)$. The only parts that contribute to ψ are those i such that $\sigma(i) = i$; in other words, those i such that $t_i^{-1} g t_i \in H$. The block $\rho(t_i^{-1} g t_i)$ has contribution $\chi(t_i^{-1} g t_i)$. It follows that

$$\psi(g) = \sum_{\substack{1 \leq i \leq m \\ t_i^{-1} g t_i \in H}} \chi(t_i^{-1} g t_i) \tag{10.4.2}$$

But χ is a class function and for $h \in H$ we have $(t_i h)^{-1} g (t_i h) \in H$ if and only if $t_i^{-1} g t_i \in H$. Therefore

$$\psi(g) = \frac{1}{|H|} \sum_{\substack{x \in G \\ x^{-1} g x \in H}} \chi(x^{-1} g x), \tag{10.4.3}$$

as claimed. ☐

This shows in particular that the isomorphism class of the G-representation $\mathrm{Ind}_H^G V$ is independent of the choice of representatives $t_1, \ldots, t_m$ of G/H, because the character of $\mathrm{Ind}_H^G V$ does not depend on this choice.

10.5 Frobenius reciprocity

Theorem 10.4 (Frobenius reciprocity)**.** *Let V be an H-representation with character χ and W a G-representation with character ψ. Then*

$$\langle \mathrm{Ind}_H^G(\chi), \psi \rangle_G = \langle \chi, \mathrm{Res}_H^G(\psi) \rangle_H. \tag{10.5.1}$$

The inner product on the left is taken in the space of class functions on G; the inner product on the right is taken in the space of class functions on H.

Proof. There is an abstract proof of this theorem using the group algebra, see Section 10.6. For a more hands on approach, we simply calculate both sides and compare the result. Indeed, by Proposition 10.3 we have

$$\langle \operatorname{Ind}_H^G(\chi), \psi \rangle_G = \frac{1}{|G||H|} \sum_{\substack{g,x\in G \\ x^{-1}gx\in H}} \chi(x^{-1}gx)\overline{\psi(g)} \tag{10.5.2}$$

Setting $h = x^{-1}gx$, then in the above sum we may instead sum over $h \in H$ and $x \in G$, so the is expression equals

$$\frac{1}{|G||H|} \sum_{h\in h, x\in G} \chi(h)\overline{\psi(xhx^{-1})} \tag{10.5.3}$$

Since ψ is a class function, summing $\overline{\psi(xhx^{-1})}$ over all $x \in G$ equals $|G|\overline{\psi(h)}$. We conclude that the above expression equals

$$\frac{1}{|H|} \sum_{h\in H} \chi(h)\overline{\psi(h)} \tag{10.5.4}$$

which equals $\langle \chi, \operatorname{Res}_H^G(\psi) \rangle_H$, as desired. □

Example 10.5. *Let V be the 2-dimensional irreducible representation of S_3 and let W be the induction of V along $S_3 \leq S_4$. Then we can determine the decomposition of W into S_4-irreducibles.*

Frobenius reciprocity is very useful when trying to determine wether the induction $\operatorname{Ind}_H^G(\chi)$ of a character of H is irreducible. Indeed, we know that

$$\langle \operatorname{Ind}_H^G(\chi), \operatorname{Ind}_H^G(\chi) \rangle_G = \langle \chi, \operatorname{Res}_H^G \operatorname{Ind}_H^G \chi \rangle \tag{10.5.5}$$

It therefore suffices to compute how often χ occurs in the character $\operatorname{Res}_H^G \operatorname{Ind}_H^G \chi$. This can be done explicitly in examples and in general leads to a subject called *Mackey theory*.

10.6 Induction using the group algebra

This section was not covered in class, but might be of interest to the more algebraically inclined. It explains how induction can be naturally viewed in the context of representations of algebras.

Let A be an (associative, unital, finite-dimensional) $\mathbb{C}$-algebra and $B \subset A$ a subalgebra. Let V be a left B-module.

Definition 10.6. *The induction of V to A, written $A \otimes_B V$, is the quotient of the tensor product $A \otimes V$ by the subspace spanned by $\{(ab) \otimes v - a \otimes (bv) \mid a \in A, b \in B, v \in V\}$. The assignment $a \cdot (a' \otimes v) = (aa') \otimes v$ induces a left A-module structure of $A \otimes_B V$.*

Let V be a B-module and W an A-module. Then W is also naturally a B-module by restriction, and the map

$$\operatorname{Hom}_A(A \otimes_B V, W) \to \operatorname{Hom}_B(V, W) \tag{10.6.1}$$
$$\phi \mapsto (v \mapsto \phi(1 \otimes v)) \tag{10.6.2}$$

is an isomorphism of vector spaces. In other words, B-homomorphisms $V \to W$ correspond to A-homomorphisms $A \otimes_B V \to W$. This is the universal property of $A \otimes_B V$.

Now let G be a finite group and $H \leq G$ a subgroup. Then we know that representations of these groups correspond to representations of the associated group algebras. So given an H-representation V, which corresponds to a $\mathbb{C}[H]$-module V, we may define the $\mathbb{C}[G]$-module

$$\mathbb{C}[G] \otimes_{\mathbb{C}[H]} V, \tag{10.6.3}$$

which corresponds to a G-representation. It can be checked that this indeed coincides with the concrete definition of $\mathrm{Ind}_H^G(V)$ given above. This gives an elegant (but not very concrete way) of defining the induction of a representation.

To see why this is useful, note that by Proposition 6.27 Frobenius reciprocity is equivalent to the statement that $\dim \mathrm{Hom}_G(\mathrm{Ind}_H^G V, W) = \dim \mathrm{Hom}_H(V, \mathrm{Res}_H^G W)$. But this follows from (10.6.1)! This gives an alternative proof of Theorem 10.4.

11 Compact groups and the Peter–Weyl theorem

Lecture 12 starts here

We leave the safe world of finite groups and take our first dive into the representation theory of infinite groups.

Most groups that arise in mathematical nature carry some extra structure, like a topology, or a differential structure, or an algebraic structure. In that case it is only natural to study those representations that are compatible with this structure. In this lecture we will see this in action for the first time for *compact* topological groups. Later we will study the representation theory of compact Lie groups in more detail via their associated Lie algebras.

11.1 Basic definitions

Definition 11.1. *A topological group is a topological space G that is also a group such that the multiplication map $G \times G \to G, (x, y) \mapsto xy$ and inversion map $G \to G, x \mapsto x^{-1}$ are both continuous.*

A homomorphism of topological groups is a continuous group homomorphism, and we say it is an isomorphism if the inverse is also a continuous group homomorphism.

As a convention, we will assume that all topological spaces (and thus topological groups) are Hausdorff, usually without further mention.

Examples 11.2.

1. *Any group can be made into a topological group by endowing it with the discrete topology (i.e. every set is open).*
2. *The group $G = \mathbb{R}$ under addition is a topological group (with the Euclidean topology on $\mathbb{R}$). So is $G = \mathbb{R}_{>0}$ under multiplication. The exponential map $\mathbb{R} \to \mathbb{R}_{>0}, t \mapsto e^t$ is an isomorphism of topological groups. Similarly $\mathbb{R}^n$ and $\mathbb{C}^n$ are topological groups.*
3. *The group $G = \mathrm{GL}_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$ and the induced topology on G turns it into a topological group. Indeed, multiplication of matrices is given by polynomials in the entries so is continuous. Inversion is continuous because of Cramer's rule: if $A \in \mathrm{GL}_n(\mathbb{R})$ then $A^{-1} = (\det(A))^{-1}adj(A)$, where the adjunct matrix $adj(A)$ has entries that are polynomial in the entries of A. Similarly $\mathrm{GL}_n(\mathbb{C})$ is a topological group.*

4. *Let $F = \mathbb{R}$ or $\mathbb{C}$ and V a (finite-dimensional) vector space over F. Then V (under addition) and $\mathrm{GL}(V)$ (under composition) are naturally topological groups. Indeed, choosing a basis of V defines group isomorphisms $V \simeq F^n$ and $\mathrm{GL}(V) \simeq \mathrm{GL}_n(F)$, and we endow V and $\mathrm{GL}(V)$ with the unique topologies such that these maps are homeomorphisms. These topologies don't depend on the choice of basis, since every other basis changes these isomorphisms by multiplication/conjugation by a matrix $X \in \mathrm{GL}_n(F)$, which is a homeomorphism.*

5. *Every subgroup of a topological group, endowed with the subspace topology, is a topological group itself. (Exercise!)*

6. *The subgroup $S^1 = \{z \in \mathbb{C}^\times \mid |z| = 1\} \leq \mathbb{C}^\times$ is closed being the pre-image of $\{1\}$ under the continuous map $\mathbb{C}^\times \to \mathbb{R}_{>0}, z \mapsto |z|$. Therefore S^1 is a topological group too. Since it is a closed and bounded subset of $\mathbb{R}^2$, it is in fact a* compact *topological group. You will show on the problem set that there is an isomorphism of topological groups $S^1 \simeq \mathbb{R}/\mathbb{Z}$.*

7. *Generalizing the previous example, we see that $\mathrm{SL}_n(\mathbb{R}), \mathrm{SL}_n(\mathbb{C}), \mathrm{SO}(n), \mathrm{U}(n), \mathrm{SU}(n), \ldots$ are all topological groups. We will consider all of these examples in detail later.*

8. *There are some more exotic topological groups out there. For example, $\mathbb{Z}_p$, $\mathrm{GL}_n(\mathbb{Z}_p)$, $\mathrm{GL}_n(\mathbb{Q}_p)$, $\mathrm{Gal}(\bar{\mathbb{Q}}|\mathbb{Q})$ are all topological groups. They are very important in number theory and the Langlands program (especially their representation theory), but we won't consider them in this course.*

In one of the above examples we have defined the topological group $\mathrm{GL}(V)$ for every finite-dimensional $\mathbb{C}$-vector space.

Definition 11.3. *A (finite-dimensional) representation of G is a continuous homomorphism $G \to \mathrm{GL}(V)$, where V is a finite-dimensional $\mathbb{C}$-vector space.*

There are obvious notions of: G-homomorphism, G-isomorphism, subrepresentation, trivial representation and irreducible representation.

It also makes sense (and is often fruitful) to think about infinite-dimensional representations. We will briefly mention this later.

11.2 Integration on topological groups

In general, topological groups can be huge and difficult, and their representation theory can be extremely complicated. However, if G is *compact* then many basis results in rep theory of finite groups will carry over! The main observation our very useful averaging procedure $\frac{1}{|G|}\sum_{g \in G}$ carries over to the compact group setting, by replacing the sum by an integral.

Let $\mathcal{C}(G, \mathbb{R})$ be the set of continuous functions $G \to \mathbb{R}$.

Theorem 11.4 (Existence of Haar integral)**.** *Let G be a compact (and Hausdorff) group. Then there exists a linear map $\int_G : \mathcal{C}(G, \mathbb{R}) \to \mathbb{R}$ satisfying the following properties:*

1. *If $f \geq 0$, then $\int_G f(g)dg \geq 0$;*
2. *For every $f \in \mathcal{C}(G, \mathbb{R})$, $\left|\int_G f\right| \leq \sup_{g \in G} |f(g)|$;*
3. *If $f = 1_G$ the function with constant value 1, then $\int_G 1_G = 1$;*

4. *We have left-translation invariance: if $t_g^* f$ denotes the function $(t_g^* f)(x) = f(gx)$, then $\int_G f = \int_G t_g^* f$ for all $f \in \mathcal{C}(G, \mathbb{R})$.*

Moreover, $\int_G$ is the unique linear map $\mathcal{C}(G, \mathbb{R}) \to \mathbb{R}$ satisfying these properties.

The integral $\int_G f$ (sometimes written $\int_G f(g)\,\mathrm{d}g$) is called the Haar integral of $f \in \mathcal{C}(G, \mathbb{R})$. By considering real and imaginary parts, we may also define the Haar integral of a continuous function $f\colon G \to \mathbb{C}$.

Remark 11.5 (For those who know about measure theory)**.** *Theorem 11.4 is equivalent to the existence of a measure μ (called the Haar measure), defined on the Borel σ-algebra of G, that is left-translation invariant (so $\mu(gS) = \mu(S)$ for all Borel sets S) and satisfies $\mu(G) = 1$. This Haar measure then allows us to define the integral over all measurable functions.*

Remark 11.6. *There is an analogue of Theorem 11.4for locally compact (but not compact) Hausdorff groups: there exists a unique Haar integral $\int_G$ from* compactly supported *functions on G to $\mathbb{R}$ satisfying Part 1,2 and 4. Since 1_G is not in general compactly supported, we need to replace Part 3 with $\int_G 1_K < \infty$ for every compact subset $K \subset G$. The integral will then only be unique up to a positive scalar multiple.*

Lemma 11.7. *Let G be a compact Hausdorff group. Then the Haar integral is also right-translation invariant.*

Proof. For every $x \in G$, the map $f \mapsto \int_G f(gx)\,\mathrm{d}g$ satisfies Parts 1-4 of the theorem so must equal $\int_G f(g)\,\mathrm{d}g$ by the uniqueness of the Haar integral. ☐

Examples 11.8. 1. *If G is finite and endowed with the discrete topology, then the Haar integral is simply $\int_G f = \frac{1}{|G|}\sum_{g\in G} f(g)$. It is easy to check that Parts 1-4 are satisfied.*

2. *Let $G = S^1 = \mathbb{R}/\mathbb{Z}$. When identifying functions on G with periodic functions on $\mathbb{R}$, the Haar integral is just given by the usual 'Riemann integral' on $[0, 1]$. In other words, if $f\colon S^1 \to \mathbb{R}$ then*

$$\int_G f(g) = \int_0^1 f(e^{2\pi i t})\,\mathrm{d}t. \tag{11.2.1}$$

We can also write this as $\frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\mathrm{d}t$.

3. *If $G = \mathrm{SU}(2)$, then the Haar integral gives you a way to integrate functions on the 3-sphere. It is quite complicated to describe in general, but for class functions If $f \in \mathcal{C}(G, \mathbb{R})$ is a class function, then Weyl's integration formula says that*

$$\int_G f(g) = \int_0^1 f\left(\begin{pmatrix} e^{2\pi i t} & 0 \\ 0 & e^{-2\pi i t}\end{pmatrix}\right)(e^{\pi i t} - e^{-\pi i t})\,\mathrm{d}t. \tag{11.2.2}$$

11.3 Basic results representation theory compact groups

Let G be a compact (Hausdorff) group. For example, G might be a finite group equipped with the discrete topology. Then many basic results in the representation theory of G are identical to the finite group setting. For example:

Proposition 11.9 (Weyl's unitary trick)**.** *Let G be a compact group and V a representation of G. Then there exists a G-invariant Hermitian inner product on V.*

Proof. The proof is identical to the finite groups case! Let $\langle -,-\rangle_0 \colon V\times V\to\mathbb{C}$ be any Hermitian inner product on V. For fixed $v,w\in V$, note that $g\mapsto\langle g\cdot v, g\cdot w\rangle_0$ is a continuous function $G\to\mathbb{C}$. We may therefore define, for every $v,w\in V$,

$$\langle v,w\rangle=\int_G\langle g\cdot v,g\cdot w\rangle_0\,\mathrm{d}g. \tag{11.3.1}$$

We claim that $\langle -,-\rangle$ is a G-invariant Hermitian inner product. Since $\int_G$ is linear, $\langle -,-\rangle$ is linear in the first variable. We have $\langle w,v\rangle=\overline{\langle v,w\rangle}$ since $\langle -,-\rangle_0$ has the same property. If $v\in V$ is nonzero, the map $g\mapsto\langle gv,gw\rangle_0$ is a continuous map $G\to\mathbb{R}_{>0}$. Since G is compact, there exists an $\epsilon>0$ such that $\langle gv,gv\rangle\geq\epsilon$ for all $g\in G$. Therefore $\int_G\langle gv,gv\rangle_0\geq\int_G\epsilon\,\mathrm{Id}_G\geq\epsilon>0$. This shows that $\langle -,-\rangle$ is a Hermitian inner product; it remains to show that it is G-invariant. This follows from the right-translation invariance of $\int_G$: for every $h\in G$, we have

$$\langle hv,hw\rangle=\int_G\langle (gh)v,(gh)w\rangle_0\,\mathrm{d}g \tag{11.3.2}$$

Since $\int_G$ is right-translation invariant, this equals $\int_G\langle gv,gw\rangle_0\,\mathrm{d}g=\langle v,w\rangle$, as desired. $\square$

Corollary 11.10. *Every finite-dimensional G-representation is a direct sum of irreducible representations.*

Here are some other results, with identical proofs as in the finite group case:

Proposition 11.11 (Schur's lemma). *If V,W are irreducible G-representations, then every element of $\mathrm{Hom}_G(V,W)$ is zero or invertible. Moreover $\mathrm{End}_G(V)=\mathbb{C}\cdot\mathrm{Id}_V$.*

Corollary 11.12. *If G is abelian, every irreducible representation is one-dimensional.*

There is also a satisfactory analogue of character theory. If $\rho\colon G\to\mathrm{GL}(V)$ is a representation, $\chi_V(g)=\mathrm{tr}(\rho(g))$ is a continuous function $G\to\mathbb{C}$ called the character of V. It is a class function. If $\phi,\psi\colon G\to\mathbb{C}$ are continuous class functions, set

$$\langle\phi,\psi\rangle:=\int_G\phi(g)\overline{\psi(g)}\,\mathrm{d}g. \tag{11.3.3}$$

We then have an analogue of the first projection formula:

$$\langle\chi_V,1\rangle=\dim V^G \tag{11.3.4}$$

and consequently of orthogonality of characters: if V,W are irreducible then

$$\langle\chi_V,\chi_W\rangle=\begin{cases}1 & \text{if } V\simeq W,\\ 0 & \text{if } V\not\simeq W.\end{cases} \tag{11.3.5}$$

This is all rather formal, and we didn't construct any interesting representations of G yet. Do characters span the space of all continuous class functions? What is the analogue of the regular representation? This is the content of the Peter–Weyl theorem. To state it in its full glory, we need some preliminaries, but we can already state one part of it here: consider the space of continuous class functions $G\to\mathbb{C}$ equipped with the uniform norm $||f||=\max_{g\in G}|f(g)|$. Then the Peter–Weyl implies that the $\mathbb{C}$-span of the irreducible characters form a dense subspace in this space with respect to this topology.

11.4 Representations on infinite-dimensional vector spaces

We recall some notions from topology and functional analysis. These are included for completeness to state the Peter–Weyl theorem but we will not use these notions in later lectures.

Recall from Definition 4.6 that a Hermitian inner product $\langle -,-\rangle\colon V\times V\to\mathbb{C}$ on a $\mathbb{C}$-vector space is a pairing that satisfies $\langle y,x\rangle=\overline{\langle x,y\rangle}$, that is $\mathbb{C}$-linear in the first variable, and that satisfies $\langle x,x\rangle>0$ if $x\neq 0$. Note that any such Hermitian inner product on V induces a norm on V by setting $||x||:=\sqrt{\langle x,x\rangle}$. The norm $||\cdot||$ induces a metric on V via $||x-y||$ and hence V has the structure of a metric (and so topological) space.

Definition 11.13. *A Hilbert space is a (possibly infinite-dimensional) $\mathbb{C}$-vector space H together with a Hermitian inner product $\langle -,-\rangle\colon H\times H\to\mathbb{C}$ such that H is complete with respect to the norm $||x||=\sqrt{\langle x,x\rangle}$. This means that every Cauchy sequence in H has a limit.*

If H is a Hilbert space, write $\mathrm{U}(H)$ for the set of unitary linear maps $\phi\colon H\to H$, namely those such that $\langle\phi(v),\phi(w)\rangle=\langle v,w\rangle$ for all $v,w\in V$.

Definition 11.14. *Let G be a topological group. A unitary Hilbert space representation of G is a group homomorphism $G\to\mathrm{U}(H)$ where H is a Hilbert space such that the associated map $G\times H\to H$ is continuous.*

Definition 11.15. *Let $\{H_i\}_{i\in I}$ be a family of Hilbert spaces. We define the Hilbert direct sum of the $\{H_i\}$ as*

$$\widehat{\bigoplus}_{i\in I}H_i:=\left\{(x_i)\in\prod_{i\in I}H_i \mid \sum_{i\in I}|x_i|^2<\infty\right\}. \tag{11.4.1}$$

Here $\sum_{i\in I}||x_i||^2<\infty$ means that only a countable number of the x_i are nonzero and we interpret convergence over this countable set in the usual way.

11.5 The Peter–Weyl theorem

Recall that if G is a finite group, the regular representation $\mathbb{C}G$ could also be seen (Problem Set 1) as the space of functions $G\to\mathbb{C}$ where G acts on such a function via $(g\cdot f)(x)=f(g^{-1}x)$. When G is a topological group, the latter generalizes better, and the regular representation should be like a space of functions $G\to\mathbb{C}$. We should not consider *all* functions on G though, as this would be bad from a topological viewpoint. We will only consider functions that 'lie' in L^2.

Definition 11.16. *Let G be a compact group. We define $L^2(G)$ as the space of measurable functions $f\colon G\to\mathbb{C}$ such that $\int_G|f|^2<\infty$, modulo the subspace of those f such that $\int_G|f|^2=0$. Then $L^2(G)$ is a Hilbert space with Hermitian inner product $\langle f,f'\rangle=\int_G f(g)\overline{f'(g)}\,\mathrm{d}g$.*

We need to take the quotient to ensure that nonzero elements in $L^2(G)$ have $\langle f,f\rangle>0$.

The group G acts on $L^2(G)$ via $(g\cdot f)(x)=f(g^{-1}x)$; this is called the left-regular representation.

Lemma 11.17. *The above association defines a unitary representation of G on $L^2(G)$.*

Proof. This follows from the left-translation invariance of $\int_G$. □

The following theorem summarizes all the properties analogous to 'completenses of characters' in the group setting.

Theorem 11.18 (Peter–Weyl theorem). *Let G be a compact group.*

1. *Every irreducible unitary Hilbert space representation of G is finite-dimensional.*
2. *The set of characters is dense in the space of continuous class functions. (Equipped with the uniform norm.)*
3. *There exists a Hilbert direct sum decomposition*

$$L^2(G) \simeq \widehat{\bigoplus_{\rho}} V_\rho^{\oplus \dim V_\rho}, \tag{11.5.1}$$

where $\rho\colon G \to \mathrm{GL}(V_\rho)$ ranges over all irreducible finite-dimensional representations of G.

We will not prove this theorem: the main input is the spectral theory of compact self-adjoint operators on Hilbert spaces.

Example 11.19. *Let $G = S^1$. On Problem Set 4, you will show that every irreducible representation of S^1 is of the form $S^1 \to \mathbb{C}^\times, z \mapsto z^n$. Then the Peter–Weyl theorem says that every periodic L^2-function on $\mathbb{R}$ equals a convergent sum of $e^{2\pi int}$ almost everywhere. This is the main result in the theory of Fourier series.*

Note that this doesn't really give an alternative proof in the theory of Fourier series: it just generalizes it massively to the context of compact groups. The proof will still rely on sophisticated results in functional analysis.

Example 11.20. *Let $G = \mathbb{R}$. Then G is not a compact group. It turns out that its irreducible unitary representations are given by $\mathbb{R} \to \mathbb{C}^\times, x \mapsto e^{2\pi ixy}$ for some $y \in \mathbb{R}$. Moreover, the theory of the Fourier transform shows that if $f \in L^2(\mathbb{R})$ then $f(x) = \int_{\mathbb{R}} \hat{f}(y)e^{2\pi ixy}\,\mathrm{d}y$. So in some sense, $L^2(\mathbb{R})$ is a like a 'direct integral' over its irreducible unitary representations, contrary to the compact case.*

Remark 11.21. *For non-compact groups, like $G = \mathrm{SL}_2(\mathbb{R})$, there will usually be interesting infinite-dimensional unitary irreducible representations.*

12 Differential geometry background

Lecture 13 starts here

We assume some prior exposure to the basic notions in manifold theory. But here are some reminders to refresh your memory and set up some notation.

12.1 Basic definitions

Let M be a topological space.

Definition 12.1. *A chart (U, ϕ) of M is a homeomorphism $\phi\colon U \to \phi(U)$, where $U \subset M$ is an open subset and $\phi(U) \subset \mathbb{R}^n$ is an open subset for some $n \geq 0$.*

Definition 12.2. *We say two charts (U, ϕ) and (V, ψ) are compatible if the composition*

$$\psi \circ \phi^{-1}\colon \phi(U \cap V) \to U \cap V \to \psi(U \cap V) \tag{12.1.1}$$

is smooth. (I.e. all partial derivatives up to every order are continuous.)

Definition 12.3. *An atlas on M is a collection of charts $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}$ such that*

1. *$M = \bigcup_{\alpha\in\mathcal{A}} U_\alpha$;*
2. *For every $\alpha, \beta \in \mathcal{A}$, the charts (U_α, ϕ_α) and (U_β, ϕ_β) are compatible.*

We say two atlases $\mathcal{A}$ and $\mathcal{A}'$ are equivalent if their union $\mathcal{A} \cup \mathcal{A}'$ is again an atlas, in other words if every chart in $\mathcal{A}$ is compatible with every chart in $\mathcal{A}'$.

Example 12.4. *Let $M = S^1 = \{(x, y) \mid x^2 + y^2 = 1\} \subset \mathbb{R}^2$. Then M is a topological space with the subspace topology of $\mathbb{R}^2$. We can define charts on S^1 by slicing up the circle along the x or y-axis. Let $U^+ = \{(x, y) \in S^1 \mid y > 0\}$ and $U^- = \{(x, y) \in S^1 \mid y < 0\}$. Then $\phi^+ \colon U^+ \to \mathbb{R}^2, (x, y) \mapsto x$ is a homeomorphism onto the open unit disk, hence defines a chart of S^1. Similarly we have a chart (U^-, ϕ^-), and if $V^\pm = \{(x, y) \in S^1 \mid \pm x > 0\}$ we can define charts $(V^\pm, \psi^\pm)$. The four charts $\{(U^\pm, \phi^\pm), (V^\pm, \psi^\pm)\}$ form an atlas on S^1. Indeed, the open sets clearly cover S^1. We just check that the charts (U^+, ϕ^+) and (V^+, ψ^+) are compatible, leaving the others to the reader. But $U^+ \cap V^+ = \{(x, y) \in S^1 \mid x, y > 0\}$, and the transition map $\phi^+(U^+ \cap V^+) \to \psi^+(U^+ \cap V^+)$ is the map $(0, 1) \to (0, 1)$ sending t to $\sqrt{1 - t^2}$, which is indeed smooth with smooth inverse.*

Of course, in the above example, there are many other natural choices of atlases on S^1: we could use stereographic projection to define charts, or the angle parametrization of the circle. However, all the atlases will turn out to be equivalent, and for the purposes of doing calculus on S^1 this choice shouldn't matter.

Definition 12.5. *A (smooth or differentiable) manifold is a topological space M together with an equivalence class of smooth atlases on M, that is moreover Hausdorff and second countable.*

We recall that M is Hausdorff if for every two distinct $x, y \in M$, there exist open subsets U, V of M with $x \in U$ and $y \in V$ such that $U \cap V$. ('Any two points are separated by open sets'.) This is a very desirable property, and excludes some weird spaces like the line with two origins. Recall that M is second-countable if there exists a countable basis $\{U_n\}_{n\in\mathbb{N}}$ of open sets: for every point $p \in M$ and open set V containing p, there exists an $n \geq 1$ such that $p \in U_n$ and $U_n \subset V$. This condition is slightly more technical, but is satisfies by all topological spaces that are not absurdly big (every subset of $\mathbb{R}^n$ is second-countable). It ensures the existence of partitions of unity for example.

Remark 12.6. *Instead of taking equivalence classes of atlases, one could also use a maximal atlas. Given an atlas $\mathcal{A}$, the union of all atlases equivalent to $\mathcal{A}$ is again an atlas $\mathcal{A}'$; it is the unique maximal atlas equivalent to $\mathcal{A}$, and it consists of all charts compatible to every chart in $\mathcal{A}$. Then one could equivalently define a manifold as a (Hausdorff, second-countable) topological space with a maximal atlas.*

If M is a manifold, a chart on M is by definition a chart in some atlas in the given equivalence class.

Examples 12.7.

1. *$M = S^1$ with the atlas given in Example 12.4 defines a manifold.*
2. *$M = \mathbb{R}^n$ with atlas $\{(\mathbb{R}^n, \mathrm{Id})\}$ defines a manifold.*
3. *Every open subset of a manifold M has the structure of a manifold, by taking the restriction of every chart in an atlas of M.*
4. *Not every closed subset S of a manifold M can be given a manifold structure: a necessary condition for this is that S is locally Euclidean, that is every point $p \in S$ has an open neighbourhood $p \in U \subset S$ such that U is homeomorphic to an open subset of $\mathbb{R}^n$ for some n. If $M = \mathbb{R}^2$ and S is the union of the coordinate axis, then S is not locally Euclidean at the origin hence cannot be given an atlas.*

If M is a manifold and $p \in M$, we say M has dimension n at p if there exists a chart (U, ϕ) containg p such that $\phi(U)$ is an open subset of $\mathbb{R}^n$. It is an exercise to check that this definition does not depend on the choice of chart and defines a locally constant function $M \to \mathbb{Z}_{\geq 0}$. If this function is constant (which is always the case if M is connected) and equal to n, we say that M has dimension n.

12.2 Smooth maps

Definition 12.8. *Let M, N be manifolds. A continuous map $F\colon M \to N$ is smooth at $p \in M$ if there exists charts (U, ϕ) of M and (V, ψ) of N containing p and $F(p)$ respectively such that the composition*

$$\psi \circ f \circ \phi^{-1}\colon \phi(U \cap F^{-1}(V)) \to U \cap F^{-1}(V) \to V \to \psi(V) \tag{12.2.1}$$

is smooth at p. We say F is smooth if it is smooth at every point of M. If F is smooth and has a smooth inverse, we say F is a diffeomorphism.

A map $F\colon M \to N$ is smooth at $p \in M$ if and only if the above condition is satisfied for all charts (U, ϕ) and (V, ψ) with $p \in U$ and $F(p) \in V$. (Exercise.)

We reiterate that when we use the word 'chart' on a manifold M we mean a chart compatible with the smooth structure, i.e. lying in an atlas in the equivalence class.

Example 12.9. *A continuous map $f\colon M \to \mathbb{R}$ is smooth if for every $p \in M$ there exists a chart (U, ϕ) with $p \in U$ such that $f \circ \phi^{-1}\colon \phi(U) \to \mathbb{R}$ is smooth.*

For manifolds M, N, write

$$\mathcal{C}^\infty(M, N) = \{\text{Smooth functions } M \to N\}, \tag{12.2.2}$$

$$\mathcal{C}^\infty(M) = \mathcal{C}^\infty(M, \mathbb{R}). \tag{12.2.3}$$

Since we can add and multiply functions $M \to \mathbb{R}$ pointwise, $\mathcal{C}^\infty(M)$ is a commutative $\mathbb{R}$-algebra.

Definition 12.10. *A curve on a manifold M is a smooth map $\gamma\colon I \to M$ for some open interval $I \subset \mathbb{R}$.*

12.3 Tangent spaces

Let M be a manifold that has dimension n at $p \in M$. Then we would like to define an n-dimensional vector space T_pM of 'tangent vectors at p'. Since we think of M as an abstract manifold, not as a subset of $\mathbb{R}^N$, it's not obvious a priori how to make the right definition. The key idea is to view tangent vectors as 'directions in which you can differentiate functions'.

To illustrate this, let us assume that $M = \mathbb{R}^n$ for a moment. Given a smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$ (or one only defined on an open subset of $\mathbb{R}^n$ containing p) and a vector $v \in \mathbb{R}^n$, we can differentiate f at p in the direction v:

$$\left.\frac{\partial}{\partial v}\right|_p f := \lim_{t \mapsto 0} \frac{f(p + tv) - f(p)}{t}. \tag{12.3.1}$$

So every $v \in \mathbb{R}^n$ defines a map $X\colon \mathcal{C}^\infty(\mathbb{R}^n) \to \mathbb{R}$, $f \mapsto \left.\frac{\partial}{\partial v}\right|_p f$. This map is $\mathbb{R}$-linear and satisfies the Leibniz rule: $X(fg) = X(f)g(p) + f(p)X(g)$ for all smooth functions $f, g \in \mathcal{C}^\infty(\mathbb{R}^n)$.

Proposition 12.11. *The map $v \mapsto \frac{\partial}{\partial v}\big|_p$ induces an $\mathbb{R}$-linear isomorphism:*

$$\mathbb{R}^n \xrightarrow{\sim} \left\{ \begin{matrix} \mathbb{R}\text{-linear maps } X\colon \mathcal{C}^\infty(\mathbb{R}^n) \to \mathbb{R} \text{ satisfying} \\ X(fg) = X(f)g(p) + f(p)X(g) \text{ for all } f, g \in \mathcal{C}^\infty(\mathbb{R}^n) \end{matrix} \right\} \tag{12.3.2}$$

Proof. See any textbook on manifolds. This uses first-order Taylor series. □

It is the right hand side of the isomorphism in the above proposition that generalizes well. Let M be a manifold.

Definition 12.12. *A derivation on M at p is a $\mathbb{R}$-linear map $X\colon \mathcal{C}^\infty(M) \to \mathbb{R}$ such that $X(fg) = f(p)X(g) + X(f)g(p)$ for all $f, g \in \mathcal{C}^\infty(M)$. Write $\mathrm{Der}_p(M)$ for the set of all $\mathbb{R}$-linear derivations on M at p.*

Note that $\mathrm{Der}_p(M)$ is an $\mathbb{R}$-subspace of the space of linear functionals on $\mathcal{C}^\infty(M)$, hence a vector space itself.

Definition 12.13. *The tangent space of M at p is defined to be the vector space $T_pM \coloneqq \mathrm{Der}_p(M)$.*

With this definition, defining the derivative of a map looks silly:

Definition 12.14. *Let $F\colon M \to N$ be a smooth map between manifolds. If $p \in M$, then the differential (or derivative) of F at p is the linear map $(dF)_p\colon T_pM \to T_{F(p)}N$ defined by sending $X \in T_pM = \mathrm{Der}_p(M)$ to $(dF)_p(X) \in \mathrm{Der}_{F(p)}(N)$ that is defined via*

$$(dF)_p(X)(g) \coloneqq X(g \circ F) \tag{12.3.3}$$

for all $g \in \mathcal{C}^\infty(N)$.

In other words, the induced map on tangent spaces is just pulling back derivations.

Proposition 12.15 (Chain rule)**.** *Let $F\colon M \to N$ and $G\colon N \to L$ be smooth maps between manifolds. Then $G \circ F\colon M \to L$ is smooth and for all $p \in M$ we have*

$$(d(G \circ F))_p = (dG)_{F(p)} \circ (dF)_p \tag{12.3.4}$$

Proof. The fact that $G \circ F$ is smooth follows from the fact that a composition of smooth functions on open subsets of $\mathbb{R}^n$ is smooth. (This uses the ordinary chain rule.) The centered identity (12.3.4) can be proven formally: for $X \in \mathrm{Der}_p(M)$ and $h \in \mathcal{C}^\infty(L)$ we have

$$\begin{aligned} (d(G \circ F))_p(X)(h) &= X(h \circ G \circ F) && (12.3.5)\\ &= X((h \circ G) \circ F) && (12.3.6)\\ &= (dF)_p(X)(h \circ G) && (12.3.7)\\ &= (dG)_{F(p)}((dF)_p(X))(h). && (12.3.8) \end{aligned}$$

This looks like weird symbol manipulating, but it is just saying that precomposing functions with $(G \circ F)$ is the same as first precomposition with G and then with F. □

The chain rule implies that if $F\colon M \to N$ is a smooth map with a smooth inverse then $(dF)_p\colon T_pM \to T_{F(p)}N$ is an isomorphism for all $p \in M$. Indeed, if $G\colon N \to M$ is the smooth inverse of F then $(dG)_{F(p)} \circ (dF)_p = \mathrm{Id}$ and $(dF)_p \circ (dG)_{F(p)} = \mathrm{Id}$.

We can think concretely about tangent vectors on M at p once we choose a chart (U,ϕ) containing p. Indeed, in that case $\phi(U) \subset \mathbb{R}^n$ and the n coordinate functions $\mathbb{R}^n \to \mathbb{R}, (x_1,\dots,x_n) \mapsto x_i$ give rise to coordinate functions $x_1,\dots,x_n \colon U \to \mathbb{R}$. These are called local coordinates on M. Define for $f \in \mathcal{C}^\infty(M)$:

$$\left.\frac{\partial}{\partial x_i}\right|_p (f) := \left.\frac{\partial}{\partial x_i}\right|_{\phi(p)} (f \circ \phi^{-1}). \tag{12.3.9}$$

Then $\{\frac{\partial}{\partial x_1}\big|_p, \dots, \frac{\partial}{\partial x_n}\big|_p\}$ are derivations at p and form a basis of T_pM, by Proposition 12.11. If N is another manifold, $F \colon M \to N$ is a smooth map, and (V,ψ) is a chart of N containing $F(p)$, we have local coordinates $y_1,\dots,y_m \colon V \to \mathbb{R}$ on N and $T_{F(p)}N$ has basis $\{\frac{\partial}{\partial y_j}\big|_{F(p)}\}_{1\le j\le m}$. If we set $F_j = y_j \circ F \circ \phi^{-1} \colon \phi(U \cap F^{-1}(V)) \to \mathbb{R}$, then $F_1,\dots,F_m$ are functions in the variables $x_1,\dots,x_n$ defined on an open subset containing $\phi(p)$. The differential $(dF)_p \colon T_pM \to T_{F(p)}N$ in these explicit bases is exactly given by

$$\left.\frac{\partial}{\partial x_i}\right|_p \mapsto \sum_{j=1}^m \frac{\partial F_j}{\partial x_i}(p) \left.\frac{\partial}{\partial y_j}\right|_{F(p)}. \tag{12.3.10}$$

Therefore all the abstractly defined notions of tangent space and differential specialize to the familiar notions from multivarialbe calculus.

Finally, we connect back to the motivation for the definition of the tangent space:

Lemma 12.16. *Let U be an open subset of a finite-dimensional $\mathbb{R}$-vector space V. Then U is a manifold, and for every $p \in U$ the map $V \to T_pU, t \mapsto \gamma_v'(0)$, where $\gamma_v(t) = p + tv$, is an isomorphism.*

This is a good exercise in the definitions. We will use this lemma over and over again and will often implicitly identify T_pU with V via this map.

12.4 Vector fields

Lecture 14 starts here

Let M be a manifold. We will endow the set $T_*M = \sqcup_{p\in M} T_pM$ with the structure of a manifold, as follows. There is a projection map $\pi \colon T_*M \to M, (p,v) \mapsto p$ whose fibers are exactly all the tangent spaces of M. Let $\mathcal{A}$ be an atlas for M defining the smooth structure. Let $\phi \colon U \to \phi(U) \subset \mathbb{R}^n$ be a chart in $\mathcal{A}$, defining local coordinates $x_1,\dots,x_n \colon U \to \mathbb{R}$. For each $p \in U$, the basis $\{\frac{\partial}{\partial x_i}\big|_p\}$ determines an isomorphism $\alpha_p \colon T_pM \xrightarrow{\sim} \mathbb{R}^n$. Collecting these isomorphisms determines a bijection $\phi_{T_*M} \colon \pi^{-1}(U) \to \phi(U) \times \mathbb{R}^n, (p,v) \mapsto (\phi(p), \alpha_p(v))$. Let $\mathcal{A}' = \{(\pi{-}1(U), \phi_{T_*M}) \mid (U,\phi) \in \mathcal{A}\}$. We give T_*M the unique topology for every map in $\mathcal{A}'$ is a homeomorphism, and declare $\mathcal{A}'$ to be an atlas. One can check that this is indeed an atlas and $\pi \colon T_*M \to M$ is a smooth map. The manifold T_*M is called the tangent bundle of M.

Definition 12.17. *Let M be a manifold. A vector field on M is a smooth map $V \colon M \to T_*M$ such that $V(p) \in T_pM$. In other words, $\pi \circ V = \mathrm{Id}_M$.*

In other words, a vector field is a choice of tangent vector in T_pM for every $p \in M$ that varies smoothly in some way. This smoothness depends on the manifold structure on T_*M, but we can express it concretely as follows:

Let (U,ϕ) be a chart of M, $x_1,\dots,x_n$ the corresponding local coordinates on U, giving rise to the basis $\{\frac{\partial}{\partial x_1}\big|_p, \dots, \frac{\partial}{\partial x_n}\big|_p\}$ for each $p \in U$. If $V \colon M \to T_*M$ is any map satisfying $V(p) \in T_pM$ for all $p \in M$, then there exist functions $V_1,\dots,V_n \colon U \to \mathbb{R}$ such that

$$V(p) = \sum_{i=1}^n V_i(p) \left.\frac{\partial}{\partial x_i}\right|_p \tag{12.4.1}$$

for all $p \in U$. Then V is a vector field (in other words, V is smooth) if and only if the functions $V_i \colon U \to \mathbb{R}$ are smooth for all $1 \le i \le n$ and all charts (U, ϕ) in a given atlas of M.

We have seen (by definition) that tangent vectors at $p \in M$ are derivations of M at p. Since a vector field is a collection of tangent vectors, we can use it to differentiate a function at every point.

Definition 12.18. *A derivation on M is a $\mathbb{R}$-linear map $D \colon \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ satisfying $D(fg) = D(f)g + fD(g)$ for all $f, g \in \mathcal{C}^\infty(M)$. Write $\mathrm{Der}(M)$ for the set of all derivations on M.*

The difference between $\mathrm{Der}_p(M)$ and $\mathrm{Der}(M)$ is that an element of $\mathrm{Der}_p(M)$ is a map $\mathcal{C}^\infty(M) \to \mathbb{R}$, while an element of $\mathrm{Der}(M)$ is a map $\mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$.

Let $V \colon M \to T_*M$ be a vector field on M. Then V gives rise to a derivation $D_V \colon \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ via

$$(D_V(f))(p) = V(p)(f) \tag{12.4.2}$$

for $f \in \mathcal{C}^\infty(M)$ and $p \in M$. In other words, each $V(p)$ is a derivation at p so for each p we can derive f at p using $V(p)$.

Proposition 12.19. *The map $V \mapsto D_V$ defines an $\mathbb{R}$-linear isomorphism between the set of vector fields on M and $\mathrm{Der}(M)$.*

Why is this useful? If $D, D' \in \mathrm{Der}(M)$, then the composite maps $D \circ D' \colon M \to M$ and $D' \circ D \colon M \to M$ are not derivations, but one can check that

$$[D, D'] \coloneqq D \circ D' - D' \circ D \in \mathrm{Der}(M). \tag{12.4.3}$$

Definition 12.20. *Given two vector fields V, V' on M, there exists a unique vector field on M corresponding to the derivation $[D_V, D_{V'}]$. This vector field is called the Lie bracket of V and V'.*

12.5 Integral curves

Let M be a manifold and $\gamma \colon I \to M$ a curve, i.e. a smooth map from an open interval $I \subset \mathbb{R}$. If $t \in I$ and $\mathbb{R}$ has coordinate x, T_tI has a canonical generator, namely $\frac{\partial}{\partial x}\big|_t$.

Definition 12.21. *If $\gamma \colon I \to M$ is a curve, we define the tangent vector $\gamma'(t) \in T_{\gamma(t)}M$ at t to be $(D\gamma)_t\left(\frac{\partial}{\partial x}\big|_t\right)$.*

Definition 12.22. *Let V be a vector field on M. An integral curve for V on M is a smooth curve $\gamma \colon I \to M$ such that $\gamma'(t) = V(\gamma(t))$ for all $t \in I$.*

Example 12.23. *Let $M = \mathbb{R}$, $V = x\frac{\partial}{\partial x}$. Then an integral curve is a smooth map $\gamma \colon I \to \mathbb{R}$ such that $\gamma'(t) = \gamma(t)$. This differential equation has solutions $\gamma(t) = Ce^t$ for some $C \in \mathbb{R}$.*

Example 12.24. *Let $M = \mathbb{R}^2$ and $V = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$. Then integral curves for V are concentric circles.*

Working out these examples show that finding integral curves is like finding solutions to ordinary differential equations. Since we can always solve them locally, we have the following theorem:

Theorem 12.25. *Let M be a manifold, V a vector field on M and $p \in M$. Then there exists an open interval $I \subset \mathbb{R}$ containing 0 and a smooth map $\gamma \colon I \to M$ such that $\gamma(0) = p$ such that γ is an integral curve for V. If $J \subset \mathbb{R}$ is another open interval containing 0 and $\tilde{\gamma} \colon J \to M$ an integral curve for V with $\tilde{\gamma}(0) = p$, then $\gamma = \tilde{\gamma}$ on $I \cap J$.*

12.6 Submanifolds

Let M be a manifold. We now define what it means to be a submanifold of M. There are actually two competing notions for this, and their difference is subtle. We start with the notion that has the best properties.

Definition 12.26. *Let $S \subset M$ be a subset. We say S is an embedded submanifold of dimension k if for every $p \in S$, there exists a chart $\phi\colon U \to \phi(U) \subset \mathbb{R}^n$ of M such that $\phi(S \cap U) = \{(x_1, \ldots, x_n) \in \phi(U) \mid x_{k+1} = \cdots = x_n = 0\}$. Such a chart is called a slice chart.*

If S is an embedded submanifold of M, equipped with the subspace topology, then there exists a unique smooth structure on S (i.e. manifold structure) such that the inclusion $S \hookrightarrow M$ is smooth. Indeed, we may simply take the restriction of slice charts to the first k coordinates.

Example 12.27. *Any vector subspace S of $\mathbb{R}^n$ is an embedded submanifold.*

Example 12.28. *The subset $S^1 \subset \mathbb{R}^2$ is an embedded submanifold. This can be seen using polar coordinates.*

To produce examples of embedded submanifolds, we will need the notion of a submersion.

Definition 12.29. *A smooth map $F\colon M \to N$ is said to be*

1. *An immersion at $p \in M$ if $(dF)_p$ is injective;*
2. *A submersion at $p \in M$ if $(dF)_p$ is surjective.*

We say F is an immersion/submersion if it is an immersion/submersion at every point.

Theorem 12.30 (Submersion theorem)**.** *Let $F\colon M \to N$ be a submersion and $c \in N$. If F is a submersion at every $p \in F^{-1}(c)$, then $F^{-1}(c)$ is an embedded submanifold of M.*

Corollary 12.31 (Level sets)**.** *Let $f\colon M \to \mathbb{R}$ be a smooth map and $c \in \mathbb{R}$ such that $(df)_p$ is nonzero for all $p \in f^{-1}(c)$. Then $f^{-1}(c)$ is an embedded submanifold of M.*

Example 12.32 (Spheres are embedded submanifolds)**.** *Let $M = \mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}, (x_1, \ldots, x_n) \mapsto x_1^2 + \cdots + x_n^2$. If $p = (a_1, \ldots, a_n)$, then $(df)_p = (2a_1, \ldots, 2a_n)$. It follows that $(df)_p$ is a submersion at every point $p \neq 0$ and so $f^{-1}(1) = S^{n-1}$ is an embedded submanifold of $\mathbb{R}^n$.*

Given an embedded submanifold $S \subset M$, we have seen that S is a manifold itself and the definition of the manifold structure shows that $S \to M$ is an injective immersion. Does this characterize embedded submanifolds, i.e. is every image of an injective immersion an embedded submanifold? This turns out to be false.

Example 12.33. *Let $M = \mathbb{R}^2$ and let S be the 'figure six' curve inside M. Then there exists an injective immersion $f\colon (0,1) \to M$ with image S. (Draw picture.) However, S is not an embedded submanifold because it is not locally Euclidean at the branch point where the ends of $(0,1)$ meet.*

Definition 12.34. *Let $S \subset M$ be a subset of M. If S is the image of an injective immersion $N \to M$ for some manifold N, we say S is an immersed submanifold of M.*

So every embedded submanifold is immersed, but the converse might not be true. The image of an injective immersion $f\colon N \to M$ is an embedded submanifold if the subspace topology of $f(N)$ induces a homeomorphism $f\colon N \to f(N)$.

Example 12.35. *Let $M = S^1 \times S^1$ and let $\alpha \in \mathbb{R}$. Let $N = \mathbb{R}$ and consider the map $f\colon N \to M, t \mapsto (e^{2\pi i t}, e^{2\pi i \alpha t})$. If α is rational, the image of f is an embedded submanifold diffeomorphic to S^1. If α is irrational, f is an injective immersion and the image is an immersed submanifold. It turns out that in this case the image of f is dense in M and so is certainly not an embedded submanifold of M.*

12.7 Exercises on manifolds

Here are some additional (non hand-in) exercises for those who want to practice their manifold skills. You're very welcome to discuss any of these during office hours!

1. Let M be a manifold with atlas $\mathcal{A}$. Show that a chart (U, ϕ) of the topological space M lies in an atlas of the equivalence class of $\mathcal{A}$ if and only if (U, ϕ) is compatible with every chart in $\mathcal{A}$.
2. Let $f\colon M \to N$ be a smooth map of manifolds. Show that for every chart (U, ϕ) of M and (V, ψ) of N with $U \cap f^{-1}(V) \neq \varnothing$, the induced map $\psi \circ f \circ \phi^{-1}\colon \phi(U \cap f^{-1}(V)) \to V \to \psi(V)$ is smooth.
3. Let (U, ϕ) be a chart on a manifold M. Show that $\phi\colon U \to \phi(U)$ is a diffeomorphism (with the manifold structures on U and $\phi(U)$).
4. Define an atlas on real projective space $\mathbb{P}^n(\mathbb{R}) = (\mathbb{R}^{n+1} \setminus \{0\})/\mathbb{R}^\times$ and define a smooth map $S^n \to \mathbb{P}^n(\mathbb{R})$.
5. Let M and N be manifolds. Equip $M \times N$ with the structure of a manifold.
6. Verify that the chain rule in local coordinates is equivalent to the familiar chain rule from analysis.
7. Let $F\colon M \to N$ be a smooth map with M connected and $(dF)_p = 0$ for all $p \in M$. Show that F is constant.
8. Let p be a point of a manifold M and $v \in T_pM$. Show that there exists a curve $\gamma\colon I \to M$ with $0 \in I \subset \mathbb{R}$ such that $\gamma(0) = p$ and $\gamma'(0) = v$. Let $F\colon M \to N$ be a smooth map. Show that $(dF)_p(v) = (F \circ \gamma)'(0)$, so derivatives can be computed derivatives of curves.
9. Let V be a finite-dimensional $\mathbb{R}$-vector space and $U \subset V$ an open subset. Then U has the structure of a manifold, and for every $p \in U$ there exists a 'canonical' isomorphism $V \simeq T_pU$ given by sending v to the derivative of the curve $p + tv$ at zero.
10. For which values of $c \in \mathbb{R}$ is the zero locus in $\mathbb{R}^3$ of $z^2 - (x^2 + y^2)^2 + c$ an embedded submanifold of $\mathbb{R}^3$? When is it an immersed submanifold?

13 Lie groups

13.1 Basic definitions

Definition 13.1. *A Lie group is a manifold G whose underlying set has a group structure such that the maps $G \times G \to G, (x, y) \mapsto x \cdot y$ and $G \to G, x \mapsto x^{-1}$ are smooth. A map between Lie groups $\phi\colon G \to H$ is a Lie group homomorphism (or simply a homomorphism) if it is smooth and a group homomorphism. If ϕ has an inverse that is again a Lie group homomorphism, we say that ϕ is an isomorphism.*

Since every Lie group is in particular a group and a topological space, we can add properties to Lie groups that put restrictions on the group structure or topology. For example, a Lie group G is connected if the underlying topological space is connected, and a Lie group G is commutative if the underlying group is commutative.

Definition 13.2. *Let G be a Lie group and $g \in G$. Write $L_g\colon G \to G, x \mapsto gx$ and $R_g\colon G \to G, x \mapsto xg$ for left/right multiplication by g.*

The maps L_g, R_g are diffeomorphisms of G and provide G with an extraordinary amount of symmetry. For example, $d(L_g)_e$ induces an isomorphism $T_eG \to T_gG$ for all $g \in G$, so all the tangent spaces T_gG are canonically identified.

13.2 Examples of Lie groups

Examples 13.3. *1.* $G = (\mathbb{R}^n, +)$ *is a Lie group under addition.*

2. *The group* $G = \mathrm{GL}_n(\mathbb{R})$ *is a Lie group under matrix multiplication. Indeed, since* G *is an open subset of* $\mathbb{R}^{n^2}$*, it inherits a manifold structure from* $\mathbb{R}^{n^2}$*. The multiplication map is smooth since it is polynomial, and the inversion map is smooth by Cramer's rule. Similarly,* $\mathrm{GL}_n(\mathbb{C})$ *is a Lie group.*

3. *If* $F = \mathbb{R}$ *or* $\mathbb{C}$ *and* V *is a finite-dimensional* F*-vector space,* $\mathrm{GL}(V)$ *has the structure of a Lie group, by an argument similar to Example 4 in §11.1.*

4. *The subset* $B_n(\mathbb{R})$ *of upper triangular matrices*

$$B_n(\mathbb{R}) = \left\{ \begin{pmatrix} * & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & * \end{pmatrix} \right\} \subset \mathrm{GL}_n(\mathbb{R}) \tag{13.2.1}$$

is a Lie group, because it is an embedded submanifold of $\mathrm{GL}_n(\mathbb{R})$ *so a manifold, and the multiplication/inversion map are smooth. Similarly the subset of strictly upper triangular matrices*

$$U_n(\mathbb{R}) = \left\{ \begin{pmatrix} 1 & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \right\} \subset \mathrm{GL}_n(\mathbb{R}) \tag{13.2.2}$$

is a Lie group.

5. *If* G, H *are Lie groups, the direct product* $G \times H$ *(with the product manifold structure and product group structure) is a Lie group.*

6. $G = S^1 = \{z \in \mathbb{C}^\times \mid |z| = 1\}$ *is a Lie group under multiplication. (The fact that multiplication/inversion is smooth can be seen by taking angle parametrizations.)*

7. $G = \mathrm{GL}_n(\mathbb{R})^+ = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid \det(A) > 0\}$ *is a Lie group, being an open subgroup of* $\mathrm{GL}_n(\mathbb{R})$*. In fact,* $\mathrm{GL}_n(\mathbb{R})^+$ *is connected. (See problem set 5.)*

We will be able to write down more examples after the next section.

13.3 Lie subgroups

Let G be a Lie group.

Definition 13.4. *An immersed Lie subgroup (or simply Lie subgroup) of* G *is the image of a Lie group homomorphism* $H \to G$ *that is an injective immersion.*

Definition 13.5. *An embedded Lie subgroup of* G *is a subgroup* $H \leq G$ *that is also an embedded submanifold of* G*.*

If H is an embedded Lie subgroup of G, then H is a manifold itself (by restricting slice charts) and with respect to this manifold structure and subgroup structure H is a Lie group itself. (Exercise.)

Example 13.6. S^1 *is an embedded Lie subgroup of* $\mathbb{C}^\times$*.*

Example 13.7. *A line with irrational slope in $S^1 \times S^1$ (Example 12.35) is a Lie subgroup that is immersed but not embedded.*

The following nontrivial theorem (which we won't prove) gives an easy way of generating new Lie groups:

Theorem 13.8 (Closed subgroup theorem)**.** *The following are equivalent for a subgroup H of a Lie group G:*

1. *H is closed;*
2. *H is an embedded Lie subgroup of G.*

This theorem implies that every closed subgroup of a Lie group is a Lie group itself.

Example 13.9. *Since $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid \det(A) = 1\}$ is closed in $\mathrm{GL}_n(\mathbb{R})$ (being the preimage of $\{1\}$ under $\det$), the closed subgroup theorem implies that $\mathrm{SL}_n(\mathbb{R})$ is an embedded Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$ hence a Lie group itself. Similarly $\mathrm{SL}_n(\mathbb{C})$ is a Lie group.*

Example 13.10. *Since the group $\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A\bar{A}^t = I\}$ is closed in $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{U}(n)$ is a Lie group.*

See the problem set for many more examples of Lie groups.

13.4 Lie algebras

It turns out that the tangent space of a Lie group at the identity has the structure of a Lie algebra, and that this Lie algebra remembers a lot about the corresponding Lie group. Before we define this we start with some generalities on Lie algebras.

Definition 13.11. *A Lie algebra over a field F is a vector space $\mathfrak{g}$ over F together with a bilinear pairing $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ (called the Lie bracket) such that*

1. *$[-,-]$ is alternating: $[x,x] = 0$ for all $x \in \mathfrak{g}$. (This implies that $[x,y] = -[y,x]$ for all $x, y \in \mathfrak{g}$.)*
2. *$[-,-]$ satisfies the Jacobi identity: for all $x, y, z \in \mathfrak{g}$, $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$.*

A morphism of Lie algebras is a linear map $f\colon \mathfrak{g} \to \mathfrak{h}$ such that $f([x,y]) = [f(x), f(y)]$ for all $x, y \in \mathfrak{g}$.

Example 13.12. *Consider the vector space $\mathrm{Mat}_n(F)$ of $n \times n$-matrices with Lie bracket $[A,B] \coloneqq AB - BA$. This bracket is clearly bilinear and alternating, and a calculation shows that it satisfies the Jacobi identity. We write this Lie algebra as $\mathfrak{gl}_n(F)$. (Of course, as a vector space it is the same as $\mathrm{Mat}_n(F)$, but we write $\mathfrak{gl}_n(F)$ to emphasize that we think of it as a Lie algebra.) Similarly, we have a Lie algebra $\mathfrak{gl}(V)$ for any F-vector space, where the underlying vector space is $\mathrm{End}(V)$ and the bracket is $[f,g] = f \circ g - g \circ f$.*

Warning 13.13. *A Lie algebra is not an algebra in the sense of Section 8.1. In particular, a Lie algebra does not necessarily have a unit, nor do we require it to be associative.*

The following lemma might demystify the Jacobi identity a little bit:

Lemma 13.14. *Let $\mathfrak{g}$ be a vector space with an alternating bilinear bracket $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Then $\mathfrak{g}$ is a Lie algebra (i.e. satisfies the Jacobi identity) if and only if the map $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}), x \mapsto [x,y]$ respects the brackets on both sides.*

Proof. The condition that ad preserves the bracket means that $\mathrm{ad}_{[x,y]} = \mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x$ for all $x, y \in \mathfrak{g}$. Evaluating this at $z \in \mathfrak{g}$ and using the fact that $[-,-]$ is alternating exactly recovers the Jacobi identity. □

13.5 Lie algebra of a Lie group

Lecture 15 starts here

Let G be a Lie group with identity $e \in G$.

Definition 13.15. *The tangent space T_eG at e of G, denoted by $\mathfrak{g}$ or $\mathrm{Lie}(G)$, is called the Lie algebra of G.*

In the next paragraphs we will define a bracket $[-,-]$ on $\mathfrak{g}$ and show that $\mathfrak{g}$ is a Lie algebra over $\mathbb{R}$. This will take a few steps.

We start by considering for every $g \in G$ the conjugation map

$$\Psi_g\colon G \to G, x \mapsto gxg^{-1} \tag{13.5.1}$$

Then Ψ_g is a Lie group homomorphism sending e to e. We can therefore take its derivative, and get a homomorphism $\mathrm{Ad}(g) := (d\Psi_g)_e\colon\ \mathfrak{g} \to \mathfrak{g}$. Since $\Psi_g \circ \Psi_h = \Psi_{gh}$, the chain rule implies that $\mathrm{Ad}(g) \circ \mathrm{Ad}(h) = \mathrm{Ad}(gh)$ for all $g, h \in G$. In other words, the map $g \mapsto \mathrm{Ad}(g)$ defines a homomorphism

$$\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g}). \tag{13.5.2}$$

This homomorphism is called the adjoint representation of G. Let us first prove that this map is smooth. (This is a technical detail that you can skip over on first reading.) It suffices to show that for every $X \in \mathfrak{g}$, the map $g \mapsto \mathrm{Ad}(g)(X)$ is smooth. (Indeed, if that's the case then after choosing a basis of $\mathfrak{g}$ and taking X to be the basis vectors, we can conclude that all the matrix entries are smooth.) Let $\gamma\colon I = (-\epsilon, \epsilon) \to G$ be a curve with $\gamma(0) = e$ and $\gamma'(0) = X$. Then $\mathrm{Ad}(g)(X) = \frac{d}{dt}\big|_{t=0}[g\gamma(t)g^{-1}]$. Since the map $G \times G \to G, (g, x) \mapsto gxg^{-1}$ is smooth, the map $G \times I \to G, (g, t) \mapsto g\gamma(t)g^{-1}$ is smooth as well. Since taking the partial derivative of a smooth map is smooth, $(g, t) \mapsto \frac{d}{dt}[g\gamma(t)g^{-1}]$ is smooth. Since the restriction of a smooth function is smooth, $g \mapsto \frac{d}{dt}\big|_{t=0}[g\gamma(t)g^{-1}] = \mathrm{Ad}(g)(X)$ is smooth too, as desired.

Since Ad is smooth and sends e to Id, we can take the derivative of Ad at e. Using the fact that $T_e \mathrm{GL}(\mathfrak{g}) = \mathrm{End}(\mathfrak{g})$ (Lemma 12.16), we get an $\mathbb{R}$-linear map

$$\mathrm{ad} := (d\,\mathrm{Ad})_e\colon\ \mathfrak{g} \to \mathrm{End}(\mathfrak{g}), \tag{13.5.3}$$

called the adjoint representation of $\mathfrak{g}$.

Definition 13.16. *The Lie bracket on $\mathfrak{g}$ is defined by setting for $x, y \in \mathfrak{g}$:*

$$[x, y] := \mathrm{ad}(x)(y) \in \mathfrak{g}\,. \tag{13.5.4}$$

Since ad is linear and lands in $\mathrm{End}(\mathfrak{g})$, $[-,-]$ is bilinear. Before we prove that $[-,-]$ defines a Lie algebra structure on $\mathfrak{g}$, we start with some lemmas.

Lemma 13.17. *Let $G = \mathrm{GL}_n(\mathbb{R})$. Then $[A, B] = AB - BA$ for all $A, B \in \mathfrak{g} = \mathfrak{gl}_n(\mathbb{R})$. Hence $\mathrm{Lie}(G)$ is a Lie algebra.*

Proof. It will be useful in this proof (and later) to calculate derivatives using curves: if $F\colon M \to N$ is a smooth map between manifolds, $v \in T_pM$ and $\gamma\colon (-\epsilon, \epsilon) \to M$ is a curve on M with $\gamma(0) = p$ and $\gamma'(0) = v$, then $(dF)_p(v) = (F \circ \gamma)'(0)$.

To apply this principle here, choose $\epsilon > 0$ small enough so that the curves $\gamma\colon (-\epsilon, \epsilon) \to \mathrm{GL}_n(\mathbb{R}), t \mapsto I + tA$

and $\psi(t) = I + tB$ are well-defined (i.e. $I + tA, I + tB$ are invertible for $|t| < \epsilon$). Then

$$[A, B] = \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}(\gamma(t))(B) \tag{13.5.5}$$
$$= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} \mathrm{Ad}(I + tA)(I + sB) \tag{13.5.6}$$
$$= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} (I + tA)(I + sB)(I + tA)^{-1}. \tag{13.5.7}$$

So we need to derive the matrix valued function $(t, s) \mapsto (I + tA)(I + sB)(I + tA)^{-1}$. For this we can use the matrix-valued product rule: if $t \mapsto X(t), Y(t) \in \mathrm{Mat}_n(\mathbb{R})$ are curves, then $\frac{d}{dt}\big|_{t=0}(X(t)Y(t)) = X'(t)Y(t) + X(t)Y'(t)$. Applied to the variable s first, we see that $[A, B] = \frac{d}{dt}\big|_{t=0}((I + tA)B(I + tA)^{-1})$. Applied to $X(t) = (I+tA)$ and $Y(t) = (I+tA)^{-1}$, we see that $\frac{d}{dt}\big|_{t=0}(X(t)Y(t)) = 0$ so $\frac{d}{dt}\big|_{t=0}(I+tA)^{-1} = -A$. So

$$\frac{d}{dt}\bigg|_{t=0}((I + tA)B(I + tA)^{-1}) = \frac{d}{dt}\bigg|_{t=0}(I + tA)B + B\frac{d}{dt}\bigg|_{t=0}((I + tA)^{-1}) \tag{13.5.8}$$
$$= AB - BA. \tag{13.5.9}$$

□

The next lemma shows that the Lie bracket behaves well with respect to group homomorphisms.

Lemma 13.18. *Let $\Phi\colon G \to H$ be a Lie group homomorphism with $\phi = \mathrm{Lie}(\Phi) = (d\Phi)_e\colon \mathfrak{g} \to \mathfrak{h}$ the induced linear map. Then*

$$\phi([x, y]) = [\phi(x), \phi(y)] \tag{13.5.10}$$

for all $x, y \in \mathfrak{g}$.

Proof. Since $\Phi(gxg^{-1}) = \Phi(g)\Phi(x)\Phi(g)^{-1}$ we have (using the notation of (13.5.1)) that $\Phi \circ \Psi_g = \Psi_{\Phi(g)} \circ \Phi$ for all $g \in G$. Taking the derivative at e shows that $\phi \circ \mathrm{Ad}(g) = \mathrm{Ad}(\Phi(g)) \circ \phi$ for all $g \in G$. Let $X, Y \in \mathfrak{g}$ and choose a curve $\gamma\colon I = (-\epsilon, \epsilon) \to G$ with $\gamma(0) = e$ and $\gamma'(0) = X$. Then $[X, Y] = \frac{d}{dt}\big|_{t=0} \mathrm{Ad}(\gamma(t))(Y)$ and by the above calculation

$$\phi(\mathrm{Ad}(\gamma(t))(Y)) = \mathrm{Ad}(\Phi(\gamma(t)))(\phi(Y)). \tag{13.5.11}$$

The derivative of the left-hand side at $t = 0$ is $\phi([X, Y])$. Since $t \mapsto \Phi(\gamma(t))$ is a curve on H with derivative $\phi(X)$, the right-hand side equals $\mathrm{ad}(\phi(X))(\phi(Y)) = [\phi(X), \phi(Y)]$. Therefore these two are equal, concluding the proof. □

Combining the above two lemmas already proves that $[-, -]$ defines a Lie algebra structure on $\mathfrak{g}$ if G is an immersed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Indeed, in that case the inclusion $G \hookrightarrow \mathrm{GL}_n(\mathbb{R})$ induces an injective map $\mathfrak{g} \hookrightarrow \mathfrak{gl}_n(\mathbb{R})$ respecting the Lie bracket. So when we view $X, Y \in \mathfrak{g}$ as $n \times n$-matrices, Lemma 13.18 shows that $[X, Y] = XY - YX$, which can be explicitly seen to be a Lie algebra.

To prove that the bracket $[-, -]$ defines a Lie bracket in general, we postpone the proof to the next section.

Example 13.19. *You will show on the problem set that the derivative of* $\det\colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ *at the identity equals the trace* $\mathrm{tr}\colon \mathfrak{gl}_n(\mathbb{R}) \to \mathbb{R}$*. Therefore* $\mathfrak{sl}_n(\mathbb{R}) = \mathrm{Lie}(\mathrm{SL}_n(\mathbb{R})) \subset \mathfrak{gl}_n(\mathbb{R})$ *equals* $\{X \in \mathfrak{gl}_n(\mathbb{R}) \mid \mathrm{tr}(X) = 0\}$.

Warning 13.20. *In general, XY does not make sense as an element of $\mathfrak{g}$, it is only $[X, Y]$ that we can define. For example, if $X, Y \in \mathfrak{sl}_n(\mathbb{R})$, then the matrix XY does not necessarily lie in $\mathfrak{sl}_n(\mathbb{R})$. This is why we focus on the Lie algebra structure.*

Example 13.21. *A calculation similar to Lemma 13.17 shows that the Lie bracket on $\mathfrak{gl}_n(\mathbb{C}) = \mathrm{Lie}(\mathrm{GL}_n(\mathbb{C}))$ is $[A, B] = AB - BA$.*

Example 13.22. *The Lie algebra of $\mathrm{SO}(n) \subset \mathrm{GL}_n(\mathbb{R})$ turns out to be $\mathfrak{so}(n) = \{X \in \mathfrak{gl}_n(\mathbb{R}) \mid X + X^t = 0\}$. Indeed, the map $\mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{GL}_n(\mathbb{R}), A \mapsto AA^t$ has derivative $\mathfrak{gl}_n(\mathbb{R}) \to \mathfrak{gl}_n(\mathbb{R}), X \mapsto X + X^t$.*

Example 13.23. *Suppose that G is a commutative Lie group. Then $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. Indeed, in that case the conjugation maps Ψ_g are all trivial!*

13.6 One-parameter subgroups

Let G be a Lie group, $\mathfrak{g}$ its Lie algebra and $X \in \mathfrak{g}$. Even though X just defines a single element in the tangent space of G at e, we can use the group structure on G to extend it to a vector field. Recall that for $g \in G$, $L_g \colon G \to G, x \mapsto gx$ denotes the multiplication by g map.

Definition 13.24. *We say a vector field $V \colon G \to T_*G$ is left-invariant if $V(gh) = (dL_g)_h(V(h))$ for all $g, h \in G$.*

Proposition 13.25. *If $X \in \mathfrak{g}$, the association $V_X(g) \coloneqq (dL_g)_e(X)$ defines a vector field. The map $X \mapsto V_X$ is an isomorphism between $\mathfrak{g}$ and the set of all left-invariant vector fields.*

Proof. We first check that V_X is smooth. To prove this, it suffices to show (why?) that for every open $U \subset G$ and $f \in \mathcal{C}^\infty(U)$, $D_{V_X}(f) \in \mathcal{C}^\infty(U)$. But $D_{V_X}(f)(g) = V_X(g)(f) = (dL_g)_e(X)(f) = X(f \circ L_g)$. Let $\gamma \colon (-\epsilon, \epsilon) \to G$ be a curve with $\gamma(0) = e$ and $\gamma'(0) = X$. Then

$$X(f \circ L_g) = \frac{d}{dt}\bigg|_{t=0} f(g \cdot \gamma(t)) \tag{13.6.1}$$

Since the map $U \times (-\epsilon, \epsilon) \to \mathbb{R}, (g, t) \mapsto f(g \cdot \gamma(t))$ is smooth, taking the derivative of the t-variable and evaluating at 0 still gives a smooth map, so $g \mapsto X(f \circ L_g)$ is smooth, as desired.

To check that V_X is left-invariant, we compute for $g, h \in G$ that

$$V_X(gh) = (dL_{gh})(V_X(e)) = (dL_g)(dL_h)(V_X(e)) = dL_g(V_X(h)). \tag{13.6.2}$$

(We have omitted the subscripts from dL_g etcetera for ease of notation.)

We conclude that V_X is a left-invariant vector field. If V is any left-invariant vector field and $X = V(e) \in \mathfrak{g}$, then $V(g) = (dL_g)(V(e)) = (dL_g)(X) = V_X(g)$ for all $g \in G$ so $V = V_X$. So $X \mapsto V_X$ and $V \mapsto V(e)$ give mutually inverse bijections between $\mathfrak{g}$ and the set of left-invariant vector fields. $\square$

Remark 13.26. *It turns out that the Lie bracket of two left-invariant vector fields is again left-invariant, and that the map $X \mapsto V_X$ from Proposition 13.25 respects Lie brackets on both sides. This can be used to give an alternative definition of the Lie bracket on $\mathfrak{g}$.*

Lecture 16 starts here

The next proposition determines what the integral curves for the vector fields V_X look like: they turn out to be extremely important.

Proposition 13.27. *Let $X \in \mathfrak{g}$. Then there exists an integral curve $\varphi_X \colon \mathbb{R} \to G$ with $\varphi_X(0) = e$ for V_X. Moreover, it is the unique Lie group homomorphism $\varphi_X \colon \mathbb{R} \to G$ satisfying $\varphi_X'(0) = X$.*

Proof. By Theorem 12.25, there exists an open interval containing 0 and an integral curve for V through e. Moreover, any two integral curves for V_X through the same point agree on their domain of definition. Therefore, if $I \subset \mathbb{R}$ is the union of all open intervals containing 0 for which an integral curve through $e \in G$ exists, then we obtain a 'maximal' integral curve $\varphi_X : I \to G$ through e for V_X. The content of the proposition is that $I = \mathbb{R}$ and that φ_X is a homomorphism. The idea will be to use the group structure to 'copy' a small segment of the curve and extend it.

We first show that if $s \in I$ is fixed then $\gamma(t) = \varphi_X(s)\varphi_X(t)$ is an integral curve $I \to G$ for V_X through $g_0 := \varphi_X(s)$. Indeed, using the fact that φ_X is an integral curve and V_X is left-invariant, we have

$$\begin{aligned}
\gamma'(t) &= \frac{d}{dt}[L_{g_0}(\varphi_X(t))] && (13.6.3)\\
&= (dL_{g_0})_{\varphi_X(t)}(\varphi_X'(t)) && (13.6.4)\\
&= (dL_{g_0})_{\varphi_X(t)} V_X(\varphi_X(t)) && (13.6.5)\\
&= V_X(\varphi_X(s)\varphi_X(t)) && (13.6.6)\\
&= V_X(\gamma(t)). && (13.6.7)
\end{aligned}$$

Therefore $\gamma\colon I \to G$ is indeed an integral curve through φ_X. By the uniqueness of integral curves, we see that $\varphi_X(s+t) = \varphi_X(s)\varphi_X(t)$ whenever $s, t, s+t \in I$. We can therefore paste together $\varphi_X(t)$ and $\varphi_X(t-s)\varphi_X(s)$ and get a curve defined on the union of I and $s+I$. Since I is the maximal interval over which integral curves are defined, we see that $s + I \subset I$. Since I is non-empty by Theorem 12.25, we may take $s \in I$ nonzero, which implies that $I = \mathbb{R}$. In the course of this proof, we have also showed that $\varphi_X(s+t) = \varphi_X(s)\varphi_X(t)$ for all $s, t \in \mathbb{R}$, in other words φ_X is a Lie group homomorphism.

It remains to show that if $\phi\colon \mathbb{R} \to G$ is a Lie group homomorphism with $\phi'(0) = X \in \mathfrak{g}$, then $\phi = \phi_X$. To prove this, it suffices to show (by the uniqueness of integral curves) that ϕ is an integral curve of V_X. But this follows from the fact that $\phi'(t) = \frac{d}{ds}\big|_{s=0}\phi(t+s) = \frac{d}{ds}\big|_{s=0}(\phi(t)\phi(s)) = (dL_{\phi(t)})_e(X) = V_X(\phi(t))$. □

Definition 13.28. *If $X \in \mathfrak{g}$, the homomorphism $\varphi_X : \mathbb{R} \to G$ from Proposition 13.27 is called the one-parameter subgroup associated to G and $X \in \mathfrak{g}$.*

We emphasize that a one-parameter subgroup is, despite its name, a homomorphism $\mathbb{R} \to G$, and not just its image.

We can use the above proposition to show that $[-,-]$ is a well-defined Lie bracket for all Lie groups G (not necessarily embedded in $\mathrm{GL}_n(\mathbb{R})$), as promised in the previous section.

Proposition 13.29. *The bracket $[-,-]$ defines a Lie algebra structure on $\mathfrak{g}$.*

Proof. We just need to check that $[-,-]$ is alternating and satisfies the Jacobi identity.

To check that it is alternating, let $X \in \mathfrak{g}$ and let $\varphi_X : \mathbb{R} \to G$ be the one-parameter subgroup associated to X. Then

$$\begin{aligned}
[X, X] &= \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}(\varphi_X(t))(X) && (13.6.8)\\
&= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} \left[\varphi_X(t)\varphi_X(s)\varphi_X(t)^{-1}\right]. && (13.6.9)
\end{aligned}$$

Since $\varphi_X : \mathbb{R} \to G$ is a homomorphism, the elements in the image of φ_X commute with each other. So $\varphi_X(t)\varphi_X(s)\varphi_X(t)^{-1} = e$, and hence the derivative $[X, X]$ is zero.

To prove the Jacobi identity, apply Lemma 13.18 to $G = G$, $H = \mathrm{GL}(\mathfrak{g})$ and $\Phi = \mathrm{Ad}$. We get that $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ preserves the Lie bracket, that is $\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X$ for all $x, y \in \mathfrak{g}$. Plugging in $z \in \mathfrak{g}$ we get $[[x,y],z] = [x,[y,z]] - [y,[x,z]]$. Rearranging and using the skew-symmetry, this gives exactly the Jacobi identity! □

13.7 The exponential map

The exponential map is a crucial tool to relate the Lie group to its Lie algebra.

Definition 13.30. *The exponential map $\exp\colon \mathfrak{g} \to G$ is defined via $\exp(X) = \varphi_X(1) \in G$, where $\varphi_X : \mathbb{R} \to G$ is the one-parameter subgroup associated to G.*

We sometimes simply write e^X for $\exp(X)$ if $X \in \mathfrak{g}$. We also write $\varphi_X(t) = e^{tX}$ for the corresponding one-parameter subgroup.

Example 13.31. *Let $G = \mathrm{GL}_n(\mathbb{R})$. Then for $X \in \mathfrak{g}$, the expression $e^X := \sum_{n\geq 0} X^n/n!$ converges with respect to the norm $||X|| = \sqrt{\mathrm{tr}(XX^t)}$ on $\mathfrak{g} = \mathrm{Mat}_n(\mathbb{R})$ and defines an element of $\mathrm{GL}_n(\mathbb{R})$. The assignment $t \mapsto e^{tX} =$ is a Lie group homomorphism $\mathbb{R} \to G$ with derivative X. By the uniqueness of one-parameter subgroups, we see that $\varphi_X(t) = e^{tX}$ and so the exponential map in this case is exactly $\exp(X) = e^X = I + X + X^2/2 + X^3/3! + \cdots \in \mathrm{GL}_n(\mathbb{R})$, which explains the name.*

The following theorem summarizes the most important properties of the exponential map.

Theorem 13.32.

1. *$\exp$ is a smooth map.*
2. *$(d\exp)_0\colon T_0\,\mathfrak{g} = \mathfrak{g} \to T_eG = \mathfrak{g}$ is the identity map.*
3. *$\exp$ is a local diffeomorphism around $0 \in \mathfrak{g}$: there exists open subsets $U \subset \mathfrak{g}$ and $V \subset G$ with $0 \in U$ and $e \in V$ such that $\exp$ restricts to a diffeomorphism $\exp|_U\colon U \xrightarrow{\sim} V$.*
4. *$\exp$ is natural: if $F\colon G \to H$ is a Lie group homomorphism with derivative f then the following diagram is commutative:*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\exp} & G \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle F} \\ \mathfrak{h} & \xrightarrow{\exp} & H \end{array}$$

   In other words, $F \circ \exp = \exp \circ f$.

Proof.

1. This follows from the fact that integral curves (i.e. solutions to ODEs) depend smoothly on the initial parameters, so the map $\mathfrak{g} \times \mathbb{R}, (X,t) \mapsto \varphi_X(t)$ is smooth. (Detailed proof omitted.)
2. We have $(d\exp)_0(X) = \frac{d}{dt}\big|_{t=0} \exp(tX) = \frac{d}{dt}\big|_{t=0}\varphi_{tX}(1) = \frac{d}{dt}\big|_{t=0}\varphi_X(t) = X$.
3. Follows from the previous part and the inverse function theorem. (The inverse function theorem says that if $F\colon M \to N$ is smooth and $(dF)_p\colon T_pM \to T_qN$ is an isomorphism, there exist open subsets $U \subset M$ and $V \subset N$ containing p and q such that $F|_U\colon U \to V$ is a diffeomorphism.)
4. This follows from the fact that $\varphi_{f(X)} = F \circ \varphi_X$ for all $X \in \mathfrak{g}$, as both are one-parameter subgroups $\mathbb{R} \to H$ for $f(X)$ and the fact that one-parameter subgroups are unique. Therefore $F(\exp(X)) = F(\varphi_X(1)) = \varphi_{f(X)}(1) = \exp(f(X))$ for all $X \in \mathfrak{g}$.

□

Warning 13.33. *The exponential map is not a homomorphism in general. For example, for matrices $A, B \in \mathfrak{gl}_n(\mathbb{R})$ it is not necessarily true that $e^{A+B} = e^A \cdot e^B$. The latter identity does hold when A and B commute, and you will show on the problem set that more generally $e^{X+Y} = e^X \cdot e^Y$ if $X, Y \in \mathfrak{g}$ satisfy $[X, Y] = 0$.*

We can use the exponential map to show that a map between Lie groups (with connected domain) is determined by its Lie algebra. We need the following lemma, whose proof is an exercise on Problem Set 5.

Lemma 13.34. *Let $U \subset G$ be an open subset containing the identity in a connected Lie group G. Then U generates G.*

Proposition 13.35. *Let $F, F' \colon G \to H$ be Lie group homomorphisms with G connected, and let $f, f' \colon \mathfrak{g} \to \mathfrak{h}$ be the induced map on Lie algebras. Then $F = F'$ if and only if $f = f'$.*

Proof. If $F = F'$ then clearly $f = (dF)_e = (dF')_e = f'$, so assume for the converse that $f = f'$. By Lemma 13.34, it suffices to show that F and F' agree on an open neighbourhood of the identity of G. By Part 4 of Theorem 13.32, $F(\exp(X)) = \exp(f(X)) = \exp(f'(X)) = F'(\exp(X))$, so F and F' agree on the image of the exponential map $\exp\colon \mathfrak{g} \to G$. Since the image of the exponential map contains an open neighbourhood of the identity (Part 3 of Theorem 13.32), we conclude that $F = F'$. □

Remark 13.36. *We cannot drop the assumption that G is connected. For example, every finite group (with the discrete topology) defines a Lie group with Lie algebra the zero vector space. In fact, if $G^\circ \subset G$ is the connected component of the identity of G, then $\mathrm{Lie}(G^\circ) = \mathrm{Lie}(G)$, so the Lie algebra only 'detects' the identity component of G.*

13.8 Covering spaces of Lie groups

Lecture 17 starts here

In this section we consider covering spaces: these will turn out to be maps that 'don't change the Lie algebra'. We start by recalling some topological preliminaries.

Definition 13.37. *A continuous map $p\colon Y \to X$ between topological spaces is called a covering space (or covering map) if for every $x \in X$, there exists an open $x \in U \subset X$, such that $p^{-1}(U)$ is homeomorphic to a (possibly infinite) disjoint union of open sets $\{U_i\}_{i \in I}$, such that p restricts to a homeomorphism $p|_{U_i} \colon U_i \to U$ for each $i \in I$.*

We say a covering space $p\colon Y \to X$ is trivial if Y is homeomorphic to a disjoint union of open sets X_i, each mapping homeomorphically onto X. We note that every covering space is 'locally trivial', but there definitely are nontrivial covering spaces; this is what makes the theory interesting! We say a connected X is simply connected if every covering of X is trivial, equivalently every connected covering of X is a homeomorphism. It turns out that this is equivalent to the usual definition of simply connected, i.e. $\pi_1(X, x)$ is trivial for every $x \in X$.

Examples 13.38. *The topological spaces $\mathbb{R}^n$, S^n ($n \geq 2$) and their products are simply connected. The space S^1 is not simply connected: it has fundamental group $\mathbb{Z}$ and connected coverings of the form $S^1 \to S^1, z \mapsto z^n$ and $\mathbb{R} \to S^1$.*

Even if our space we started with is not simply connected, we can cover it by a simply connected space:

Proposition 13.39. *Let X be a connected and 'reasonable' topological space. Then there exists a covering space $p\colon \tilde{X} \to X$ with $\tilde{X}$ connected and simply connected. Moreover, there is a bijection between $p^{-1}(x)$ and $\pi_1(X, x)$ for every $x \in X$. We call $p\colon \tilde{X} \colon X$ a universal cover of X. If $p' \colon \tilde{X}' \to X$ is another universal cover of X, then there exists a homeomorphism $\phi\colon \tilde{X} \to \tilde{X}'$ such that $p' \circ \phi = p$.*

Reasonable above means: locally simply connected. Any manifold is reasonable.

Examples 13.40. *1.* $\mathbb{R} \to S^1, t \mapsto e^{2\pi i t}$ *is the universal cover of* S^1*, and the preimage of* $1 \in S^1$ *is* $\mathbb{Z} = \pi_1(S^1, 1)$.

2. $\mathbb{C} \to \mathbb{C}^\times, z \mapsto e^z$ *is the universal cover of* $\mathbb{C}^\times$.

3. If you know about real projective space $\mathbb{P}^n(\mathbb{R}) = (\mathbb{R}^{n+1} \setminus \{(0)\})/\mathbb{R}^\times$*, then you might appreciate that the natural map* $S^n \to \mathbb{P}^n(\mathbb{R})$ *is the universal cover for* $n \geq 2$*, so* $\mathbb{P}^n(\mathbb{R})$ *has fundamental group of order* 2.

For our Lie group purposes, we will state the following two facts without proof. (Their proofs are not very hard, but I don't want to get too much into the topology details.)

Proposition 13.41. *Let* G *be a Lie group.*

1. *Let* $\tilde{G} \to G$ *be a covering space. Let* $\tilde{e} \in \tilde{G}$ *be an element with* $\pi(\tilde{e}) = e$*. Then there exists a unique group structure and manifold structure on* $\tilde{G}$ *such that* $\tilde{G}$ *is a Lie group with identity* $\tilde{e}$ *and* π *is a Lie group homomorphism. The kernel* π *is a discrete (closed) subgroup of* $\tilde{G}$ *contained in the center of* $\tilde{G}$.

2. *Conversely, if* $\Gamma \leq G$ *is a discrete and central (i.e. contained in the center) subgroup, then* G/Γ *has the structure of a manifold and* $\pi\colon G \to G/\Gamma$ *is a Lie group homomorphism that is also a covering space.*

Proof. (Sketch of ideas) The idea of Part 2 is that the discreteness assumption shows that we can find around each $g \in G$ a sufficiently small chart (U, ϕ) with $g \in U \subset G$, so that the image of U in G/Γ is homeomorphic to U. Such a U can be used to define a chart on G/Γ around $g \mod \Gamma$, and it can be checked that this defines a manifold structure with the desired properties. For Part 1, it suffices to treat the case that $\tilde{G}$ is simply connected, for then we can obtain Lie group structures on intermediate coverings by Part 2. But in the simply connected case, there exists a unique continuous map $\tilde{G} \times \tilde{G} \to \tilde{G}$ lifting the multiplication map on G and mapping $(\tilde{e}, \tilde{e})$ to $\tilde{e}$. One then checks that this map defines a group structure on $\tilde{G}$. To define the manifold structure, for each $g \in G$ we can find a chart $\phi\colon U \to \phi(U)$ containing g such that $p^{-1}(U) = \sqcup U_i$ where each U_i maps homeomorphically onto U. Then we declare $U_i \xrightarrow{p} U \xrightarrow{\phi} \phi(U)$ to be charts, and one can check that this defines a manifold structure such that $\tilde{G}$ becomes a Lie group. The fact that the kernel is central follows from PSET5 Q4. □

The above proposition shows that the universal cover of a connected Lie group is again a Lie group: we call it the universal cover or simply connected cover of G and denote it by G^{sc} (or sometimes $\tilde{G}$).

Proposition 13.42. *Let* G, H *be connected Lie groups. Then a Lie group homomorphism* $\pi\colon G \to H$ *is a covering space if and only if* $(d\pi)_e\colon \mathfrak{g} \to \mathfrak{h}$ *is an isomorphism.*

Proof. (Sketch of ideas) On Problem Set 5 you have shown that π covering space $\Rightarrow (d\pi)_e$ is an isomorphism. For the converse, assume that $(d\pi)_e$ is an isomorphism. By left-translating, it follows that $(d\pi)_g$ is an isomorphism for all $g \in G$. By the inverse function theorem, π is a local diffeomorphism at every point $g \in G$. Therefore the image of π contains an open neighbourhood of the identity of H, and hence by Lemma 13.34, π is surjective. The kernel of π is a discrete and normal (and hence central by PSET5 Q4) subgroup of G. By Part (2) of Proposition 13.41, $G/\ker(\pi)$ is a Lie group and the induced map $G/\ker(\pi) \to H$ is a bijective Lie group homomorphism whose derivative at every point is an isomorphism. By the inverse function theorem, the inverse of this map is smooth and so $G/\ker(\pi) \simeq H$ and we conclude that $G \to G/\ker(\pi) = H$ is a covering space. □

Definition 13.43. *Let* G, H *be connected Lie groups. A Lie group homomorphism* $G \to H$ *that is also a covering space is called an* isogeny*. We say two Lie groups* G *and* H *are* isogenous *if there exists an isogeny* $G \to H$ *or*

$H \to G$. We say G and H are in same isogeny class if there exists a third Lie group K that is isogenous to both G and H.

Being isogenous is not an equivalence relation, but being in the same isogeny class is (it is the equivalence relation generated by isogeny). Two Lie groups are in the same isogeny class if and only if their universal covers are isomorphic as Lie groups. (Exercise!)

Proposition 13.42 shows that two connected Lie groups lying in the same isogeny class have isomorphic Lie algebras. We will show in the next section that the converse is also true.

Example 13.44. *The map $\mathbb{R} \to S^1, t \mapsto e^{2\pi i t}$ is a covering space and a Lie group homomorphism.*

Example 13.45. *On Problem Set 5 you will show that there exists a Lie group homomorphism $\pi\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ with kernel $\{\pm 1\}$. This implies that π is a covering map. Indeed, by Proposition 13.42 it suffices to show that $(d\pi)_e$ is an isomorphism. Since both $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are three-dimensional, it suffices to show that $(d\pi)_e$ is injective. But if $X \in \mathrm{Lie}(\mathrm{SU}(2))$ lies in the kernel of $(d\pi)_e$, then by naturality of the exponential map the one-parameter subgroup $\varphi_X : \mathbb{R} \to \mathrm{SU}(2)$ lands in the kernel of π, that is $\{\pm 1\}$. Since this kernel is discrete and $\mathbb{R}$ is connected, $\varphi_X(t)$ is constant, so $\varphi_X'(0) = X = 0$. This shows that $(d\pi)_e$ is an isomorphism and hence π is a covering space.*

Since $\mathrm{SU}(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1 \right\}$ is homeomorphic to the unit sphere S^3 in $\mathbb{C}^2 = \mathbb{R}^4$, it is simply connected (indeed, it is a fact from topology that S^n is simply connected for all $n \geq 2$). So $\mathrm{SU}(2)$ does not have any nontrivial connected coverings. It turns out that the center of $\mathrm{SU}(2)$ is exactly $\{\pm I\}$, so $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are the only two members in their isogeny class by Proposition 13.41.

Example 13.46. *Combining the above two examples, we see that the Lie groups $\mathbb{R} \times \mathrm{SO}(3)$ and $S^1 \times \mathrm{SU}(2)$ are not isogenous but are in the same isogeny class: they are both covered by $\mathbb{R} \times \mathrm{SU}(2)$.*

Example 13.47. *The group $\mathrm{SO}(2)$ is isomorphic to S^1 and so has universal cover $\mathbb{R}$. The groups $\mathrm{SO}(n)$ for $n \geq 3$ have fundamental group $\mathbb{Z}/2\mathbb{Z}$ (this is not obvious) and their universal covers are double covers $\mathrm{Spin}(n) \to \mathrm{SO}(n)$, called* spin groups. *Apparently they are important in physics.*

Example 13.48. *Proposition 13.41 shows that quotients of Lie groups by discrete central subgroups are still Lie groups. For example, if $G = \mathrm{SL}_2(\mathbb{R})$ and $\Gamma = \{\pm I\}$ then $G/\Gamma = \mathrm{PSL}_2(\mathbb{R})$ is a Lie group.*

We can use the exponential map and the notion of a covering space to classify connected commutative Lie groups.

Proposition 13.49. *Let G be a connected commutative Lie group. Then G is isomorphic to a product of copies of S^1 and $\mathbb{R}$.*

Proof. On Problem Set 5, you have shown that if $X, Y \in \mathfrak{g}$ then $\exp(X + Y) = \exp(X)\exp(Y)$. In particular, since G is commutative the Lie bracket on $\mathfrak{g}$ is trivial and so $\exp\colon \mathfrak{g} \to G$ is a Lie group homomorphism. Since the image of $\exp$ contains an open neighbourhood of the identity and since such an open neighbourhood generates G, we see that $\exp$ is surjective. Since the derivative at 0 of $\exp$ is an isomorphism, we conclude that $\exp$ is a covering space. It follows that G is isomorphic to the quotient of $\mathfrak{g} \simeq \mathbb{R}^n$ by a discrete subgroup Γ. It turns out that such a subgroup is always of the form $\mathbb{Z}v_1 + \cdots + \mathbb{Z}v_k$, where $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ is a linearly independent set of vectors. (We omit the proof of this standard fact.) It follows that $G \simeq \mathbb{R}^n/\Gamma \simeq (S^1)^k \times \mathbb{R}^{n-k}$. $\square$

13.9 Lie group-Lie algebra correspondence: statements

Lecture 18 starts here

Let G, H be Lie groups with G connected. Proposition 13.35 shows that the map $F \mapsto (dF)_e$

$$\mathrm{Hom}_{\mathsf{LieGrp}}(G, H) \to \mathrm{Hom}_{\mathsf{LieAlg}/\mathbb{R}}(\mathfrak{g}, \mathfrak{h}) \tag{13.9.1}$$

is an injection. This raises some natural questions:

1. What is the image of this map? In other words, which Lie algebra homomorphisms $\mathfrak{g} \to \mathfrak{h}$ can be realized as derivatives of Lie group homomorphisms?
2. If $\mathfrak{g} \simeq \mathfrak{h}$, does this imply that $G \simeq H$?
3. Is every (abstract) Lie algebra isomorphic to the Lie algebra of a Lie group?

The short answers: 1) the map (13.9.1) is bijective when G is simply connected; 2) no: $\mathfrak{g} \simeq \mathfrak{h}$ if and only if G and H lie in the same isogeny class; 3) Yes!

In more detail, we will state the following theorems and prove them in the next section.

Theorem 13.50 (The homomorphism theorem)**.** *Let G, H be Lie groups, and assume that G is connected and simply connected. Then* (13.9.1) *is bijective. In other words, every Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{h}$ is of the form $(dF)_e$ for some unique Lie group homomorphism $F\colon G \to H$.*

Remark 13.51. *The simply connectedness assumption cannot be dropped. For example, on Problem Set 5 you have shown that every Lie group homomorphism $S^1 \to S^1$ is of the form $z \mapsto z^n$. But there are way more Lie algebra homomorphisms $\mathrm{Lie}(S^1) = \mathbb{R} \to \mathbb{R}$.*

The second theorem identifies Lie subgroups with Lie subalgebras.

Theorem 13.52 (Subgroup-subalgebra correspondence)**.** *Let G be a connected Lie group. The map $(H \hookrightarrow G) \mapsto (\mathfrak{h} \subset \mathfrak{g})$ induces a bijection between the set of connected immersed Lie subgroups of G and the set of Lie subalgebras of $\mathfrak{g}$. (I.e. subspaces preserved under the Lie bracket.)*

Remark 13.53. *Let $G = S^1 \times S^1$, so $\mathfrak{g} = \mathbb{R}^2$ with the trivial Lie bracket. Then Lie subalgebras of $\mathfrak{g}$ are just subspaces $\mathfrak{h} \subset \mathbb{R}^2$. Such a proper nonzero subspace is simply a line in $\mathbb{R}^2$. The subgroup in $S^1 \times S^1$ corresponding to this line is simply the image of this line under the covering map $\mathbb{R}^2 \to S^1 \times S^1$. For example, if the line has irrational slope, then we get an immersed subgroup $\mathbb{R} \hookrightarrow S^1 \times S^1$, see Example 12.35. This shows that we really need to consider* immersed *subgroups in Theorem 13.52, not just embedded ones.*

The third theorem says that Lie algebras really are the right linear algebraic objects to consider when studying Lie groups.

Theorem 13.54 (Lie's third theorem)**.** *Every finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is of the form $\mathrm{Lie}(G)$ for some Lie group G.*

Corollary 13.55. *Let G, H be connected Lie groups. Then $\mathfrak{g} \simeq \mathfrak{h}$ (as Lie algebras) if and only if G and H are in the same isogeny class, if and only if the simply connected covers of G and H are isomorphic as Lie groups.*

Proof. Write $G^{sc} \to G$ and $H^{sc} \to H$ for the universal covers of G and H (they are unique up to isomorphism). Then G and H lie in the same isogeny class if and only if $G^{sc} \simeq H^{sc}$ as Lie groups. If this holds, then $\mathfrak{g} \simeq \mathrm{Lie}(G^{sc}) \simeq \mathrm{Lie}(H^{sc}) \simeq \mathfrak{h}$. Conversely, suppose that $\mathfrak{g} \simeq \mathfrak{h}$ and let $\phi\colon \mathfrak{g} \to \mathfrak{h}$ and $\psi\colon \mathfrak{h} \to \mathfrak{g}$ be mutually inverse Lie algebra isomorphisms. By Theorem 13.50, there exist (unique) Lie group homomorphisms $\Phi\colon G^{sc} \to H^{sc}$ and $\Psi\colon H^{sc} \to G^{sc}$ with $d\Phi = \phi$ and $d\Psi = \psi$. Since $\Psi \circ \Phi\colon G^{sc} \to G^{sc}$ has differential equal to $\psi \circ \phi = \mathrm{Id}$, Proposition 13.35 shows that $\Psi \circ \Phi = \mathrm{Id}_G$. Similarly $\Phi \circ \Psi = \mathrm{Id}_H$. Therefore Φ and Ψ are mutually inverse Lie group homomorphisms, so $G^{sc} \simeq H^{sc}$ so G and H are in the same isogeny class. $\square$

Remark 13.56. *Let G, H be connected and simply connected Lie groups. Then the above corollary says:* $\mathrm{Lie}(G) \simeq \mathrm{Lie}(H)$ *if and only if* $G \simeq H$.

13.10 Lie group-Lie algebra correspondence: proofs

We indicate the proofs of the three theorems from the previous section. It will turns out the the subgroup-subalgebra correspondence will imply both the homomorphism theorem and Lie's third theorem. So assume the subgroup-subalgebra correspondence (Theorem 13.52) for now.

Proof of homomorphism Theorem 13.50. Let G, H be Lie groups with G connected and simply connected. Let $\phi\colon \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. We will use Theorem 13.52 to construct a homomorphism $\Phi\colon G \to H$ with differential ϕ. The key idea is to encode a function $f\colon X \to Y$ using its graph $\{(x, y) \in X \times Y \mid f(x) = y\}$. In this contex, let $\mathfrak{k} = \{(x, y) \in \mathfrak{g} \times \mathfrak{h} \mid \phi(x) = y\}$. Since ϕ is a Lie algebra homomorphism, $\mathfrak{k}$ is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$ (where the Lie bracket is given by $[(x, y), (x', y')] = ([x, x'], [y, y'])$). By Theorem 13.52, there exists a connected immersed Lie subgroup $K \hookrightarrow G \times H$ whose Lie algebra equals $\mathfrak{k}$. Let $p\colon K \to G$ be the projection of K onto the G-component. Then the derivative $(dp)_e\colon \mathfrak{k} \to \mathfrak{g}$ equals the projection of $\mathfrak{k}$ onto the $\mathfrak{g}$ component. Since $\mathfrak{k}$ is the graph of a function, this projection map $\mathfrak{k} \to \mathfrak{g}$ is an isomorphism. It follows that $(dp)_e$ is an isomorphism. By Proposition 13.42, this implies that p is a covering space. But since G is assumed to be simply connected, it has no nontrivial connected coverings! In other words, p must be an isomorphism $K \to G$. It follows that the composite map $G \xrightarrow{p^{-1}} K \hookrightarrow G \times H \to H$ is a Lie group homomorphism whose derivative at e is ϕ. ☐

To prove Lie's third Theorem 13.54, we will use the following purely algebraic result (no proof given):

Theorem 13.57 (Ado). *Every finite-dimensional Lie algebra $\mathfrak{g}$ over a field F of characteristic zero is isomorphic to a subalgebra of $\mathfrak{gl}_n(F)$.*

Proof of Lie's third Theorem 13.54. Let $\mathfrak{g}$ be a (finite-dimensional) Lie algebra over $\mathbb{R}$, and choose an embedding $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{R})$. By Theorem 13.52, there exists an immersed subgroup $G \hookrightarrow \mathrm{GL}_n(\mathbb{R})$ with Lie algebra $\mathfrak{g}$. ☐

Remark 13.58. *It is not true that every Lie group is an immersed Lie subgroup of* $\mathrm{GL}_n(\mathbb{R})$. *For example, it turns out (but is certainly not obvious) that the universal cover of* $\mathrm{SL}_2(\mathbb{R})$ *is not an immersed subgroup of* $\mathrm{GL}_n(\mathbb{R})$. *However, most Lie groups that we encounter in 'nature' are closed subgroups of* $\mathrm{GL}_n(\mathbb{R})$.

Remark 13.59. *The proof of Theorem 13.54 shows that every Lie group lies in the same isogeny class as an (immersed) Lie subgroup of* $\mathrm{GL}_n(\mathbb{R})$ *for some $n \geq 1$. Indeed, let $\mathfrak{g} = \mathrm{Lie}(G)$. By Ado's theorem, $\mathfrak{g}$ is isomorphic to a subalgebra of $\mathfrak{gl}_n(\mathbb{R})$. By the subgroup-subalgebra correspondence, there exists a connected Lie subgroup $G' \subset \mathrm{GL}_n(\mathbb{R})$ with* $\mathrm{Lie}(G') \simeq \mathfrak{g}$. *Therefore* $\mathrm{Lie}(G) \simeq \mathrm{Lie}(G')$. *By Corollary 13.55, G and G' lie in the same isogeny class.*

The last remaining piece is Theorem 13.52, i.e. the correspondence between Lie subgroups and Lie subalgebras. So given a connected Lie group G and a subalgebra $\mathfrak{h} \subset \mathfrak{g}$, we need to somehow produce a Lie subgroup H in G. Recall that $\exp(\mathfrak{g})$ contains an open neighbourhood of the identity in G, so by Lemma 13.34, $\exp(\mathfrak{g})$ generates G. Therefore a natural candidate for H would be the subgroup generated by $\exp(\mathfrak{h})$. But how can we guarantee that this is not bigger than expected, i.e. that $\mathrm{Lie}(H) = \mathfrak{h}$? Here is where we need to use that $\mathfrak{h}$ is a *subalgebra* of $\mathfrak{g}$, not merely a subspace. This subalgebra property will turn out to imply that $\exp(\mathfrak{h})$ is a 'subgroup' near the identity of G: this will be a consequence of the amazing Baker–Campbell–Hausdorff formula. This 'subgroup near the identity' property will turn out to be enough to prove that H has the desired properties.

First, some preliminaries. Let G be a connected Lie group, $\exp\colon \mathfrak{g} \to G$ be the exponential map, and choose by Theorem 13.32 open neighborhoods $U \subset \mathfrak{g}$ and $V \subset G$ of 0 and e such that $\exp|_U \colon U \to V$ is a diffeomorphism. Write $\log\colon V \to U$ for the inverse of this diffeomorphism. Since the map $U \times U \to \mathfrak{g}, (X,Y) \mapsto \exp(X)\cdot\exp(Y)$ is continuous, we may shrink U and V so that $\log$ is defined on the image of this map. For $X, Y \in U \subset \mathfrak{g}$, let

$$X * Y \coloneqq \log(\exp(X)\cdot\exp(Y)) \in \mathfrak{g}\,. \tag{13.10.1}$$

In other words: we take two elements of the Lie algebra, exponentiate, multiply and take the logarithm to go back to the Lie algebra. We emphasize that $\log$ is only defined on a small open neighborhood of the identity in G, and $X * Y$ is only defined for X, Y sufficiently close to 0 in $\mathfrak{g}$. Lecture 19 starts here

Example 13.60. *If $G = \mathrm{GL}_n(\mathbb{R})^+$, then*

$$\log(g) = (g-I) - \frac{(g-I)^2}{2} + \frac{(g-I)^3}{3} - \dots \tag{13.10.2}$$

defines an element of $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{R})$ for every $g \in \mathrm{GL}_n(\mathbb{R})^+$ sufficiently close to the identity I, and where it is defined it is an inverse to the exponential map. If $X, Y \in \mathfrak{gl}_n(\mathbb{R})$, what is $X * Y$? It is the logarithm of

$$\exp(X)\cdot\exp(Y) = \left(I + X + X^2/2! + \dots\right)\cdot\left(I + Y + Y^2/2! + \dots\right) \tag{13.10.3}$$

$$= I + (X+Y) + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right) + \cdots \tag{13.10.4}$$

where the $\cdots$ involve at least 3 terms X, Y and where we have to be careful to remember that we cannot assume that $XY = YX$! Set $g = \exp(X)\cdot\exp(Y)$. Ignoring terms of order ≥ 3, we have that $X * Y = \log(g)$ equals

$$(g-I) - \frac{(g-I)^2}{2} + \cdots = (X+Y) + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right) - \frac{1}{2}\left((X+Y) + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right)\right)^2 + \cdots$$

Working this out and forgetting about order ≥ 3 terms, we end up at

$$\begin{aligned} X * Y &= X + Y + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right) - \frac{1}{2}(X+Y)\cdot(X+Y) + \cdots \\ &= X + Y + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right) - \frac{1}{2}\left(X^2 + XY + YX + Y^2\right) \\ &= X + Y + \frac{1}{2}\left(XY - YX\right) \\ &= (X+Y) + \frac{1}{2}[X,Y] + \cdots \end{aligned}$$

So after some calculations, we see that the first terms only involve addition and the Lie bracket! Note that there is no clear a priori reason why this should be true: the terms X^2 and Y^2 just magically cancel out and we're left with $\frac{1}{2}(XY - YX)$. Going further, it turns out that the degree 3 term equals

$$\frac{1}{12}\left(X^2Y + XY^2 - 2XYX + Y^2X + YX^2 - 2YXY\right). \tag{13.10.5}$$

Magically, this turns out to be $\frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]]$, i.e. only involves the Lie brackets. The degree 4 term equals

$$\frac{1}{24}\left(X^2Y^2 - 2XYXY - Y^2X^2 + 2YXYX\right), \tag{13.10.6}$$

which turns out to be simply $\frac{1}{24}[Y,[X,[X,Y]]]$. It follows that for X, Y sufficiently small matrices, we have

$$X * Y = (X+Y) + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \frac{1}{24}[Y,[X,[X,Y]]] + \cdots \tag{13.10.7}$$

It turns out that the dots in (13.10.7) only involve taking more and more commutators of X and Y. This remarkable fact and its generalization to an arbitrary Lie algebra is the content of the BCH formula.

Proposition 13.61 (Baker–Campbell–Hausdorff formula). *After possible shrinking U to an even smaller open neighborhood of 0, we have for all $X, Y \in U$:*

$$X * Y = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdots \tag{13.10.8}$$

where the dots only involve (infinitely many) iterated brackets of X and Y.

Implicit in the statement is that the right hand side of (13.10.8) converges for all $X, Y \in U$. It is possible to give an explicit formula for the iterated brackets appearing in this formula (this is due to Dynkin), but it is quite complicated. For most applications the mere existence of this formula is already interesting enough. Indeed, it implies that

The group law on G 'near' the identity is entirely determined by the Lie algebra!

Indeed, if $g, h \in G$ are 'very close to' e, we may write $X = \log(g), Y = \log(h)$, and then $g \cdot h = \exp(X * Y)$, and $X * Y$ is expressible only using the Lie algebra, i.e. a sum of iterated Lie brackets of X and Y.

We omit the proof of the BCH formula; it follows from computing the derivative of the exponential map at all points of $\mathfrak{g}$, see Hall's Lie groups book for a nice exposition. We will use the BCH formula to prove the Lie subgroup-subalgebra correspondence. The following lemma will be useful, whose proof is largely manifold theoretic and I will only sketch (it is certainly non-examinable).

Lemma 13.62. *In the above notation, let $S \subset \mathfrak{g}$ be a subspace, let $H_0 = \exp(S \cap U)$ and suppose that H_0 is a 'local subgroup': $H_0^{-1} = H_0$ and $H_0 \cdot H_0 \cap \exp(U) = H_0$. Then the subgroup H generated by H_0 is an immersed Lie subgroup of G and H_0 is an open neighbourhood of the identity of H.*

Proof sketch. We have $H = \cup_{h \in H} h \cdot H_0$. For each $h \in H$ we have bijections $U \cap S \xrightarrow{\exp} H_0 \xrightarrow{h \cdot (-)} h \cdot H_0$; denote the composite $U \cap S \to h \cdot H_0$ by ϕ_h. We endow H with the unique topology such that $h \cdot H_0$ is open and $\phi_h \colon U \cap S \to h \cdot H_0$ is a homeomorphism for each $h \in H$. This topology on H is well defined, and will in general differ from the subspace topology on H. Upon choosing a basis of S to identify S with $\mathbb{R}^n$, each $\phi_h^{-1} \colon h \cdot H_0 \to U \cap S \subset S = \mathbb{R}^n$ defines a chart of H. These charts are all compatible and turn out to define a manifold structure on H. (Details omitted.) This gives H the structure of a Lie group, and the map $H \to G$ is an injective immersion. Therefore H is an immersed Lie subgroup and $H_0 \subset H$ is by definition an open neighbhourhood containing the identity. □

Proof of subgroup-subalgebra correspondence 13.52. Let $\mathfrak{h} \subset \mathfrak{g}$ be a subalgebra. Let $U \subset \mathfrak{g}$ be a sufficiently small open neighbourhood of $0 \in \mathfrak{g}$, more precisely an open neighbourhood for which the conclusion of Proposition 13.61 holds. By possibly replacing U by $U \cap (-U)$, we may assume that $U = -U$, i.e. $-X \in U$ for all $X \in U$. Let H be the subgroup of G generated by $H_0 := \exp(\mathfrak{h} \cap U)$. We will show that H is an immersed subgroup of G with Lie algebra $\mathfrak{h}$. By Lemma 13.62, it suffices to show that H_0 is a 'local subgroup' of G, i.e. that $H_0^{-1} = H_0$ and $H_0 \cdot H_0 \cap \exp(U) = H_0$. Indeed, in that case H_0 is an open neighbourhood of the identity in H so $\mathrm{Lie}(H) = T_e H = T_e H_0 = \mathfrak{h}$. (The latter equality follows from $\exp$ being a diffeomorphism when restricted to U.)

We have $H_0^{-1} = \{\exp(X)^{-1} \mid X \in \mathfrak{h} \cap U\} = \{\exp(-X) \mid X \in \mathfrak{h} \cap U\}$. Since $U = -U$ by assumption and $-\mathfrak{h} = \mathfrak{h}$ since $\mathfrak{h}$ is a subspace, $\{\exp(-X) \mid X \in \mathfrak{h} \cap U\} = \{\exp(X) \mid X \in \mathfrak{h} \cap U\} = H_0$. This shows that $H_0^{-1} = H_0$.

It remains to show that $H_0 \cdot H_0 \cap \exp(U) = H_0$. But if $X, Y \in U \cap \mathfrak{h}$ then $\exp(X) \cdot \exp(Y) = \exp(X * Y)$ by definition of $X * Y$. So we need to show that $X * Y \in \mathfrak{h}$. But this follows from the BCH formula! Indeed, since $\mathfrak{h}$ is a subalgebra, $[X, Y] \in \mathfrak{h}$ and all repeated commutators like $[X, [X, Y]]$ and $[Y, [X, Y]]$ again lie in $\mathfrak{h}$. So by Formula (13.10.8), each term on the right lies in $\mathfrak{h}$. Since the expression converges to $X * Y$, the latter also lies in $\mathfrak{h}$. □

13.11 Representation theory of Lie groups

Recall that if V is a $\mathbb{C}$-vector space, then $\mathrm{GL}(V)$ has the structure of a Lie group: a choice of basis gives a Lie group isomorphism with $\mathrm{GL}_n(\mathbb{C})$.

Definition 13.63. *A representation a Lie group G is a Lie group homomorphism*

$$R\colon G \to \mathrm{GL}(V), \tag{13.11.1}$$

where V is a finite-dimensional $\mathbb{C}$-vector space.

In other words, after choosing a basis a representation of G is nothing but a Lie group homomorphism

$$G \to \mathrm{GL}_n(\mathbb{C}), \tag{13.11.2}$$

so the representation theory of Lie groups is the same as that of finite groups, except that we add the word 'smooth' everywhere. As usual, given a representation $\rho\colon G \to \mathrm{GL}(V)$, we will often view V as a G-module and simply write $g \cdot v$ for $\rho(g) \cdot v$. We will use the same terminology as in the representation theory of finite groups: subrepresentation, irreducible representation, direct sum of representations, G-homomorphism or G-equivariant homomorphism, isomorphisms of representations,...

What happens if we take the derivative of a representation? If $\rho\colon G \to \mathrm{GL}(V)$ is a representation then $(dR)_e$ defines a homomorphism of Lie algebras

$$\rho := (dR)_e\colon\ \mathfrak{g} \to \mathfrak{gl}(V). \tag{13.11.3}$$

(Recall that $\mathfrak{gl}(V)$ is the Lie algebra $\mathrm{End}(V)$ with Lie bracket given by the commutator.) Explicitly, if $X \in \mathfrak{g}$ then $\rho(X) = \frac{d}{dt}\big|_{t=0} R(e^{tX})$. This motivates the following definition:

Definition 13.64. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$. A representation of $\mathfrak{g}$ is Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{gl}(V)$ for some finite-dimensional $\mathbb{C}$-vector space V.*

Note that $\mathfrak{gl}(V)$ is actually a vector space and Lie algebra over $\mathbb{C}$, but here we merely see it as a Lie algebra over $\mathbb{R}$. In other words, a representation of a Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is an $\mathbb{R}$-linear map $\rho\colon\ \mathfrak{g} \to \mathfrak{gl}(V)$ preserving the Lie bracket, i.e. such that $\rho([x, y]) = \rho(x) \circ \rho(y) - \rho(y) \circ \rho(x)$ for all $x, y \in \mathfrak{g}$. We will usually simply write $X \cdot v$ for $\rho(X)(v)$ if $X \in \mathfrak{g}$ and $v \in V$.

Similarly to the case of finite groups and Lie groups, there is the notion of a subrepresentation of a $\mathfrak{g}$-representation V: a subspace $W \leq V$ such that $X \cdot w \in W$ for all $X \in \mathfrak{g}$ and $w \in W$. Correspondingly, there is the notion of irreducible representation. If V and W are $\mathfrak{g}$-representations, we say a $\mathbb{C}$-linear map $f\colon V \to W$ is a $\mathfrak{g}$-homomorphism (or $\mathfrak{g}$-equivariant) if $f(X \cdot v) = X \cdot f(v)$ for all $v \in V$. We say two $\mathfrak{g}$-representations are isomorphic if there exists a $\mathfrak{g}$-equivariant isomorphism between them.

Lecture 20 starts here

Example 13.65. *Let V be a representation of a Lie group G. Then $V^* = \mathrm{Hom}(V, \mathbb{C})$ is again a representation of G via $(g \cdot \phi)(v) = \phi(g^{-1} \cdot v)$. What does the $\mathfrak{g}$-representation structure on V^* look like? For $X \in \mathfrak{g}$ and $\phi \in V^*$ and $v \in V$, we compute $(X \cdot \phi)(v) = \frac{d}{dt}\big|_{t=0} \phi((e^{tX})^{-1} \cdot v) = \frac{d}{dt}\big|_{t=0} \phi((e^{-tX}) \cdot v)$. By the chain rule and the fact that the derivative of the linear map ϕ equals ϕ itself (after identifying the tangent spaces with the ambient space), this equals $\phi(-X \cdot v)$.*

Example 13.66. *On PSET6, you have shown that if V, W are G-representations, the G-representation $V \otimes W$ induces a $\mathfrak{g}$-representation with formula $X \cdot (v \otimes w) = (Xv) \otimes w + v \otimes (Xw)$ for all $X \in \mathfrak{g}$, $v \in V$ and $w \in W$.*

There is a very tight relation between the representation theory of a connected Lie group G and that of its Lie algebra. For example:

Lemma 13.67. *Let G be a connected Lie group and $R\colon G \to \mathrm{GL}(V)$ a representation. Let $\rho = (dR)_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be the derivative of R, a representation of $\mathfrak{g}$. Then*

1. *A subspace $W \subset V$ is stable under G (i.e. $g \cdot w \in W$ for all $g \in G, w \in W$) if and only if W is stable under $\mathfrak{g}$ (i.e. $X \cdot w \in W$ for all $X \in \mathfrak{g}, w \in W$).*
2. *V is irreducible as a G-representation if and only if V is irreducible as a $\mathfrak{g}$-representation.*
3. *An element $v \in V$ is a G-fixed vector if and only if $X \cdot v = 0$ for all $X \in \mathfrak{g}$.*
4. *If W is another G-representation, a $\mathbb{C}$-linear map $V \to W$ is a G-homomorphism if and only if it is a $\mathfrak{g}$-homomorphism.*
5. *Let V and W be G-representations. Then V and W are isomorphic if and only if the corresponding $\mathfrak{g}$-representations are isomorphic.*

Proof.

1. Recall that for $X \in \mathfrak{g}$ and $v \in V$, $X \cdot v (= \rho(X)(v))$ equals $\frac{d}{dt}\big|_{t=0} R(e^{tX})(v) = \frac{d}{dt}\big|_{t=0} e^{tX} \cdot v$. Now suppose that W is G-stable. Then the above formula shows that for all $X \in \mathfrak{g}$ and $w \in W$, $X \cdot w = \frac{d}{dt}\big|_{t=0} e^{tX} \cdot w \in W$, so W is $\mathfrak{g}$-stable. Conversely, suppose that W is $\mathfrak{g}$-stable. By naturality of the exponential map applied to $R\colon G \to \mathrm{GL}(V)$, $R(e^X) = e^{\rho(X)}$ for all $X \in \mathfrak{g}$. Here e^X denotes the exponential map in $\mathfrak{g}$, and $e^{\rho(X)}$ denotes the exponential map in $\mathfrak{gl}(V)$. Therefore if $w \in W$ then
$$e^X \cdot w = \sum_{n \geq 0} \frac{\rho(X)^n}{n!} \cdot w. \tag{13.11.4}$$
Each term on the right hand side lies in W and W is a closed subspace of V, since any subspace of a finite-dimensional $\mathbb{C}$-vector space is closed. It follows that $e^X \cdot w \in W$ for all $w \in W$ and $X \in \mathfrak{g}$. Since G is connected, the image of the exponential map generates G, so $g \cdot w \in W$ for all $w \in W$ and $g \in G$, so W is G-stable.
2. Follows from part 1.
3. Exercise.
4. Let $U = \mathrm{Hom}_{\mathbb{C}}(V, W)$. Then in analogy with the finite groups case (Definition 6.23), the association $(g \cdot F)(v) = g \cdot F(g^{-1} \cdot v)$ gives U the structure of a G-representation. Moreover, a G-fixed vector in U is the same as a G-equivariant homomorphism $V \to W$. By combining examples 13.65 and 13.66 and recalling the isomorphism $U \simeq V^* \otimes W$ from Proposition 7.8, we compute that the derivative of this G-representation on U is a $\mathfrak{g}$-representation with explicit formula $(X \cdot F)(v) = X \cdot F(v) - F(X \cdot v)$ for all $X \in \mathfrak{g}$ and $v \in V$. It follows that $X \cdot F = 0$ if and only if F is $\mathfrak{g}$-equivariant. The proof now follows from part 3 applied to U.
5. Follows from part 5.

$\square$

The next proposition is a consequence of the homomorphism theorem:

Proposition 13.68. *Let G be a connected, simply connected Lie group. Then taking the derivative induces a bijection:*

$$\mathrm{Hom}_{\mathit{LieGrp}}(G, \mathrm{GL}_n(\mathbb{C})) \xrightarrow{\sim} \mathrm{Hom}_{\mathit{LieAlg}/\mathbb{R}}(\mathfrak{g}, \mathfrak{gl}_n(\mathbb{C})). \tag{13.11.5}$$

The left hand side of (13.11.5) is the set of all Lie group homomorphisms $G \to \mathrm{GL}_n(\mathbb{C})$; the right hand side equals the set of $\mathbb{R}$-linear maps $\mathfrak{g} \to \mathfrak{gl}_n(\mathbb{C})$ preserving the Lie bracket. This proposition reduces the study of representations of Lie groups (essentially) to the study of representations of Lie algebras!

Example 13.69. *Let $G = \mathrm{SU}(2)$. We have seen that G is connected and simply connected (since as a topological space it is homeomorphic to S^3). Therefore by Proposition 13.68 the representations of G correspond to the representations of $\mathfrak{g} = \mathfrak{su}(2)$, and by Proposition 13.67 the isomorphism classes of irreducible representations of* $\mathrm{SU}(2)$ *correspond to the isomorphism classes irreducible representations of $\mathfrak{su}(2)$. In the remaining lectures we will determine the irreducible representations of $\mathfrak{su}(2)$.*

14 Lie algebras

In the last few lectures we will say a few words about the representation theory of Lie algebras, mostly focusing on $\mathfrak{sl}_2(\mathbb{C})$. We will first show why it suffices to consider representations of *complex* Lie algebras.

14.1 Complex Lie algebras

If G is a Lie group, its Lie algebra is a Lie algebra over $\mathbb{R}$. However, we can also consider Lie algebras over $\mathbb{C}$: these are called complex Lie algebras.

Example 14.1. *The following are all Lie algebras over $\mathbb{C}$:*

$$\mathfrak{gl}_n(\mathbb{C}) = \mathrm{Mat}_n(\mathbb{C}) \tag{14.1.1}$$

$$\mathfrak{sl}_n(\mathbb{C}) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid \mathrm{tr}(A) = 0\} \tag{14.1.2}$$

$$\mathfrak{so}_n(\mathbb{C}) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid A + A^t = 0\}. \tag{14.1.3}$$

Example 14.2. *Even though $\mathfrak{su}(n) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid A + \bar{A}^t = 0\}$ is defined as a subset of $\mathfrak{gl}_n(\mathbb{C})$, it is not a Lie algebra over $\mathbb{C}$, since $\mathfrak{su}(n)$ is not necessarily preserved by multiplication by an element of $\mathbb{C}$.*

Definition 14.3. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$. A representation of $\mathfrak{g}$ is a homomorphism of $\mathbb{C}$-Lie algebras $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$, where V is a finite-dimensional $\mathbb{C}$-vector space. In other words, a representation is a $\mathbb{C}$-linear map $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ satisfying $\rho([x,y])(v) = \rho(x)(\rho(y)(v)) - \rho(y)(\rho(x)(v))$ for all $x, y \in \mathfrak{g}$ and $v \in V$.*

Example 14.4. *For any complex Lie algebra $\mathfrak{g}$, $V = \mathbb{C}$ with $X \cdot v = 0$ is called the trivial representation.*

Example 14.5. *For any finite-dimensional complex Lie algebra $\mathfrak{g}$, $V = \mathfrak{g}$ with $X \cdot v = [X, v]$ is called the adjoint representation. The fact that this is a representation is equivalent to the Jacobi identity.*

14.2 Complexification

If we start with a real Lie algebra, we can make a complex Lie algebra in a natural way, called its complexification. To describe this process, we first define it for vector spaces.

Definition 14.6. *Let V be an $\mathbb{R}$-vector space. The complexification of V, denoted by $V_{\mathbb{C}}$, is set of formal linear combinations of the form $v_1 + iv_2$ with $v_1, v_2 \in V$.*

More formally, we can write an element $v_1 + iv_2$ as a pair $(v_1, v_2) \in V \times V$, and $V_{\mathbb{C}}$ is the set of all such pairs. The set $V_{\mathbb{C}}$ is in fact a $\mathbb{C}$-vector space. Indeed, it is a $\mathbb{R}$-vector space by componentwise addition and multiplication $\lambda \cdot (v_1 + iv_2) = \lambda v_1 + i(\lambda v_2)$. The assignment $i \cdot (v_1 + iv_2) = -v_2 + iv_1$ defines the action of $\mathbb{C}$ on $V_{\mathbb{C}}$. As an $\mathbb{R}$-vector space, $V_{\mathbb{C}}$ is simply $V \oplus V$, but the point is that we have defined a $\mathbb{C}$-vector space structure on $V_{\mathbb{C}}$.

There is a very concrete way of thinking about complexification. Let $\{v_1, \ldots, v_n\}$ be an $\mathbb{R}$-basis of V. Then $V_{\mathbb{C}}$ has $\mathbb{C}$-basis $\{v_1, \ldots, v_n\}$, i.e. every element of $V_{\mathbb{C}}$ is a unique $\mathbb{C}$-linear combination of $v_1, \ldots, v_n$. Therefore $V_{\mathbb{C}}$ is just the space with the same coordinates as V but we also 'allow' $\mathbb{C}$-coefficients instead of just $\mathbb{R}$-coefficients.

Example 14.7. *The complexification of $\mathbb{R}^n$ is $\mathbb{C}^n$.*

Remark 14.8. *Every $\mathbb{C}$-vector space W is also an $\mathbb{R}$-vector space by restricting the $\mathbb{C}$-action to $\mathbb{R}$. So this gives a way of going from $\mathbb{C}$-vector spaces to $\mathbb{R}$-vector spaces. The complexification goes the other way around: it sends an $\mathbb{R}$-vector space to a $\mathbb{C}$-vector space.*

Remark 14.9. *More abstractly, $V_{\mathbb{C}}$ can also be defined as the tensor product $\mathbb{C} \otimes_{\mathbb{R}} V$.*

The next lemma is the universal property of complexifcation.

Lemma 14.10. *Let V be an $\mathbb{R}$-vector space and W a $\mathbb{C}$-vector space. Then the assignment*

$$\mathrm{Hom}_{\mathbb{R}-linear}(V, W) \to \mathrm{Hom}_{\mathbb{C}-linear}(V_{\mathbb{C}}, W) \tag{14.2.1}$$

$$(f\colon V \to W) \mapsto (f_{\mathbb{C}}\colon V_{\mathbb{C}} \to W, v_1 + iv_2 \mapsto f(v_1) + if(v_2)) \tag{14.2.2}$$

is a bijection.

Proof. Let us first check that $f_{\mathbb{C}}$ is $\mathbb{C}$-linear. To do this, it suffices to check for all $v = v_1 + iv_2 \in V_{\mathbb{C}}$ that $f_{\mathbb{C}}(\lambda v) = \lambda f_{\mathbb{C}}(v)$ for all $\lambda \in \mathbb{R}$ and that $f_{\mathbb{C}}(iv) = if_{\mathbb{C}}(v)$. We check that

$$f_{\mathbb{C}}(\lambda v) = f_{\mathbb{C}}(\lambda v_1 + i(\lambda v_2)) = f(\lambda v_1) + if(\lambda v_2) = \lambda f(v_1) + i\lambda f(v_2) = \lambda f_{\mathbb{C}}(v). \tag{14.2.3}$$

using the fact that f is $\mathbb{R}$-linear, and

$$f_{\mathbb{C}}(iv) = f_{\mathbb{C}}(-v_2 + iv_1) = -f(v_2) + if(v_1) = i \cdot (f(v_1) + if(v_2)) = if_{\mathbb{C}}(v). \tag{14.2.4}$$

We conclude that $f_{\mathbb{C}}$ is $\mathbb{C}$-linear and that the map (14.2.10) is well defined. To prove that it is bijective, we construct an inverse. Given a $\mathbb{C}$-linear map $g\colon V_{\mathbb{C}} \to W$, let $g_0\colon V \to W$ be the map $g_0(v) = g(v + i \cdot 0)$ (so just restrict g to $V \subset V_{\mathbb{C}}$). We leave the verification that this is an inverse to $f \mapsto f_{\mathbb{C}}$ as an exercise. □

Remark 14.11. *After choosing bases, this proposition becomes very concrete: $\mathbb{R}$-linear maps $\mathbb{R}^n \to \mathbb{C}^m$ correspond to $\mathbb{C}$-linear maps $\mathbb{C}^n \to \mathbb{C}^m$. Indeed, both are given by choosing n elements in $\mathbb{C}^m$ (the images of the basis vectors).*

We can also complexify Lie algebras:

Lemma 14.12. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$. Then the association $[x_1 + ix_2, y_1 + iy_2] = ([x_1, y_1] - [x_2, y_2]) + i([x_1, y_2] + [x_2, y_1])$ defines a Lie bracket on $\mathfrak{g}_{\mathbb{C}}$ and endows $\mathfrak{g}_{\mathbb{C}}$ with the structure of a Lie algebra over $\mathbb{C}$.*

Proof. Exercise! □

It follows that the complexification of a Lie algebra is again a Lie algebra!

Example 14.13. *Here are some examples of complexifications:*

$$\mathfrak{gl}_n(\mathbb{R})_{\mathbb{C}} \simeq \mathfrak{gl}_n(\mathbb{C}) \tag{14.2.5}$$
$$\mathfrak{sl}_n(\mathbb{R})_{\mathbb{C}} \simeq \mathfrak{sl}_n(\mathbb{C}) \tag{14.2.6}$$
$$\mathfrak{so}(p,q)_{\mathbb{C}} \simeq \mathfrak{so}_{p+q}(\mathbb{C}). \tag{14.2.7}$$

In particular, the complexifications of $\mathfrak{so}(p,q)$ only depends on $p+q$.

Example 14.14. *You will show on the example sheet that $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{R})_{\mathbb{C}}$ (they are both isomorphic to $\mathfrak{sl}_2(\mathbb{C})$).*

Proposition 14.15. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$ and $\mathfrak{h}$ a Lie algebra over $\mathbb{C}$. Then the assignment*

$$\mathrm{Hom}_{\mathbb{R}-\mathit{LieAlg}}(\mathfrak{g},\mathfrak{h}) \to \mathrm{Hom}_{\mathbb{C}-\mathit{LieAlg}}(\mathfrak{g}_{\mathbb{C}},\mathfrak{h}) \tag{14.2.8}$$
$$(f\colon\ \mathfrak{g}\to\mathfrak{h}) \mapsto (f_{\mathbb{C}}\colon\ \mathfrak{g}_{\mathbb{C}}\to\mathfrak{h}, v_1+iv_2 \mapsto f(v_1)+if(v_2)) \tag{14.2.9}$$

is a bijection.

Proof. By Lemma 14.10, it suffices to show that f preserves the Lie bracket on $\mathfrak{g}$ if and only if $f_{\mathbb{C}}$ preserves the Lie bracket on $\mathfrak{g}_{\mathbb{C}}$. This is a simple exercise. □

Lecture 21 starts here

Applying the above proposition to $\mathfrak{gl}_n(\mathbb{C})$, we get:

Corollary 14.16. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$. There is a bijection*

$$\mathrm{Hom}_{\mathbb{R}-\mathit{LieAlg}}(\mathfrak{g},\mathfrak{gl}_n(\mathbb{C})) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{C}-\mathit{LieAlg}}(\mathfrak{g}_{\mathbb{C}},\mathfrak{gl}_n(\mathbb{C})). \tag{14.2.10}$$

In other words, representations of $\mathfrak{g}$ correspond bijectively to representations of $\mathfrak{g}_{\mathbb{C}}$.

Example 14.17. *Let $G = \mathrm{SU}(2)$. Proposition 13.68 shows that representations of G correspond bijectively to representations of $\mathfrak{g} = \mathfrak{su}(2)$. Corollary 14.16 shows that representations of $\mathfrak{g}$ correspond bijectively to representations of the complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$. Since $\mathfrak{g}_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$, it suffices to study the representations of the $\mathbb{C}$-Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. This is what we will do in the remaining lectures.*

14.3 Represenations of $\mathfrak{sl}_2(\mathbb{C})$

After all our efforts in relating Lie groups to Lie algebras, we have reduced the study of representations of Lie groups (more or less) to the representation theory of their complexified Lie algebras. At this point, it is thus very natural to study complex Lie algebras and their representation theory in detail. This is a beautiful topic, quite elementary because it is essentially linear algebra, but would take a whole other course to understand. Since we only have a few lectures left, we will content ourselves to studying a single example in detail: $\mathfrak{sl}_2(\mathbb{C})$. So in this section, let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$.

Recall that $\mathfrak{g}$ consists of 2×2-matrices over $\mathbb{C}$ of trace zero, with Lie bracket $[X,Y] = XY - YX$. To study it, we will use the following basis:

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{14.3.1}$$

We can explicitly compute that

$$[H,X] = 2X, \quad [H,Y] = -2Y \quad [X,Y] = H. \tag{14.3.2}$$

These will be the only matrix calculations we will do.

A representation of $\mathfrak{g}$ is a finite-dimensional vector space $V/\mathbb{C}$ together with a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{gl}(V)$. By considering the images of the generators X, H, Y and after choosing a basis of V, a representation is nothing else but a triple of matrices (A_x, A_h, A_y) satifying $A_h A_x - A_x A_h = 2A_x$, $A_h A_y - A_y A_h = -2A_y$ and $A_x A_y - A_y A_x = A_h$. This is a very concrete linear algebraic problem, and we will solve it completely in what follows.

Here are some examples of representations of $\mathfrak{g}$:

- Trivial representation: $V = \mathbb{C}$ with $T \cdot v = 0$ for all $T \in \mathfrak{g}$.
- Defining representation: $V = \mathbb{C}^2$ and $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C}) \hookrightarrow \mathfrak{gl}_2(\mathbb{C})$ the natural inclusion. Picking the standard basis e_1, e_2 of V, we have for example $H \cdot e_1 = e_1$, $H \cdot e_2 = -e_2$, $X \cdot e_1 = 0$ etcetera.
- Adjoint representation: $V = \mathfrak{g} \simeq \mathbb{C}^3$ and $\mathfrak{g} \to \mathfrak{gl}(V), T \cdot v = [T, v]$. This is a representation because of the Jacobi identity. In the basis $\{X, H, Y\}$, we may calculate that the actions of X, H, Y are given by

$$\begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}. \tag{14.3.3}$$

Note that in these examples, the image of H in $\mathfrak{gl}(V)$ is always diagonalizable. This will be true in general, and the element H will play a crucial role in our classification of irreducible representations of $\mathfrak{g}$.

14.4 Irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$

Let V be a (finite-dimensional) irreducible representation of $\mathfrak{g}$. Let's try to say as much as possible about V. For each $\lambda \in \mathbb{C}$, let

$$V_\lambda = \{v \in V \mid Hv = \lambda v\} \tag{14.4.1}$$

be the λ-eigenspace of V. We know that $V_\lambda \neq 0$ for some $\lambda \in \mathbb{C}$, but we don't know yet that $V = \oplus V_\lambda$ since the H-action on V might not be diagonalizable. How do the elements X, Y interact with these eigenspaces?

Lemma 14.18. *For every $\lambda \in \mathbb{C}$, $X \cdot V_\lambda \subset V_{\lambda+2}$ and $Y \cdot V_\lambda \subset V_{\lambda-2}$.*

Proof. Let $v \in V_\lambda$, so $Hv = \lambda v$. Since $HX - XH = [H, X] = 2X$ and $\mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism, we compute that

$$H(Xv) = (XH + [H, X])(v) = X(Hv) + 2Xv \tag{14.4.2}$$
$$= X(\lambda v) + 2Xv \tag{14.4.3}$$
$$= (\lambda + 2)(Xv). \tag{14.4.4}$$

Therefore $Xv \in V_{\lambda+2}$. The calculation for Y is similar, using that $[H, Y] = -2Y$. □

Lemma 14.18 implies that $W = \oplus_{\lambda \in \mathbb{C}} V_\lambda$ is a subspace of V that is stable under X, H, Y so is $\mathfrak{g}$-stable. It is also nonzero, since H has at least one eigenvalue. Since V is irreducible, this implies that $V = W$, so H is diagonalizable.

The following picture summarizes the situation:

$$\begin{tikzcd} \cdots \arrow[r, bend left, "X\cdot"] & V_{\lambda-2} \arrow[l, bend left, "Y\cdot"] \arrow[r, bend left, "X\cdot"] & V_{\lambda} \arrow[l, bend left, "Y\cdot"] \arrow[r, bend left, "X\cdot"] & V_{\lambda+2} \arrow[l, bend left, "Y\cdot"] \arrow[r, bend left, "X\cdot"] & \cdots \arrow[l, bend left, "Y\cdot"] \end{tikzcd}$$

Let $S = \{\lambda \in \mathbb{C} \mid V_\lambda \neq 0\}$. Lemma 14.18 implies that if $\mu \in S$, then $\oplus_{\lambda \in \mu + 2\mathbb{Z}} V_\lambda$ is $\mathfrak{g}$-stable. It follows that this subspace must equal V by irreducibility, so every two elements in S differ by an even integer.

How do we proceed further? By the finite dimensionality of V, in the above picture you can't keep hopping to the right. In other words, there exists a $\lambda \in S$ such that $V_\lambda \neq 0$ and $V_{\lambda+2} = 0$. Let $v \in V_\lambda$ be a nonzero element. This is called a highest weight vector.

Lemma 14.19. *The subspace $W = span\{v, Yv, Y^2v, \dots\}$ is $\mathfrak{g}$-stable.*

Proof. It suffices to check that W is stable under X, H and Y. We have $Y \cdot W \subset W$ by construction. By Lemma 14.18, $H \cdot (Y^n v) = (\lambda - 2n)(Y^n v)$ so $H \cdot W \subset W$. It suffices to check that W is stable under X. Let's check that $X \cdot (Y^n v) \in W$ for small values of n. If $n = 0$, then $X \cdot v = 0$ since $X \cdot v \in V_{\lambda+2} = 0$. If $n = 1$, then

$$\begin{aligned} X(Yv) &= (YX + [X, Y])v && (14.4.5)\\ &= Y(Xv) + Hv && (14.4.6)\\ &= \lambda v. && (14.4.7) \end{aligned}$$

If $n = 2$, we compute

$$\begin{aligned} X(Y^2 v) &= (XY)(Yv) && (14.4.8)\\ &= (YX + [X, Y])(Yv) && (14.4.9)\\ &= Y(XYv) + HYv. && (14.4.10) \end{aligned}$$

Since $XYv = \lambda v$ and $Yv \in V_{\lambda-2}$, we get $X(Y^2 v) = (2\lambda - 2)Yv$. We can see a pattern here, and prove by induction (exercise) that for all $n \geq 0$ we have

$$X \cdot (Y^n v) = n(\lambda - n + 1)(Y^{n-1} v). \qquad (14.4.11)$$

This proves that $X \cdot W \subset W$, concluding the lemma. ☐

Corollary 14.20. *All the H-eigenspaces are one-dimensional: if $\lambda \in S$ then $\dim V_\lambda = 1$.*

Proof. Let $v \in V_\lambda$ be a highest weight vector. Then $W = \mathrm{span}\{v, Yv, Y^2 v, \dots\}$ is $\mathfrak{g}$-stable so by irreducibility $V = W$. It follows that $S = \{\lambda, \lambda - 2, \dots\}$ and $V_{\lambda - 2n}$ is spanned by $Y^n v$ (which might be zero). ☐

Corollary 14.21. *There exists a nonnegative integer $n \geq 0$ such that $S = \{-n, -n+2, \dots, n-2, n\}$ and V has dimension $n + 1$.*

Proof. Let $v \in V_\lambda$ be a highest weight vector. Let $n \geq 0$ be the largest nonnegative integer such that $Y^n v \neq 0$. Then $W = \mathrm{span}\{v, Yv, Y^2 v, \dots, Y^n v\}$. Formula (14.4.11) applied to $Y^{n+1} v$ shows that

$$X(Y^{n+1} v) = (n + 1)(\lambda - n)Y^n v. \qquad (14.4.12)$$

Since $Y^{n+1} v = 0$ and $Y^n v \neq 0$, $(n + 1)(\lambda - n) = 0$. Since $n + 1 \geq 1$, this implies that $\lambda = n$. ☐

Theorem 14.22. *Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ and $n \in \mathbb{Z}_{\geq 0}$. Then up to isomorphism $\mathfrak{g}$ has a unique representation of dimension $n + 1$, denoted by $V(n)$.*

Proof. We first show uniqueness. Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be an $(n+1)$-dimensional irreducible representation of $\mathfrak{g}$. Let $v \in V$ be a highest weight vector. The above analysis shows that $Hv = nv$ and $V = \operatorname{span}\{v, Yv, \ldots, Y^n v\}$. The proof of Lemma 14.19 explicitly describes the action of $\mathfrak{g}$ on V. Indeed, in the basis $\{v, Yv, \ldots, Y^n v\}$ this action is given by

$$\rho(H) = \begin{pmatrix} n & & & \\ & n-2 & & \\ & & \ddots & \\ & & & -n \end{pmatrix} \tag{14.4.13}$$

$$\rho(X) = \begin{pmatrix} 0 & n & & & \\ & 0 & 2n-2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & n \\ & & & & 0 \end{pmatrix} \tag{14.4.14}$$

$$\rho(Y) = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \tag{14.4.15}$$

The matrix $\rho(X)$ has coefficient $i(n-i+1)$ at entry $(i, i+1)$. This shows that every irreducible representation of $\mathfrak{g}$ of dimension $n+1$ must be isomorphic to one given by these matrices. To prove existence, we can merely check that the above big matrices indeed define a Lie algebra homomorphism $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}_{n+1}(\mathbb{C})$, by checking that the commutation relations $[\rho(H), \rho(X)] = 2\rho(X)$, $[\rho(H), \rho(Y)] = -2\rho(Y)$ and $[\rho(X), \rho(Y)] = \rho(H)$. We still need to verify that this representation is irreducible. But if $W \subset \mathbb{C}^{n+1}$ is a $\mathfrak{g}$-stable subspace, acting by X repeatedly shows that W contains a highest weight vector. Applying Y repeatedly to this highest weight vector shows that W generates V, as in Lemma 14.19. □

Remark 14.23. *Another way to prove existence is to explicitly exhibit $V(n)$: it is given by $\operatorname{Sym}^n V$, where $V = \mathbb{C}^2$ is the defining representation of $\mathfrak{sl}_2(\mathbb{C})$. (Problem set 7)*

The H-eigenspaces on $V(n)$ are $-n, -n+2, \ldots, n-2, n$ and n is the largest eigenvalue of H. That's why we denote this representation by $V(n)$ and call it the representation of highest weight n.

So we have completely classified all irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$: there is exactly one of them in each dimension. But what if we start with a representation that is not irreducible? Is every representation a direct sum of irreducibles? In the finite group case, this was a consequence of Maschke's theorem. For a general Lie algebra, this need not be true. For our specific example of $\mathfrak{sl}_2(\mathbb{C})$, we're in luck:

Proposition 14.24 (Complete reducibility)**.** *Every (finite-dimensional) representation of $\mathfrak{sl}_2(\mathbb{C})$ is a direct sum of irreducible representations.*

There's a purely algebraic proof of this fact, using the so-called 'Casimir element'. We will take a different route, and use the connection with the representation theory of the compact group $\mathrm{SU}(2)$. We therefore defer the proof of this proposition after the discussion of $\mathrm{SU}(2)$. Lecture 22 starts here

14.5 Connection to representations of $\mathrm{SU}(2)$

Recall that $\mathrm{SU}(2)$ is a connected, simply connected Lie group with Lie algebra $\mathfrak{su}(2)$. Recall that $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$. Combining the homomorphism theorem 13.50 with Corollary 14.16, we obtain bijections

$$\mathrm{Hom}(\mathrm{SU}(2), \mathrm{GL}_n(\mathbb{C})) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{R}-\text{LieAlg}}(\mathfrak{su}(2), \mathfrak{gl}_n(\mathbb{C})) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{C}-\text{LieAlg}}(\mathfrak{su}(2)_{\mathbb{C}}, \mathfrak{gl}_n(\mathbb{C})). \tag{14.5.1}$$

By Lemma 13.67, these bijections preserve the notion of subrepresentation, irreducibility and isomorphism. (Strictly speaking, that lemma only talks about the first bijection, but its analogue for the second bijection can be easily seen to also hold.) It follows that the representation theory of $\mathrm{SU}(2)$ is equivalent to that of $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$. Therefore, as a consequence of Theorem 14.22 we have:

Theorem 14.25. *For every $n \in \mathbb{Z}_{\geq 0}$, up to isomorphism there exists a unique irreducible representation of $\mathrm{SU}(2)$ of dimension $n+1$.*

How can we describe this irreducible representation more explicitly? Let $V = \mathbb{C}^2$ be the defining representation of $\mathrm{SU}(2)$, corresponding to the inclusion $\mathrm{SU}(2) \hookrightarrow \mathrm{GL}_2(\mathbb{C})$. Then $\mathrm{Sym}^n V \subset V^{\otimes n}$ is a representation of $\mathrm{SU}(2)$, defined in the same way as for finite groups in §7.5. Under the bijections (14.5.1), this representation induces the representation $\mathrm{Sym}^n V$ of $\mathfrak{sl}_2(\mathbb{C})$, which on Problem Set 7 you have shown to be the irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$ of dimension $n+1$. By Lemma 13.67, it follows that $\mathrm{Sym}^n V$ is irreducible as a $\mathrm{SU}(2)$-representation.

In more detail, let $\mathcal{P}_n$ be the set of homogeneous degree n polynomials in $\mathbb{C}[x,y]$:

$$\mathcal{P}_n = \{\sum_{i=0}^{n} a_i x^{n-i} y^i \mid a_i \in \mathbb{C}\}. \tag{14.5.2}$$

The vector space $\mathcal{P}_n$ receives an action of $\mathrm{SU}(2)$ via

$$(g \cdot P)(x,y) = P((x,y) \cdot g) \tag{14.5.3}$$

for all $g \in \mathrm{SU}(2)$ and $P \in \mathcal{P}_n$. Here we interpret $(x,y) \cdot A$ as the multiplication of a 1×2-matrix and 2×2-matrix. Then $\mathcal{P}_n = \mathrm{Sym}^n V$ is the unique irreducible representation of $\mathrm{SU}(2)$ of dimension $n+1$.

It is instructive to take the derivative of this representation and see what we get on the Lie algebra side. If $H \in \mathfrak{sl}_2(\mathbb{C})$ is our diagonal basis element, then $iH \in \mathfrak{su}(2)$ and $e^{t(iH)} = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$. We can then compute that

$$\begin{aligned}
(iH) \cdot (x^{n-k} y^k) &= \frac{d}{dt}\Big|_{t=0} \left(\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \cdot (x^{n-k} y^k) \right) && (14.5.4)\\
&= \frac{d}{dt}\Big|_{t=0} (e^{(n-2k)it} x^{n-k} y^k) && (14.5.5)\\
&= (n-2k) i x^{n-k} y^k. && (14.5.6)
\end{aligned}$$

Therefore $x^{n-k}y^k$ is an eigenvector for H with eigenvalue $n-2k$. This is in agreement with the fact that $\mathcal{P}_n$ should have H-eigenvalues $n, n-2, \ldots, -n$.

14.6 Complete reducibility of $\mathfrak{sl}_2(\mathbb{C})$-representations

Using the connection to $\mathrm{SU}(2)$, we now prove the complete reducibility of representations of $\mathfrak{sl}_2(\mathbb{C})$.

Proof of Proposition 14.24. Let V be a representation of $\mathfrak{sl}_2(\mathbb{C})$. Under the bijections of (14.5.1), V can also be seen as a representation of $\mathrm{SU}(2)$. But $\mathrm{SU}(2)$ is a compact group! In particular, by Weyl's unitary trick (averaging a Hermitian inner product), we have shown that every (finite-dimensional) representation of $\mathrm{SU}(2)$ is a direct sum of irreducible representations (Corollary 11.10). Therefore we may write $V = V_1 \oplus \cdots \oplus V_k$, where each V_i is an irreducible $\mathrm{SU}(2)$-representation. (We may have $V_i \simeq V_j$ for different i, j.) We claim that this is also a decomposition of V into a direct sum of irreducible $\mathfrak{sl}_2(\mathbb{C})$-representations. This essentially follows from Lemma 13.67. In more detail, write $G = \mathrm{SU}(2)$. Since each V_i is G-stable we know that each V_i is also $\mathfrak{g}$-stable, hence also $\mathfrak{g}_{\mathbb{C}}$-stable since W is a $\mathbb{C}$-subspace. To show that each V_i is irreducible, we go in the other way: suppose that $W \subset V_i$ is a $\mathfrak{g}_{\mathbb{C}}$-stable subspace. Then W is clearly also $\mathfrak{g}$-stable. By Lemma 13.67, it is also G-stable. But since V_i is irreducible, this implies that $W = 0$ or V_i. □

This is an example how the representation theory of Lie algebras and Lie groups inform each other: usually we use Lie algebras to say interesting things about Lie groups, but here we have used Lie groups to say something interesting about Lie algebras!

Remark 14.26. *Let $G = \mathrm{SL}_2(\mathbb{R})$ and let V be a G-representation. Then we claim that V is a direct sum of irreducible G-representations. Indeed, we know that $\mathfrak{g}_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$ and that every $\mathfrak{g}_{\mathbb{C}}$-representation is completely reducible. So when we view V as a $\mathfrak{g}_{\mathbb{C}}$-representation, we can decompose V into a direct sum of irreducible $\mathfrak{g}_{\mathbb{C}}$-representations $W_1 \oplus \cdots \oplus W_k$. By an argument very similar to the above proof using Lemma 13.67, this is also a decomposition of V into G-irreducibles. Therefore, even though $\mathrm{SL}_2(\mathbb{R})$ is not compact, every (finite-dimensional) representation of G is still completely reducible.*

14.7 Representations of related Lie groups

We will use our knowledge of the representation theory of $\mathrm{SU}(2)$ to classify irreducible representations for other related groups: $\mathrm{SO}(3), \mathrm{U}(2)$ and $\mathrm{SO}(4)$.

14.7.1 Representations of $\mathrm{SO}(3)$

Recall that there is a double cover $\pi\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$, which is surjective with kernel $\{\pm I\}$. Therefore the irreducible representations of $\mathrm{SO}(3)$ correspond to the irreducible representations of $\mathrm{SU}(2)$ on which $-I$ acts trivially. Looking at the representation $\mathcal{P}_n$ of $\mathrm{SU}(2)$, we see that

$$(-I) \cdot P(x, y) = P(-x, -y) = (-1)^n P(x, y) \tag{14.7.1}$$

since P is homogeneous of degree n. It follows that $-I$ acts trivially on $\mathcal{P}_n$ if and only if n is even! We obtain:

Theorem 14.27. *Every irreducible representation of $\mathrm{SO}(3)$ has odd dimension, and moreover for every odd $k \in \mathbb{Z}_{\geq 1}$ there exists a unique irreducible representation of $\mathrm{SO}(3)$ of dimension k.*

Remark 14.28. *The representation $V(n)$ of $\mathrm{SU}(2)$ is sometimes called the representation with spin $n/2$. Representations with n even are said to have integer spin. In this language, we can say that representation of $\mathrm{SO}(3)$ correpond to representations of $\mathrm{SU}(2)$ with integer spin.*

Can we give a more direct description of the irreducible representations of $\mathrm{SO}(3)$? It turns out we can! The group $G = \mathrm{SO}(3)$ acts on $\mathbb{R}^3$ in a norm preserving way, so G acts on the 2-sphere S^2. It follows that G acts on functions on the 2-sphere:

$$L^2(S^2) = \{L^2 - \text{functions } f\colon S^2 \to \mathbb{C}\}, \tag{14.7.2}$$

$$(g \cdot f)(x) = f(g^{-1}x). \tag{14.7.3}$$

It turns out that in this way $L^2(S^2)$ becomes a unitary Hilbert space representation of G, albeit infinite-dimensional. What can we say about this representation? For example, the G-fixed points are exactly the functions on S^2 that have the same value at every point: in other words, the constant functions. It turns out that

$$L^2(S^2) = \bigoplus_{\ell \geq 0} W_\ell \tag{14.7.4}$$

where W_ℓ is the unique irreducible representation of dimension $2\ell + 1$. It turns out that a basis for W_ℓ is given by the spherical harmonics of degree ℓ!

Remark 14.29. *You might be wondering how to study representations of disconnected groups like $O(3)$. In general, if G is a (possibly disconnected) Lie group with identity component G°, then there is a way to define the induction of a G°-representation to a G-representation (similarly to the finite group case) and show that every irreducible representation of G is a subrepresentation of the induction of some irreducible representation of G°. Therefore the representation theory of G is some kind of 'mix' between that of the connected Lie group G° and the finite group $\pi_0(G) = G/G^\circ$.*

14.7.2 Representations of $\mathrm{U}(2)$

Lemma 14.30. *The map $(\lambda, A) \mapsto \lambda A$ induces an isomorphism of Lie groups $(S^1 \times \mathrm{SU}(2))/\langle(-1, -I)\rangle \simeq \mathrm{U}(2)$.*

Proof. The induced map $S^1 \times \mathrm{SU}(2) \to \mathrm{U}(2)$ is surjective with kernel $\langle(-1, -I)\rangle$. □

The following proposition is the analogue of Theorem 7.25.

Proposition 14.31. *Let G_1, G_2 be compact Lie groups. Then every irreducible representation of $G_1 \times G_2$ is of the form $V_1 \boxtimes V_2$ for V_i is an irreducible representation of G_i, uniquely determined up to isomorphism.*

Proof. This follows from considering characters of representations and the Peter–Weyl theorem; we omit the details. □

It follows that if $G = (G_1 \times G_2)/N$, then the irreducible representations of G are of the form $\rho = \rho_1 \boxtimes \rho_2$, where ρ_i is an irrep of G_i, that additionally satisfy $\rho(n) = \mathrm{Id}$ for all $n \in N$. In the case $\mathrm{U}(2) = (S^1 \times \mathrm{SU}(2))/\langle(-1, I)\rangle$, this gives the following classification. Let $V = \mathbb{C}^2$ be the defining representation of $\mathrm{U}(2)$ and $\rho_n = \mathrm{Sym}^n V$ its n-th symmetric power, with action exactly as described in §14.5.

Theorem 14.32. *Every irreducible representation of $\mathrm{U}(2)$ is of the form $(\det)^{\otimes k} \otimes \mathrm{Sym}^n V$ for some unique integers $k, n \in \mathbb{Z}$ with $n \geq 0$.*

14.7.3 Representations of $\mathrm{SO}(4)$

Lemma 14.33. *There is an isomorphism of Lie groups $\mathrm{SO}(4) \simeq (\mathrm{SU}(2) \times \mathrm{SU}(2))/\langle(-I, -I)\rangle$.*

Proof. Recall that $\mathrm{SU}(2)$ can be identified with norm 1 quaternions $\mathbb{H}$. Then the map $\mathbb{H} \times \mathbb{H} \to \mathrm{GL}(\mathbb{H})$, $(p, q) \mapsto (x \mapsto pxq^{-1})$ restricts to a map $\pi\colon\ \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. The kernel of this map is the set of $(p, q) \in \mathbb{H}_{Nm=1}$ such that $pxq^{-1} = x$ for all $x \in \mathbb{H}$. Setting $x = 1$ implies that $p = q$. Then $pxp^{-1} = x$ for all $x \in \mathbb{H}$ implies that p lies in the center of $\mathbb{H}$, which is $\mathbb{R} \cdot 1$. Since $Nm(p) = 1$, $p = \pm 1$. Comparing dimensions, we see that π is a double cover. □

Let ρ_n be the irreducible representation of $\mathrm{SU}(2)$ of dimension $n+1$.

Theorem 14.34. *Every irreducible representation of $\mathrm{SO}(4)$ is of the form $\rho_n \boxtimes \rho_m$ with $m, n \in \mathbb{Z}_{\geq 0}$ and $m \equiv n \bmod 2$.*

Proof. The representation $\rho_n \boxtimes \rho_m$ has $(-I, -I)$ in its kernel if and only if $m - n$ is even. $\square$

Lecture 23 starts here

15 Overview of further topics

In this last lecture (which is non-examinable) we will briefly mention where you could go from here if you wanted to know more about Lie groups and representation theory. No proofs are given.

15.1 Root systems

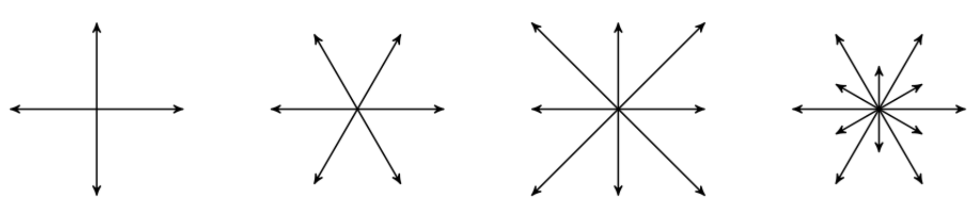
Let E be a finite dimensional $\mathbb{R}$-vector space with an inner product $(\cdot,\cdot)\colon E \times E \to \mathbb{R}$. For every nonzero $v \in E$, write $w_v(x) = x - \frac{2(x,v)}{(v,v)}v$ for the reflection through the hyperplane orthogonal to v.

Definition 15.1. *A root system in E is a finite subset $\Phi \subset E \setminus \{0\}$ with the property that:*

1. *the $\mathbb{R}$-span of Φ is E;*
2. *if $\alpha \in \Phi$ and $c \in \mathbb{R}$, then $c\alpha \in \Phi$ if and only if $c = \pm 1$;*
3. *$w_\alpha(\Phi) = \Phi$ for all $\alpha \in \Phi$;*
4. *for all $\alpha, \beta \in \Phi$, the quantity $2(\alpha,\beta)/(\alpha,\alpha)$ is an integer.*

Note that if θ is the angle between two roots $\alpha, \beta \in \Phi$, then the last condition implies $4\cos(\theta)^2 = \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} \in \mathbb{Z}$ so $\cos(\theta)^2 \in \frac{1}{4}\mathbb{Z}$. This means that θ must be $0, 30, 45, 60, 90, 120, 135, 150$ or 180 degrees.

The simplest root system is the A_1 root system: $\Phi = \{\pm 1\} \subset \mathbb{R}$. The root systems in $\mathbb{R}^2$ are (up to a suitable notion of isomorphism) given by the following four pictures:

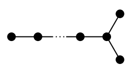
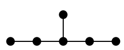
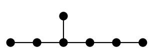
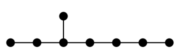


$\mathrm{A}_1 \times \mathrm{A}_1$ $\qquad$ A_2 $\qquad$ $\mathrm{B}_2 \cong \mathrm{C}_2$ $\qquad$ G_2

Root systems can be classified by certain graphs called 'Dynkin diagrams', as follows. Given a root system (E, Φ), let $v \in E$ be any nonzero element such that the hyperplane $H = \{x \in E \mid (x,v) = 0\}$ is disjoint from Φ. Using this choice, we may define the subsets of positive roots $\Phi^+ = \{\alpha \in \Phi \mid (v,\alpha) > 0\}$ and analogously negative roots Φ^-, giving rise to a decomposition $\Phi = \Phi^+ \sqcup \Phi^-$. One can prove that there exists a unique subset $\Delta \subset \Phi^+$ (called simple roots) such that every element of Φ^+ is expressible as an integer linear combination of simple roots with nonnegative coordinates. Define the graph $D(\Phi)$ as follows. The vertices of $D(\Phi)$ are indexed by the elements of Δ. Between vertices α and β we draw $4(\alpha,\beta)^2/(\alpha,\alpha)(\beta,\beta)$ (a positive integer by the fourth axiom of a root system) edges; if α and β do not have equal lengths we draw an arrow on the edges pointing towards the smaller vector among α and β. This procedure produces a graph which does not depend on the choice of hyperplane H, called the Dynkin diagram of Φ.

To state the classification, note that $(E \oplus E', \Phi \sqcup \Phi')$ is a root system when (E, Φ) and (E', Φ') are, and we say Φ is irreducible if it is not the sum of two nonzero root systems.

Theorem 15.2. *The map $\Phi \mapsto D(\Phi)$ induces a bijection between isomorphism classes of irreducible root systems and the following graphs, called Dynkin diagrams:*

- (A_n):
- (B_n):
- (C_n):
- (D_n): 
- (E_6): 
- (E_7): 
- (E_8): 
- (F_4):
- (G_2):

15.2 Classification of semisimple complex Lie algebras

Let $\mathfrak{g}$ be a (finite-dimensional) complex Lie algebra.

Definition 15.3. *An ideal of a Lie algebra $\mathfrak{g}$ is a subspace I with the property that $[x, I] \subset I$ for all $x \in \mathfrak{g}$. A Lie algebra is said to be simple if the Lie bracket is not identically zero and it has no nonzero proper ideal, and semisimple if it is isomorphic to a direct sum of simple ones.*

Example 15.4. *The Lie algebras $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{so}_n(\mathbb{C})$ are semisimple. They are simple, except $\mathfrak{sl}_1(\mathbb{C}), \mathfrak{so}_1(\mathbb{C}), \mathfrak{so}_2(\mathbb{C})$ and $\mathfrak{so}_4(\mathbb{C})$.*

We briefly explain how to classify semisimple Lie algebras in terms of root systems. Let $\mathfrak{g}$ be a semisimple Lie algebra. There exists a subalgebra $\mathfrak{t} \subset \mathfrak{g}$ with the property that ad_x is semisimple for each $x \in \mathfrak{t}$ and the dimension of $\mathfrak{t}$ is maximal with respect to this property; such a subalgebra is called a Cartan subalgebra or CSA. For example, if $\mathfrak{g} = \mathfrak{sl}_n$ then a choice for $\mathfrak{t}$ is the subset of diagonal matrices (of trace zero). Fix a choice of CSA $\mathfrak{t} \subset \mathfrak{g}$. It turns out that $[x, y] = 0$ for all $x, y \in \mathfrak{t}$, so the ad_x are mutually commuting semisimple linear maps for $x \in \mathfrak{t}$. It follows that they can be simultaneously diagonalised, i.e. there is a decomposition

$$\mathfrak{g} = \bigoplus_{f\colon\ \mathfrak{g} \to \mathbb{C}} \{x \in \mathfrak{g} \mid [t, x] = f(t)x \,\forall t \in \mathfrak{t}\}, \tag{15.2.1}$$

where the sum runs over all linear functionals of $\mathfrak{g}$. It turns out that the part corresponding to the zero functional is exactly $\mathfrak{t}$, i.e. the only elements of $\mathfrak{g}$ commuting with $\mathfrak{t}$ are the elements of $\mathfrak{t}$ itself. Writing $\Phi \subset \mathfrak{t}^* = \{f\colon \mathfrak{g} \to \mathbb{C}\}$ for the set of nonzero functionals for which the corresponding eigenspace $\mathfrak{g}_f$ is nonzero, we may write

$$\mathfrak{g} = \mathfrak{t} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha\,.$$

There exists a canonical perfect pairing on $\mathfrak{g}$, called the Killing form, which induces a bilinear form on $\mathfrak{t}^*$. If E denotes the $\mathbb{R}$-span of Φ inside $\mathfrak{t}^*$, then equipped with the restriction of this pairing it turns out that (E, Φ) is a root system!

Theorem 15.5 (Cartan–Killing)**.** *The association $\mathfrak{g} \mapsto (E, \Phi)$ is, up to isomorphism, independent of any choices and induces a bijection between simple complex Lie algebras and irreducible root systems. Consequently, every simple complex Lie algebra is isomorphic to one of the following:*

- $\mathfrak{sl}_n(\mathbb{C})$
- $\mathfrak{so}_n(\mathbb{C})$
- $\mathfrak{sp}_{2n}(\mathbb{C})$
- *Five exceptional Lie algebras: $\mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8, \mathfrak{f}_4$ or $\mathfrak{g}_2$.*

Example 15.6. *Let $\mathfrak{g} = \mathfrak{sl}_2$ and let $\mathfrak{t} \subset \mathfrak{g}$ be the subspace of diagonal matrices. Then $\Phi = \{\pm\alpha\}$, where $\alpha\colon \mathfrak{t} \to \mathbb{C}$ sends $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ to 2. The associated root system is A_1. More generally, $\mathfrak{sl}_{n+1}$ corresponds to the A_n root system.*

15.3 Classification of compact Lie groups

One of the reasons why we care about semisimple Lie algebras is the following proposition:

Proposition 15.7. *Let G be a connected compact Lie group with discrete center. Then $\mathfrak{g}_{\mathbb{C}}$ is a semisimple complex Lie algebra. Conversely, every semisimple complex Lie algebra is of this form for some unique connected simply connected compact Lie group.*

This proposition shows that the Cartan–Killing classification immediately implies a classification for connected compact Lie groups up to isogeny.

Theorem 15.8. *Let G be a connected compact Lie group with no nontrivial co Then G lies in the same isogeny class as one of the following Lie groups:*

- $\mathrm{SU}(n)$
- $\mathrm{SO}(n)$
- $\mathrm{USp}(2n) = \mathrm{U}(2n) \cap \mathrm{Sp}_{2n}(\mathbb{C})$
- *Five exceptional compact Lie groups.*

The classification of *all* (not necessarily compact) Lie groups is more complicated, since non-semisimple Lie algebras are hard to classify, and various real Lie algebras can give rise to the same complexified Lie algebra. See Fulton–Harris for an overview.