261A: Lie Groups

Nir Elber

Fall 2024

Contents

How strange to actually have to see the path of your journey in order to make it.

—Neal Shusterman, [Shu16]

Theme 1
Topological Background

Hold tight to your geometric motivation as you learn the formal structures which have proved to be so effective in studying fundamental questions.

—Ravi Vakil, [Vak17]

1.1 August 28

Today we review differential topology. Here are some logistical notes.

- There will be weekly homeworks, of about 5 problems.
- There will be a final take-home exam.
- This course has a bCourses page.
- We will mostly follow Kirillov's book [Kir08].

1.1.1 Group Objects

The goal of this class is to study symmetries of geometric objects. As such, we are interested in studying (infinite) groups with some extra geometric structure, such as a real manifold or a complex manifold or a scheme structure. Speaking generally, we will have some category $\mathcal{C}$ of geometric objects, equipped with finite products (such as a final object), which allows us to have group objects in $\mathcal{C}$.

Definition 1.1 (group object)**.** Fix a category $\mathcal{C}$ with finite products, such as a final object $*$. A *group object* is the data (G, m, e, i) where $G \in \mathcal{C}$ is an object and $m\colon G \times G \to G$ and $e\colon * \to G$ and $i\colon G \to G$ are morphisms. We require this data to satisfy some associativity, identity, and inverse coherence laws.

For concreteness, we go ahead and write out the coherence diagrams, but they are not so interesting.

- Associative: the following diagram commutes.

$$\begin{array}{ccc} G \times G \times G & \xrightarrow{\mathrm{id}_G \times m} & G \times G \\ {\scriptstyle m \times \mathrm{id}_G}\big\downarrow & & \big\downarrow{\scriptstyle m} \\ G \times G & \xrightarrow{m} & G \end{array}$$

- Identity: the following diagram commutes.

$$\begin{tikzcd}
G \arrow[r, "\mathrm{id}_G \times e"] \arrow[dr, equal] & G \times G \arrow[d, "m"] & G \arrow[l, "e \times \mathrm{id}_G"'] \arrow[dl, equal] \\
& G &
\end{tikzcd}$$

- Inverses: the following diagram commutes.

$$\begin{tikzcd}
& G \times G \arrow[dr, "m"] & \\
G \arrow[ur, "i \times \mathrm{id}_G"] \arrow[r] \arrow[dr, "\mathrm{id}_G \times i"'] & * \arrow[r, "e"] & G \\
& G \times G \arrow[ur, "m"'] &
\end{tikzcd}$$

Example 1.2. In the case where $\mathcal{C} = \mathrm{Set}$, we recover the notion of a group, where G is the set, m is the multiplication law, e is the identity, and i is the inverse.

Example 1.3. Group objects in the category of manifolds will be Lie groups.

1.1.2 Review of Topology

This course requires some topology as a prerequisite, but let's review these notions for concreteness. We refer to [Elb22] for most of these notions.

Definition 1.4 (topological space)**.** A *topological space* is a pair $(X, \mathcal{T})$ of a set X and collection $\mathcal{T} \subseteq \mathcal{P}(X)$ of open subsets of X, which we require to satisfy the following axioms.

- $\varnothing, X \in \mathcal{T}$.
- Finite intersection: for $U, V \in \mathcal{T}$, we have $U \cap V \in \mathcal{T}$.
- Arbitrary unions: for a subcollection $\mathcal{U} \subseteq \mathcal{T}$, we have $\bigcup_{U \in \mathcal{U}} U \in \mathcal{T}$.

We will suppress the notation $\mathcal{T}$ from our topological space as much as possible.

Example 1.5. The set $\mathbb{R}$ equipped with its usual (metric) topology is a topological space.

Example 1.6. Given a topological space X and a subset $Z \subseteq X$, we can make Z into a topological space with open subsets given by $U \cap Z$ whenever $U \subseteq Z$ is open.

Definition 1.7 (closed)**.** A subset Z of a topological space X is *closed* if and only if $X \setminus Z$ is open.

One way to describe topologies is via a base.

Definition 1.8 (base)**.** Given a topological space X, a *base* $\mathcal{B} \subseteq \mathcal{P}(X)$ for the topology such that any open subset $U \subseteq X$ is the union of a subcollection of $\mathcal{B}$. Equivalently, for any open subset $U \subseteq X$ and $x \in U$, there is $B \in \mathcal{B}$ such that $x \in B \subseteq U$.

Example 1.9. The collection of open intervals $(a, b) \subseteq \mathbb{R}$ generates the usual topology. In fact, one can even restrict ourselves to open intervals (a, b) where $a, b \in \mathbb{Q}$, so $\mathbb{R}$ has a countable base.

Our morphisms are continuous maps.

Definition 1.10 (continuous)**.** A function $f\colon X \to Y$ between topological spaces is *continuous* if and only if $f^{-1}(V) \subseteq X$ is open for each open subset $V \subseteq Y$.

Thus, we can define Top as the category of topological spaces equipped with continuous maps as its morphisms. Thinking categorically allows us to make the following definition.

Definition 1.11 (homeomorphism)**.** A *homeomorphism* is an isomorphism in Top. Namely, a function $f\colon X \to Y$ between topological spaces which is continuous and has a continuous inverse.

Remark 1.12. There are continuous bijections which are not homeomorphisms! For example, one can map $[0, 2\pi) \to S^1$ by sending $x \mapsto e^{ix}$, which is a continuous bijection, but the inverse is discontinuous at $1 \in S^1$.

Earlier, we wanted to have finite products in our category. Here is how we take products of pairs.

Definition 1.13 (product topology)**.** Given topological spaces X and Y, we define the topological space $X \times Y$ as having $X \times Y$ as its set and open subsets given by arbitrary unions of sets of the form $U \times V$ where $U \subseteq X$ and $V \subseteq Y$ are open.

Remark 1.14. Alternatively, we can say that the topology $X \times Y$ has a base given by the "rectangles" $U \times V$ where $U \subseteq X$ and $V \subseteq Y$ are open. In fact, if $\mathcal{B}_X$ and $\mathcal{B}_Y$ are bases for X and Y, respectively, then we can check that the open subsets

$$\{U \times V : U \in \mathcal{B}_X, V \in \mathcal{B}_Y\}$$

is a base for $X \times Y$.

Remark 1.15. The final object in Top is the singleton space.

Now, group objects in Top are called topological groups, which are interesting in their own right. For example, locally compact topological groups have a good Fourier analysis theory.

Example 1.16. The group $\mathbb{R}$ under addition is a topological group. In fact, $\mathbb{Q}$ under addition is also a topological group, though admittedly a more unpleasant one.

Example 1.17. The group $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ is a topological group.

1.1.3 Review of Differential Topology

However, in this course, we will be more interested in manifolds, so let's define these notions. We refer to [Elb24] for (a little) more detail, and we refer to [Lee13] for (much) more detail. To begin, we note arbitrary topological spaces are pretty rough to handle; here are some niceness requirements. The following is a smallness assumption.

Definition 1.18 (separable)**.** A topological space X is *separable* if and only if it has a countable base.

The following says that points can be separated.

Definition 1.19 (Hausdorff)**.** A topological space X is *Hausdorff* if and only if any pair of distinct points $p, q \in X$ have disjoint open neighborhoods.

The following is another smallness assumption, which we will use frequently but not always.

Definition 1.20 (compact)**.** A topological space X is *compact* if and only if any open cover $\mathcal{U}$ (i.e., each $U \in \mathcal{U}$ is open, and $X = \bigcup_{U \in \mathcal{U}} U$) has a finite subcollection which is still an open cover.

We are now ready for our definition.

Definition 1.21 (topological manifold)**.** A *topological manifold of dimension* n is a topological space X satisfying the following.

- X is Hausdorff.
- X is separable.
- Locally Euclidean: X has an open cover $\{U_\alpha\}_{\alpha \in \kappa}$ such that there are open subsets $V_\alpha \subseteq \mathbb{R}^n$ and homeomorphisms $\varphi_\alpha : U_\alpha \to V_\alpha$.

Remark 1.22. By passing to open balls, one can require that all the V_α are open balls. By doing a little more yoga with such open balls (noting $B(0,1) \cong \mathbb{R}^n$), one can require that $V_\alpha = \mathbb{R}^n$ always.

Remark 1.23. It turns out that open subsets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ can only be homeomorphic if and only if $n = m$. This implies that the dimension of a connected component of X is well-defined without saying what n is in advance. However, we should say what n is in advance in order to get rid of pathologies like $\mathbb{R} \sqcup \mathbb{R}^2$.

To continue, we must be careful about our choice of U_αs and φ_αs.

Definition 1.24 (chart, atlas, transition function)**.** Fix a topological manifold X of dimension n.

- A *chart* is a pair (U, φ) of an open subset $U \subseteq X$ and homeomorphism φ of U onto an open subset of $\mathbb{R}^n$.
- An *atlas* is a collection of charts $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in \kappa}$ such that $\{U_\alpha\}_{\alpha \in \kappa}$ is an open cover of X.
- The *transition function* between two charts (U, φ) and (V, ψ) is the composite homeomorphism
$$\varphi(U \cap V) \xleftarrow{\varphi} (U \cap V) \xrightarrow{\psi} \psi(U \cap V).$$
Note that there is also an inverse transition map going in the opposite direction.

Let's see some examples.

Example 1.25. The space $\mathbb{R}^n$ is a topological manifold of dimension n. It has an atlas with the single chart $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$.

Example 1.26. The singleton $\{*\}$ is a topological manifold of dimension 0. In fact, $\{*\} = \mathbb{R}^0$.

Example 1.27. The hypersurface $S^n \subseteq \mathbb{R}^{n+1}$ cut out by the equation

$$x_0^2 + \cdots + x_n^2 = 1$$

is a topological manifold of dimension n. It has charts given by stereographic projection out of some choice of north and south poles. Alternatively, it has charts given by the projection maps $\mathrm{pr}_i \colon S^n \to \mathbb{R}^n$ given by deleting the ith coordinate, defined on the open subsets

$$U_i^{\pm} \coloneqq \{(x_0, \ldots, x_n) \in \mathbb{R}^n : \pm x_i > 0\}$$

for choice of index i and sign in $\{\pm\}$.

Calculus on our manifolds will come from our transition maps.

Definition 1.28. An atlas $\mathcal{A}$ on a topological manifold X is C^k, real analytic, or complex analytic (if $\dim X$ is even) if and only if the transition maps have the corresponding condition.

1.2 August 30

Today we finish our review of smooth manifolds. Once again, we refer to [Elb24] for a few more details and [Lee13] for many more details.

Notation 1.29. We will use the word *regular* to refer to one of the regularity conditions C^k, smooth, real analytic, or complex analytic. We may abbreviate complex analytic to "complex" when no confusion is possible. We use the field $\mathbb{F}$ to denote the "ground field," which is $\mathbb{C}$ when considering the complex analytic case and $\mathbb{R}$ otherwise.

1.2.1 Smooth Manifolds

We now define a regular manifold.

Definition 1.30 (regular manifold)**.** A *regular manifold* of dimension n is a pair $(M, \mathcal{A})$ of a topological manifold M and a maximal regular atlas $\mathcal{A}$; a chart is called regular if and only if it is in $\mathcal{A}$. We will eventually suppress the $\mathcal{A}$ from our notation as much as possible.

The reason for using a maximal atlas is to ensure that it is more or less unique.

Remark 1.31. Here is perhaps a more "canonical" way to deal with atlas confusion. One can say that two regular atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are compatible if and only if the transition maps between them are also regular; this is the same as saying that $\mathcal{A}_1 \cup \mathcal{A}_2$ is regular. Compatibility forms an equivalence relation, and each equivalence class $[\mathcal{A}]$ has a unique maximal element, which one can explicitly define as

$$\mathcal{A}_{\max} \coloneqq \{(U, \varphi) : \mathcal{A} \text{ and } (U, \varphi) \text{ are compatible}\}.$$

This explains why it is okay to just work with maximal atlases.

Example 1.32. One can give the topological manifold $\mathbb{R}^2$ many non-equivalent complex structures. For example, one has the usual choice of $\mathbb{R}^2 \cong \mathbb{C}$, but one can also make $\mathbb{R}^2$ homeomorphic to $B(0,1) \subseteq \mathbb{C}$.

Example 1.33. There are "exotic" smooth structures on S^7.

Example 1.34. Given regular manifolds $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, one can form the product manifold $X \times Y$. It should have maximal atlas compatible with the atlas

$$\{(U \times V, \varphi \times \psi) : (U, \varphi) \in \mathcal{A} \text{ and } (V, \psi) \in \mathcal{B}\}.$$

1.2.2 Regular Functions

With any class of objects, we should have morphisms.

Definition 1.35. A function $f\colon X \to Y$ of regular manifolds $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ is regular if and only if any $p \in X$ has a choice of charts $(U, \varphi) \in \mathcal{A}$ and $(V, \psi) \in \mathcal{B}$ such that $p \in U$ and $f(U) \subseteq V$ and the composite

$$\varphi(U) \overset{\varphi}{\leftarrow} U \overset{f}{\to} V \overset{\psi}{\to} \psi(V)$$

is a regular function between open subsets of Euclidean space.

Remark 1.36. One can replace the single choice of charts above with any choice of charts satisfying $p \in U$ and $f(U) \subseteq V$.

Remark 1.37. Here is another way to state this: for any open $V \subseteq Y$ and smooth function $h\colon V \to \mathbb{F}$, the composite

$$f^{-1}(U) \overset{f}{\to} V \overset{h}{\to} \mathbb{F}$$

succeeds in being smooth (in any local coordinates).

Definition 1.38 (diffeomorphism)**.** A *diffeomorphism* of regular manifolds $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ is a regular map $f\colon X \to Y$ with regular inverse.

Remark 1.39. Alternatively, one can say that the charts in $\mathcal{A}$ and the charts in $\mathcal{B}$ are in natural bijection via f. Checking that these notions align is not too hard.

The above definition of regular map is a little rough to handle, so let's break it down into pieces.

Definition 1.40 (local coordinates)**.** Fix a regular manifold $(X, \mathcal{A})$ of dimension n. Then a system of *local coordinates* around some point $p \in X$ is a choice of regular chart $(U, \varphi) \in \mathcal{A}$ for which $\varphi(p) = 0$. From here, our local coordinates $(x_1, \ldots, x_n)$ are the composite of φ with a coordinate projection to the ground field. (In the complex analytic case, we want the ground field to be $\mathbb{C}$; otherwise, the ground field is $\mathbb{R}$.)

Now, we are able to see that a function $f\colon X \to \mathbb{R}$ is regular if and only if it becomes regular in local coordinates. One can even define regularity with respect to a subset of X.

Regularity allows us to produce lots of manifolds, as follows.

Theorem 1.41. Given regular maps $f_1, \ldots, f_m\colon X \to \mathbb{F}$, the subset

$$\{p \in \mathbb{F}^n : f_1(p) = \cdots = f_m(p) = 0 \text{ and } \{df_1(p), \ldots, df_n(p)\} \text{ are linearly independent}\}$$

is a manifold of dimension $n - m$.

Sketch. This is more or less by the implicit function theorem; for the $\mathbb{F} = \mathbb{R}$ cases, one can essentially follow [Lee13, Corollary 5.14]. ■

Example 1.42. The function $\mathbb{R}^{n+1} \to \mathbb{R}$ given by $(x_0, \ldots, x_n) \mapsto x_0^2 + \cdots + x_n^2$ is real analytic and sufficiently regular at the value 1, which establishes that S^n defined in Example 1.27 succeeds at being a real analytic manifold.

Functions to $\mathbb{F}$ have a special place in our hearts, so we take the following notation.

Notation 1.43. Give a regular manifold X and any open subset $U \subseteq X$, we let $\mathcal{O}_X(U)$ denote the set of regular functions $U \to \mathbb{F}$

Remark 1.44. One can check that the data $\mathcal{O}_X$ assembles into a sheaf. Namely, an inclusion $U \subseteq V$ produces restriction maps $\mathcal{O}_X(V) \to \mathcal{O}_X(U)$.

Remark 1.45. Once we have all of our regular functions out of X, we note that some Yoneda-like philosophy explains that the sheaf of X determines its full regular structure. Here is an explicit statement: given a manifold X and two maximal regular atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ determining sheaves of regular functions $\mathcal{O}_1$ and $\mathcal{O}_2$, having $\mathcal{O}_1 = \mathcal{O}_2$ forces $\mathcal{A}_1 = \mathcal{A}_2$. Indeed, it is enough to show the inclusion $\mathcal{A}_1 \subseteq \mathcal{A}_2$, so suppose (U, φ) is a regular chart in $\mathcal{A}_1$. Then the corresponding local coordinates $(x_1, \ldots, x_n)$ all succeed at being regular for $\mathcal{A}_1$, so they are smooth functions in $\mathcal{O}_1$, so they live in $\mathcal{O}_2$ also, so (U, φ) will succeed at being a regular local diffeomorphism for $\mathcal{A}_2$ and hence be a regular chart.

Sheaf-theoretic notions tell us that we should be interested in germs.

Definition 1.46 (germ)**.** Fix a point p on a regular manifold X. A *germ* of a regular function $f \in \mathcal{O}_X(U)$ (where $p \in U$) is the equivalence class of functions $g \in \mathcal{O}_X(V)$ (for a possibly different open subset V containing p) such that $f|_{U\cap V} = g|_{U\cap V}$. The collection of equivalence classes is denoted $\mathcal{O}_{X,p}$ and is called the stalk at p.

1.3 September 4

Today we hope to finish our review of differential topology.

Convention 1.47. For the remainder of class, our manifolds will be smooth, real analytic, or complex analytic.

1.3.1 Tangent Spaces

Now that we are thinking locally about our functions via germs, we can think locally about our tangent spaces.

Definition 1.48 (derivation)**.** Fix a point p on a regular manifold X. A *derivation* at p is an $\mathbb{F}$-linear map $D\colon \mathcal{O}_{X,p} \to \mathbb{F}$ satisfying the Leibniz rule

$$D(fg) = g(p)D(f) + f(p)D(g).$$

Definition 1.49 (tangent space)**.** Fix a point p on a regular manifold X. Then the *tangent space* T_pX is the $\mathbb{F}$-vector space of all derivations on $\mathcal{O}_{X,p}$.

As with everything in this subject, one desires a local description of the tangent space.

Lemma 1.50. Fix an n-dimensional regular manifold X and a point $p \in X$. Equip p with a chart (U, φ) giving local coordinates $(x_1, \ldots, x_n)$. Then the maps $D_i \colon \mathcal{O}_{X,p} \to \mathbb{F}$ given by

$$D_i \colon [(V, f)] \mapsto \left.\frac{\partial f|_{U\cap V}}{\partial x_i}\right|_p$$

provide a basis for T_pX.

Proof. Checking that this is a derivation follows from the Leibniz rule on the chart. Linear independence of the $D_\bullet$s can also be checked locally by plugging in the germs $[(U, x_i)]$ into any linear dependence.

It remains to check that our derivations span. Well, fix any other derivation D which we want to be in the span of the $D_\bullet$s. By replacing D with $D - \sum_i D(x_i)D_i$, we may assume that $D(x_i) = 0$ for all i. We now want to show that $D = 0$. This amounts to some multivariable calculus. Fix a germ $[(V, f)]$, and shrink U and V enough so that f is defined on U; we want to show $D(f) = 0$. The fundamental theorem of calculus implies

$$f(x_1, \ldots, x_n) = f(0) + \int_0^1 \frac{d}{dt} f(tx_1, \ldots, tx_n)\, dt.$$

However, one can expand out the derivative on the right by the chain rule to see that

$$f(x_1, \ldots, x_n) = f(0) + \sum_{i=1}^n x_i h_i(x_1, \ldots, x_n)$$

for some regular functions $h_1, \ldots, h_n \colon X \to \mathbb{F}$. Applying D, we see that

$$D(f) = \sum_{i=1}^n \underbrace{D(x_i)}_{0} h_i(p) + \underbrace{x_i(p)}_{0} D(h_i) = 0,$$

as required. ■

Tangent spaces have a notion of functoriality.

Definition 1.51. Fix a regular map $F \colon X \to Y$ of regular manifolds. Given $p \in X$, the *differential map* is the linear map $dF_p \colon T_pX \to T_{F(p)}Y$ defined by

$$dF_p(v)(g) \coloneqq v(g \circ F)$$

for any $v \in T_pX$ and germ $g \in \mathcal{O}_{X,p}$. We may also denote $dF_p(v)$ by F_*v.

One has to check that dF_p is linear (which does not have much to check) and satisfies the Leibniz rule (which is a matter of expansion); we will omit these checks.

Remark 1.52. One also has a chain rule: for regular maps $F \colon X \to Y$ and $G \colon Y \to Z$, one has $d(G \circ F)_p = dG_{F(p)} \circ dF_p$.

1.3.2 Immersions and Submersions

This map at the tangent space is important enough to give us other definitions.

Definition 1.53 (submersion, immersion, embedding)**.** Fix a regular function $F\colon X \to Y$.

- The map F is a *submersion* if and only if dF_p is surjective for all $p \in X$.
- The map F is an *immersion* if and only if dF_p is injective for all $p \in X$.
- The map F is an *embedding* if and only if F is an immersion and a homeomorphism onto its image.

Remark 1.54. One can check that submersions $F\colon X \to Y$ have local sections $Y \to X$. Explicitly, for $Q \in Y$, the fiber $F^{-1}(\{Q\}) \subseteq X$ is a manifold, and if $Q \in \operatorname{im} F$, the fiber has codimension $\dim Y$.

Remark 1.55. If $F\colon X \to Y$ is an embedding, then the image $F(X) \subseteq Y$ inherits a unique manifold structure so that the inclusion $F(X) \subseteq Y$ is smooth.

Example 1.56. The projection map $\pi\colon \mathbb{R}^2 \to \mathbb{R}$ given by $\pi(x,y) := x$ is a submersion.

Example 1.57 (lemniscate)**.** The function $F\colon S^1 \to \mathbb{R}^2$ given by

$$F(\theta) := \left(\frac{\cos\theta}{1+\sin^2\theta}, \frac{\cos\theta\sin\theta}{1+\sin^2\theta}\right)$$

can be checked to be an immersion (namely, $F'(\theta) \neq 0$ always), but it fails to be injective because $F(\pi/4) = F(3\pi/4) = (0,0)$, so it is not an embedding.

Example 1.58. The map $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) := x^3$ is a smooth homeomorphism onto its image, but it is not an immersion.

Example 1.59. For any open subset $U \subseteq X$ of a manifold, the inclusion map $U \to X$ is an embedding. (In fact, it is also a submersion.)

We will want to distinguish between embeddings, notably to get rid of open embeddings.

Definition 1.60 (closed)**.** An embedding $F\colon X \to Y$ of regular manifolds is *closed* if and only if $F(X) \subseteq Y$ is closed.

Example 1.61. Fix a submersion $F\colon X \to Y$. A point $Q \in Y$ gives rise to a fiber $Z := F^{-1}(\{Q\})$, which Remark 1.54 explains is a closed submanifold of X of codimension $\dim Y$. One can check that T_pZ is exactly the kernel of $dF_p\colon T_pX \to T_pY$; see [Lee13, Proposition 5.37].

1.3.3 Lie Groups

We now may stop doing topology.

Definition 1.62 (Lie group)**.** A regular *Lie group* is a group object in the category of regular manifolds. For brevity, we may call (real) smooth Lie groups simply "Lie groups" or "real Lie groups," and we may call complex analytic Lie groups simply "complex Lie groups."

As with any object, we have a notion of morphisms.

Definition 1.63 (homomorphism)**.** A homomorphism of regular Lie groups is a regular map of the underlying manifolds and a homomorphism of the underlying groups; an isomorphism of regular Lie groups is a homomorphism with an inverse which is also a homomorphism.

Remark 1.64. If X is already a regular manifold, and we are equipped with continuous multiplication and inverse maps, to check that X becomes a regular Lie group, it is enough to check that merely the multiplication map is regular. See [Lee13, Exercise 7-3].

Remark 1.65. Hilbert's 5th problem asks when C^0 Lie groups can give rise to real Lie groups, and there is a lot of work in this direction. As such, we will content ourselves to focus on real Lie groups instead of any weaker regularity.

Remark 1.66. Any complex Lie group is also a real Lie group.

Here is a basic check which allows one to translate checks to the identity.

Lemma 1.67. Fix a regular Lie group G. For any $g \in G$, the maps $L_g \colon G \to G$ and $R_g \colon G \to G$ defined by $L_g(x) := gx$ and $R_g(x) := xg$ are regular diffeomorphisms.

Proof. Regularity follows from regularity of multiplication. Our inverses of L_g and R_g are given by $L_{g^{-1}}$ and $R_{g^{-1}}$, which verifies that we have defined regular diffeomorphisms. ■

1.4 September 6

Last time we defined a Lie group. Today and for the rest of the course, we will study them.

1.4.1 Examples of Lie Groups

Here are some examples of Lie groups and isomorphisms.

Example 1.68. For our field $\mathbb{F}$, the $\mathbb{F}$-vector space $\mathbb{F}^n$ is a Lie group over $\mathbb{F}$.

Example 1.69. Any finite (or countably infinite) group given the discrete topology becomes a real and complex Lie group.

Example 1.70. The groups $\mathbb{R}^\times$ and $\mathbb{R}^+$ (under multiplication) are real Lie groups. In fact, one has an isomorphism $\{\pm 1\} \times \mathbb{R}^+ \to \mathbb{R}^\times$ of real Lie groups given by $(\varepsilon, r) \mapsto \varepsilon r$.

Example 1.71. The group $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ (under multiplication) is a real Lie group.

Example 1.72. The group $\mathbb{C}^\times$ is a real Lie group. In fact, one has an isomorphism $S^1 \times \mathbb{R}^+ \to \mathbb{C}^\times$ of real Lie groups given again by $(\varepsilon, r) \mapsto \varepsilon r$.

Example 1.73. Over our field $\mathbb{F}$, the set $\mathrm{GL}_n(\mathbb{F})$ of invertible $n \times n$ matrices is a Lie group. Indeed, it is an open subset of $\mathbb{F}^{n^2}$ and thus a manifold, and one can check that the inverse and multiplication maps are rational and hence smooth.

Example 1.74. Consider the collection of matrices

$$\mathrm{SU}_2 := \left\{A \in \mathrm{GL}_2(\mathbb{C}) : \det A = 1 \text{ and } AA^\dagger = 1_2\right\},$$

where $A^\dagger$ is the conjugate transpose. Then SU_2 is an embedded submanifold of $\mathrm{GL}_2(\mathbb{C})$ (cut out by the given equations) and also a subgroup. By writing out $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, one can write out our equations on the coefficients as

$$\begin{cases} ad - bc = 1, \\ a\overline{a} + b\overline{b} = 1, \\ a\overline{c} + b\overline{d} = 0, \\ c\overline{c} + d\overline{d} = 1. \end{cases}$$

In particular, we see that the vector $(a, b) \in \mathbb{C}^2$ is orthogonal to the vector $(\overline{c}, \overline{d})$, so we can solve for this line as providing some $\lambda \in \mathbb{C}$ such that $(\overline{c}, \overline{d}) = \lambda(-b, a)$. But then the determinant condition requires $\lambda = 1$ from $|a|^2 + |b|^2 = 1$. By expanding out $a = w + ix$ and $b = y + iz$, one finds that SU_2 is diffeomorphic to S^3.

The classical groups provide many examples of Lie groups over our field $\mathbb{F}$.

- One has $\mathrm{GL}_n(\mathbb{F})$ and $\mathrm{SL}_n(\mathbb{F})$, which are subsets of matrices cut out by the conditions $\det A \neq 0$ and $\det A = 1$, respectively.

- Orthogonal: fix a non-degenerate symmetric 2-form Ω on $\mathbb{F}^n$. One can always adjust our basis of $\mathbb{F}^n$ so that Ω is diagonal, and by adjusting our basis by squares, we may assume that Ω has only $+1$ or -1s on the diagonal. If $\mathbb{F} = \mathbb{C}$, we can in fact assume that $\Omega = 1_n$, and then we find that our group is

  $$\mathrm{O}_n(\mathbb{C}) := \{A : A^\intercal A = 1_n\}.$$

  Otherwise, if $\mathbb{F} = \mathbb{R}$, then our adjustment (and rearrangement) of the basis allows us to assume that Ω takes the form $\Omega_{k,n-k} := \operatorname{diag}(+1, \ldots, +1, -1, \ldots, -1)$ with k copies of $+1$ and $n - k$ copies of -1, and we define

  $$\mathrm{O}_{k,n-k}(\mathbb{R}) := \{A : A^\intercal \Omega A = \Omega\}$$

- Special orthogonal: one can add the condition that $\det A = 1$ to all the above orthogonal groups, which makes the special orthogonal groups.

- Symplectic: for $\mathbb{F}^{2n}$, one can fix a non-degenerate symplectic 2-form Ω. It turns out that, up to basis, we find that $\Omega = \left[\begin{smallmatrix} 0 & -1_n \\ 1_n & 0 \end{smallmatrix}\right]$, and we define

  $$\mathrm{Sp}_{2n}(\mathbb{F}) := \{A : A^\intercal \Omega A = 1_{2n}\}.$$

- Unitary: using the non-degenerate Hermitian forms, we can similarly define

  $$\mathrm{U}_n(\mathbb{C}) := \left\{A : A^\dagger A = 1_n\right\}$$

  is a real Lie group. (Conjugation is not complex analytic, so this is not a complex Lie group!)

1.4.2 Connected Components

We will want to focus on connected Lie groups in this class, so we spend a moment describing why one might hope that this is a reasonable reduction. The main point is that it is basically infeasible to classify finite groups, and allowing for disconnected Lie groups forces us to include all these groups in our study by Example 1.69.

Quickly, recall our notions of connectivity; we refer to [Elb22, Appendix A.1] for details.

Definition 1.75 (connected)**.** A topological space X is *disconnected* if and only if there exists disjoint nonempty open subsets $U, V \subseteq X$ covering X. If there exists no such pair of open subsets, then X is *connected*; in other words, the only subsets of X which are both open and closed are $\varnothing$ and X.

Definition 1.76 (connected component)**.** Given a topological space X and a point $p \in X$, the *connected component* of $p \in X$ is the union of all connected subspaces of X containing p.

Remark 1.77. One can check that the connected component is in fact connected and is thus the maximal connected subspace.

Definition 1.78 (path-connected)**.** A topological space X is *path-connected* if and only if any two points $p, q \in X$ have some (continuous) path $\gamma\colon [0,1] \to X$ such that $\gamma(0) = p$ and $\gamma(1) = q$.

Definition 1.79. Given a topological space X and a point $p \in X$, the *path-connected component* of $p \in X$ is the collection of all $q \in X$ for which there is a path connecting p and q.

Remark 1.80. One can check that having a path connecting two points of X is an equivalence relation on the points of X. Then the path-connected components are the equivalence classes for this equivalence relation. From this, one can check that the path-connected components are the maximal path-connected subsets of a topological space.

One has the following lemmas.

Lemma 1.81. Fix a topological space X.

(a) If X is path-connected, then X is connected.

(b) If X is a connected topological n-manifold, then X is path-connected.

Proof. Part (a) is [Elb22, Lemma A.16]. Part (b) is [Elb24, Proposition 1.39]. ■

Lemma 1.82. Fix a continuous surjection $f\colon X \to Y$ of topological spaces. If X is connected, then Y is connected.

Proof. This is [Elb22, Lemma A.8]. ■

Anyway, we are now equipped to return to our discussion of Lie groups.

Lemma 1.83. Fix a Lie group G, and let $G^\circ \subseteq G$ be the connected component of the identity $e \in G$. For any $g \in G$, we see that gG° is the connected component of g.

Proof. Certainly gG° is a connected subset containing g by Lemma 1.82 (note multiplication is continuous), so it is contained in the connected component of g. On the other hand, any connected subset U around g must have $g^{-1}U$ be a connected subset around e, so $g^{-1}U \subseteq G^\circ$, so $U \subseteq gG^\circ$. In particular, the connected component of g is also contained in gG°. ■

Proposition 1.84. Fix a Lie group G, and let $G^\circ \subseteq G$ be the connected component of the identity $e \in G$.

(a) Then G° is a normal subgroup of G.

(b) The quotient $\pi_0(G) := G/G^\circ$ given the quotient topology from the surjection $G \twoheadrightarrow \pi_0(G)$ is a discrete countable group.

Proof. We show the parts independently.

(a) We check this in parts.

- Of course G° is a subgroup: it contains the identity, and the images of the maps $i\colon G^\circ \to G$ and $m\colon G^\circ \times G^\circ \to G$ must land in connected subsets of G containing the identity by Lemma 1.82, so we see that G° is contained
- We now must check that G° is normal. Fix some $g \in G$, and we want to show that $gG^\circ g^{-1} \subseteq G^\circ$. Well, define the map $G^\circ \to G$ by $a \mapsto gag^{-1}$, which we note is continuous because multiplication and inversion are continuous. Lemma 1.82 tells us that the image must be connected, and we see $e \mapsto e$, so the image must actually land in G°.

(b) One knows that $\pi_0(G)$ is a group because G° is normal, and it is discrete because connected components are both closed and open in G, so the corresponding points are closed and open in $\pi_0(G)$. (We are implicitly using Lemma 1.83.) This is countable because a separated topological space must have countably many connected components. ■

Remark 1.85. One can restate the above result as providing a short exact sequence

$$1 \to G^\circ \to G \to \pi_0(G) \to 1$$

of Lie groups. In this way, we can decompose our study of G into connected Lie groups and discrete countable groups. In this course, we will ignore studying discrete countable groups because they are too hard.

1.5 September 9

Today we talk more about subgroups and coverings.

1.5.1 Closed Lie Subgroups

Arbitrary subgroups of Lie groups may not inherit a manifold structure, so we add an adjective to acknowledge this.

Definition 1.86 (closed Lie subgroup)**.** Fix a Lie group G. A *closed Lie subgroup* is a subgroup $H \subseteq G$ which is also an embedded submanifold.

Remark 1.87. On the homework, we will show that closed Lie subgroups are in fact closed subsets of G. It is a difficult theorem (which we will not prove nor use in this class) that being a closed subset and a subgroup implies that it is an embedded submanifold.

Here are some checks on subgroups.

Lemma 1.88. Fix a connected topological group G. Given an open neighborhood U of G of the identity $e \in G$, the group G is generated by U is all of G.

Proof. Let H be the subgroup generated by U. For each $h \in H$, we see that $hU \subseteq H$, which is an open neighborhood (see Lemma 1.67), so $H \subseteq G$ is open. However, we also see that

$$G = \bigsqcup_{[g]} gH,$$

where $[g]$ varies over representatives of cosets. Thus, $G \setminus H$ is again the union of open subsets, so H is also closed, so $G = H$ because G is connected. ■

Lemma 1.89. Fix a homomorphism $f\colon G_1 \to G_2$ of connected Lie groups. If $df_e\colon T_eG_1 \to T_eG_2$ is surjective, then f is surjective.

Proof. By translating around (by Lemma 1.67), we see that f is a submersion. (Explicitly, for each $g \in G_1$, we see that $R_{f(g)} \circ f$ must continue to be a submersion at the identity, but this equals $f \circ R_g$, so f is a submersion at g too.) Because submersions are open [Lee13, Proposition 4.28], we see that f being a submersion means that its image is an open subgroup of G_2, which is all of G_2 by Lemma 1.88. ■

Here is a check to be a closed Lie subgroup.

Lemma 1.90. Fix a regular Lie group G of dimension n. A subgroup $H \subseteq G$ is a closed Lie subgroup of dimension k if and only if there is a single regular chart (U, φ) with $e \in U$ such that

$$U \cap H = \{g \in U : \varphi_{k+1}(g) = \cdots = \varphi_n(g) = 0\}$$

for some.

Proof. We have constructed a slice chart for the identity $e \in H$. We will translate this slice chart around to produce a slice chart for arbitrary $h \in H$, which will complete the proof by [Lee13, Theorem 5.8]. In particular, for any $h \in H$, we know that left translation $L_{h^{-1}}\colon G \to G$ is a diffeomorphism, so the composite

$$hU \overset{L_{h^{-1}}}{\to} U \overset{\varphi}{\to} \varphi(U)$$

continues to be a chart of G with $h \in hU$. Furthermore, we see that $g \in hU$ lives in H if and only if $L_{h^{-1}}g \in U \cap H$, which by hypothesis is equivalent to

$$\varphi_{k+1}\left(L_{h^{-1}}g\right) = \cdots = \varphi_n\left(L_{h^{-1}}g\right) = 0.$$

Thus, we have constructed the desired slice chart. ■

We may want some more flexibility with our subgroups.

Example 1.91. Fix an irrational number $\alpha \in \mathbb{R}$. Then there is a Lie group homomorphism $f\colon \mathbb{R} \to (\mathbb{R}/\mathbb{Z})^2$ defined by $f(t) := (t, \alpha t)$. One can check that $\operatorname{im} f \subseteq (\mathbb{R}/\mathbb{Z})^2$ is a dense subgroup, but it is not closed!

So we have the following definition.

Definition 1.92 (Lie subgroup)**.** Fix a Lie group G. A *Lie subgroup* is a subgroup $H \subseteq G$ which is an immersed submanifold.

1.5.2 Quotient Groups

Along with subgroups, we want to be able to take quotients.

Definition 1.93 (fiber bundle)**.** A *fiber bundle* with fiber F on a smooth manifold X is a surjective continuous map $\pi\colon Y \to X$ such that there is an open cover $\mathcal{U}$ of X and (local) homeomorphisms making the following diagram commute for all $U \in \mathcal{U}$.

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\sim"] \arrow[dr, "\pi"'] & & F \times U \arrow[dl, "\mathrm{pr}_2"] \\ & U & \end{tikzcd}$$

Fiber bundles are the correct way to discuss quotients.

Theorem 1.94. Fix a closed Lie subgroup H of a Lie group G.

(a) Then G/H is a manifold of dimension $\dim G - \dim H$ equipped with a quotient map $q\colon G \twoheadrightarrow G/H$.

(b) In fact, q is a fiber bundle with fiber H.

(c) If H is normal in G, then G/H is actually a Lie group (with the usual group structure).

(d) We have $T_e(G/H) \cong T_eG/T_eH$.

Proof. We construct the manifold structure on G/H as follows: for each $g \in G$, we produce a coset $\overline{g} \in G/H$, which we note as $q^{-1}(\overline{g}) = gH$. Now, $gH \subseteq G$ is an embedded submanifold because H is (we are using Lemma 1.67), so one can locally find a submanifold $M \subseteq G$ around g intersecting gH transversally, meaning that

$$T_gG = T_gM \oplus T_g(gH).$$

By shrinking M, we can ensure that the above map continues to be an isomorphism in a neighborhood of g, so the multiplication map $M \times H \to UH$ is a diffeomorphism. Now, MH is an open neighborhood of $g \in G$, and M projects down to an open subset of G/H, so $M \cong q(\overline{M})$ provides our chart.

Now, (a) and (d) follows by inspection of the construction. We see that (b) follows because we built our projection map $G \twoheadrightarrow G/H$ so that it locally looks like $U \times H \twoheadrightarrow \overline{U}$, so we get our fiber bundle. Lastly, (d) follows by the equality $T_gG = T_gM \oplus T_g(gH)$. ■

Remark 1.95. Writing the above out in detail would take several pages; see [Lee13, Theorem 21.10].

Access to quotients permits an isomorphism theorem, which we will prove later when we have talked a bit about Lie algebras.

Theorem 1.96 (Isomorphism)**.** Fix a Lie group homomorphism $f\colon G_1 \to G_2$.

(a) The kernel $\ker f$ is a normal closed Lie subgroup of G_1.

(b) The quotient $G_1/\ker f$ is a Lie subgroup of G_2.

(c) The image $\operatorname{im} f$ is a Lie subgroup of G_2. If $\operatorname{im} f$ is further closed, then $G_1/\ker f \to \operatorname{im} f$ is an isomorphism of Lie subgroups.

1.5.3 Actions

Groups are known by their actions, so let's think about how our actions behave.

Definition 1.97 (action)**.** Fix a Lie group G and regular manifold X. Then a *regular action* of G on X is a regular map $\alpha\colon G \times X \to X$ satisfying the usual constraints, as follows.

(a) Identity: $\alpha(e, x) := x$.

(b) Composition: $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$.

This allows us to define the usual subsets.

Definition 1.98 (orbit, stabilizer)**.** Fix a regular action of a Lie group G on a regular manifold X.

(a) The *orbit* of $x \in X$ is the subspace $Gx := \{gx : g \in G\}$.

(b) The *stabilizer* of $x \in X$ is the subgroup $G_x := \{g \in G : gx = x\}$.

Here are some examples.

Example 1.99. The group $\mathrm{GL}_n(\mathbb{F})$ acts on the vector space $\mathbb{F}^n$.

Example 1.100. The group $\mathrm{SO}_3(\mathbb{R})$ preserves distances in its action on $\mathbb{R}^3$, so its action descends to an action on S^2.

Representations are special kinds of actions.

Definition 1.101 (representation)**.** Fix a Lie group G over $\mathbb{F}$. Then a *representation* of G is the (regular) linear action of G on a finite-dimensional vector space V over $\mathbb{F}$; namely, the map $v \mapsto g \cdot v$ for each $g \in G$ must be a linear map $V \to V$. A *homomorphism* of representations V and W is a linear map $A\colon V \to W$ such that $A(gv) = g(Av)$. These objects and morphisms make the category $\mathrm{Rep}_{\mathbb{F}}(G)$.

Remark 1.102. Equivalently, we may ask for the induced map $G \to \mathrm{GL}(V)$, given by sending $g \in G$ to the map $v \mapsto gv$, to be a Lie group homomorphism.

Remark 1.103. The category $\operatorname{Rep}_{\mathbb{F}}(G)$ comes with many nice operations.

(a) Duals: given a representation $\pi\colon G \to \operatorname{GL}(V)$, we can induce a G-action on $V^* := \operatorname{Hom}(V, \mathbb{F})$ by

$$\left((\pi^* g)(v^*)\right)(v) := v^*\left(g^{-1}v\right).$$

(Here, $\pi^* g$ should be a map $V^* \to V^*$, so it takes a linear functional $v^* \in V^*$ as input and produces the linear functional $(\pi^* g)(v^*)$ as output.)

(b) Tensor products: given representations $\pi\colon G \to \operatorname{GL}(V)$ and $\pi'\colon G \to \operatorname{GL}(V')$, we can induce a G-action on $V \otimes W$ by

$$\left((\pi \otimes \pi')(g \otimes g')\right)(v \otimes v') := \pi(g)v \otimes \pi'(g')v'.$$

(c) Hom sets: given representations $\pi\colon G \to \operatorname{GL}(V)$ and $\pi'\colon G \to \operatorname{GL}(V')$, we can induce a G-action on $\operatorname{Hom}(V, W)$ by

$$(g\varphi)(v) := \pi'(g)\left(\varphi(\pi(g)^{-1}v)\right).$$

(d) Quotients: given representations $\pi\colon G \to \operatorname{GL}(V)$ and $\pi'\colon G \to \operatorname{GL}(V')$, where $V \subseteq V'$ is a G-representation, then we can induce a G-action on V'/V by

$$\pi'(g)(v' + V) := gv' + V.$$

One can check that these operations make $\operatorname{Rep}_{\mathbb{F}}(G)$ into a symmetric monoidal abelian category. Checking that these actually form actions is a matter of writing out the definitions, so we will omit it. (Notably, all of these actions are algebraic combinations of previous actions, so all regularity is inherited.)

Returning to group actions on manifolds, we remark that Theorem 1.96 can be seen as a version of the Orbit–stabilizer theorem.

Theorem 1.104 (Orbit–stabilizer). Fix a regular action of a Lie group G on a regular manifold X. Further, fix $x \in X$.

(a) The orbit Gx is an immersed submanifold of X.

(b) The stabilizer G_x is a closed Lie subgroup of G.

(c) The quotient map $f\colon G/G_x \to X$ given by $g \mapsto gx$ is an injective immersion.

(d) If Gx is an embedded submanifold, then the map f of (c) is a diffeomorphism.

1.6 September 11

Today we talk more about homogeneous spaces.

1.6.1 Homogeneous Spaces

Let's see some applications of Theorem 1.104.

Example 1.105. Suppose a regular Lie group G acts smoothly and transitively on a regular manifold X. For each $x \in X$, we see that $G/G_x \to X$ is a bijective immersion. In particular, Sard's theorem implies that $\dim G/G_x = \dim X$, so we conclude that this map is in fact a bijective local diffeomorphism, which of course is just a diffeomorphism. Thus, Theorem 1.94 tells us that the map $G \to X$ given by $g \mapsto gx$ is a fiber bundle with fiber G_x.

The above situation is so nice that it earns a name.

Definition 1.106 (homogeneous space)**.** Fix a regular Lie group G acting smoothly and transitively on a regular manifold X. If the action of G is transitive, we say that X is a *homogeneous space* of G.

Here are many examples.

Example 1.107. Continuing from Example 1.100, we recall that $\mathrm{SO}_3(\mathbb{R})$ acts on S^2. In fact, one can check that the stabilizer of any $x \in S^3$ is isomorphic to S^1, so Example 1.105 tells us that $\mathrm{SO}_3(\mathbb{R}) \to S^2$ is a fiber bundle with fiber S^1. In general, we find that $\mathrm{SO}_n(\mathbb{R}) \to S^n$ is a fiber bundle with fiber $\mathrm{SO}_{n-1}(\mathbb{R})$.

Example 1.108. The group SU_2 acts on $\mathbb{CP}^1$ by matrix multiplication. We see that the stabilizer of some line in $\mathbb{CP}^1$ consists of the matrices in SU_2 with a nonzero eigenvector on the line. For example, using the computation of Example 1.74, we see that trying to stabilizer $[1:0]$ gives rise to the matrices $\left[\begin{smallmatrix} a & -b \\ \overline{b} & \overline{a} \end{smallmatrix}\right]$ where we require $b=0$. Thus, we see that our stabilizer is isomorphic to U_1. In particular, our orbits are compact immersed submanifolds of $\mathbb{CP}^1$ of dimension $\dim \mathrm{SU}_2 - \dim \mathrm{U}_1 = \dim \mathbb{CP}^1$, so the action must be transitive in order for orbits to be closed and the correct dimension.

Example 1.109. One can check that SU_n acts on $S^{2n-1} \subseteq \mathbb{C}^n$ with stabilizer isomorphic to SU_{n-1}.

Example 1.110 (flag varieties)**.** Let $\mathcal{F}_n$ be the set of "flags" of $\mathbb{F}^n$, which is an ascending chain of subspaces

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1} \subsetneq V_n = \mathbb{F}^n.$$

Then we see that $\mathrm{GL}_n(\mathbb{F})$ acts on $\mathcal{F}_n$ by matrix multiplication. On the homework, we check that this action is transitive with stabilizer (of the standard flag $\{\operatorname{span}_{\mathbb{F}}(e_1, \ldots, e_i)\}_{i=0}^n$) given by the matrix subgroup $B_n(\mathbb{F}) \subseteq \mathrm{GL}_n(\mathbb{F})$ of upper-triangular matrices. Thus, we see that we can realize $\mathcal{F}_n$ as the manifold quotient $\mathrm{GL}_n(\mathbb{F})/B_n(\mathbb{F})$, providing a manifold structure.

Example 1.111 (Grassmannians)**.** Let $\mathrm{Gr}_k(\mathbb{F}^n)$ be the set of vector subspaces $V \subseteq \mathbb{F}^n$ of dimension k. Then we see that $\mathrm{GL}_n(\mathbb{F})$ acts transitively on $\mathrm{Gr}_k(\mathbb{F}^n)$ with stabilizer of $\operatorname{span}(e_1, \ldots, e_k)$ given by matrices of the form

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

where $A \in \mathbb{F}^{k \times k}$ and $B \in \mathbb{F}^{k \times (n-k)}$ and $D \in \mathbb{F}^{(n-k)\times(n-k)}$. Thus, we can realize $\mathrm{Gr}_k(\mathbb{F}^n)$ as the manifold quotient of $\mathrm{GL}_n(\mathbb{F})$, providing a manifold structure.

Example 1.112. There are many regular actions of G on itself.

- Regular left: define our action $R_\ell \colon G \times G \to G$ by $(g, x) \mapsto gx$.
- Regular right: define our action $R_r \colon G \times G \to G$ by $(g, x) \mapsto xg^{-1}$.
- Adjoint: define our action $\mathrm{Ad} \colon G \times G \to G$ by $(g, x) \mapsto gxg^{-1}$. (This action is rarely transitive!)

Example 1.113 (adjoint). Fix a regular Lie group G. Note that $\operatorname{Ad}_g(1) = 1$, so we may take the differential to provide a map $(d\operatorname{Ad}_g)_e : T_eG \to T_eG$, so we get an adjoint representation

$$\begin{array}{rcl} G \times T_eG & \to & T_eG \\ (g\ ,\ \ v) & \mapsto & (d\operatorname{Ad}_g)_e(v) \end{array}$$

which we will frequently abuse notation to abbreviate the above as providing some $\operatorname{Ad}_g \in \operatorname{GL}(T_eG)$. Expanding everything in sight into coordinates reveals that this action is smooth; in fact, one can check (again in coordinates) that the map $G \to \operatorname{GL}(T_eG)$ given by $g \mapsto (d\operatorname{Ad}_g)_e$ is smooth. Taking the differential of this last map produces a map $T_eG \to \operatorname{End}(T_eG)$, which is sometimes called the adjoint representation of T_eG.

1.6.2 Covering Spaces

It will help to recall some theory around covering spaces. See [Elb23] for (some) more detail about this theory or [Hat01] for (much) more detail.

Definition 1.114 (covering space). A *covering space* is a fibration $p: Y \to X$ with discrete fiber S. The *degree* of p equals $\#S$.

In more words, we are asking for each $x \in X$ to have an open neighborhood U such that the restriction $p^{-1}(U) \to U$ is homeomorphic (over U) to $\bigsqcup_{s\in S} p^{-1}(U) \to U$ for some discrete set S.

Remark 1.115. If X is a regular manifold and $\deg p \le |\mathbb{N}|$, then Y is also a regular manifold. Indeed, being a manifold is checked locally, so one can find neighborhoods as in the previous remark to witness the manifold structure.

We are interested in paths in topological spaces, but there are too many. To make this set smaller, we consider it up to homotopy.

Definition 1.116 (homotopy). Fix a topological space X. Two paths $\gamma_0, \gamma_1 : [0,1] \to X$ are *homotopic* relative to their endpoints if and only if there is a continuous map $H_\bullet : [0,1]^2 \to X$ such that $H_0 = \gamma_0$ and $H_1 = \gamma_1$ and $H_s(0) = \gamma_0(0) = \gamma_1(0)$ and $H_s(1) = \gamma_0(1) = \gamma_1(1)$ for all s. The map H is called a *homotopy*.

Definition 1.117 (simply connected). A topological space X is *simply connected* if and two paths with the same endpoints are homotopic relative to those endpoints.

Example 1.118. One can check that S^1 fails to be simply connected because the path going around the circle is not homotopic to the constant path.

It is important to know that one can lift paths.

Theorem 1.119. Fix a covering space $p: Y \to X$. Fix some point $x \in X$ and a path $\gamma: [0,1] \to X$ with $\gamma(0) = x$. Then each $\widetilde{x} \in p^{-1}(\{x\})$ has a unique path $\widetilde{\gamma}: [0,1] \to Y$ such that $\widetilde{\gamma}(0) = \widetilde{x}$ and making the following diagram commute.

$$\begin{tikzcd} {[0,1]} \arrow[r, "\widetilde{\gamma}"] \arrow[dr, "\gamma"'] & \widetilde{X} \arrow[d, "p"] \\ & X \end{tikzcd}$$

Remark 1.120. One can further check that having two homotopic paths $\gamma_1 \sim \gamma_2$ downstairs produce homotopic paths $\widetilde{\gamma}_1 \sim \widetilde{\gamma}_2$.

Remark 1.121. More generally, fix a simply connected topological space Z. Then given a map $f: Z \to X$ and a choice of $\widetilde{x} \in p^{-1}(\{x\})$ and $z \in f^{-1}(\{x\})$, there will be a unique lift $\widetilde{f}: Z \to Y$ such that $\widetilde{f}(z) = \widetilde{x}$. In short, given any $z' \in Z$, find a path connecting z and z', send this path into X and then lift it up to Y. Because Z is simply connected (and the above theorem), the choice of path from z to z' does not really matter.

Anyway, we now define our collection of paths.

Definition 1.122 (fundamental group)**.** Fix a point x of a topological space X. Then the set of paths both of whose endpoints are x forms a monoid with operation given by composition (i.e., concatenation). If we take the quotient of this monoid by homotopy classes of paths, then we get a group of path homotopy classes, which we call $\pi_1(X, x)$. This is the *fundamental group*.

Remark 1.123. For any two $x, y \in X$ in the same path-connected component, the path $\alpha: [0,1] \to X$ connecting x to y produces an isomorphism $\pi_1(X, x) \cong \pi_1(X, y)$ by $\gamma \mapsto \alpha \cdot \gamma \cdot \alpha^{-1}$, where $\cdot$ denotes path composition.

Remark 1.124. The above remark allows us to verify that X is simply connected if and only $\pi_1(X, x)$ is trivial for all x. In fact, we only have to check this for one x in each path-connected component.

1.7 September 13

Today we continue our discussion of coverings.

1.7.1 The Universal Cover

There is more or less one covering space which produces all the other ones.

Definition 1.125 (universal cover)**.** Fix a path-connected topological space X. Then a covering space $p: Y \to X$ is the *universal cover* if and only if Y is connected and simply connected.

We now discuss an action of $\pi_1(X, b)$ on covering spaces in order to better understand this universal cover. Fix a covering space $p: Y \to X$ and a basepoint $x \in X$. Then we note that $\pi_1(X, x)$ acts on the fiber $p^{-1}(\{x\})$ as follows: for any $[\gamma] \in \pi_1(X, x)$ and $\widetilde{x} \in p^{-1}(\{x\})$, we define $\widetilde{\gamma}: [0,1] \to Y$ by lifting the path $\gamma: [0,1] \to X$ up to Y so that $\widetilde{\gamma}(0) = \widetilde{x}$; then

$$[\gamma] \cdot \widetilde{x} := \widetilde{\gamma}(1).$$

One can check that this action is well-defined (namely, it does not depend on the representative γ and does provide a group action). Here are some notes.

- If Y is path-connected, then the action is transitive: and $\widetilde{x}, \widetilde{x}' \in p^{-1}(\{x\})$ admit a path $\widetilde{\gamma}: [0,1] \to Y$ with $\widetilde{\gamma}(0) = \widetilde{x}$ and $\widetilde{\gamma}(1) = \widetilde{x}'$, so $\gamma := p \circ \widetilde{\gamma}$ has

$$[\gamma] \cdot \widetilde{x} := \widetilde{x}'$$

  by construction of γ.

- If Y is simply connected, then this action is also free. Indeed, choose two paths $\gamma_1, \gamma_2 \colon [0,1] \to X$ representing classes in $\pi_1(X,x)$. Now, suppose that $[\gamma_1] \cdot \widetilde{x} = [\gamma_2] \cdot \widetilde{x}$ for each $\widetilde{x} \in p^{-1}(\{x\})$, and we will show that $[\gamma_1] = [\gamma_2]$. Well, choosing lifts $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$, the hypothesis implies that they have the same endpoints. Thus, because Y is simply connected, we know $\widetilde{\gamma}_1 \sim \widetilde{\gamma}_2$. We now see that $\gamma_1 \sim \gamma_2$ by composing the homotopy witnessing $\widetilde{\gamma}_1 \sim \widetilde{\gamma}_2$ with p.

The conclusion is that $p^{-1}(\{x\})$ is in bijection with $\pi_1(X,x)$ when $p\colon Y \to X$ is the universal cover. Here are some examples.

Example 1.126. One has a covering space $p\colon S^n \to \mathbb{RP}^n$ given by

$$(x_0, \ldots, x_n) \mapsto [x_0 : \ldots : x_n].$$

For $n \geq 2$, we know S^n is simply connected, so it will be the universal cover, and we are able to conclude that $\pi_1(\mathbb{RP}^n)$ is isomorphic to a fiber of p, which has two elements, so $\pi_1(\mathbb{RP}^n) \cong \mathbb{Z}/2\mathbb{Z}$.

Example 1.127. One has a covering space $p\colon \mathbb{R} \to S^1$ given by $p(t) := e^{2\pi i t}$. We can see that $\mathbb{R}$ is simply connected (it's convex), so this is a universal covering. This at least tells us that $\pi_1\left(S^1\right)$ is countable, and one can track through the group law through the above bijections to see that actually $\pi_1\left(S^1\right) \cong \mathbb{Z}$.

Example 1.128. One can show that $\pi_1(\mathbb{C} \setminus \{z_1, \ldots, z_n\})$ is the free group on n generators, basically corresponding to how one goes around each point.

Now, in the context of our Lie groups, we get the following result.

Theorem 1.129. Fix a regular Lie group G, and let $p\colon \widetilde{G} \to G$ be the universal cover.

(a) Then $\widetilde{G}$ has the structure of a regular Lie group.

(b) The projection p is a homomorphism of Lie groups.

(c) The kernel $\ker p \subseteq \widetilde{G}$ is discrete, central, and isomorphic to $\pi_1(G,e)$. In particular, $\pi_1(G,e)$ is commutative.

Proof. Here we go.

(a) Remark 1.115 tells us that $\widetilde{G}$ is a regular manifold, so it really only remains to exhibit the group structure. We will content ourselves with merely describing the group structure. Fix any $\widetilde{e} \in p^{-1}(\{e\})$, which will be our identity.

Now, $\widetilde{G}$ is simply connected, so $\widetilde{G} \times \widetilde{G}$ is also simply connected. Thus, Remark 1.115 explains that the composite

$$\widetilde{G} \times \widetilde{G} \to G \times G \stackrel{m}{\to} G$$

will lift to a unique map to the universal cover as a map $\widetilde{m}$ making the following diagram commute.

$$\begin{tikzcd} \widetilde{G} \times \widetilde{G} \arrow[r, "\widetilde{m}"] \arrow[d, "{(p,p)}"'] & \widetilde{G} \arrow[d, "p"] \\ G \times G \arrow[r, "m"] & G \end{tikzcd} \qquad \begin{tikzcd} (\widetilde{e}, \widetilde{e}) \arrow[r, mapsto] \arrow[d, mapsto] & \widetilde{e} \arrow[d, mapsto] \\ (e,e) \arrow[r, mapsto] & e \end{tikzcd}$$

One can construct the inverse map similarly by lifting the map $\widetilde{G} \stackrel{p}{\to} G \stackrel{i}{\to} G$ to a map to $\widetilde{G}$ sending $\widetilde{e} \mapsto \widetilde{e}$. Uniqueness of lifting will guarantee that we satisfy the group law.

(b) We see that p is a homomorphism by construction of $\widetilde{m}$ above.

(c) This is on the homework. ■

Example 1.130. Recall that we have the fiber bundle $\mathrm{SO}_n(\mathbb{R}) \to S^{n-1}$ with fiber $\mathrm{SO}_{n-1}(\mathbb{R})$. Thus, the long exact sequence in homotopy groups produces

$$\pi_2\left(S^{n-1}\right) \to \pi_1(\mathrm{SO}_{n-1}(\mathbb{R})) \to \pi_1(\mathrm{SO}_n(\mathbb{R})) \to \pi_1\left(S^{n-1}\right) \to \pi_0\left(\mathrm{SO}_{n-1}(\mathbb{R})\right).$$

Now, for $n \geq 4$, one has that $\pi_2\left(S^{n-2}\right) = \pi_1\left(S^{n-1}\right) = 1$, so we have $\pi_1\left(\mathrm{SO}_{n-1}(\mathbb{R})\right) \cong \pi_1\left(\mathrm{SO}_n(\mathbb{R})\right)$. One can check that $\mathrm{SO}_3(\mathbb{R}) \cong \mathbb{RP}^3$, so we see that

$$\pi_1\left(\mathrm{SO}_n(\mathbb{R})\right) \cong \mathbb{Z}/2\mathbb{Z}$$

for $n \geq 4$. The universal (double) cover of $\mathrm{SO}_n(\mathbb{R})$ is called Spin_n, and Theorem 1.129 explains that we have a short exact sequence

$$1 \to \mathbb{Z}/2\mathbb{Z} \to \mathrm{Spin}_n \to \mathrm{SO}_n(\mathbb{R}) \to 1.$$

Example 1.131. More concretely, one can show that $\mathrm{SU}_2(\mathbb{C})$ has an action on $\mathbb{R}^3$ preserving distances and orientation, so we get a homomorphism $\mathrm{SU}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{R})$. One can check that this map is surjective with kernel isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

Remark 1.132. In general, Theorem 1.129 explains that we have a short exact sequence

$$1 \to \pi_1(G) \to \widetilde{G} \to G \to 1$$

for any regular Lie group G, so it does not cost us too much to pass from G to $\widetilde{G}$, allowing us to assume that the Lie groups we study are simply connected. (Note that even though $\pi_1(G)$ is discrete, the short exact sequence does not split: $\widetilde{G}$ succeeds at being connected.)

1.7.2 Vector Fields

Fix a regular manifold X of dimension n. We may be interested in thinking about all our tangent spaces at once.

Definition 1.133 (tangent bundle)**.** Fix a regular manifold X. Then we define the *tangent bundle* as

$$TX \coloneqq \{(x, v) : v \in T_x X\}.$$

Note that there is a natural projection map $TX \to X$ by $(x, v) \mapsto x$.

Remark 1.134. Locally on a chart (U, φ) of X, we see that φ provides coordinates $(x_1, \ldots, x_n)$ on U, so one has a bijection

$$U \times \mathbb{R}^n \to TU$$

by sending $(x, \partial/\partial x_i) \mapsto \left(\varphi^{-1}(x), d\varphi_x^{-1}(\partial/\partial x_i)\right)$ (In the future, we may conflate $d\varphi_x^{-1}(\partial/\partial x_i)$ with $\partial/\partial x_i$). This provides a chart for TU, and one can check that these charts are smoothly compatible by an explicit computation using the smooth compatibility of charts on X. The point is that $TX \to X$ is a vector bundle of rank n.

Vector bundles are interesting because of their sections.

Definition 1.135 (vector field)**.** Fix a regular manifold X. Then a *vector field* on X is a smooth section $\sigma\colon X \to TX$ of the natural projection map $TX \to X$.

Remark 1.136. Locally on a chart (U, φ) with coordinates $(x_1, \ldots, x_n)$, we see that we can think about a vector field σ locally as

$$\sigma(x) := \sum_{i=1}^{n} \sigma_i(x) \frac{\partial}{\partial x_i}\bigg|_x,$$

where the smoothness of σ enforces the $\sigma_\bullet$s to be smooth. Changing coordinates to (U', φ') with a coordinate expansion $\sigma(x) = \sum_i \sigma_i'(x) \frac{\partial}{\partial x_i'}$, one can change bases using the Jacobian of $\varphi' \circ \varphi^{-1}$ to find that

$$\sigma_i'(x) = \sum_{j=1}^{n} \frac{\partial x_i'}{\partial x_j'} \sigma_j(x).$$

Anyway, the point is that we can define a vector field locally on these coordinates and then going back and checking that we have actually defined something that will glue smoothly up to X.

The reason we care so much about tangent spaces in this class is because they give rise to our Lie algebras, whose representations are somehow our main focus.

Definition 1.137 (Lie algebra)**.** Fix a Lie group G. Then the *Lie algebra* of G is the vector space

$$\mathfrak{g} := T_e G.$$

We may also notate $\mathfrak{g}$ by $\operatorname{Lie}(G)$.

It is somewhat difficult to find structure in this tangent space immediately, so we note that T_eG is isomorphic with another vector space.

Definition 1.138 (invariant vector field)**.** Fix a Lie group G. Then a vector field $\xi\colon G \to TG$ is *left-invariant* if and only if

$$\xi(gx) = dL_g(\xi(x))$$

for any $x, g \in G$. One can define *right-invariant* analogously.

Remark 1.139. We claim that the vector space of left-invariant vector fields is isomorphic to $\mathfrak{g}$. Here are our maps.

- Given a left-invariant vector field ξ, one can produce the tangent vector $\xi(e) \in \mathfrak{g}$.
- Given some $\xi(e) \in \mathfrak{g}$, we define

  $$\xi(g) := dL_g(\xi(e)) \in T_gG.$$

  It is not difficult to check that $\xi\colon G \to TG$ is at least a section of the natural projection $TG \to G$. We omit the check that ξ is smooth because it is somewhat involved.

Remark 1.140. As an aside, we note that the produced left-invariant vector fields parallelizes G after providing a basis of $\mathfrak{g}$; in particular, one has a canonical isomorphism $TG \cong G \times \mathfrak{g}$. One can actually show that TG is a Lie group with Lie group structure given by functoriality of the tangent bundle applied to the group operations of G, and one finds that $TG \cong G \rtimes \mathfrak{g}$, where G acts on $\mathfrak{g}$ by the adjoint action.

Next class we will go back and argue that our classical groups are actually Lie groups and compute their Lie algebras.

1.8 September 16

Today we will talk about Lie algebras of classical groups.

1.8.1 The Exponential Map: The Classical Case

Let's work through our examples by hand. Recall that our classical groups are our subgroups of $\operatorname{GL}_n(\mathbb{F})$ cut out by equations involving $\det$ and preserving a bilinear/sesquilinear form (symmetric, symplectic, or Hermitian).

Example 1.141. We show that $\operatorname{GL}_n(\mathbb{F})$ is a Lie group over $\mathbb{F}$ and compute its Lie algebra.

Proof. Note $\operatorname{GL}_n(\mathbb{F})$ is an open submanifold of $\operatorname{M}_n(\mathbb{F}) \cong \mathbb{F}^{n\times n}$. Matrix multiplication and inversion are rational functions of the coordinates and hence smooth, so $\operatorname{GL}_n(\mathbb{F})$ succeeds at being a Lie group. Lastly, we see that being open implies that our tangent space is

$$T_e \operatorname{GL}_n(\mathbb{F}) \cong T_e \operatorname{M}_n(\mathbb{F}) \cong \mathbb{F}^n,$$

as required. ■

We will postpone the remaining computations until we discuss the exponential. For these computations, we want the exponential map.

Definition 1.142 (exponential)**.** For $X \in \mathfrak{gl}_n(\mathbb{F})$, we define the exponential map $\exp\colon \mathfrak{gl}_n(\mathbb{F}) \to \operatorname{GL}_n(\mathbb{F})$ by

$$\exp(X) := \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

Note that $\exp$ is an isomorphism at the identity, so the Inverse function theorem provides a smooth "local" inverse $\log(1_n + X)$ defined in an open neighborhood of 1_n. In fact, one can formally compute that

$$\log(1_n + X) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{X^k}{k}.$$

We run a few small checks.

Remark 1.143. Note $\exp(0) = 1$. In fact, one can check that $d\exp_0(A) = A$ for any $A \in \mathfrak{gl}_n(\mathbb{F})$ by taking the derivative term by term.

Remark 1.144. We also see that $\exp\left(AXA^{-1}\right) = A\exp(X)A^{-1}$ and $\exp(X^\intercal) = \exp(X)^\intercal$ and $\exp(X^\dagger) = \exp(X)^\dagger$ by a direct expansion.

What's important about $\exp$ is the following multiplicative property.

Lemma 1.145. Fix $X, Y \in \mathfrak{gl}_n(\mathbb{F})$ which commute. Then

$$\exp(X + Y) = \exp(X)\exp(Y).$$

Proof. We check this in formal power series. Because everything in sight converges, this is safe. The main point is to just expand everything. Indeed,

$$\begin{aligned}\exp(X+Y) &= \sum_{k=0}^{\infty} \frac{(X+Y)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\left(\sum_{a+b=k} \binom{k}{a} X^aY^b\right) \\ &= \sum_{a,b=0}^{\infty} \frac{1}{(a+b)!} \cdot \frac{(a+b)!}{a!b!} X^aY^b \\ &= \sum_{a,b=0}^{\infty} \frac{X^a}{a!} \cdot \frac{Y^b}{b!} \\ &= \exp(X)\exp(Y),\end{aligned}$$

as required. ■

Remark 1.146. For fixed X, the previous point implies that the map $\mathbb{F} \to \mathrm{GL}_n(\mathbb{F})$ given by $t \mapsto \exp(tX)$ is a Lie group homomorphism. (Smoothness is automatic by smoothness of $\exp$.) The image of this map is called the "1-parameter subgroup" generated by X.

Remark 1.147. Taking inverses shows that $\log(XY) = \log X + \log Y$ for X and Y close enough to the identity.

Here is another check which is a little more interesting.

Lemma 1.148. Fix $X \in \mathfrak{gl}_n(\mathbb{F})$. Then

$$\det\exp(X) = \exp(\operatorname{tr} X).$$

Proof. The computations do not change if we extend the base field, so we may work over $\mathbb{C}$ everywhere. Thus, we may assume that X is upper-triangular by conjugating (see Remark 1.144) say with diagonal entries $\{d_1, \ldots, d_n\}$. Now, for any $k \geq 0$, any X^k continues to be upper-triangular with diagonal entries $\{d_1^k, \ldots, d_n^k\}$. Thus, we see that $\exp(X)$ is upper-triangular with diagonal entries $\{\exp(d_1), \ldots, \exp(d_n)\}$, so

$$\begin{aligned}\det\exp(X) &= \exp(d_1)\cdots\exp(d_n) &(1.1)\\ &= \exp(d_1 + \cdots + d_n) &(1.2)\\ &= \exp(\operatorname{tr} X), &(1.3)\end{aligned}$$

as required. ■

1.8.2 The Classical Groups

For our classical groups, we will show the following result.

Theorem 1.149. For each classical group $G \subseteq \mathrm{GL}_n(\mathbb{F})$, there will exist a vector subspace $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{F})$ (which can be identified with T_eG via the embedding $G \subseteq \mathrm{GL}_n(\mathbb{F})$) and open neighborhoods of the identity $U \subseteq \mathrm{GL}_n(\mathbb{F})$ and $\mathfrak{u} \subseteq \mathfrak{g}$ such that $\exp\colon (U \cap G) \to (\mathfrak{u} \cap \mathfrak{g})$ is a local isomorphism.

Before engaging with the examples, we note the following corollary.

Corollary 1.150. For each classical group G, we see that G is a Lie group with $T_eG = \mathfrak{g}$ and $\dim G = \dim \mathfrak{g}$.

Proof. It suffices to provide a slice chart of the identity for $G \subseteq \mathrm{GL}_n(\mathbb{F})$; we then get slice charts everywhere by translation. Well, ■

Let's now proceed with our examples. We begin with some general remarks.

Lemma 1.151. Let $G \subseteq \mathrm{GL}_n(\mathbb{F})$ be a closed Lie subgroup, and let $SG \coloneqq \{g \in G : \det g = 1\}$. We show SG is a Lie subgroup and compute $T_1SG \subseteq T_1G$ as

$$T_1SG = \{g \in T_1G : \operatorname{tr} g = 0\}.$$

Proof. Let G act on $\mathbb{F}$ by $\mu\colon G \times \mathbb{F} \to \mathbb{F}$ by $\mu(g, c) \coloneqq (\det g)c$. Note that μ is a polynomial and hence regular, so this is a regular action upon checking that $\mu(1, c) = c$ and $\mu(g, \mu(h, c)) = \mu(gh, c)$, which hold because $\det$ is a homomorphism.

Now, the stabilizer of $1 \in \mathbb{F}$ consists of the $g \in G$ such that $(\det g) \cdot 1 = 1$, which is equivalent to $\det g = 1$ and hence equivalent to $g \in SG$. Thus, $SG \subseteq G$ is a closed Lie subgroup with

$$T_1SG(\mathbb{F}) = \{v \in T_1G : (d\det)_1(v) = 0\},$$

where $\det\colon G \to \mathbb{F}$ is the determinant map. To compute $(d\det)_1(v)$, we identify $T_1G \subseteq T_1\mathrm{GL}_n(\mathbb{F}) = T_1M_n(\mathbb{F}) \cong M_n(\mathbb{F})$; then for any $X \in M_n(\mathbb{F})$, we note that the path $\gamma\colon \mathbb{R} \to M_n(\mathbb{R})$ defined by $\gamma(t) \coloneqq 1 + tX$ has $\gamma(0) = 1$ and $\gamma'(0) = X$, so

$$(d\det)_1(X) = (d\det)_1(\gamma'(0)) = (\det \circ \gamma)'(0) = \frac{d}{dt}\det(1 + tX)\bigg|_{t=0}.$$

Thus, we are interested in the linear terms of the polynomial $\det(1 + tX)$. Now, writing X out in coordinates as $X = [X_{ij}]_{1 \le i,j \le n}$ and setting $A_{ij} = 1_{i=j} + tX_{ij}$, we note

$$\det(1 + tX) = \det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}.$$

Now, the only way a summand can produce linear terms is if there is at most one non-diagonal entry A_{ij}, which of course forces all entries to be diagonal. Thus,

$$\frac{d}{dt}\det(1 + tX)\bigg|_{t=0} = \frac{d}{dt}(1 + tX_{11}) \cdots (1 + tX_{nn})\bigg|_{t=0} \overset{*}{=} (X_{11} + \cdots + X_{nn}) = \operatorname{tr} X,$$

where $\overset{*}{=}$ holds by an expansion of the terms looking for linear terms. Thus,

$$T_1SG = \{X \in T_1G : \operatorname{tr} X = 0\}.$$

■

Lemma 1.152. Let $J \in M_n(\mathbb{F})$ be some matrix, and let $(-)^*$ denote either of the involutions $(-)^\intercal$ or $(-)^\dagger$. Then one has the subgroup

$$O_J(\mathbb{F}) \coloneqq \{g \in \mathrm{GL}_n(\mathbb{F}) : g^*Jg = J\}.$$

We claim that $O_J(\mathbb{F}) \subseteq \mathrm{GL}_n(\mathbb{F})$ is a closed Lie subgroup (though if $(-)^* = (-)^\dagger$ and $\mathbb{F} = \mathbb{C}$, then $O_J(\mathbb{F})$ is a group over $\mathbb{R}$) and compute that

$$T_1O_J(\mathbb{F}) = \{X \in M_n(\mathbb{F}) : X^*J + JX = 0\}.$$

Proof. Indeed, let $\mathrm{GL}_n(\mathbb{F})$ act on $M_n(\mathbb{F})$ by $\mu(g, A) \coloneqq g^*Ag$. This (right!) action is polynomial and hence regular (with the previous parenthetical in mind), and we can check that it is an action because $\mu(1, A) = A$ and $\mu(g, \mu(h, A)) = g^*h^*Ahg = \mu(hg, A)$.

Now, the stabilizer of $J \in M_n(\mathbb{F})$ is precisely $O_J(\mathbb{F})$ by definition, so $O_J(\mathbb{F}) \subseteq \mathrm{GL}_n(\mathbb{F})$ is in fact a closed Lie subgroup. We also go ahead and compute $T_1O_J(\mathbb{F})$. Letting $f\colon \mathrm{GL}_n(\mathbb{F}) \to M_n(\mathbb{F})$ be defined by $f(g) \coloneqq g^*Jg$, we see that

$$T_1O_J(\mathbb{F}) = \ker df_1,$$

so we want to compute df_1. As usual, we identify $T_1G \subseteq T_1 \mathrm{GL}_n(\mathbb{F}) = T_1M_n(\mathbb{F}) \cong M_n(\mathbb{F})$; then for any $X \in M_n(\mathbb{F})$, we note that the path $\gamma\colon \mathbb{R} \to M_n(\mathbb{R})$ defined by $\gamma(t) \coloneqq 1 + tX$ has $\gamma(0) = 1$ and $\gamma'(0) = X$, so

$$df_1(X) = df_1(\gamma'(0)) = (f \circ \gamma)'(0) = \frac{d}{dt} f(1 + tX)\bigg|_{t=0}.$$

Thus, we go ahead and compute

$$f(1 + tX) = (1 + tX)^*J(1 + tX) = J + t(X^*J + JX) + t^2X^*JX,$$

so

$$df_1(X) = \frac{d}{dt} f(1 + tX)\bigg|_{t=0} = X^*J + JX.$$

Thus,

$$T_1O_J(\mathbb{F}) = \{X \in M_n(\mathbb{F}) : X^*J + JX = 0\},$$

as required. ∎

We now execute our computations in sequence.

(a) Using the preceding remarks, we see that

$$T_1 \mathrm{U}_{p,q}(\mathbb{C}) = \{X \in M_n(\mathbb{C}) : X^*B_{p,q} + B_{p,q}X = 0\},$$

where $B_{p,q} = \begin{bmatrix} 1_p & \\ & -1_q \end{bmatrix}$ is a diagonal matrix. We now continue as in (c). Set $X \coloneqq \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ to have the appropriate dimensions, and then we compute

$$\begin{aligned} X^*B_{p,q} + B_{p,q}X &= \begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} + \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \\ &= \begin{bmatrix} A^* & -C^* \\ B^* & -D^* \end{bmatrix} + \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} \\ &= \begin{bmatrix} A^* + A & B - C^* \\ B^* - C & -D^* - D \end{bmatrix}. \end{aligned}$$

In particular, this will vanish if and only if A and D are skew-Hermitian and $B = C^*$, so

$$\begin{aligned} T_1 \mathrm{U}_{p,q}(\mathbb{C}) &= \left\{ \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} : A \in M_p(\mathbb{C}), D \in M_q(\mathbb{C}), A = -A^*, D = -D^* \right\}, \\ T_1 \mathrm{U}_n(\mathbb{C}) &= \{A \in M_n(\mathbb{C}) : A = -A^*\}. \end{aligned}$$

Now, the space of $p \times p$ skew-Hermitian matrices A (namely, satisfying $A = -A^*$) is forced to have imaginary diagonal, and then the remaining entries are uniquely determined by their values strictly above the diagonal. Thus, the real dimension of this space is $p + p(p - 1) = p^2$. We conclude that

$$\begin{aligned} \dim_{\mathbb{R}} \mathrm{U}_{p,q}(\mathbb{C}) &= p^2 + 2pq + q^2 = n^2, \\ \dim_{\mathbb{R}} \mathrm{U}_n(\mathbb{C}) &= n^2. \end{aligned}$$

From here, we address SU by recalling that

$$T_1 \mathrm{SU}_{p,q}(\mathbb{R}) = \{X \in T_1 \mathrm{U}_{p,q}(\mathbb{C}) : \operatorname{tr} X\}.$$

In particular,

$$T_1 \operatorname{SU}_{p,q}(\mathbb{C}) = \left\{ \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} : A \in M_p(\mathbb{C}), D \in M_q(\mathbb{C}), A = -A^*, D = -D^*, \operatorname{tr} A + \operatorname{tr} D = 0 \right\},$$
$$T_1 \operatorname{SU}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A = -A^*, \operatorname{tr} A = 0\}.$$

Now, tr continues is real and actually surjects onto $\mathbb{R}$ for $n \geq 1$, even for our family of matrices above (for example, for any real number r, the matrix $\operatorname{diag}(r, 0, \ldots, 0)$ has trace r and lives in the above families). Thus, the kernel has dimension one smaller than the total space, giving

$$\dim_{\mathbb{R}} \operatorname{SU}_{p,q}(\mathbb{C}) = n^2 - 1,$$
$$\dim_{\mathbb{R}} \operatorname{SU}_n(\mathbb{C}) = n^2 - 1.$$

Example 1.153. We will show that

$$\operatorname{SL}_n(\mathbb{F}) \coloneqq \{A \in \operatorname{GL}_n(\mathbb{F}) : \det A = 1\}$$

is a Lie group over $\mathbb{F}$ and compute its Lie algebra to find $\dim_{\mathbb{F}} \operatorname{SL}_n(\mathbb{F}) = n^2 - 1$.

Proof. We use Lemma 1.151. We see that

$$T_1 \operatorname{SL}_n(\mathbb{F}) = \{X \in M_n(\mathbb{F}) : \operatorname{tr} X = 0\}.$$

(Note $T_1 \operatorname{GL}_n(\mathbb{F}) = T_1 M_n(\mathbb{F}) \cong M_n(\mathbb{F})$.) As such, we note that $\operatorname{tr}\colon M_n(\mathbb{F}) \to \mathbb{F}$ is surjective (for $n \geq 1$), so $\dim_{\mathbb{F}} T_1 \operatorname{SL}_n(\mathbb{F}) = \dim_{\mathbb{F}} \ker \operatorname{tr} = \dim_{\mathbb{F}} M_n(\mathbb{F}) - 1 = n^2 - 1$. ■

We now begin our computations for bilinear forms.

Example 1.154. Let $B \coloneqq 1_n$ be the standard bilinear form. We will show that

$$\operatorname{O}_n(\mathbb{F}) \coloneqq \{A \in \operatorname{GL}_n(\mathbb{F}) : ABA^\intercal B\}$$

is a Lie group over $\mathbb{F}$ and compute its Lie algebra to find $\dim_{\mathbb{F}} \operatorname{O}_n(\mathbb{F}) = \frac{1}{2}n(n-1)$.

Proof. We use Lemma 1.152. We see that

$$T_1 \operatorname{O}_n(\mathbb{F}) = \{X \in M_n(\mathbb{F}) : X^\intercal + X = 0\},$$

which is the space of alternating matrices. Thus, we see that the diagonal of $X \in T_1 \operatorname{O}_n(\mathbb{F})$ vanishes, and the remaining entries are determined by the values strictly above the diagonal, of which there are $\frac{1}{2}n(n-1)$. Thus, $\dim \operatorname{O}_n(\mathbb{F}) = \frac{1}{2}n(n-1)$. ■

Example 1.155. We will show that

$$\operatorname{SO}_n(\mathbb{F}) \coloneqq \{A \in \operatorname{O}_n(\mathbb{F}) : \det A = 1\}$$

is a Lie group over $\mathbb{F}$ and compute its Lie algebra to find $\dim_{\mathbb{F}} \operatorname{SO}_n(\mathbb{F}) = \frac{1}{2}n(n-1)$.

Proof. Using Lemma 1.151, we see that

$$T_1 \operatorname{SO}_n(\mathbb{F}) = \{X \in \operatorname{O}_n(\mathbb{F}) : \operatorname{tr} X = 0\}.$$

However, alternating matrices already have vanishing traces, so $T_1 \operatorname{SO}_n(\mathbb{F})$ is simply the full space of alternating matrices, giving $\dim_{\mathbb{F}} \operatorname{SO}_n(\mathbb{F}) = \frac{1}{2}n(n-1)$. ■

Over $\mathbb{R}$, there are more bilinear forms.

Example 1.156. Let $B_{p,q} \coloneqq 1_p \oplus 1_q$ where $n = p + q$. We will show that

$$\mathrm{O}_{p,q}(\mathbb{R}) \coloneqq \{A \in \mathrm{GL}_n(\mathbb{R}) : AB_{p,q}A^\intercal B_{p,q}\}$$

is a Lie group over $\mathbb{R}$ and compute its Lie algebra to find $\dim_{\mathbb{R}} \mathrm{O}_{p,q}(\mathbb{R}) = \frac{1}{2}n(n-1)$.

Proof. By Lemma 1.152, we see that

$$T_1 \mathrm{O}_{p,q}(\mathbb{R}) = \{X \in M_n(\mathbb{R}) : X^\intercal B_{p,q} + B_{p,q}X = 0\},$$

where $B_{p,q} = \left[\begin{smallmatrix} 1_p & \\ & -1_q \end{smallmatrix}\right]$ is a diagonal matrix. To compute the dimension of this space of matrices, we set $X \coloneqq \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ to have the appropriate dimensions, and then we compute

$$\begin{aligned} X^\intercal B_{p,q} + B_{p,q}X &= \begin{bmatrix} A^\intercal & C^\intercal \\ B^\intercal & D^\intercal \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} + \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \\ &= \begin{bmatrix} A^\intercal & -C^\intercal \\ B^\intercal & -D^\intercal \end{bmatrix} + \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} \\ &= \begin{bmatrix} A^\intercal + A & B - C^\intercal \\ B^\intercal - C & -D^\intercal - D \end{bmatrix}. \end{aligned}$$

In particular, this will vanish if and only if A and D are both alternating, and $B = C^\intercal$, yielding

$$T_1 \mathrm{O}_{p,q}(\mathbb{R}) = \left\{ \begin{bmatrix} A & B \\ B^\intercal & D \end{bmatrix} : A \in M_p(\mathbb{R}) \text{ and } D \in M_q(\mathbb{R}) \text{ are alternating} \right\}.$$

Thus, the dimension of our space is

$$\begin{aligned} \dim_{\mathbb{R}} \mathrm{O}_{p,q}(\mathbb{R}) &= \underbrace{\frac{1}{2}p(p-1)}_{A} + \underbrace{pq}_{B=C^\intercal} + \underbrace{\frac{1}{2}q(q-1)}_{D} \\ &= \frac{1}{2}\left(p^2 + 2pq + q^2 - p - q\right) \\ &= \frac{1}{2}(p+q)(p+q-1) \\ &= \frac{1}{2}n(n-1), \end{aligned}$$

where the dimension computations for (the spaces of) A and D are as in. ■

Example 1.157. We will show that

$$\mathrm{SO}_{p,q}(\mathbb{R}) \coloneqq \{A \in \mathrm{O}_{p,q}(\mathbb{F}) : \det A = 1\}$$

is a Lie group over $\mathbb{R}$ and compute its Lie algebra to find $\dim_{\mathbb{R}} \mathrm{SO}_{p,q}(\mathbb{R}) = \frac{1}{2}n(n-1)$.

Proof. We use Lemma 1.151, we note that

$$T_1 \mathrm{SO}_{p,q}(\mathbb{R}) : \{X \in T_1 \mathrm{O}_{p,q}(\mathbb{R}) : \operatorname{tr} X = 0\},$$

but our description of $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ has A and D alternating, so $\operatorname{tr} X = \operatorname{tr} A + \operatorname{tr} D = 0$. Thus, we see $T_1 \mathrm{SO}_{p,q}(\mathbb{R}) = T_1 \mathrm{O}_{p,q}(\mathbb{R})$, so the above description of tangent space and dimension go through. ■

Example 1.158. Let $\Omega_{2n} := \left[\begin{smallmatrix} 0_n & -1_n \\ 1_n & 0_n \end{smallmatrix}\right]$ be the standard symplectic form. We will show that

$$\operatorname{Sp}_{2n}(\mathbb{F}) := \{A \in \operatorname{GL}_n(\mathbb{F}) : A\Omega A^\intercal = A\}$$

is a Lie group over $\mathbb{F}$ and compute its Lie algebra to find $\dim_{\mathbb{F}} \operatorname{Sp}_{2n}(\mathbb{F}) = 2n^2 + n$.

Proof. By Lemma 1.152, we see that

$$T_1 \operatorname{Sp}_{2n}(\mathbb{F}) = \{X \in M_n(\mathbb{F}) : X^\intercal \Omega + \Omega X = 0\},$$

where $\Omega = \begin{bmatrix} & -1_n \\ 1_n & \end{bmatrix}$ is alternating. As usual, we set $X := \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ to have the appropriate dimensions, and we compute

$$\begin{aligned} X^\intercal \Omega + \Omega X &= \begin{bmatrix} A^\intercal & C^\intercal \\ B^\intercal & D^\intercal \end{bmatrix} \begin{bmatrix} & -1 \\ 1 & \end{bmatrix} + \begin{bmatrix} & -1 \\ 1 & \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \\ &= \begin{bmatrix} C^\intercal & -A^\intercal \\ D^\intercal & -B^\intercal \end{bmatrix} + \begin{bmatrix} -C & -D \\ A & B \end{bmatrix} \\ &= \begin{bmatrix} C^\intercal - C & -D - A^\intercal \\ A + D^\intercal & B - B^\intercal \end{bmatrix}. \end{aligned}$$

Thus, we see that

$$T_1 \operatorname{Sp}_{2n}(\mathbb{F}) = \left\{ \begin{bmatrix} A & B \\ C & -A^\intercal \end{bmatrix} : A, B, C \in M_n(\mathbb{F}), B = B^\intercal, C = C^\intercal \right\},$$

and our dimension is

$$\dim_{\mathbb{F}} \operatorname{Sp}_{2n}(\mathbb{F}) = \underbrace{n^2}_{A} + \underbrace{\frac{1}{2}n(n+1)}_{B} + \underbrace{\frac{1}{2}n(n+1)}_{C} = 2n^2 + n,$$

where we compute the dimension of space of symmetric matrices exactly analogously to the case of alternating matrices, except now the diagonal is permitted to be nonzero. ■

Lastly, we handle Hermitian forms.

Example 1.159. Let $B_{p,q} := 1_p \oplus 1_q$ where $n = p + q$. We will show that

$$\operatorname{U}_{p,q}(\mathbb{C}) := \left\{A \in \operatorname{GL}_n(\mathbb{F}) : AB_{p,q}A^\dagger B_{p,q}\right\}$$

is a Lie group over $\mathbb{R}$ and compute its Lie algebra to find $\dim_{\mathbb{R}} \operatorname{U}_{p,q}(\mathbb{C}) = n^2$.

Proof. Using Lemma 1.152, we see that

$$T_1 \operatorname{U}_{p,q}(\mathbb{C}) = \{X \in M_n(\mathbb{C}) : X^* B_{p,q} + B_{p,q} X = 0\},$$

where $B_{p,q} = \left[\begin{smallmatrix} 1_p & \\ & -1_q \end{smallmatrix}\right]$ is a diagonal matrix. We now continue as in (c). Set $X := \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ to have the appropriate dimensions, and then we compute

$$\begin{aligned} X^* B_{p,q} + B_{p,q} X &= \begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} + \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \\ &= \begin{bmatrix} A^* & -C^* \\ B^* & -D^* \end{bmatrix} + \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} \\ &= \begin{bmatrix} A^* + A & B - C^* \\ B^* - C & -D^* - D \end{bmatrix}. \end{aligned}$$

In particular, this will vanish if and only if A and D are skew-Hermitian and $B = C^*$, so

$$T_1 \operatorname{U}_{p,q}(\mathbb{C}) = \left\{ \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} : A \in M_p(\mathbb{C}), D \in M_q(\mathbb{C}), A = -A^*, D = -D^* \right\},$$
$$T_1 \operatorname{U}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A = -A^*\}.$$

Now, the space of $p \times p$ skew-Hermitian matrices A (namely, satisfying $A = -A^*$) is forced to have imaginary diagonal, and then the remaining entries are uniquely determined by their values strictly above the diagonal. Thus, the real dimension of this space is $p + p(p-1) = p^2$. We conclude that

$$\dim_{\mathbb{R}} \operatorname{U}_{p,q}(\mathbb{C}) = p^2 + 2pq + q^2 = n^2,$$

as required. ■

Example 1.160. We will show that

$$\operatorname{SU}_{p,q}(\mathbb{C}) \coloneqq \{A \in \operatorname{SU}_{p,q}(\mathbb{C}) : \det A = 1_n\}$$

is a Lie group over $\mathbb{R}$ and compute its Lie algebra to find $\dim_{\mathbb{R}} \operatorname{SU}_{p,q}(\mathbb{C}) = n^2 - 1$.

Proof. By Lemma 1.151, we see

$$T_1 \operatorname{SU}_{p,q}(\mathbb{R}) = \{X \in T_1 \operatorname{U}_{p,q}(\mathbb{C}) : \operatorname{tr} X\}.$$

In particular,

$$T_1 \operatorname{SU}_{p,q}(\mathbb{C}) = \left\{ \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} : A \in M_p(\mathbb{C}), D \in M_q(\mathbb{C}), A = -A^*, D = -D^*, \operatorname{tr} A + \operatorname{tr} D = 0 \right\},$$
$$T_1 \operatorname{SU}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A = -A^*, \operatorname{tr} A = 0\}.$$

Now, tr continues is real and actually surjects onto $\mathbb{R}$ for $n \geq 1$, even for our family of matrices above (for example, for any real number r, the matrix $\operatorname{diag}(r, 0, \ldots, 0)$ has trace r and lives in the above families). Thus, the kernel has dimension one smaller than the total space, giving

$$\dim_{\mathbb{R}} \operatorname{SU}_{p,q}(\mathbb{C}) = n^2 - 1,$$

as required. ■

THEME 2

PASSING TO LIE ALGEBRAS

It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out.

—Emil Artin

2.1 September 18

Today we compute our Lie algebras.

2.1.1 The Exponential Map: The General Case

Given a Lie group G with Lie algebra $\mathfrak{g}$, we would like to define an exponential map $\exp\colon \mathfrak{g} \to G$. Recall that $\exp$ gave rise to our homomorphisms $\gamma\colon \mathbb{R} \to G$ with $\gamma(0) = e$ and $\gamma'(0)$ is specified. This will be our starting point.

Proposition 2.1. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. For each $X \in \mathfrak{g}$, there exists a unique Lie group homomorphism $\gamma_X : \mathbb{R} \to G$ such that $\gamma_X'(0) = X$.

Proof. We use the theory of integral curves; see [Lee13, Chapter 9]. In particular, we see that we must satisfy $\gamma(s+t) = \gamma(t)\gamma(s)$ for all $s, t \in \mathbb{R}$, which yields

$$\gamma'(t) = \gamma(t)\gamma'(0),$$

where this multiplication really means $dL_{\gamma(t)}(\gamma'(0))$.

Thus, we see that we want to extend $X \in T_eG$ to a left-invariant vector field, and then we let $\gamma\colon \mathbb{R} \to G$ be the integral curve of this vector field satisfying $\gamma(0) = e$. (A priori, γ can only be defined in a neighborhood of the identity, but we can translate around in the group G to get a global solution. See [Lee13, Lemma 9.15] and in particular its corollary [Lee13, Theorem 9.18].) Then

$$\gamma'(t) = X(\gamma(t)) = dL_{\gamma(t)}(X(0)) = dL_{\gamma(t)}(X)$$

for each $t \in \mathbb{R}$.

Thus far we have shown that there is at most one Lie group homomorphism $\gamma_X\colon \mathbb{R} \to G$ satisfying $\gamma_X'(0) = X$; namely, it will be the above integral curve! It remains to check that the above integral curve actually satisfies $\gamma(t+s)=\gamma(t)\gamma(s)$. Well, for $s\in\mathbb{R}$, we define $\gamma_1(t)=\gamma(t+s)$ and $\gamma_2(t)=\gamma(s)\gamma(t)$. Then we see that γ_1 and γ_2 are both integral curves satisfying the ordinary differential equation

$$\widetilde{\gamma}'(t)=dL_{\widetilde{\gamma}(t)}(\widetilde{\gamma}'(0))$$

with initial condition $\widetilde{\gamma}(0)=\gamma(s)$, so the must be equal, completing the proof. ■

Remark 2.2. Here is one way to conclude without using [Lee13, Theorem 9.18]. The last paragraph of the proof provides a path $\gamma\colon(-\varepsilon,\varepsilon)\to G$ for some $\varepsilon>0$ satisfying the homomorphism property. But then any $N>0$ allows us to define $\widetilde{\gamma}\colon(-N,N)\to G$ given by

$$\widetilde{\gamma}(t)\coloneqq\gamma(t/N)^N.$$

However, we can check that $\widetilde{\gamma}$ satisfies $\widetilde{\gamma}'(t)=dL_{\widetilde{\gamma}(t)}(X)$ with initial condition $\widetilde{\gamma}(0)=e$, so $\widetilde{\gamma}$ extends γ. Thus, we can extend γ to $\bigcup_{N>0}(-N\varepsilon,N\varepsilon)=\mathbb{R}$.

We now define $\exp$ motivated by the classical case.

Definition 2.3 (exponential)**.** Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. For each $X\in\mathfrak{g}$, define γ_X via Proposition 2.1. Then we define $\exp_G\colon\mathfrak{g}\to G$ by

$$\exp_G(X)\coloneqq\gamma_X(1).$$

We will omit the subscript from $\exp_G$ as much as possible.

Example 2.4. If $G\subseteq\mathrm{GL}_n(\mathbb{F})$ is classical, we can take $\gamma_X(t)=\exp(tX)$ where $\exp$ is defined as for GL_n. Thus, $\exp(X)$ matches with the above definition.

Example 2.5. Consider the Lie group $\mathbb{R}^n$. Then for each $X\in T_0\mathbb{R}^n$, we identify $T_0\mathbb{R}^n\cong\mathbb{R}^n$ to observe that we can take $\gamma_X(t)\coloneqq tX$. Thus, $\exp(X)=X$.

Example 2.6. For any G, we can take $\gamma_0(t)\coloneqq 0$, so $\exp(0)=1$.

Example 2.7. We can directly compute that

$$(d\exp)_0(X)=\frac{d}{dt}\exp(tX)\bigg|_0\overset{*}{=}\frac{d}{dt}\gamma_X(t)\bigg|_{t=0}=X.$$

The equality $\exp(tX)\overset{*}{=}\gamma_X(t)$ is explained as follows: we can check that $\gamma_{rX}(t)=\gamma_X(rt)$ for any $r,t\in\mathbb{R}$ by computing the derivative at 0, so $\exp(tX)=\gamma_{tX}(1)=\gamma_X(t)$ follows.

Here are some quick checks.

Proposition 2.8. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Then $\exp\colon\mathfrak{g}\to G$ is regular and a local diffeomorphism.

Proof. Note that $\exp$ solves the differential equation given by Example 2.7, for which the theory of integral curves promises that this solution must be regular. Example 2.7 also tells us that $\exp$ is an isomorphism at the identity and hence a local diffeomorphism. ■

We would like to know something like $\exp(A+B)=\exp(A)\exp(B)$ when A and B commute, but one needs to be a little careful in how to state this. Here are some manifestations.

Proposition 2.9. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Then

$$\exp((s+t)X) = \exp(sX)\exp(tX)$$

for any $s,t \in \mathbb{R}$ and $X \in \mathfrak{g}$.

Proof. This is a matter of following the definitions around. Let $\gamma_X : \mathbb{R} \to G$ be the one-parameter family for X. Then we see that $\gamma_{rX}(t) = \gamma_X(rt)$ for any $r \in \mathbb{R}$ as explained in Example 2.7, so

$$\exp((s+t)X) = \gamma_{(s+t)X}(1) = \gamma_X(s+t) = \gamma_X(s)\gamma_X(t) = \exp(sX)\exp(tX),$$

as desired. ■

Proposition 2.10. Fix a homomorphism $\varphi\colon G \to H$ of Lie groups. Then

$$\varphi(\exp_G(X)) = \exp_H(d\varphi_0(X))$$

for any $X \in T_eG$.

Proof. This follows from the definition. In particular, we claim that

$$\gamma_{d\varphi_0(X)}(t) \stackrel{?}{=} \varphi(\gamma_X(t)).$$

To see this, note that $t \mapsto \varphi(\gamma_X(t))$ is a Lie group homomorphism $\mathbb{R} \to H$, and we can compute the derivative at 0 to be $d\varphi_0(\gamma_X'(0)) = d\varphi_0(X)$, as required. Plugging in $t=1$ to the above equation completes the proof. ■

Corollary 2.11. Fix homomorphisms $\varphi_1, \varphi_2\colon G \to H$ of Lie groups. Suppose G is connected. If $d\varphi_1 = d\varphi_2$, then $\varphi_1 = \varphi_2$.

Proof. Using Proposition 2.10, we see that

$$\varphi(\exp(X)) = \exp(d\varphi_0(X))$$

produces the same answer for $\varphi \in \{\varphi_1, \varphi_2\}$. However, $\exp$ is a local diffeomorphism by Proposition 2.8, so we have determined the values of φ_1 and φ_2 on the image of $\exp$, which must contain an open neighborhood of the identity of G. Thus, because G is connected, we see that G is generated by this open neighborhood, so in fact we have fully determined the values of φ_1 and φ_2. ■

Proposition 2.12. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$, and let $\operatorname{Ad}_\bullet\colon G \to \operatorname{GL}(\mathfrak{g})$ be the adjoint representation of Example 1.113. For any $g \in G$ and $X \in \mathfrak{g}$, we have

$$g\exp(X)g^{-1} = \exp(\operatorname{Ad}_g X).$$

Proof. By Proposition 2.10, we see that

$$g\exp(X)g^{-1} = \operatorname{Ad}_g(\exp(X)) = \exp((d\operatorname{Ad}_g)_eX) = \exp(\operatorname{Ad}_g X),$$

where the last equality holds by definition of the adjoint representation. (Yes, the notation is somewhat confusing.) ■

While we are here, we note that there is a logarithm map.

Definition 2.13 (logarithm)**.** Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Because $\exp$ is a local diffeomorphism, there are open neighborhoods $U \subseteq G$ and $\mathfrak{u} \subseteq \mathfrak{g}$ of the identities so that $\log\colon U \to \mathfrak{u}$ is an inverse for $\exp$.

2.1.2 The Commutator

Define the form $\mu\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$\mu(X,Y) := \log(\exp(X)\exp(Y)).$$

(Technically, μ is a priori only defined on an open neighborhood of the identity of $\mathfrak{g} \times \mathfrak{g}$.) Expanding out everything into coordinates, we see that μ has a Taylor series expansion as

$$\mu(X,Y) = c + \alpha_1(X) + \alpha_2(Y) + Q_1(X) + Q_2(Y) + \lambda(X,Y) + \cdots,$$

where c is constant, α_1 and α_2 are linear, Q_1 and Q_2 are quadratic, λ is bilinear, and $+\cdots$ denotes cubic and higher-order terms. However, we see that $\mu(X,0) = 0$ and $\mu(0,Y) = 0$ for any $X, Y \in \mathfrak{g}$, so $c = Q_1 = Q_2 = 0$ and $\alpha_1(X) = X$ and $\alpha_2(Y) = Y$. Further, we claim that λ is skew-symmetric: it is enough to show that $\lambda(X,X) = 0$, for which we note that

$$2X = \log(\exp(2X)) = \log(\exp(X)\exp(X)) = \mu(X,X) = X + X + \lambda(X,X) + \cdots,$$

so $\lambda(X,X) = 0$ is forced.

This λ allows us to define the Lie bracket on $\mathfrak{g}$ in a purely group-theoretic way.

Definition 2.14 (Lie bracket)**.** Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Then we define the *commutator* as the skew-symmetric form $\frac{1}{2}\lambda\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, denoted $[-,-]$. In particular, we see that

$$\exp(X)\exp(Y) = \exp\left(X + Y + \frac{1}{2}[X,Y] + \cdots\right), \tag{2.1}$$

where $+\cdots$ denotes higher-order terms (as usual).

Remark 2.15. A priori, the commutator may only be defined on an open neighborhood of the identity of $\mathfrak{g} \times \mathfrak{g}$, so (2.1) only holds (a priori) for sufficiently small X and Y. However, bilinearity allows us to scale our definition of $[-,-]$ from this open neighborhood everywhere.

Example 2.16. We compute the commutator map for GL_n. We see that

$$\exp(X)\exp(Y) = 1 + X + Y + XY + \frac{1}{2}\left(X^2 + Y^2\right) + \cdots, \tag{2.2}$$

$$\exp\left(X + Y + \frac{1}{2}[X+Y] + \cdots\right) = 1 + X + Y + \frac{1}{2}\left(X^2 + Y^2\right) + \frac{1}{2}XY + \frac{1}{2}YX + \frac{1}{2}[X,Y] + \cdots, \tag{2.3}$$

giving $[X,Y] = XY - YX$ subtracting.

To compute the commutator for the classical groups, we need to check some functoriality.

Proposition 2.17. Fix a homomorphism $\varphi\colon G \to H$ of Lie groups. For any $X, Y \in T_eG$, we have

$$d\varphi_0([X,Y]) = [d\varphi_0(X), d\varphi_0(Y)].$$

Proof. We unravel the definitions. Everything in sight is linear, so we may assume that X and Y are sufficiently small, so $d\varphi_0(X)$ and $d\varphi_0(Y)$ are sufficiently small. We now compute

$$\begin{aligned}\exp\left(d\varphi_0(X)+d\varphi_0(Y)+\frac{1}{2}[d\varphi_0(X),d\varphi_0(Y)]\right) &= \exp\left(d\varphi_0(X)\right)\exp\left(d\varphi_0(Y)\right)\\ &\overset{*}{=} \varphi(\exp(X))\varphi(\exp(Y))\\ &= \varphi(\exp(X)\exp(Y))\\ &= \varphi\left(\exp\left(X+Y+\frac{1}{2}[X,Y]+\cdots\right)\right)\\ &\overset{*}{=} \exp\left(d\varphi_0(X)+d\varphi_0(Y)+\frac{1}{2}d\varphi_0([X,Y])+\cdots\right),\end{aligned}$$

where we have used Proposition 2.10 at the equalities $\overset{*}{=}$. Because $\exp$ is a diffeomorphism for X and Y sufficiently small, the desired equality follows. ■

Example 2.18. The embedding $\mathrm{SL}_n(\mathbb{F})\to\mathrm{GL}_n(\mathbb{F})$ implies by Proposition 2.17 that the Lie bracket on the Lie algebra $\mathfrak{sl}_n$ can be computed by restricting the commutator Lie bracket on $\mathfrak{gl}_n$ (given by Example 2.16). In particular, we see that $\mathfrak{sl}_n$ is closed under taking commutators, which is not totally obvious a priori! A similar operation permits computation of the Lie bracket of a Lie group G whenever given an embedding $G\subseteq\mathrm{GL}_n$ (such as for the classical groups).

Corollary 2.19. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Let $\mathrm{Ad}_\bullet\colon G\to\mathrm{GL}(\mathfrak{g})$ denote the adjoint representation. For

$$\mathrm{Ad}_g([X,Y])=[\mathrm{Ad}_g(X),\mathrm{Ad}_g(Y)].$$

Proof. We simply apply Proposition 2.17 to $\mathrm{Ad}\colon G\to\mathrm{GL}(\mathfrak{g})$, which yields

$$(d\mathrm{Ad}_g)_e([X,Y])=[(d\mathrm{Ad}_g)_eX,(d\mathrm{Ad}_g)_eT],$$

which is the original equation after enough abuse of notation. ■

Proposition 2.20. Fix a Lie group G. For sufficiently small $X,Y\in T_eG$, we have

$$\exp(X)\exp(Y)\exp(X)^{-1}\exp(Y)^{-1}=\exp([X,Y]+\cdots).$$

Proof. This is a direct computation. We compute

$$\begin{aligned}\exp(X)\exp(Y)\exp(-X)\exp(-Y) &= \exp\left(X+Y+\frac{1}{2}[X,Y]+\cdots\right)\exp\left(-X-Y+\frac{1}{2}[X,Y]+\cdots\right)\\ &= \exp\left([X,Y]+\cdots\right),\end{aligned}$$

where we get some omitted cancellation of lower-order terms in the last equality (and there is a lot of higher-order terms). ■

Corollary 2.21. If G is abelian, then $[X,Y]=0$ for any X and Y.

Proof. It suffices to assume that X and Y are sufficiently small because the conclusion is linear. Now, Proposition 2.20 implies that

$$\exp([X,Y]+\cdots)=0,$$

so because $\exp$ is a diffeomorphism for small enough X and Y, so $[X,Y]=0$ follows. ■

2.2 September 20

Today we continue discussing the Lie bracket.

2.2.1 The Adjoint Action

Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Here is the standard example of a "Lie algebra representation."

Notation 2.22. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Note that the map $(d\mathrm{Ad}_g)_1 \colon G \to \mathrm{GL}(\mathfrak{g})$ is smooth, so we can consider the differential of this map, which we label $\mathrm{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Here are some checks on this map.

Proposition 2.23. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$.

(a) For $X, Y \in \mathfrak{g}$, we have $\mathrm{ad}_X(Y) = [X, Y]$.

(b) For $X \in \mathfrak{g}$, we have $\mathrm{Ad}_{\exp(X)} = \exp(\mathrm{ad}_X)$ as operators $\mathfrak{g} \to \mathfrak{g}$.

Proof. Here we go.

(a) By definition of our differential action, we have

$$(d\mathrm{Ad}_g)_1(Y) = \frac{d}{dt} g \exp(tY) g^{-1}\bigg|_{t=0}$$

for any $g \in G$ and $Y \in \mathfrak{g}$. We would like to compute a derivative of this map with respect to g (at the identity). As such, we plug in $g = \exp(sX)$ to compute

$$\begin{aligned} \mathrm{ad}_X(Y) &= \frac{d}{ds}(d\mathrm{Ad}_{\exp(sX)})_1(Y)\bigg|_{s=0} \\ &= \frac{d}{ds}\frac{d}{dt}\exp(sX)\exp(tY)\exp(-sX)\bigg|_{t=0}\bigg|_{s=0} \\ &\overset{*}{=} \frac{d}{ds}\frac{d}{dt}\exp(tY + st[X,Y] + \cdots)\bigg|_{t=0}\bigg|_{s=0}, \end{aligned}$$

where in $\overset{*}{=}$ we have used the definition of our bracket. Upon expanding out $\exp$ as a series, we see that the lower-order terms are $1 + tY + st[X, Y] + \cdots$ (everything higher is at least quadratic) for small enough s and t, so the derivative evaluates to $[X, Y]$.

(b) This follows immediately from Proposition 2.10 upon setting $\varphi = (d\mathrm{Ad}_\bullet)_1$. ■

Here is an example computation of what all this adjoint business looks like for GL_n, more directly than appealing to the bracket.

Lemma 2.24. Identify $T\,\mathrm{GL}_n(\mathbb{F})$ with $\mathrm{GL}_n(\mathbb{F}) \times \mathfrak{gl}_n(\mathbb{F})$ via left-invariant vector fields. For $X \in \mathfrak{gl}_n(\mathbb{F})$, we have

$$\begin{cases} dL_g(X) = gX, \\ dR_g(X) = Xg^{-1}, \\ d\,\mathrm{Ad}_g(X) = gXg^{-1}. \end{cases}$$

Proof. Set $G := \mathrm{GL}_n(\mathbb{F})$ and $\mathfrak{g} := \mathfrak{gl}_n(\mathbb{F})$. Note that the adjoint is the composite of L_g and R_g, so the last equality follows from the first two. For the first equality, we are computing the differential of the maps $L_g, R_g \colon G \to G$ at some $h \in G$. Well, L_g and R_g actually extend to perfectly fine linear maps $M_n(\mathbb{F}) \to M_n(\mathbb{F})$, and the differential of any linear map is simply itself upon identifying the tangent spaces of $M_n(\mathbb{F})$ with itself, so we conclude that $dL_g(X) = gX$ and $dR_g(X) = Xg^{-1}$, as required. ■

Lemma 2.25. Fix a homomorphism $\varphi \colon G \to H$ of Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ respectively. For any $g \in G$ and $X \in \mathfrak{g}$, we have

$$(d\mathrm{Ad}_{\varphi(g)})_e(d\varphi_1(X)) = d\varphi_1((d\mathrm{Ad}_g)_e(X)).$$

Proof. Simply take the differential (at 1) of the equation $\mathrm{Ad}_{\varphi(g)} \circ \varphi = \varphi \circ \mathrm{Ad}_g$, which is true because φ is a homomorphism. ■

Example 2.26. Given any embedding $G \subseteq \mathrm{GL}_n(\mathbb{F})$ with Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{F})$, we can use Lemma 2.25 to compute the adjoint action on $\mathfrak{g}$ by conjugation (via Lemma 2.24)!

Proposition 2.27. Let $(d\mathrm{Ad}_\bullet)_1 \colon \mathrm{GL}_n(\mathbb{F}) \to \mathrm{GL}(\mathfrak{gl}_n(\mathbb{F}))$ denote the adjoint representation. Then

$$\mathrm{ad}_X(Y) = XY - YX.$$

Proof. To parse the symbols, we note that $(d(d\mathrm{Ad}_\bullet)_1)_1 \colon \mathfrak{gl}_n(\mathbb{F}) \to \mathrm{End}(\mathfrak{gl}_n(\mathbb{F}))$, so the statement at least makes sense. Now, given $X \in \mathfrak{gl}_n(\mathbb{F})$, define $\gamma \colon \mathbb{F} \to M_n(\mathbb{F})$ by $\gamma(t) := 1 + tX$. Then $\gamma'(0) = X$. As such,

$$(d(d\mathrm{Ad}_\bullet)_1)_1(X) = (d(d\mathrm{Ad}_\bullet)_1)_1(\gamma'(0)) = ((d\mathrm{Ad}_\bullet)_1 \circ \gamma)'(0).$$

In particular, plugging in some $Y \in \mathfrak{gl}_n(\mathbb{F})$, we may use Lemma 2.24 to compute that

$$\begin{aligned}
\frac{d}{dt}((d\mathrm{Ad}_\bullet)_1 \circ \gamma)(t)(Y)\bigg|_{t=0} &= \frac{d}{dt}(d\mathrm{Ad}_{1+tX})_1(Y)\bigg|_{t=0} \\
&= \frac{d}{dt}(1+tX)Y(1+tX)^{-1}\bigg|_{t=0} \\
&= \frac{d}{dt}(1+tX)Y\left(1 - tX + t^2X^2 + \cdots\right)\bigg|_{t=0} \\
&= XY - YX,
\end{aligned}$$

where the series expansion takes t small enough for the series to converge. (For example, one can take t small enough so that all eigenvalues of tX are less than 1.) ■

Example 2.28. Given any embedding $G \subseteq \mathrm{GL}_n(\mathbb{F})$ with Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{F})$, we note that the action of G on $\mathfrak{g}$ actually extends to an action of G on $\mathfrak{gl}_n(\mathbb{F})$ (still by conjugation) which happens to stabilize $\mathfrak{g}$. Then the action $G \to \mathrm{GL}(\mathfrak{gl}_n(\mathbb{F}))$ is a restriction of the adjoint action $\mathrm{GL}_n(\mathbb{F}) \to \mathrm{GL}(\mathfrak{gl}_n(\mathbb{F}))$ given by conjugation still, whose differential action $\mathfrak{gl}_n(\mathbb{F}) \to \mathfrak{gl}(\mathfrak{gl}_n(\mathbb{F}))$ we computed above to be given by $\mathrm{ad}_X \colon Y \mapsto XY - YX$. This restricts back to the subspace $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{F})$ (via the inclusion $G \subseteq \mathrm{GL}_n(\mathbb{F})$), where we know that the action must happen to stabilize $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{F})$. The point is that we have computed our adjoint representation $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is given by the commutator. (Alternatively, one can redo the computation of the above proof.)

2.2.2 Lie Algebras

Here is a standard consequence of this theory.

Proposition 2.29 (Jacobi identity)**.** Fix a Lie group G with Lie algebra $\mathfrak{g}$. Then we have the Jacobi identity

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

Proof. Doing some rearranging with Proposition 2.23 (and the skew-symmetry), we see that this is equivalent to plugging Z into the identity

$$\operatorname{ad}_{[X,Y]}\stackrel{?}{=}\operatorname{ad}_X\circ\operatorname{ad}_Y-\operatorname{ad}_Y\circ\operatorname{ad}_X.$$

To verify this, we note that the right-hand side is $[\operatorname{ad}_X,\operatorname{ad}_Y]$, where the commutator is taken in $\mathfrak{gl}(\mathfrak{g})$. Thus, we are trying to show that the adjoint preserves a commutator, which we do as follows: recall that $\operatorname{Ad}_\bullet\colon G\to\operatorname{GL}(\mathfrak{g})$ is a morphism of Lie groups, meaning that the differential map ad preserves the commutator by Proposition 2.17. ■

The Jacobi identity is important enough to earn the following definition.

Definition 2.30 (Lie algebra)**.** Fix a field F. Then a *Lie algebra* is an F-vector space $\mathfrak{g}$ equipped with a bilinear form $[-,-]\colon\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ satisfying the following.

(a) Skew-symmetric: $[X,X]=0$ for all $X\in\mathfrak{g}$.

(b) Jacobi identity: for any $X,Y,Z\in\mathfrak{g}$, we have

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

A morphism of Lie algebras is an F-linear morphism preserving the forms.

Definition 2.31 (commutative)**.** A Lie algebra $\mathfrak{g}$ is *commutative* if and only if $[X,Y]=0$ for all $X,Y\in\mathfrak{g}$.

Example 2.32. For any F-algebra A, we produce a Lie bracket on A given by

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

This map is of course linear in both X and Y (because multiplication is F-linear in an F-algebra), and $[X,X]=X^2-X^2=0$. Lastly, to see the Jacobi identity, we expand:

$$\begin{aligned}[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]&=[X,YZ-ZY]+[Y,ZX-XZ]+[Z,XY-YX]\\&=X(YZ-ZY)-(YZ-ZY)X\\&\qquad+Y(ZX-XZ)-(ZX-XZ)Y\\&\qquad+Z(XY-YX)-(XY-YX)Z\\&=0.\end{aligned}$$

For example, one can take A to be $\operatorname{End}_F(V)$ for some F-vector space V; this produces the Lie algebra $\mathfrak{gl}(V)$.

Example 2.33. Given a regular Lie group G, the tangent space at the identity $\mathfrak{g}$ is a Lie algebra according to the above definition.

The above example upgrades into a functor.

Proposition 2.34. Fix a regular Lie group G. For any morphism of Lie groups $\varphi\colon G_1 \to G_2$, the differential $d\varphi_e\colon T_eG_1 \to T_eG_2$ is a (functorial) morphism of Lie algebras. In fact, if G_1 is connected, the induced map

$$\mathrm{Hom}_{\mathrm{LieGrp}}(G_1, G_2) \to \mathrm{Hom}_{\mathrm{Lie}(k)}(T_eG_1, T_eG_2)$$

is injective. In other words, the functor $G \to T_eG$ from connected Lie groups to Lie algebras is faithful.

Proof. The differential being a homomorphism of Lie algebras follows from Proposition 2.17. Functoriality follows from the corresponding functoriality for differentials of more general smooth maps. The injectivity follows from Corollary 2.11. ■

Remark 2.35. It turns out that the functor above is also full, though we are not in a position to show this yet.

2.2.3 Subalgebras

Lie algebras are interesting enough to study on their own right, but we now note that we have sufficient motivation from Proposition 2.34.

Definition 2.36 (subalgebra, ideal)**.** Fix a Lie algebra $\mathfrak{g}$.

- A *Lie subalgebra* $\mathfrak{h} \subseteq \mathfrak{g}$ is a subspace closed under the Lie bracket of $\mathfrak{g}$; note that $\mathfrak{h}$ continues to be a Lie algebra.
- A *Lie ideal* is a subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ with the stronger property that

  $$[X, Y] \in \mathfrak{h}$$

  for any $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$.

Definition 2.37 (representation)**.** A *representation* of a Lie algebra $\mathfrak{g}$ over a field F is a morphism $\mathfrak{g} \to \mathfrak{gl}(V)$ for some (finite-dimensional) vector space V over F. The representation is *faithful* if and only if the morphism $\mathfrak{g} \to \mathfrak{gl}(V)$ is injective.

Here is how these things relate back to Lie groups.

Proposition 2.38. Fix a regular Lie subgroup H of a regular Lie group G. Let their Lie algebras be $\mathfrak{h}$ and $\mathfrak{g}$, respectively.

(a) Then $\mathfrak{h} \subseteq \mathfrak{g}$ is a Lie subalgebra.

(b) If H is normal in G, then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

(c) If G and H are connected, and $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then H is normal in G.

Proof. Here we go.

(a) Certainly $\mathfrak{h} \subseteq \mathfrak{g}$ is a subspace, so we want to check that $[X, Y] \in \mathfrak{h}$ for $X, Y \in \mathfrak{h}$, where the target bracket is taken in $\mathfrak{g}$. Consider the embedding $\varphi\colon H \to G$ so that $\mathfrak{h} = \operatorname{im} d\varphi_0$. Thus, we use Proposition 2.17 to see that

$$d\varphi_0([X, Y]) = [d\varphi_0(X), d\varphi_0(Y)].$$

Thus, for any $X, Y \in \operatorname{im} d\varphi_0$, we see that $[X, Y] \in \operatorname{im} d\varphi_0$, as required.

(b) For any $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, we want to check that $[X,Y] \in \mathfrak{h}$. By Proposition 2.23, we are asking to check that $\operatorname{ad}_X(Y) \in \mathfrak{h}$. Well, for any $g \in G$, we see that $gHg^{-1} \subseteq H$, so the adjoint $\operatorname{Ad}_g \colon G \to G$ restricts to $\operatorname{Ad}_g \colon H \to H$. In particular, by taking the differential, we see that the adjoint $(d\operatorname{Ad}_\bullet)_1 \colon G \to \operatorname{GL}(\mathfrak{g})$ restricts to $(d\operatorname{Ad}_\bullet)_1 \colon G \to \operatorname{GL}(\mathfrak{h})$. (Namely, $(d\operatorname{Ad}_g)_1(Y) \in \mathfrak{h}$ for any $Y \in \mathfrak{h}$.) Taking the differential of this, we see that we get our map $\operatorname{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{h})$, meaning that $\operatorname{ad}_X(Y) \in \mathfrak{h}$ for any $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$.

(c) Recall from Proposition 2.23 that

$$\operatorname{Ad}_{\exp(X)}(Y) = \exp(\operatorname{ad}_X Y).$$

Thus, for any $X \in \mathfrak{g}$, we see that $\operatorname{Ad}_{\exp(X)}$ is an operator $\mathfrak{h} \to \mathfrak{h}$. Thus, for $g \in G$ sufficiently close to the identity, we see that $\operatorname{Ad}_g(Y) \in \mathfrak{h}$ for $Y \in \mathfrak{h}$. Taking the exponential, Proposition 2.12 tells us that $ghg^{-1} \in H$ for $g \in G$ and $h \in H$ both sufficiently close to the identity.

Concretely, we get an open neighborhood U of the identity of G such that $ghg^{-1} \in H$ for any $g \in U$ and $h \in H \cap U$. Now, the subset of G normalizing $U \cap H$ is a subgroup of G containing U, so we see that it must be all of G because G is connected. Then the subset of H normalized by G is again a subgroup of H containing $U \cap H$, so we see that it must be all of H because H is connected. Thus, H is normal in G. ■

Here is some motivation for our definition of ideal.

Lemma 2.39. Fix a morphism $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ of Lie algebras.

(a) The kernel $\ker \varphi \subseteq \mathfrak{g}$ is a Lie ideal.

(b) The image $\operatorname{im} \varphi \subseteq \mathfrak{h}$ is a Lie subalgebra.

Proof. Here we go.

(a) For any $X \in \ker \varphi$ and $Y \in \mathfrak{g}$, we need to check that $[X,Y] \in \ker \varphi$. Well,

$$\varphi([X,Y]) = [\varphi(X), Y] = [0, Y] = 0$$

by the bilinearity of $[-,-]$.

(b) For any $X, Y \in \operatorname{im} \varphi$, we must check that $[X,Y] \in \operatorname{im} \varphi$. Well, find $X_0, Y_0 \in \mathfrak{g}$ such that $X = \varphi(X_0)$ and $Y = \varphi(Y_0)$, and then we see that

$$[X,Y] = [\varphi(X_0), \varphi(Y_0)] = \varphi([X_0, Y_0])$$

is in the image of φ, as required. ■

Here are some more ways to build Lie ideals.

Remark 2.40. Fix a collection $\{\mathfrak{g}_\alpha\}_{\alpha \in \kappa}$ of Lie ideals of $\mathfrak{g}$. Then we claim that the intersection $\bigcap_{\alpha \in \kappa} \mathfrak{g}_\alpha$ is still a Lie ideal of $\mathfrak{g}$. Indeed, for any $X \in \bigcap_{\alpha \in \kappa} \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}$, we see that $X \in \mathfrak{g}_\alpha$ and hence $[X,Y] \in \mathfrak{g}_\alpha$ for all $\alpha \in \kappa$; thus, $[X,Y] \in \bigcap_{\alpha \in \kappa} \mathfrak{g}_\alpha$.

Remark 2.41. For two Lie ideals I and J of a Lie algebra $\mathfrak{g}$, we claim that

$$[I,J] \coloneqq \operatorname{span}\{[X,Y] : X \in I, Y \in J\}$$

is also a Lie ideal of $\mathfrak{g}$. Indeed, this is certainly a subspace (because it is a span). To check that $[\mathfrak{g}, [I,J]] \subseteq [I,J]$, we note that it is enough to check this for a spanning subset of I, so we pick up $Z \in \mathfrak{g}$ and $[X,Y] \in [I,J]$ and compute

$$[Z,[X,Y]] = -[X,[Y,Z]] - [[X,Z],Y] \in [I,J]$$

by the Jacobi identity, so we are done.

Lemma 2.42. Fix a Lie ideal $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$. Then the quotient space $\mathfrak{g}/\mathfrak{h}$ is a Lie algebra with bracket given by

$$[X+\mathfrak{h}, Y+\mathfrak{h}]_{\mathfrak{g}/\mathfrak{h}} := [X,Y]_{\mathfrak{g}} + \mathfrak{h}.$$

Proof. The main issue is checking that the bracket is well-defined. Well, if $X, Y \in \mathfrak{g}$ and $X', Y' \in \mathfrak{h}$, we must check that

$$[X+X', Y+Y'] + \mathfrak{h} \stackrel{?}{=} [X,Y] + \mathfrak{h},$$

where the bracket is taken in $\mathfrak{g}$. This is a matter of expanding with the bilinearity: note

$$\begin{aligned}[X+X', Y+Y'] &= [X+X', Y] + [X+X', Y'] \\ &= [X,Y] + [X',Y] + [X,Y'] + [X',Y'],\end{aligned}$$

and now we see that the last three terms live in $\mathfrak{h}$ because $\mathfrak{h} \subseteq \mathfrak{g}$ is an ideal.

Now, note that we have a canonical surjective linear map $\pi\colon \mathfrak{g} \twoheadrightarrow \mathfrak{g}/\mathfrak{h}$ which satisfies

$$\pi([X,Y]) = [\pi(X), \pi(Y)].$$

Thus, the bilinearity, skew-symmetry, and Jacobi identity for $\mathfrak{g}/\mathfrak{h}$ are immediately inherited from the corresponding checks on $\mathfrak{g}$. Rigorously, perhaps one should note that (for example) the Jacobi identity corresponds to showing that some linear functional on $(\mathfrak{g}/\mathfrak{h})^3$ vanishes; however, this linear functional can be checked to vanish on the level of $\mathfrak{g}^3$. ■

Proposition 2.43. Fix a morphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ of Lie algebras. Then the induced quotient map

$$\overline{\varphi}\colon \mathfrak{g}/\ker\varphi \to \operatorname{im}\varphi$$

is an isomorphism.

Proof. Linear algebra implies that $\overline{\varphi}$ is already an isomorphism of vector spaces. Thus, it merely remains to check that $\overline{\varphi}$ is a morphism of Lie algebras. Well, for $X, Y \in \mathfrak{g} =$, we see

$$\begin{aligned}\overline{\varphi}([X+\ker\varphi, Y+\ker\varphi]) &= \overline{\varphi}([X,Y]+\ker\varphi) \\ &= \varphi([X,Y]) + \ker\varphi \\ &= [\varphi(X), \varphi(Y)] + \ker\varphi \\ &= [\overline{\varphi}(X), \overline{\varphi}(Y)],\end{aligned}$$

as required. ■

2.2.4 Lie Algebra of a Vector Field

One can in general provide a Lie algebra of a vector field. Fix a regular vector field $\xi\colon X \to TX$ on a regular manifold X. For any regular function f on an open subset $U \subseteq X$, we may define

$$(\xi f)(x) := \xi_x(f_x),$$

where we recall that $\xi_x \in T_xX$ is some derivation which outputs a number when fed a germ f_x. The point is that ξf is itself a regular function $X \to \mathbb{F}$! We are now able to define a bracket.

Proposition 2.44. Fix a regular manifold X. Given vector fields $\xi, \eta\colon X \to TX$, we define the Lie bracket

$$[\xi, \eta] := \xi\eta - \eta\xi.$$

Then $[-,-]$ is a Lie bracket.

Proof. At each $x \in X$, we have certainly defined a map taking regular functions f on X and outputting an element of $\mathbb{F}$ given by

$$[\xi,\eta]_x(f) \coloneqq \xi_x(\eta f)_x - \eta_x(\xi f)_x.$$

This is certainly linear in f because ξ and η are. Further, the value of $[\xi,\eta]_x(f)$ only depends on the germ f_x because having $f_x = g_x$ for functions f and g implies $(f-g)_x = 0_x$, and then $\eta(f-g)$ and $\xi(f-g)$ both vanish in a neighborhood of x, so $[\xi,\eta]_x(f-g) = 0$.

It remains to check the product rule. Well, for regular functions f and g and some $y \in X$, we compute

$$(\eta fg)(y) = \eta_y(f_y g_y) = f(y)\eta_y(g_y) + g(y)\eta_y(g_y) = (f\cdot\eta g + g\cdot\eta f)(y),$$

and a similar computation works for ξ. Thus,

$$\begin{aligned}\xi(\eta fg)(x) &= \xi(f\eta g + g\eta f)(x)\\ &= \xi(f\eta g)(x) + \xi(g\eta f)(x)\\ &= f(x)\xi(\eta g)(x) + (\xi f)(x)(\eta g)(x) + g(x)\xi(\eta f)(x) + (\xi g)(x)(\eta f)(x),\end{aligned}$$

and a similar computation holds for $\eta(\xi fg)(x)$. Thus, we see that

$$\begin{aligned}[\xi,\eta]_x(fg) &= f(x)\xi(\eta g)(x) + (\xi f)(x)(\eta g)(x) + g(x)\xi(\eta f)(x) + (\xi g)(x)(\eta f)(x)\\ &\qquad - (f(x)\eta(\xi g)(x) + (\eta f)(x)(\xi g)(x) + g(x)\eta(\xi f)(x) + (\eta g)(x)(\xi f)(x))\\ &= f(x)[\xi,\eta]_x g + g(x)[\xi,\eta]_x f\end{aligned}$$

after sufficient cancellation and rearranging. ■

Example 2.45. Fix regular functions f and g on some open subset of $U \subseteq \mathbb{R}^m$, and let x_i and x_j be two coordinates. Then we compute

$$\left[f\frac{\partial}{\partial x_i}, g\frac{\partial}{\partial x_j}\right] = f\frac{\partial g}{\partial x_i}\frac{\partial}{\partial x_j} - g\frac{\partial f}{\partial x_j}\frac{\partial}{\partial x_i}.$$

Proof. Fixing some $p \in U$ and regular germ h, we see

$$\begin{aligned}\left[f\frac{\partial}{\partial x_i}, g\frac{\partial}{\partial x_j}\right]_p(h) &= f(p)\frac{\partial}{\partial x_i}g\frac{\partial h}{\partial x_j}\bigg|_p - g(p)\frac{\partial}{\partial x_j}f\frac{\partial h}{\partial x_i}\bigg|_p\\ &= f(p)\frac{\partial g}{\partial x_i}\bigg|_p\frac{\partial h}{\partial x_j}\bigg|_p + f(p)g(p)\frac{\partial h}{\partial x_i\partial x_j}\bigg|_p - g(p)\frac{\partial f}{\partial x_j}\bigg|_p\frac{\partial h}{\partial x_i}\bigg|_p - f(p)g(p)\frac{\partial h}{\partial x_i\partial x_j}\bigg|_p\\ &= f(p)\frac{\partial g}{\partial x_i}\bigg|_p\frac{\partial h}{\partial x_j}\bigg|_p - g(p)\frac{\partial f}{\partial x_j}\bigg|_p\frac{\partial h}{\partial x_i}\bigg|_p,\end{aligned}$$

as required. ■

Remark 2.46. In local coordinates in some chart (U, φ) with $\varphi = (x_1, \ldots, x_m)$ of our regular manifold M, one can write vector fields as

$$\xi = \sum_{i=1}^{m} a_i \frac{\partial}{\partial x_i} \qquad \text{and} \qquad \eta = \sum_{i=1}^{m} b_i \frac{\partial}{\partial x_i},$$

where a_i and b_i are regular functions. Then one can expand the bilinearity to see that

$$[\xi, \eta] = \sum_{i,j=1}^{m} \left(a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j} \right) \frac{\partial}{\partial x_i}.$$

Indeed, after applying bilinearity, the main point is to compute $\left[f \frac{\partial}{\partial x}, g \frac{\partial}{\partial y} \right]$ for regular functions f and g and coordinates x and y, which we did in the previous example.

Remark 2.47. For example, if ξ and η are tangent to a regular submanifold $N \subseteq M$ of dimension k, then $[\xi, \eta]$ continues to be tangent. One can check this using a local slice chart, where the condition that ξ is tangent to Y is equivalent to having $a_i = 0$ for $i > k$. Combining this with the computation of the previous remark completes the argument.

2.3 September 23

Today we continue talking about vector fields.

2.3.1 Vector Fields on Lie Groups

Let's return to Lie groups.

Lemma 2.48. Fix a regular Lie group G. A vector field ξ on G is left-invariant if and only if

$$\xi(f \circ L_g) = \xi f \circ L_g$$

for any germ f defined in a neighborhood of g.

Proof. We show the two implications separately.

- If ξ is left-invariant, then $\xi_{gh} = (dL_g)_h(\xi_h)$ for any $g, h \in G$. Thus, for any $h \in G$, we see that

$$\begin{aligned}(\xi f \circ L_g)(h) &= \xi_{gh} f \\ &= ((dL_g)_h \xi_h) f \\ &= \xi_h(f \circ L_g),\end{aligned}$$

  as required.

- Suppose $\xi(f \circ L_g) = \xi f \circ L_g$ for any f. Then plugging in the identity tells us that

$$\xi_g f = (\xi f \circ L_g)(e) = \xi_e(f \circ L_g) = ((dL_g)_e(\xi_e))(f).$$

  Thus, $\xi_g = (dL_g)_e \xi_e$, as required. ■

Lemma 2.49. Fix a left-invariant vector field ξ on a regular Lie group G. Then for a germ f at a point $g \in G$, one has
$$\xi_g f = \frac{d}{dt} f(g\exp(t\xi_e))\bigg|_{t=0}.$$

Proof. This is more or less the chain rule. For our $g \in G$, Lemma 2.48 tells us that
$$\xi_g f = \xi_e(f \circ L_g).$$
Now, the path $\gamma\colon \mathbb{F} \to G$ given by $\gamma(t) := \exp(t\xi_e)$ has $\gamma'(0) = \xi_e$, so
$$\xi_e(f \circ L_g) = d(f \circ L_g \circ \gamma)'(0) = \frac{d}{dt} f(g\exp(t\xi_e))\bigg|_{t=0},$$
as required. ■

Proposition 2.50. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Then the collection of left-invariant vector fields $\operatorname{Vect}^L(G)$ is a Lie subalgebra of $\operatorname{Vect}(G)$ which is isomorphic to $\mathfrak{g}$.

Proof. By Remark 1.139, one certainly has an isomorphism $\operatorname{Vect}^L(G) \to \mathfrak{g}$ given by $\xi \mapsto \xi_e$, with inverse given by $X \mapsto \xi_X$, where ξ_X is the vector field $\xi_X(g) := dL_g(X)$. Now, by Lemma 2.48, ξ is left-invariant if and only if
$$\xi(f \circ L_g) = \xi f \circ L_g$$
for any germ f defined in a neighborhood of g. Thus, we see that $\operatorname{Vect}^L(G)$ is preserved by the commutator of $\operatorname{Vect}(G)$.

It remains to check that our isomorphism with $\mathfrak{g}$ is a morphism of Lie algebras. Fix $X, Y \in \mathfrak{g}$, and we would like to show that $[\xi_X, \xi_Y] = \xi_{[X,Y]}$. It is enough to check this equality after mapping back down to $\mathfrak{g}$, so we want to check that $[\xi_X, \xi_Y]_e = [X, Y]$. This is a direct computation: by Lemma 2.49, any germ f at e has
$$\begin{aligned}
[\xi_X, \xi_Y]_e f &= \frac{d}{dt}\left(\xi_Y f(\exp(tX)) - \xi_X f(\exp(tY))\right)\bigg|_{t=0} \\
&= \frac{\partial^2}{\partial s \partial t}\frac{d}{ds}\left(f(\exp(tX)\exp(sY)) - f(\exp(tY)\exp(sX))\right)\bigg|_{(s,t)=(0,0)} \\
&= \frac{\partial^2}{\partial s \partial t}\left(f\exp\left(tX + sY + \frac{1}{2}st[X,Y] + \cdots\right) - f\exp\left(tX + sY - \frac{1}{2}st[X,Y] + \cdots\right)\right)\bigg|_{(s,t)=(0,0)}.
\end{aligned}$$
Now, one can imagine taking a Taylor series expansion of $f \circ \exp\colon \mathfrak{g} \to \mathbb{R}$ in terms of Z, in which we see that the above derivative will only depend on the st term of the relevant expansion. More precisely, write $(f \circ \exp)(Z) = f(e) + \lambda(Z) + Q(Z) + C(Z)$, where λ is linear, Q is quadratic, and C has vanishing first- and second-order derivatives. Then, after cancellation within λ, we see that
$$\begin{aligned}
[\xi_X, \xi_Y]_e f = \frac{\partial^2}{\partial s \partial t} st\lambda(st[X,Y])\bigg|_{(s,t)=(0,0)} \\
+ \frac{\partial^2}{\partial s \partial t} Q\left(tX + sY + \frac{1}{2}st[X,Y] + \cdots\right)\bigg|_{(s,t)=(0,0)} \\
+ \frac{\partial^2}{\partial s \partial t} Q\left(tX + sY - \frac{1}{2}st[X,Y] + \cdots\right) + \cdots\bigg|_{(s,t)=(0,0)},
\end{aligned}$$

where $+\cdots$ denotes higher-order terms which will not affect the current derivative (for example, containing C). Now, the linear terms inside Q will produce cancelling terms after expansion, so the only term we are left to care about is

$$[\xi_X, \xi_Y]_e f = \lambda([X,Y]) = \frac{d}{dt}(f \circ \exp)(t[X,Y])\bigg|_{t=0} = (\xi_{[X,Y]})_e f,$$

as required. ■

2.3.2 Group Actions via Lie Algebras

In general, if G acts on a regular manifold M via the action $a\colon G \times X \to X$, one can define an action of $\mathfrak{g}$ on $\operatorname{Vect}(X)$ by analogy with Lemma 2.49.

Definition 2.51. Fix a left action $a\colon G \times M \to M$ of a regular Lie group G on a regular manifold M. Then we define $a_*\colon \mathfrak{g} \to \operatorname{Vect}(M)$ by

$$(a_*X)_p f := \frac{d}{dt} f(a(\exp(-tX), p))\bigg|_{t=0}$$

for any $p \in M$ and germ f at p.

Remark 2.52. Let's explain the sign in the above definition: the action of G on M induces a natural action of G on the regular functions $\mathcal{O}(M)$ by $(g \cdot f)(p) := f\left(g^{-1} \cdot p\right)$. It is this action of G on $\mathcal{O}(M)$ which motivates the above definition.

Let's run our checks on this definition.

Lemma 2.53. Let $a\colon G \times M \to M$ be an action of a regular Lie group G on a regular manifold M. Then that a_* is a homomorphism of Lie algebras.

Proof. We run our many checks in sequence. Throughout, $p, q \in M$ and $X, Y \in \mathfrak{g}$ and $s, t \in \mathbb{F}$ and f and g are regular functions on an open neighborhood of p.

1. For any regular function f defined in an open neighborhood of a point $p \in M$, we claim that

$$da_{(e,p)}(-X, 0)(f_p) \stackrel{?}{=} (a_*X)_p(f).$$

This is a matter of computation. Define $\gamma\colon \mathbb{F} \to G \times M$ by $\gamma(t) := (\exp(-tX), p)$. Then we see that $\gamma'(0) = (-X, 0)$ by definition of $\exp$. Thus, using the chain rule, we see that

$$\begin{aligned} da_{(e,p)}(-X,0)(f_p) &= da_{(e,p)}(-X,0)(f_p) \\ &= d(f \circ a)_{(e,p)}(-X,0) \\ &= d(f \circ a)_{(e,p)}(\gamma'(0)) \\ &= (f \circ a \circ \gamma)'(0) \\ &= \frac{d}{dt}(f \circ a \circ \gamma)(t)\bigg|_{t=0} \\ &= \frac{d}{dt} f(a(\exp(-tX), p))\bigg|_{t=0} \\ &= (a_*X)_p(f), \end{aligned}$$

as required.

2. We check that $(a_*X)_p$ is a derivation T_pM. This follows essentially immediately from the previous step. We enumerate the checks for clarity.
   - Note that $(a_*X)_p(f)$ only depends on the germ f_p because it equals $da_{(e,p)}(-X,0)(f)$, and
   $$da_{(e,p)}(-X,0) \in T_pM$$
   only depends on the germ f_p. Thus, we may redefine $(a_*X)_p$ as taking germs as input.[1]
   - Now, we see that $(a_*X)_p$ is a function taking input as germs at p and outputting elements of $\mathbb{F}$; in particular, it equals the differential $da_{(e,p)}(-X,0)$, so $(a_*X)_p$ immediately becomes a linear map and satisfies the product rule, making it a derivation.
3. We check that a_*X is a vector field. Thus, far we know that we have $(a_*X)_p \in T_pM$ for each $p \in M$, so we have a section $a_*X\colon M \to TM$. It remains to check that a_*X is smooth. Well, the first step tells us that $(a_*X)_p = da_{(e,p)}(-X,0)$, so we see that a_*X equals the composite
   $$\begin{array}{ccccccccc} M & \to & TM & \to & TG \times TM & \simeq & T(G\times M) & \overset{da}{\to} & TM \\ p & \mapsto & (p,0) & \mapsto & ((e,-X),(p,0)) & \mapsto & ((e,p),(-X,0)) & \mapsto & da_{(e,p)}(-X,0) \end{array}$$
   of smooth maps, so $a_*X\colon M \to TM$ is smooth.
4. Thus far, we know that we have a well-defined map $a_*\colon \mathfrak{g} \to \operatorname{Vect}(X)$. It remains to check that this is a homomorphism of Lie algebras. We begin by checking that it is $\mathbb{F}$-linear. Well, for $X, Y \in \mathfrak{g}$ and $c, d \in \mathbb{F}$, we are asking to check that $a_*(cX+dY) = ca_*X + da_*Y$. For this, we check the equality of derivations at some point $p \in M$, for which the first step verifies
   $$\begin{aligned} a_*(cX+dY)_p &= da_{(e,p)}(-cX-dY,0) \\ &= c\cdot da_{(e,p)}(-X,0) + d\cdot da_{(e,p)}(-Y,0) \\ &= (ca_*X + da_*Y)_p, \end{aligned}$$
   as required.
5. We check that a_* is a homomorphism of Lie algebras. We already know that $a_*\colon \mathfrak{g} \to \operatorname{Vect}(M)$ is linear (and we know that everything in sight is a Lie algebra from class), so it only remains to check that a_* preserves the Lie bracket. Explicitly, we would like to show that $a_*[X,Y] = [a_*X, a_*Y]$ for given $X, Y \in \mathfrak{g}$. For this, we choose a germ f_p represented by a regular function f defined in an open neighborhood of p.

   To run our computations, we employ a trick motivated by one in the Etingof book. Namely, define $F\colon \mathfrak{g} \to \mathbb{F}$ by $F(Z) \coloneqq f(a(\exp(Z),p))$. Now, we compute
   $$\begin{aligned} (a_*X)_p(a_*Yf) &= \frac{d}{dt}(a_*Yf)(a(\exp(-tX),p))\bigg|_{t=0} \\ &= \frac{d}{dt}\frac{d}{ds}f\big(a(\exp(-sY), a(\exp(-tX),p))\big)\bigg|_{s=0}\bigg|_{t=0} \\ &= \frac{d}{dt}\frac{d}{ds}f\big(a(\exp(-sY)\exp(-tX),p)\big)\bigg|_{s=0}\bigg|_{t=0} \\ &= \frac{d}{dt}\frac{d}{ds}f\left(a\left(\exp\left(-sY - tX + \frac{1}{2}st[Y,X] + \cdots\right),p\right)\right)\bigg|_{s=0}\bigg|_{t=0} \\ &= \frac{\partial^2}{\partial s\partial t}F\left(-sY - tX + \frac{1}{2}st[Y,X] + \cdots\right)\bigg|_{(s,t)=(0,0)} \\ &= \frac{\partial^2}{\partial s\partial t}F\left(-tX - sY - \frac{1}{2}st[X,Y] + \cdots\right)\bigg|_{(s,t)=(0,0)}. \end{aligned}$$

[1] One can also check this directly: regular local functions f and g with $f_p = g_p$ has $f_p - g_p$ vanish in a neighborhood of p, permitting us to compute $(a_*X)_p(f) = (a_*X)_p(g)$.

By reversing the roles of X and Y in the above argument, we see that

$$(a_*Y)_p(a_*Xf) = \frac{\partial^2}{\partial s \partial t} F\left(-tX - sY + \frac{1}{2}st[X,Y] + \cdots\right)\bigg|_{(s,t)=(0,0)}.$$

Thus, we see that we want to compute some particular derivatives of F. Now, f is regular, so F is a regular function $\mathfrak{g} \to \mathbb{F}$, so it will be approximately equal its Taylor expansion in a neighborhood of 0 as

$$F(Z) = f(0) + \lambda(Z) + Q(Z) + \cdots,$$

where λ is a linear functional, Q is a quadratic form, and $+\cdots$ refers to higher-order terms (with vanishing first- and second-derivatives). Plugging everything in and expanding, we see that

$$\begin{aligned}(a_*X)_p(a_*Yf) - (a_*Y)_p(a_*Xf) = \frac{\partial^2}{\partial s \partial t} - st\lambda([X,Y])\bigg|_{(s,t)=(0,0)} \\ + \frac{\partial^2}{\partial s \partial t} Q\left(-tX - sY + \frac{1}{2}st[X,Y] + \cdots\right)\bigg|_{(s,t)=(0,0)} \\ - \frac{\partial^2}{\partial s \partial t} Q\left(-tX - sY - \frac{1}{2}st[X,Y] + \cdots\right) + \cdots\bigg|_{(s,t)=(0,0)},\end{aligned}$$

where $+\cdots$ continues to denote higher-order terms, but now we see that we are only going to care about $-\lambda([X,Y])$ when computing $\frac{\partial}{\partial s \partial t}\big|_{(s,t)=(0,0)}$. (Notably, the last two terms cancel out as a derivative of $Q(-tX - sY + \cdots) - Q(-tX - sY + \cdots)$.) But now we see that

$$(a_*X)_p(a_*Yf) - (a_*Y)_p(a_*Xf) = \lambda(-[X,Y]) = \frac{d}{dt} F(-t[X,Y])\bigg|_{t=0} = a_*[X,Y]_pf,$$

as required. ■

We can now prove the Orbit–stabilizer theorem (Theorem 2.54) in the following more precise form.

Theorem 2.54 (Orbit–stabilizer)**.** Fix a left action $a\colon G \times M \to M$ of a regular Lie group G on a regular manifold M. Fix some $p \in M$.

(a) For all $p \in M$, the stabilizer G_p is a closed Lie subgroup with Lie algebra

$$\operatorname{Lie} G_p = \{X \in \mathfrak{g} : (a_*X)_p = 0\}.$$

(b) The induced map $G/G_p \to M$ given by $g \mapsto g \cdot p$ is an injective immersion. In particular, the orbit Go is an immersed submanifold.

(c) If the induced map $G/G_p \to M$ is an embedding, then G/G_p is diffeomorphic to Gp.

Proof. We begin with the proof of (a), which we do in steps.

1. Set

$$\mathfrak{g}_p := \{X \in \mathfrak{g} : (a_*X)_p = 0\}$$

for brevity. We claim that $\mathfrak{g}_p \subseteq \mathfrak{g}$ is a Lie subalgebra. Certainly $X \mapsto (a_*X)_p$ is a linear map $\mathfrak{g} \to \operatorname{Vect}(M) \to T_pM$, so $\mathfrak{g}_p$ is a linear subspace.

It remains to check that $\mathfrak{g}_p$ is preserved by the bracket. Fix $X, Y \in \mathfrak{g}_p$, and we want to check $[X,Y] \in \mathfrak{g}_p$. Well, because a_* is a homomorphism of Lie algebras, we see

$$a_*[X,Y]_pf = \underbrace{(a_*X)_p}_{0}(a_*Yf) - \underbrace{(a_*Y)_p}_{0}(a_*Xf) = 0$$

for any germ f at p. Thus, $a_*[X,Y] = 0$, so $[X,Y] \in \mathfrak{g}_p$.

2. For $X \in \mathfrak{g}_p$, we check that $\exp(X) \in G_p$. Indeed, we claim the two curves $\gamma_1(t) := \exp(-tX) \cdot p$ and $\gamma_2(t) := p$ are both integral curves for a_*X with the same initial condition at 0. This completes the check because it implies that $\exp(X) \cdot p = \gamma_1(-1) = \gamma_2(-1) = p$ by uniqueness of integral curves.

   To prove the claim, we note that γ_2 is constant, so there is nothing to check there. For γ_1, we must check that
$$\gamma_2'(t) \stackrel{?}{=} (a_*X)_{\gamma_2(t)}$$
   in $T_{\gamma_2(t)}M$. To check this, we pass through an arbitrary germ f to see that
$$\gamma_2'(t)f = (f \circ \gamma_2)'(t) = \frac{d}{ds} f(\exp(-sX - tX) \cdot p)\bigg|_{s=0},$$
   and
$$(a_*X)_{\gamma_2(t)} f = \frac{d}{ds} f(\exp(-tX - sX) \cdot p)\bigg|_{s=0},$$
   as required.

3. We attempt to control $\mathfrak{g}/\mathfrak{g}_p$. Choose a complement $\mathfrak{u}$ of $\mathfrak{g}_p \subseteq \mathfrak{g}$ so that $\mathfrak{g} = \mathfrak{g}_p \oplus \mathfrak{u}$. (We do not require that $\mathfrak{u}$ is a Lie subalgebra, despite the font.) Then the map $f \colon \mathfrak{u} \to T_pM$ given by $Z \mapsto (a_*Z)_p$ has kernel $\mathfrak{g}_p \cap \mathfrak{u} = 0$ and hence is injective. Thus, the Implicit function theorem tells us that the map $F \colon \mathfrak{u} \to M$ given by $v \mapsto \exp(-V) \cdot p$ must be an injective immersion for small v because $df_p(V) = dF_p(V)$.

   Instead of using the Implicit function theorem, we can argue using local diffeomorphisms as follows: fix a basis $\{e_1, \ldots, e_k\}$ of $\mathfrak{u}$, and extend the linearly independent set $\{dF_p(e_1), \ldots, df_p(e_k)\} \subseteq T_pM$ to a basis $\{dF_p(e_1), \ldots, dF_p(e_k)\} \sqcup \{e'_{k+1}, \ldots, e'_m\}$. Then define a local map $\widetilde{F} \colon \mathfrak{u} \times \mathbb{F}^{m-k} \to M$ by
$$\widetilde{F}(a_1e_1 + \cdots + a_me_m) = F(a_1e_1 + \cdots + a_ke_k) + a_{k+1}e'_{k+1} + \cdots + a_me'_m,$$
   where the addition on the right-hand side is defined in a local chart of M around p. (Technically, $\widetilde{F}$ is only defined in a neighborhood of $0 \in \mathfrak{u}$.) Then $\widetilde{F}$ is a local diffeomorphism at 0 by construction, so F is an injective immersion in this same neighborhood of 0.

4. We construct a slice chart for $G_p \subseteq G$ at the identity, which will complete the proof (a) by Lemma 1.90. Note that the map $\exp^{\oplus} \colon \mathfrak{g}_p \oplus \mathfrak{u} \to G$ given by $(V, X) \mapsto \exp(V)\exp(X)$ is a local diffeomorphism at 0 (because the differential is simply the identity by checking what happens on each piece $\mathfrak{g}_p$ and $\mathfrak{u}$ separately). Thus, for $g \in G$ sufficiently close to e, we can write g uniquely as in the image of e and thus as $g = \exp(V)\exp(X)$ where $V \in \mathfrak{u}$ and $X \in \mathfrak{g}_p$. Now, we see that $g \in G_p$ if and only if $\exp(V) \in G_p$, which for small enough V is equivalent to $V \in \mathfrak{g}_p$ by the previous step.

   In total, we have constructed a very small open neighborhood $U \subseteq \mathfrak{g}_p \oplus \mathfrak{u}$ of the identity such that $e|_U$ is a diffeomorphism onto its image $\exp^{\oplus}(U) \subseteq G$ and
$$G_p \cap \exp^{\oplus}(U) = \{(V, X) \in \mathfrak{g}_p \oplus \mathfrak{u} : V = 0\},$$
   which is a slice chart.

We now proceed with (b). Let $\overline{\varphi}$ denote the induced map $G/G_p \to M$ given by $\overline{\varphi}(g) := g \cdot p$, which we want to see is an injective immersion. Injectivity follows by definition of G_p: if $\varphi(g_1) = \varphi(g_2)$, then $g_1 \cdot p = g_2 \cdot p$, so $g_1^{-1}g_2 \in G_p$, so $g_1G_p = g_2G_p$. Being an immersion more or less follows from the proof. By translation, it suffices to show that $d\overline{\varphi}_e$ is injective.[2] Well, the Lie algebra of G/G_p is the quotient $\mathfrak{g}/\mathfrak{g}_p$ by Theorem 1.94, which is isomorphic to $\mathfrak{u}$ by construction of $\mathfrak{u}$. But we know that the action map is injective on $\mathfrak{u}$ by the third step above, so we are done.

Lastly, we note that (c) follows immediately from (b) because embeddings are diffeomorphic onto their images by the uniqueness of the smooth structure of embedded submanifolds. ■

[2] Once $d\overline{\varphi}_e$ is injective, we note that $\overline{\varphi} \circ L_g = L_g \circ \overline{\varphi}$ (where the first L_g is a map $G \to G$ and the second is a map $M \to M$, but both are diffeomorphisms), so $d\overline{\varphi}_g \circ d(L_g)_e = d(L_g)_p \circ d\overline{\varphi}_e$ verifies that $d\overline{\varphi}_g$ is injective.

Remark 2.55. We also remark that Theorem 1.96 follows quickly from the above result. Indeed, let G act on H via the homomorphism $\varphi\colon G \to H\colon g \cdot h := \varphi(g)h$. Then the stabilizer of any $h \in H$ is given by $\ker\varphi$, proving $\ker\varphi$ is in fact a closed Lie subgroup. Now, passing to $\overline{\varphi}$ as in the above proof shows that $G/\ker\varphi \to \operatorname{im}\varphi$ is an injective immersion.

2.4 September 25

We began class by finishing the proof of Theorem 2.54 and giving an example.

2.4.1 The Orbit–Stabilizer Theorem for Fun and Profit

Let's see an example of Theorem 2.54.

Example 2.56. Fix a finite-dimensional representation V of a regular Lie group G given by $\rho\colon G \to \mathrm{GL}(V)$. For $v \in V$, its stabilizer G_v has Lie algebra given by

$$\mathfrak{g}_v = \{X \in \mathfrak{g} : (\rho_* X)_v = 0\}.$$

Example 2.57. Fix a finite-dimensional algebra A over a field $\mathbb{F}$. Then we claim that $\mathrm{Aut}_k(A)$ is a closed Lie subgroup of $\mathrm{GL}(A)$, and we claim that

$$\mathrm{Lie}(\mathrm{Aut}\,A) = \mathrm{Der}(A) \subseteq \mathrm{End}(A).$$

Proof. Note that $\varphi \in \mathrm{GL}(A)$ is an automorphism if and only if φ also preserves the multiplication map $\mu\colon A \otimes A \to A$ of A. Now, $\mathrm{GL}(A)$ has a natural action $\rho\colon \mathrm{GL}(A) \to \mathrm{GL}(\mathrm{Hom}(A \otimes A, A))$ by

$$(\rho(g)\varphi)(x \otimes y) := g\varphi\left(g^{-1}x \otimes g^{-1}y\right).$$

Precisely speaking, this is the composite of the actions of G on the various pieces by Remark 1.103, so this is in fact a representation of G. Now, $g \in \mathrm{GL}(A)$ preserves the multiplication map μ if and only if

$$g(\mu(a \otimes b)) = \mu(g(a) \otimes g(b))$$

for all $a, b \in A$, which is equivalent to

$$(\rho(g)\mu)(a \otimes b) = g\mu\left(g^{-1}a \otimes g^{-1}b\right) = \mu(a \otimes b)$$

for all $a, b \in A$. Thus, $\mathrm{Aut}(A) \subseteq \mathrm{GL}(A)$ is the stabilizer of $\mu \in \mathrm{Hom}(A \otimes A, A)$ and hence a closed Lie subgroup by Theorem 2.54.

It remains to compute the Lie algebra, which Theorem 2.54 tells us is

$$\mathfrak{gl}(A)_\mu = \{X \in \mathfrak{gl}(A) : (\rho_* X)_\mu = 0\}.$$

Thus, we want to compute $(\rho_* X)_\mu$. Note that $\mathrm{Hom}(A \otimes A, A)$ is some finite-dimensional $\mathbb{F}$-vector space, so for any germ f defined around μ, we may use the chain rule to compute

$$\begin{aligned}(\rho_* X)_\mu f &= \frac{d}{dt} f(\rho(\exp(-tX), \mu))\bigg|_{t=0} \\ &= df_\mu\left(\frac{d}{dt}\rho(\exp(-tX), \mu)\bigg|_{t=0}\right).\end{aligned}$$

Thus, we see that $X \in \mathfrak{gl}(A)_\mu$ if and only if $\frac{d}{dt}\rho(\exp(-tX),\mu)\big|_{t=0} = 0$. Now, linear operators pass through derivatives, and evaluation is a linear operator on $\operatorname{Hom}(A\otimes A, A)$, so it suffices to check when

$$\frac{d}{dt}\rho(\exp(-tX),\mu)(a\otimes b)\big|_{t=0}$$

vanishes, for arbitrary $a,b\in A$. Thus, we compute

$$\begin{aligned}\rho(\exp(-tX),\mu)(a\otimes b) &= \exp(-tX)\mu(\exp(tX)a,\exp(tX)b)\\ &= \sum_{n=0}^\infty \frac{(-1)^n t^n}{n!}X^n\left(\sum_{k=0}^\infty \frac{t^k}{k!}X^k a\cdot\sum_{\ell=0}\frac{t^\ell}{\ell!}X^\ell b\right)\\ &= \sum_{n,k,\ell=0}^\infty \frac{(-1)^n t^{n+k+\ell}}{n!k!\ell!}X^n\left(X^k a\cdot X^\ell b\right),\end{aligned}$$

where we have rearranged the sums with impunity because everything in sight converges absolutely. Furthermore, we can differentiate term-by-term to see that

$$\begin{aligned}\frac{d}{dt}\rho(\exp(-tX),\mu)(a\otimes b)\bigg|_{t=0} &= \frac{(-1)^1t^{1+0+0}}{1!0!0!}X^1\left(X^0a\cdot X^0b\right)\\ &\quad+\frac{(-1)^0t^{0+1+0}}{0!1!0!}X^0\left(X^1a\cdot X^0b\right)\\ &\quad+\frac{(-1)^0t^{0+0+1}}{0!0!1!}X^0\left(X^0a\cdot X^1b\right)\\ &= -X(a\cdot b)+Xa\cdot b+a\cdot Xb.\end{aligned}$$

Thus,

$$\operatorname{Lie}(\operatorname{Aut} A) = \{X\in\mathfrak{gl}(A) : X(a\cdot b) = Xa\cdot b + a\cdot Xb\},$$

which of course is the set of derivations. ■

Remark 2.58. A close examination of the above proof finds that we only need μ to be an element of $\operatorname{Hom}(A\otimes A, A)$ for the argument to go through. Notably, we may replace (A,μ) above with a Lie algebra $(\mathfrak{g},[-,-])$ to find that $\operatorname{Aut}_{\mathrm{LieAlg}}(\mathfrak{g})\subseteq\operatorname{GL}(\mathfrak{g})$ is a Lie subgroup with Lie algebra given by the derivations

$$\{\varphi\in\mathfrak{gl}(\mathfrak{g}) : \varphi([X,Y]) = [\varphi(X),Y]+[X,\varphi(Y)] \text{ for all } X,Y\in\mathfrak{g}\}.$$

Remark 2.59. The adjoint map $\operatorname{ad}\colon\mathfrak{g}\to\mathfrak{gl}(\mathfrak{g})$ actually lands in $\operatorname{Der}(\mathfrak{g})$: checking this is tantamount to checking that $X,Y,Z\in\mathfrak{g}$ has

$$\operatorname{ad}_X[Y,Z]\stackrel{?}{=}[\operatorname{ad}_X Y,Z]+[Y,\operatorname{ad}_X Z],$$

which one can check is equivalent to the Jacobi identity of Proposition 2.29.

Remark 2.60. Similarly, the adjoint action $\operatorname{Ad}\colon G\to\operatorname{GL}(\mathfrak{g})$ actually lands in $\operatorname{Aut}_{\mathrm{LieAlg}}(\mathfrak{g})$. Indeed, for $g\in G$ and $X,Y\in\mathfrak{g}$, this amounts to checking that

$$\operatorname{Ad}_g[X,Y]\stackrel{?}{=}[\operatorname{Ad}_g X,\operatorname{Ad}_g Y],$$

which is Corollary 2.19.

Here is another application.

Definition 2.61 (center)**.** Fix a group G. Then the *center* of G is the subset

$$Z(G) := \{z \in G : zg = gz \text{ for all } g \in G\}.$$

Similarly, fix a Lie algebra $\mathfrak{g}$, then the *center* of $\mathfrak{g}$ is

$$\mathfrak{z}(\mathfrak{g}) := \{X \in \mathfrak{g} : [X, Y] = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

Remark 2.62. We will not bother to check that $Z(G)$ is a subgroup because this is a standard result of group theory. However, in order to do something, let's check that $\mathfrak{z}(\mathfrak{g})$ is a Lie ideal of $\mathfrak{g}$. Note that $\mathfrak{z}(\mathfrak{g})$ is the kernel of the collection of linear maps $X \mapsto [X, Y]$ as $Y \in \mathfrak{g}$ varies, so $\mathfrak{z}(\mathfrak{g})$ is an intersection of Lie ideals (by Lemma 2.39) and hence a Lie ideal by Remark 2.40.

Proposition 2.63. Fix a connected regular Lie group G. Then $Z(G)$ is a closed Lie subgroup with Lie algebra

$$\mathfrak{z}(\mathfrak{g}) := \{X \in \mathfrak{g} : [X, Y] = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

Proof. We would like to see that $Z(G)$ is the kernel of the adjoint map $\operatorname{Ad}: G \to \operatorname{Aut} G$, but it is difficult to make sense of this argument because $\operatorname{Aut} G$ is not a manifold.

Instead, we note that $g \in G$ if and only if g commutes with an open neighborhood U of the identity: indeed, commuting with U implies commuting with the subgroup generated by U, but G is connected, so commuting with U is equivalent to commuting with G. Now, we can take U to be some neighborhood of the identity in the image of the local diffeomorphism $\exp: \mathfrak{g} \to G$, so $g \in Z(G)$ if and only if

$$g \exp(X) g^{-1} = \exp(X)$$

for all $X \in \mathfrak{g}$ in an open neighborhood of 0. Now, $\operatorname{Ad}_g \exp(X) = \exp(\operatorname{Ad}_g X)$ by Proposition 2.12, so the above equality is equivalent to having $\operatorname{Ad}_g X = X$ for X in a neighborhood of 0.

Thus, Remark 2.55 tells us that $Z(G)$ is the kernel of the representation $\operatorname{Ad}_\bullet : G \to \operatorname{GL}(\mathfrak{g})$. We conclude that its Lie algebra is the kernel of the differential of $\operatorname{Ad}_\bullet$, which of course is $\operatorname{ad}_\bullet : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. Thus,

$$\operatorname{Lie} Z(G) = \{X \in \mathfrak{g} : \operatorname{ad}_X = 0\},$$

but Proposition 2.23 explains that $\operatorname{ad}_X = [X, -]$, so we see that this is simply $\mathfrak{z}(\mathfrak{g})$, as required. ■

Definition 2.64 (adjoint)**.** Fix a connected regular Lie group G. Then the *adjoint group* of G is $G^{\text{ad}} := G/Z(G)$.

Example 2.65. For $G = \operatorname{GL}_n(\mathbb{F})$, one can check that $Z(G)$ is the subgroup $\{cI : c \in \mathbb{F}\}$. The adjoint group is then $\operatorname{PGL}_n(\mathbb{F})$.

2.4.2 The Baker–Campbell–Hausdorff Formula

For completeness, we mention the Baker–Campbell–Hausdorff formula. We will not need this result, so we will not prove it, and the discussion in this subsection will be quite terse. Fix a Lie group G with Lie algebra $\mathfrak{g}$. We would like to understand the group law on G purely in terms of $\mathfrak{g}$. As in our discussion of the commutator, we note that

$$\mu(X, Y) = \log(\exp(X) \exp(Y)),$$

defined in an open neighborhood of $\mathfrak{g}$ can be expanded out as

$$\mu(X,Y) = X + Y + \frac{1}{2}[X,Y] + \mu_3(X,Y) + \mu_4(X,Y) + \cdots = \sum_{n=1}^{\infty} \mu_n(X,Y),$$

where $\mu_n(X,Y)$ consists of the order-n terms in this Taylor expansion. Here is the main result. For example, as above, $\mu_1(X,Y) = X+Y$ and $\mu_2(X,Y) = \frac{1}{2}[X,Y]$.

Theorem 2.66 (Baker–Campbell–Hausdorff)**.** The polynomials μ_n above are independent of G.

One proves this basically by solving differential equations for the μ_n inductively in n.

Example 2.67. One could compute that

$$\mu_3(X,Y) = \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]).$$

2.4.3 The Fundamental Theorems of Lie Theory

To wrap up our transition to Lie algebras, we state the fundamental theorems of Lie theory, which we will mostly not prove.

Theorem 2.68. For a connected regular Lie group G with Lie algebra $\mathfrak{g}$, there is a bijection between Lie subgroups $H \subseteq G$ and Lie subalgebras $\mathfrak{h} \subseteq \mathfrak{g}$. This bijection sends $H \subseteq G$ to $\mathfrak{h} \coloneqq \operatorname{Lie} H$.

Theorem 2.69. Fix a simply connected regular Lie group G with Lie algebra $\mathfrak{g}$. Then for any regular Lie group H with Lie algebra $\mathfrak{h}$, the map

$$\operatorname{Hom}_{\text{LieGrp}}(G,H) \to \operatorname{Hom}_{\text{LieAlg}}(\mathfrak{g},\mathfrak{h}),$$

given by taking the differential at the identity, is a bijection.

Theorem 2.70. Any finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to the Lie algebra of some simply connected regular Lie group.

Here is the consequence.

Corollary 2.71. The (full subcategory) of simply connected regular Lie groups is equivalent to the category of finite-dimensional Lie algebras, given by the Lie algebra functor.

Proof. Theorem 2.69 shows that this functor is fully faithful, and Theorem 2.70 shows that this functor is essentially surjective. This completes the proof. ■

We will begin with the proof of Theorem 2.68 next class. This requires the theory of distributions.

2.5 September 27

Today we continue our discussion of the fundamental theorems of Lie theory.

2.5.1 Distributions and Foliations

Here is the main definition.

Definition 2.72 (distribution)**.** Fix a regular manifold M. Then a *k-dimensional distribution* $\mathcal{D}$ on X is a k-dimensional (local) subbundle $\mathcal{D} \subseteq TX$.

Remark 2.73. Locally at a point $p \in M$, we can think about $\mathcal{D}_p$ as being spanned by k linearly independent differentials which spread out over a neighborhood.

Definition 2.74 (integrable)**.** A distribution $\mathcal{D}$ of dimension k on a regular manifold M is *integrable* if and only if each $p \in M$ has a regular chart (U, φ) with local coordinates $\varphi = (x_1, \ldots, x_n)$ such that

$$D|_U = \operatorname{span}\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}\right\}.$$

Here is a more coordinate-free check for being integrable.

Definition 2.75 (foliation)**.** A distribution $\mathcal{D}$ of dimension k on a regular manifold M is a *foliation* if and only if $p \in M$ has an "integral" immersed submanifold $S_p \subseteq M$, meaning that $T_qS_p = D_q$ for all $q \in S_p$.

Foliations give rise to partitions of the manifold, called leaves.

Definition 2.76 (leaf)**.** Fix a foliation $\mathcal{D}$ of rank k on a smooth manifold M. Given $p \in M$, a *leaf* of $\mathcal{D}$ is the collection of points $q \in M$ such that there is a path γ connecting p and q with $\gamma'(t) \in \mathcal{D}_{\gamma(t)}$ for all t.

Note that the leaves of $\mathcal{D}$ are connected and partition M.

Example 2.77. Orbits of the action of $\mathbb{R}$ on $\mathbb{R}^2/\mathbb{Z}^2$ by $r\colon (x, y) \mapsto (x + r, y)$ are leaves.

Example 2.78. For a fiber bundle with connected fibers, the leaves are fibers.

Example 2.79. For any connected closed Lie subgroup H of a regular Lie group G, the quotient $G \twoheadrightarrow G/H$ is a fiber bundle and hence produces leaves given by H.

Example 2.80. If $\mathcal{D}$ is a vector field (i.e., has dimension 1), then the leaves are integral submanifolds, which are the integral curves.

Anyway, here is our main theorem.

Theorem 2.81 (Frobenius)**.** A distribution $\mathcal{D}$ on a smooth manifold M is integrable if and only if $\mathcal{D}$ is closed under the Lie bracket.

Proof. The forward direction is not so bad: note $\mathcal{D}$ being integrable means that each $p \in M$ has an open neighborhood where $\mathcal{D}$ is just given by tangent spaces, and vector fields living in tangent spaces will be preserved by the Lie bracket. For the converse, see [Lee13, Theorem 19.12]. It proceeds by induction. ■

2.5.2 Sketches of the Fundamental Theorems

We begin with Theorem 2.68.

Proof of Theorem 2.68. The main point is to produce the reverse map producing Lie subgroups from Lie subalgebras. As such, fix some Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. For each $X \in \mathfrak{g}$, let ξ_X be the corresponding left-invariant vector field. We then let

$$\mathcal{D}^{\mathfrak{h}} := \operatorname{span}\{\xi_Y : Y \in \mathfrak{h}\}.$$

Here are some checks on $\mathcal{D}^{\mathfrak{h}}$.

- Quickly, we claim that $\mathcal{D}^{\mathfrak{h}}$ is integrable. By Theorem 2.81, it is enough to check that $\mathcal{D}^{\mathfrak{h}}$ is closed under the bracket. This is a matter of computation: for two vector fields $\sum_i f_i\xi_{Y_i}$ and $\sum_j g_j\xi_{Y_j}$ contained in $\mathcal{D}^{\mathfrak{h}}$, we find

$$\begin{aligned}\left[\sum_i f_i\xi_{Y_i}, \sum_j g_j\xi_{Y_j}\right] &= \sum_{i,j} \left(f_i(\xi_{Y_i}g_j)\xi_{X_i} - g_j(\xi_{Y_j}f_i)\xi_{h_i}\right) \\ &= \cdots .\end{aligned}$$

- In fact, we note that $\mathcal{D}^{\mathfrak{h}}$ is left-invariant.

We now let S_g be the integral submanifold corresponding to $g \in G$, and we note that we can take $H := S_1$ to complete the proof. ■

We now proceed with Theorem 2.69.

Proof of Theorem 2.69. Injectivity follows from Corollary 2.11, so we merely need to get the surjectivity. The point is to pass to the graph in order to produce morphisms when we already know how to produce objects (via Theorem 2.68).

Fix some homomorphism $\psi\colon \mathfrak{g} \to \mathfrak{h}$ of Lie algebras. Well, define $\theta\colon \mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{k}$ by $\theta(X) := (X, \psi(X))$; note that this is still a Lie algebra homomorphism because it is the sum of Lie algebra homomorphisms. Now, $\operatorname{im}\theta$ is a Lie subalgebra of $\operatorname{Lie}(G \times H) = \mathfrak{g} \times \mathfrak{h}$ by Lemma 2.39, so Theorem 2.68 tells us that we can find some subgroup connected $\Gamma \subseteq G \times H$ with

$$\operatorname{Lie}\Gamma = \operatorname{im}\theta.$$

Let $\mathrm{pr}_1\colon \Gamma \to G$ and $\mathrm{pr}_2\colon \Gamma \to H$ be the projections. Note that $d(\mathrm{pr}_1)_e \circ \theta = \mathrm{id}_{\mathfrak{g}}$ by definition of θ, and $\theta \circ d(\mathrm{pr}_1)_e = \mathrm{id}_{\operatorname{Lie}\Gamma}$ by construction of Γ. Thus, we see that $d(\mathrm{pr}_1)_e$ is a bijection and hence a local isomorphism; in particular, $\mathrm{pr}_1\colon G \to \Gamma$ must be a covering space map, so we conclude that pr_1 is actually an isomorphism. We thus recover a map

$$G \overset{\mathrm{pr}_1}{\leftarrow} \Gamma \overset{\mathrm{pr}_2}{\to} H$$

which is ψ on the level of Lie algebras, as required. ■

We will largely omit the proof of Theorem 2.70. It follows from strong structure theory of Lie algebras. For example, one wants the following result.

Theorem 2.82 (Ado)**.** Any finite-dimensional Lie algebra $\mathfrak{g}$ has a faithful representation. In other words, there exists a finite-dimensional vector space V and an injective Lie algebra homomorphism $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$.

We will not show this, but we remark on a special case.

Remark 2.83. Suppose that $\mathfrak{g}$ is a Lie algebra with $\mathfrak{z}(\mathfrak{g}) = 0$. Then the adjoint representation $\operatorname{ad}_\bullet\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ given by $X \mapsto [X, -]$ is a faithful representation.

With this in hand, we can prove Theorem 2.70.

Proof of Theorem 2.70. By passing to the universal cover of the connected component, it suffices to produce some regular Lie group G with $\operatorname{Lie}G = \mathfrak{g}$. Well, embed $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ for some finite-dimensional vector space V, and then we are done by Theorem 2.68 after noticing $\mathfrak{gl}(V) = \operatorname{Lie}\operatorname{GL}(V)$. ■

2.5.3 Complexifications

In the sequel, we will want to focus on Lie algebras of $\mathbb{C}$ instead of $\mathbb{R}$. For this, we make the following definition.

Definition 2.84 (complexification)**.** Fix a real Lie algebra $\mathfrak{g}$. Then we define the *complexification* as

$$\mathfrak{g}_{\mathbb{C}} := \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}.$$

Then $\mathfrak{g}_{\mathbb{C}}$ is a Lie algebra over $\mathbb{C}$.

Example 2.85. One sees that $\mathfrak{gl}_n(\mathbb{R})_{\mathbb{C}}$ is simply $\mathfrak{gl}_n(\mathbb{C})$. However, $\mathfrak{gl}_n(\mathbb{C})$ is also $\mathfrak{u}_n(\mathbb{C})_{\mathbb{C}}$ after some care.

Example 2.86. One sees that $\mathfrak{so}_{k,\ell}(\mathbb{R})_{\mathbb{C}}$ is just $\mathfrak{so}_{\mathbb{C}}(\mathbb{C})$.

Definition 2.87 (complexification)**.** Fix a simply connected real Lie group H over $\mathbb{R}$. Then we let G be the unique simply connected complex Lie group G such that

$$\operatorname{Lie} G = \operatorname{Lie} H \otimes_{\mathbb{R}} \mathbb{C}.$$

Note that G certainly exists by Theorem 2.70.

Example 2.88. Note that $\mathrm{SL}_2(\mathbb{R})$ has a two-sheeted cover, which on the level of Lie algebras is given by $\mathfrak{sl}_2(\mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C}$.

With care, one is able to go in the reverse direction.

Definition 2.89 (real form)**.** Fix a connected complex Lie group G. A real Lie subgroup $H \subseteq G$ is a *real form* of G such that the natural map

$$\operatorname{Lie} H \otimes_{\mathbb{R}} \mathbb{C} \to \operatorname{Lie} G$$

is an isomorphism of Lie algebras over $\mathbb{C}$.

Remark 2.90. It is not technically obvious that real forms always exist. One thing one may try for simply connected G is to take the fixed points of the map $G \to G$ induced by the complex conjugation morphism $\mathfrak{g} \to \mathfrak{g}$.

Let's see a few examples.

Example 2.91. Of course, for any real Lie algebra $\mathfrak{g}$, we see that $\mathfrak{g}$ is a real form of $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. For example, $\mathfrak{sl}_n(\mathbb{R})$ is a real form of $\mathfrak{sl}_n(\mathbb{C})$.

Example 2.92. Note that $\mathfrak{su}_n$ is a real form of $\mathfrak{sl}_n(\mathbb{C})$. Indeed, define the map $\mathfrak{su}_n \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{sl}_n(\mathbb{C})$ by $X \otimes z \mapsto zX$. Certainly this map is well-defined because $\mathfrak{su}_n$ consists of traceless matrices already, it is linear by construction, and it preserves the Lie bracket because the Lie bracket is the matrix commutator everywhere. To check that we have an isomorphism, we note that our dimensions are equal by Example 1.160. Thus, for example, it is enough to note that our mapping is injective: any element of $\mathfrak{su}_n \otimes_{\mathbb{R}} \mathbb{C}$ can be written as $X \otimes 1 + Y \otimes i$, but if this goes to 0 in $\mathfrak{sl}_n(\mathbb{C})$, then $X + iY = 0$, and we conclude that $X = Y = 0$ by taking real and imaginary parts on the coordinates.

Theme 3
Building Representations

3.1 September 30

Today we will talk about representations.

3.1.1 Representations

Fix a ground field F, which is usually an extension of $\mathbb{F}$. To review, recall that a representation of a regular Lie group G is a morphism $\rho_V : G \to \mathrm{GL}(V)$ of Lie groups; given the data of only the k-vector space V, we will assume that the representation is called ρ_V. A morphism $\varphi\colon V \to W$ of representations is one respecting the G-actions: we require φ to be linear and satisfying

$$\rho_W(g) \circ \varphi = \varphi \circ \rho_V(g)$$

for all $g \in G$. The category here is called $\mathrm{Rep}_k(G)$.

Similarly, for a Lie algebra $\mathfrak{g}$, a representation is a morphism $\rho_V : \mathfrak{g} \to \mathfrak{gl}(V)$ of Lie algebras. A morphism $\varphi\colon V \to W$ of representations is one respecting the G-action again: again, we need

$$\rho_W(g) \circ \varphi = \varphi \circ \rho_V(g).$$

The category here is called $\mathrm{Rep}_k(\mathfrak{g})$.

Remark 3.1. As a quick aside, we note that a bijective morphism $\varphi\colon V \to W$ will be an isomorphism. Indeed, the inverse map $\psi\colon W \to V$ is an isomorphism of vector spaces by linear algebra, and we see that it is invariant under our action as follows: for any $w \in W$ and operator g in G or $\mathfrak{g}$, write $w = \varphi(v)$ for some unique $v \in V$ so that

$$\psi(gw) = \psi(g\varphi(v)) = \psi(\varphi(gv)) = gv = g\psi(w).$$

Note that if $\mathfrak{g} = \operatorname{Lie} G$, then we have a functor taking $\rho\colon G \to \mathrm{GL}(V)$ to $d\rho_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$. Let's explain this.

Lemma 3.2. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$.

(a) One has a functor $F\colon \mathrm{Rep}(G) \to \mathrm{Rep}(\mathfrak{g})$ sending a representation $\rho\colon G \to \mathrm{GL}(V)$ to $d\rho_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

(b) The functor F is faithful.

Proof. For (a), we explain that $F\colon \operatorname{Rep}(G) \to \operatorname{Rep}(\mathfrak{g})$ is a functor. Here is the data.

- On objects, we send $\rho\colon G \to \operatorname{GL}(V)$ to the map $d\rho_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$, which we know is a morphism of Lie algebras because ρ is a group homomorphism.
- Further, we send morphisms $\varphi\colon V \to W$ of G-representations (namely, satisfying $\varphi \circ \rho_V(g) = \rho_W(g) \circ \varphi$ for all $g \in G$) to the morphism $d\varphi_0\colon V \to W$, which of course can be identified with the original map because φ is linear. For this to make sense, we should check that $\varphi\colon V \to W$ preserves the $\mathfrak{g}$-action if it preserves the G-action. Well, for $X \in \mathfrak{g}$ and $v \in V$, we must check that

  $$\varphi(d(\rho_V)_e(X)v) \stackrel{?}{=} d(\rho_W)_e(X)\varphi(v).$$

  Well, define $\gamma\colon \mathbb{F} \to G$ by $\gamma(t) := \exp(tX)$. Then we note that linear maps (such as evaluation at v) pass through derivatives by their definition as a limit, so

  $$\begin{aligned}\varphi(d(\rho_V)_e(X)v) &= \varphi(d(\rho_V)_e(\gamma'(0))v) \\ &= \varphi\left((\rho_V \circ \gamma)'(0)v\right) \\ &= \varphi\left(\frac{d}{dt}\rho_V \circ \gamma(t)\bigg|_{t=0} v\right) \\ &= \frac{d}{dt}\varphi(\rho_V(\gamma(t))(v))\bigg|_{t=0} \\ &= \frac{d}{dt}\rho_W(\gamma(t))(\varphi(v))\bigg|_{t=0} \\ &= (\rho_W \circ \gamma)'(0)(\varphi(v)) \\ &= d\rho_e(X)(\varphi(v)),\end{aligned}$$

  as required.

Here are the coherence checks.

- Identity: note that the identity map $\mathrm{id}_V\colon V \to V$ on G-representations (which is the identity linear map) gets sent to the identity linear map $V \to V$ on $\mathfrak{g}$-representations.
- Associativity: for morphisms $\varphi\colon V \to V'$ and $\varphi'\colon V' \to V''$ of G-representations, we note that we get the exact same maps out as $\mathfrak{g}$-representations, so $\psi \circ \varphi$ as a G-representation gets sent to $F(\psi \circ \varphi) = \psi \circ \varphi = F\psi \circ F\varphi$.

The previous point has given us our functor, so we now need to check that it is faithful for (b). Well, a G-invariant map $\varphi\colon V \to W$ goes to the same map $\varphi\colon V \to W$ as a $\mathfrak{g}$-representation by definition of F. Thus, given two maps $\varphi_1, \varphi_2\colon V \to W$ of G-representations, we see that $F\varphi_1 = F\varphi_2$ implies that

$$\varphi_1 = F\varphi_1 = F\varphi_2 = \varphi_2,$$

as required. ■

Remark 3.3. It will be helpful to remember in the sequel that

$$d\rho_e(X)v = \frac{d}{dt}\rho(\exp(tX))v\bigg|_{t=0},$$

which was proved in the argument above. Note that this derivative makes sense because it takes place in some Euclidean space.

3.1.2 Operations on Representations

We present some operations on representations of G and $\mathfrak{g}$. Note that these should always be related by the ambient functor $\operatorname{Rep}(G) \to \operatorname{Rep}(\mathfrak{g})$ which is an equivalence when G is simply connected by Proposition 3.28. As basic examples, here are some trivial representations.

Lemma 3.4. Fix a Lie group G and Lie algebra $\mathfrak{g}$ and a vector space V.

(a) We can make V into a "trivial" G-representation by $\rho_V(g) = \operatorname{id}_V$ for all $g \in G$.

(b) We can make V into a "trivial" $\mathfrak{g}$-representation by $\rho_V(X) := 0$ for all $X \in \mathfrak{g}$.

(c) Suppose $\mathfrak{g} = \operatorname{Lie} G$. Making V into a trivial G-representation, we see that $F(V)$ is the trivial $\mathfrak{g}$-representation.

Proof. Here we go.

(a) We have indeed defined a homomorphism $G \to \mathrm{GL}(V)$ because this is the trivial homomorphism. It is also regular because constant maps are regular.

(b) We have indeed defined a map $\rho_V : \mathfrak{g} \to \mathfrak{gl}(V)$, and it is linear map of vector spaces. It remains to check that we have defined a map of Lie algebras, for which we note that

$$[\rho_V(X), \rho_V(Y)] = \rho_V(X) \circ \rho_V(Y) - \rho_V(Y) \circ \rho_V(X) = 0 = \rho_V([X, Y]).$$

(c) Fix the trivial representation as $\rho\colon G \to \mathrm{GL}(V)$. Then the induced map $d\rho_1 \colon \mathfrak{g} \to \mathfrak{gl}(V)$ is given by

$$d\rho_1(X)v = \frac{d}{dt}\rho(\exp(-tX))v\bigg|_{t=0},$$

but of course $\rho(\exp(-tX))v = v$ for all $t \in \mathbb{R}$, so this derivative vanishes. Thus, $d\rho_1 \colon \mathfrak{g} \to \mathfrak{gl}(V)$ is the zero map, as required. ■

Example 3.5. We always have the trivial representation on the zero-dimensional vector space.

As something else easy to do, we note that there are complex conjugate representations.

Lemma 3.6. Fix a Lie group G and Lie algebra $\mathfrak{g}$.

(a) Given a representation $V \in \operatorname{Rep}_{\mathbb{C}}(G)$, we can make the complex conjugate vector space $\overline{V}$ into a representation of $\overline{G}$ by

$$\rho_{\overline{V}}(g)(\overline{v}) := \overline{\rho_V(g)\overline{v}}.$$

(b) Given a representation $V \in \operatorname{Rep}_{\mathbb{C}}(\mathfrak{g})$, we can make the complex conjugate vector space $\overline{V}$ into a representation of $\overline{\mathfrak{g}}$ by

$$\rho_{\overline{V}}(X)(v) := \overline{\rho_V(X)v}.$$

(c) Suppose $\mathfrak{g} = \operatorname{Lie}(G)$. Given a representation $V \in \operatorname{Rep}_{\mathbb{C}}(G)$, then $F\overline{V} = \overline{FV}$ as representations in $\operatorname{Rep}_{\mathbb{C}}(\mathfrak{g})$.

Proof. Here we go.

(a) For each $g \in G$, we note that $\rho_{\overline{V}}(g) \colon \overline{V} \to \overline{V}$ is $\mathbb{C}$-linear: for any $a, a' \in \mathbb{C}$ and $\overline{v}, \overline{v}' \in \overline{V}$, we see

$$\begin{aligned}\rho_{\overline{V}}(g)(a\overline{v} + a'\overline{v}') &= \rho_{\overline{V}}(g)(\overline{\overline{a}v + \overline{a}'v'}) \\ &= \overline{\rho_V(g)(\overline{a}v + \overline{a}'v')} \\ &= \overline{\overline{a}\rho_V(g)v + \overline{a}'\rho_V(g)v'} \\ &= a\rho_{\overline{V}}(g)(\overline{v}) + a'\rho_{\overline{V}}(g)\overline{v}'.\end{aligned}$$

To show that we have defined a group homomorphism, we see that

$$\rho_{\overline{V}}(gh)\overline{v} = \overline{\rho_V(gh)v} = \overline{\rho_V(g)\rho_V(h)v} = \rho_{\overline{V}}(g)\rho_{\overline{V}}(h)\overline{v}.$$

Lastly, we note that the map $\rho_{\overline{V}} \colon G \to \operatorname{GL}(\overline{V})$ is a regular map by expanding it on a basis: upon picking a $\mathbb{C}$-basis of V (which is also a $\mathbb{C}$-basis of $\overline{V}$), we see that the matrix $\rho_{\overline{V}}(g)$ is simply the complex conjugate of the matrix of $\rho(g)$, which will continue to be a regular map after keeping track of all of our conjugations.

(b) The same check as in (a) explains that $\rho_{\overline{V}}(X)$ is at least a $\mathbb{C}$-linear map for all $X \in \mathfrak{g}$. This map is also of course linear in X given by the linearity of ρ_V. Lastly, this is a homomorphism of Lie algebras by taking the conjugate of the identity

$$\rho_V([X,Y]) = \rho_V(X)\rho_V(Y) - \rho_V(Y) - \rho_V(X).$$

(c) Simply take the conjugate everywhere in sight. ■

To begin doing something with content, we handle direct sums.

Lemma 3.7. Fix a Lie group G and Lie algebra $\mathfrak{g}$.

(a) Given representations $V, W \in \operatorname{Rep}_k(G)$, we can make $V \otimes W$ into a representation of G via the coordinate-wise action

$$\rho_{V\oplus W}(g)(v \oplus w) := \rho_V(g)v \oplus \rho_W(g)w.$$

(b) Given representations $V, W \in \operatorname{Rep}_k(\mathfrak{g})$, we can make $V \oplus W$ into a representation of G via the coordinate-wise action

$$\rho_{V\oplus W}(X)(v \oplus w) = \rho_V(X)v \oplus \rho_W(X)w.$$

(c) Suppose $\mathfrak{g} = \operatorname{Lie}(G)$. Given representations $V, W \in \operatorname{Rep}_k(G)$, then $F(V \otimes W)$ is the direct sum representation in $\operatorname{Rep}_k(\mathfrak{g})$.

Proof. Here we go.

(a) By taking the direct sum of the homomorphisms $\rho_V \colon G \to \operatorname{GL}(V)$ and $\rho_W \colon G \to \operatorname{GL}(W)$, we obtain a regular homomorphism $G \to \operatorname{GL}(V) \oplus \operatorname{GL}(W)$. To finish, we note that $\operatorname{GL}(V) \oplus \operatorname{GL}(W)$ embeds into $\operatorname{GL}(V \oplus W)$ by sending (φ, ψ) to the linear map $V \oplus W \to V \oplus W$ acting by (φ, ψ) on the coordinates. To see that this last map is a regular homomorphism, we note that fixing an ordered basis of both V and W allows us to identify these GL groups with invertible matrices, in which case our map is given by

$$(A, B) \mapsto \begin{bmatrix} A & \\ & B \end{bmatrix}.$$

In particular, this map is regular in coordinates and hence regular; one can check that it is a homomorphism directly because $(A, B) \cdot (A', B')$ goes to the block-diagonal matrix $\operatorname{diag}(AA', BB')$. In total, we have obtained a composite of regular homomorphisms $G \to \operatorname{GL}(V) \oplus \operatorname{GL}(W) \to \operatorname{GL}(V \oplus W)$.

(b) We will simply proceed directly. We define a map $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V \oplus W)$ by

$$\rho(X) := \begin{bmatrix} \rho_V(X) & \\ & \rho_W(X) \end{bmatrix},$$

where we are thinking about endomorphisms of $V \oplus W$ in the above block-diagonal format. Linearity of ρ_V and ρ_W gives linearity of ρ. To check the bracket, we compute

$$\begin{aligned} [\rho(X), \rho(Y)] &= \begin{bmatrix} \rho_V(X) & \\ & \rho_W(X) \end{bmatrix} \begin{bmatrix} \rho_V(Y) & \\ & \rho_W(Y) \end{bmatrix} - \begin{bmatrix} \rho_V(Y) & \\ & \rho_W(Y) \end{bmatrix} \begin{bmatrix} \rho_V(X) & \\ & \rho_W(X) \end{bmatrix} \\ &= \begin{bmatrix} \rho_V(X) \circ \rho_V(Y) - \rho_V(Y) \circ \rho_V(X) & \\ & \rho_W(X) \circ \rho_W(Y) - \rho_W(Y) \circ \rho_W(X) \end{bmatrix} \\ &= \begin{bmatrix} [\rho_V(X), \rho_V(Y)] & \\ & [\rho_W(X), \rho_W(Y)] \end{bmatrix} \\ &= \begin{bmatrix} \rho_V([X,Y]) & \\ & \rho_W([X,Y]) \end{bmatrix} \\ &= \rho([X,Y]). \end{aligned}$$

(c) This is a direct computation. Given the representations $\rho_V : G \to \mathrm{GL}(V)$ and $\rho_W : G \to \mathrm{GL}(W)$ with direct sum $\rho_{V\oplus W}$, we need to compute the direct sum of the representations $d\rho_V$ and $d\rho_W$. Well, for any $X \in \mathfrak{g}$ and $(v,w) \in V \oplus W$, we note that evaluation at (v,w) is a linear map and hence passes through derivative computations (in Euclidean space!), so

$$\begin{aligned} d\rho_{V\oplus W}(X)(v,w) &= \frac{d}{dt}\rho_{V\oplus W}(\exp(tX))\bigg|_{t=0}(v,w) \\ &= \frac{d}{dt}\begin{bmatrix} \rho_V(\exp(tX)) & \\ & \rho_W(\exp(tX)) \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix}\bigg|_{t=0} \\ &= \frac{d}{dt}\begin{bmatrix} \rho_V(\exp(tX))v & \\ & \rho_W(\exp(tX))w \end{bmatrix}\bigg|_{t=0}. \end{aligned}$$

Now, because we are in a Euclidean space, we can compute the derivative on each coordinate separately, which we see to be $\operatorname{diag}(d\rho_V(X), d\rho_W(X))$, as needed. ■

Next we handle the tensor product.

Lemma 3.8. Fix a Lie group G and Lie algebra $\mathfrak{g}$.

(a) Given representations $V, W \in \mathrm{Rep}_k(G)$, we can make $V \otimes W$ into a representation of G via the coordinate-wise action

$$\rho_{V\otimes W}(g)(v \otimes w) := \rho_V(g)v \otimes \rho_W(g)w.$$

(b) Given representations $V, W \in \mathrm{Rep}_k(\mathfrak{g})$, we can make $V \otimes W$ into a representation of G via the product rule action

$$\rho_{V\otimes W}(X)(v \otimes w) = \rho_V(X)v \otimes w + v \otimes \rho_W(X)w.$$

(c) Suppose $\mathfrak{g} = \mathrm{Lie}(G)$. Given representations $V, W \in \mathrm{Rep}_k(G)$, then $F(V \otimes W)$ is the tensor product representation in $\mathrm{Rep}_k(\mathfrak{g})$.

Proof. Here we go.

(a) For each $g \in G$, we need to provide a bilinear map $\rho(g)\colon (V \times W) \to (V \otimes W)$, for which we take

$$\rho(g)(v,w) := \rho_V(g)v \otimes \rho_W(g)w.$$

Linearity of $\rho_V(g)$ and $\rho_W(g)$ (and properties of the tensor product) verify that we have in fact defined a bilinear map, so we have in fact defined a map $G \to \mathrm{End}(V \otimes W)$. Here are our checks to make this map a representation.

- Group action: for the identity check, we note that

$$\rho(e)(v \otimes w) = (v \otimes w)$$

for any pure tensor $v \otimes w \in V \otimes W$. Thus, because maps out of $V \otimes W$ are determined by their action on pure tensors, we see that $\rho(e) = \mathrm{id}$. Similarly, for $g, h \in G$, we see that

$$\rho(gh)(v \otimes w) = (\rho_V(g)\rho_V(h)v \otimes \rho_W(g)\rho_W(h)w) = \rho(g)\rho(h)(v \otimes w),$$

so $\rho(gh)$ and $\rho(g) \circ \rho(h)$ are equal on pure tensors and hence equal as maps $V \otimes W \to V \otimes W$.

- Regular: we expand everything on a basis. Fix a basis $\{e_1, \ldots, e_m\}$ of V and $\{f_1, \ldots, f_m\}$ on W so that $\{e_i \otimes f_j\}_{i,j}$ is a basis of $V \otimes W$; let $\mathrm{pr}_\bullet$ be the appropriate projection whenever it appears. The previous step verifies that we have a group homomorphism $\rho\colon G \to \mathrm{GL}(V \otimes W)$, which we must now show to be regular. Notably, the matrix coefficients $\rho(g)_{i_1j_1,i_2j_2}$ of $\rho(g)$ are now computable as

$$\mathrm{pr}_{i_2j_2}\,\rho(g)(e_{i_1} \otimes f_{j_1}) = \mathrm{pr}_{i_2j_2}(\rho_V(g)e_{i_1} \otimes \rho_W(g)f_{j_1}) = \rho_V(g)_{i_1i_2}\rho_W(g)_{j_1j_2},$$

which is a product of regular functions and hence regular. Thus, ρ is regular on coordinates and hence regular.

(b) For each X, we need to provide a bilinear map $\rho(X)\colon (V \otimes W) \to (V \otimes W)$, for which we take

$$\rho(X)(v, w) \coloneqq \rho_V(X)v \otimes w + v \otimes \rho_W(X)w.$$

Linearity of $\rho_V(X)$ and $\rho_W(X)$ (and properties of the tensor product) verify that we have in fact defined a bilinear map, so we have in fact defined a map $\mathfrak{g} \to \mathfrak{gl}(V \otimes W)$. Here are our checks to make this a representation.

- Linear: for $a, b \in \mathbb{F}$ and $X, Y \in \mathfrak{g}$, we should check that $\rho(aX+bY) = a\rho(X)+b\rho(Y)$. Because pure tensors span $V \otimes W$, it is enough to check this equality on pure tensors, for which we compute

$$\begin{aligned}\rho(aX + bY)(v \otimes w) &= \rho_V(aX + bY)v \otimes w + v \otimes \rho_W(aX + bY)w \\ &= a(\rho_V(X)v \otimes w + v \otimes \rho_W(X)w) + b(\rho_V(Y)v \otimes w + v \otimes \rho_W(Y)w) \\ &= (a\rho(X) + b\rho(Y))(v \otimes w).\end{aligned}$$

- Lie bracket: for $X, Y \in \mathfrak{g}$, we need to check $[\rho(X), \rho(Y)] = \rho([X, Y])$. It is enough to check this on pure tensors, for which we compute

$$\begin{aligned}[\rho(X), \rho(Y)](v \otimes w) &= (\rho(X)\rho(Y) - \rho(Y)\rho(X))(v \otimes w) \\ &= \rho(X)\rho(Y)(v \otimes w) - \rho(Y)\rho(X)(v \otimes w) \\ &= \rho_V(X)\rho_V(Y)v \otimes w - \rho_V(Y)v \otimes \rho_W(X)w \\ &\qquad - \rho_V(X)v \otimes \rho_W(Y) + v \otimes \rho_W(X)\rho_W(Y)w \\ &\qquad - \rho_V(Y)\rho_V(X)v \otimes w + \rho_V(X)v \otimes \rho_W(Y)w \\ &\qquad + \rho_V(Y)v \otimes \rho_W(X)w - v \otimes \rho_W(Y)\rho_W(X)w \\ &= (\rho_V(X)\rho_V(Y) - \rho_V(Y)\rho_V(X))v \otimes w \\ &\qquad + v \otimes (\rho_W(X)\rho_W(Y) - \rho_W(Y)\rho_W(X))w \\ &= \rho([X, Y])(v \otimes w).\end{aligned}$$

(c) This is a direct computation. Given the representations $\rho_V\colon G \to \mathrm{GL}(V)$ and $\rho_W\colon G \to \mathrm{GL}(W)$, we would like to compute $(d\rho_{V\otimes W})_e(X) \in \mathfrak{gl}(V \otimes W)$ for some $X \in \mathfrak{g}$. Well, it is enough to compute this on pure tensors $v \otimes w$, for which we note that evaluation is a linear map and hence can be moved inside a derivative in the computation

$$\begin{aligned}(d\rho_{V\otimes W})_e(X)(v\otimes w) &= \frac{d}{dt}\rho_{V\otimes W}(\exp(tX))\bigg|_{t=0}(v\otimes w)\\ &= \frac{d}{dt}\rho_{V\otimes W}(\exp(tX))(v\otimes w)\bigg|_{t=0}\\ &= \frac{d}{dt}\rho_V(\exp(tX))v\otimes\rho_W(\exp(tX))w\bigg|_{t=0}\\ &= \frac{d}{dt}(1+td\rho_V(X)+\cdots)v\otimes\rho_W(1+td\rho_W(X)+\cdots)w\bigg|_{t=0}\\ &= d\rho_V(X)v\otimes w+v\otimes d\rho_W(X)w,\end{aligned}$$

as required. Notably, we expanded our the Taylor series in order to the computation of the derivative at $t = 0$, but one can also indirectly apply some product rule after working more explicitly with coordinates. ■

Remark 3.9. By induction, we see that we can also define a tensor representation

$$V_1\otimes\cdots\otimes V_k$$

for any finite number of representations $V_1,\ldots,V_k$. One can compute the actions by simply extending the above ones to more terms inductively.

Example 3.10. We explain how to twist by a character.

- Fix a regular Lie group G. Given a representation $\rho\colon G\to\mathrm{GL}(V)$ and a character $\chi\colon G\to\mathrm{GL}_1(\mathbb{F})$, we see that we have a representation $\chi\otimes\rho$ on $\mathbb{F}\otimes V$. However, $\mathbb{F}\otimes V$ can be identified with V by the map $c\otimes v\mapsto cv$ (on pure tensors), so we have really defined a representation $\chi\rho\colon G\to\mathrm{GL}(V)$ given by
$$(\chi\rho)(g)\coloneqq\chi(g)\rho(g).$$
- Fix a Lie algebra $\mathfrak{g}$. Given a representation $\rho\colon\mathfrak{g}\to\mathfrak{gl}(V)$ and a character $\chi\colon\mathfrak{g}\to\mathfrak{gl}(\mathbb{F})$, we again see that we have a representation $\chi\otimes\rho$ on $\mathbb{F}\otimes V$. Identifying $\mathbb{F}\otimes V$ with V as before, we see that we have defined a representation $\chi\rho\colon G\to\mathrm{GL}(V)$ by
$$(\chi\rho)(X)=\chi(X)+\rho(X).$$

We also have Hom sets.

Lemma 3.11. Fix a Lie group G and Lie algebra $\mathfrak{g}$.

(a) Given representations $V,W\in\mathrm{Rep}_k(G)$, we can make $\mathrm{Hom}(V,W)$ into a representation of G via
$$\rho_{\mathrm{Hom}(V,W)}(g)\varphi\coloneqq\rho_W(g)\cdot\varphi\circ\rho_V(g)^{-1}.$$

(b) Given representations $V,W\in\mathrm{Rep}_k(\mathfrak{g})$, we can make $\mathrm{Hom}(V,W)$ into a representation of G via
$$\rho_{\mathrm{Hom}(V,W)}(X)\varphi\coloneqq\rho_W(X)\circ\varphi-\varphi\circ\rho_V(X).$$

(c) Suppose $\mathfrak{g}=\mathrm{Lie}(G)$. Given representations $V,W\in\mathrm{Rep}_k(G)$, then $F(\mathrm{Hom}(V,W))$ is the corresponding in $\mathrm{Rep}_k(\mathfrak{g})$.

Proof. Here we go.

(a) Given finite-dimensional representations $\rho_V \colon G \to \mathrm{GL}(V)$ and $\rho_W \colon G \to \mathrm{GL}(W)$, we explain how to build a representation $\rho\colon G \to \mathrm{Hom}(V,W)$. Indeed, for $g \in G$ and $\varphi \in \mathrm{Hom}(V,W)$, define

$$(\rho(g)\varphi)(v) \coloneqq \rho_W(g)\varphi\left(\rho_V(g)^{-1}v\right).$$

In other words, $\rho(g)\varphi = \rho_W(g) \circ \varphi \circ \rho(g)^{-1}$. Here are our checks.

- Group action: for the identity check, we note

$$\rho(e)\varphi = \rho_W(e) \circ \varphi \circ \rho_V(e)^{-1} = \mathrm{id}_W \circ \varphi \circ \mathrm{id}_V^{-1},$$

  as required. For the associativity check, we choose $g, h \in G$ and note

$$\rho(g)\rho(h)\varphi = \rho_W(g) \circ \rho_W(h) \circ \varphi \circ \rho(h)^{-1} \circ \rho(g)^{-1} = \rho(gh)\varphi.$$

- Regular: it is enough to show that we have given a regular map $G \times \mathrm{Hom}(V,W) \to \mathrm{Hom}(V,W)$ by considering component-wise formulations of matrix entries. Well, our map is simply the composite

$$\begin{array}{ccccc} G \times \mathrm{Hom}(V,W) & \to & \mathrm{GL}(W) \times \mathrm{Hom}(V,W) \times \mathrm{GL}(V) & \mapsto & \mathrm{Hom}(V,W) \\ (g,\varphi) & \mapsto & \left(\rho_W(g), \varphi, \rho_V(g)^{-1}\right) & \mapsto & \rho_W(g) \circ \varphi \circ \rho_V(g)^{-1} \end{array}$$

  which is regular as the composite of (products of) regular maps. For example, the last map is regular because it is simply matrix multiplication, which is polynomial on coordinates and hence regular.

(b) For any two Lie algebra representations V and W of $\mathfrak{g}$, we note that $\mathrm{Hom}(V,W)$ also has a Lie algebra representation structure given by

$$\rho_{\mathrm{Hom}(V,W)}(X)\varphi \coloneqq \rho_W(X) \circ \varphi - \varphi \circ \rho_V(X).$$

Anyway, we now run our checks. Certainly $\rho_{\mathrm{Hom}(V,W)}(X)$ is a linear map $\mathrm{Hom}(V,W) \to \mathrm{Hom}(V,W)$ (namely, our construction is linear in φ) because composition distributes over addition. Additionally, our construction is linear in X because $\rho_W \colon \mathfrak{g} \to \mathfrak{gl}(W)$ and $\rho_V \colon \mathfrak{g} \to \mathfrak{gl}(V)$ should be linear. Lastly, we must check preservation of the bracket of our map $\mathfrak{g} \to \mathfrak{gl}(\mathrm{Hom}(V,W))$. Well, given $\varphi, \psi \in \mathrm{Hom}(V,W)$ and $X, Y \in \mathfrak{g}$, we compute

$$\begin{aligned} [\rho_{\mathrm{Hom}(V,W)}(X), \rho_{\mathrm{Hom}(V,W)}(Y)](\varphi) &= \rho_{\mathrm{Hom}(V,W)}(X) \circ \rho_{\mathrm{Hom}(V,W)}(Y)(\varphi) \\ &\quad - \rho_{\mathrm{Hom}(V,W)}(Y) \circ \rho_{\mathrm{Hom}(V,W)}(X)(\varphi) \\ &= \rho_{\mathrm{Hom}(V,W)}(X)(Y \circ \varphi - \varphi \circ Y) \\ &\quad - \rho_{\mathrm{Hom}(V,W)}(Y)(X \circ \varphi - \varphi \circ X) \\ &= X \circ Y \circ \varphi - Y \circ \varphi \circ X - X \circ \varphi \circ Y + \varphi \circ Y \circ X \\ &\quad - Y \circ X \circ \varphi + X \circ \varphi \circ Y + Y \circ \varphi \circ X - \varphi \circ X \circ Y \\ &= (X \circ Y - Y \circ X) \circ \varphi - \varphi \circ (X \circ Y - Y \circ X) \\ &= \rho_W([X,Y]) \circ \varphi - \varphi \circ \rho_V([X,Y]) \\ &= \rho_{\mathrm{Hom}(V,W)}([X,Y])(\varphi), \end{aligned}$$

where we have frequently but not always omitted our ρ_Vs and ρ_Ws.

(c) This is a direct computation. If $\varphi\colon V \to W$ were already a morphism of G-representations, then the action of (b) is simply $d\rho_{\mathrm{Hom}(V,W)}(X)(\varphi)$: indeed, the action should be

$$\begin{aligned} d\rho_{\mathrm{Hom}(V,W)}(X)(\varphi) &= \frac{d}{dt}\rho_{\mathrm{Hom}(V,W)}(\exp(tX))\varphi\bigg|_{t=0} \\ &= \frac{d}{dt}\rho_W(\exp(tX)) \circ \varphi \circ \rho_V(\exp(-tX))\bigg|_{t=0} \\ &= \frac{d}{dt}(1 + d\rho_W(tX) + \cdots) \circ \varphi \circ (1 - d\rho_W(tX) + \cdots)\bigg|_{t=0} \\ &= d\rho_W(X) \circ \varphi - \varphi \circ d\rho_V(X). \end{aligned}$$

As usual, $+\cdots$ denotes higher-order terms which cannot affect our derivative. Notably, we are using the fact that the linear term of a Taylor expansion (into some Euclidean space) is given by the derivative. ■

Example 3.12. Taking $W = k$ to be the trivial representation, we obtain duals as a special case of Lemma 3.11.

Here is also a good notion of subobjects.

Definition 3.13 (suprepresentation)**.** Fix a regular Lie group or Lie algebra. A *subrepresentation* of a representation is a subspace preserved by the G-action.

Remark 3.14. Let's make this notion more precise.

- For a regular Lie group G, we see that a subspace $U \subseteq V$ preserved by the G-action on a representation $\rho_V : G \to \mathrm{GL}(V)$ means that we can restrict the linear action map $G \times U \to V$ to an action $G \times U \to U$. Thus, we do indeed have a regular map $\rho_U : G \to \mathrm{GL}(U)$ by computing coordinates of matrices component-wise, and the natural inclusion map $U \hookrightarrow V$ is a morphism in $\mathrm{Rep}_k(G)$.
- For a Lie algebra $\mathfrak{g}$, we see that a subspace $U \subseteq V$ preserved by the $\mathfrak{g}$-action on a representation $\rho_V : \mathfrak{g} \to \mathfrak{gl}(V)$ means that we can restrict this linear map to $\rho_U : \mathfrak{g} \to \mathfrak{gl}(U)$. Notably, the Lie bracket of $\mathfrak{gl}(U)$ is more or less the restriction of the Lie bracket on $\mathfrak{gl}(V)$, so ρ_U continues to be a Lie algebra representation, and we see that the natural inclusion $U \hookrightarrow V$ is a morphism in $\mathrm{Rep}_k(\mathfrak{g})$.

Example 3.15. Let $\varphi \colon V \to W$ be a morphism in $\mathrm{Rep}_k(G)$. Then $\ker\varphi$ is a subrepresentation of V. Indeed, $\ker\varphi \subseteq V$ is certainly a linear subspace, and for the G-invariance, we note that any $v \ker\varphi$ has

$$\varphi(\rho_V(g)v) = \rho_W(g)(\varphi(v)) = 0$$

for any $g \in G$, so $\rho_V : G \to \mathrm{GL}(V)$ restricts to a subrepresentation $\rho_{\ker\varphi} \colon G \to \mathrm{GL}(\ker\varphi)$. (More precisely, we have restricted our regular action $G \times \ker\varphi \to V$ to a regular action $G \times \ker\varphi \to \ker\varphi$.)

Example 3.16. Let $\varphi \colon V \to W$ be a morphism in $\mathrm{Rep}_k(\mathfrak{g})$. Again, we see that $\ker\varphi \subseteq V$ is a subrepresentation for essentially the same reason: certainly $\ker\varphi \subseteq V$ is a linear subspace, and $v \in \ker\varphi$ has $\varphi(Xv) = X(\varphi(v)) = 0$ for any $X \in \mathfrak{g}$, so $\ker\varphi$ is closed under the G-action.

Example 3.17. Let $\varphi \colon V \to W$ be a morphism in $\mathrm{Rep}_k(G)$. Then $\operatorname{im}\varphi$ is a subrepresentation of W. Again, it is certainly a linear subspace, and it is preserved by the G-action because any $g \in G$ and $\varphi(v) \in \operatorname{im}\varphi$ has

$$\rho_W(g)(\varphi(v)) = \varphi(\rho_V(g)v) \in \operatorname{im}\varphi.$$

Example 3.18. Let $\varphi \colon V \to W$ be a morphism in $\mathrm{Rep}_k(\mathfrak{g})$. Then $\operatorname{im}\varphi$ is a subrepresentation of W. As usual, we have a linear subspace, and it is fixed by the G-action because $X \in \mathfrak{g}$ and $\varphi(v) \in \operatorname{im}\varphi$ has $X \cdot \varphi(v) = \varphi(X \cdot v) \in \operatorname{im}\varphi$.

We take a moment to remark that one can also construct quotients of representations, but we won't bother to run all the checks.

Invariants provide an important example of subrepresentations.

Definition 3.19 (invariants)**.** Fix a regular Lie group G or Lie algebra $\mathfrak{g}$.

- We denote the G-invariants of a representation $V \in \operatorname{Rep}_k(G)$ by
$$V^G := \{v \in V : \rho_V(g)v = v \text{ for all } g \in G\}.$$
- We denote the $\mathfrak{g}$-invariants of a representation $V \in \operatorname{Rep}_k(\mathfrak{g})$ by
$$V^{\mathfrak{g}} := \{v \in V : \rho_V(g)v = 0 \text{ for all } X \in \mathfrak{g}\}.$$

Remark 3.20. We won't bother to check that invariants provide subrepresentations right now. It follows from the more general Lemma 3.24.

Example 3.21. Note that $\operatorname{Hom}(V,W)^G = \operatorname{Hom}_G(V,W)$ for G-representations V and W. Indeed, a linear map $\varphi \in \operatorname{Hom}(V,W)$ is fixed by the G-action if and only if
$$g^{-1} \cdot \varphi(g \cdot v) = \left(g^{-1}\varphi\right)(v) = \varphi(v)$$
for all $g \in G$, which of course rearranges into φ being G-equivariant.

Example 3.22. Note again note that $\operatorname{Hom}(V,W)^{\mathfrak{g}} = \operatorname{Hom}_{\mathfrak{g}}(V,W)$ for this Lie algebra representation structure. Namely, we can see that $X \cdot \varphi(v) = \varphi(X \cdot v)$ for any X and v if and only if $X\varphi = 0$.

Example 3.23. Let V be a vector space, and fix a nonnegative integer $k \geq 0$. Then S_k acts on the tensor power $V^{\otimes k}$ by permuting the coordinates. Explicitly, permuting the coordinates provides a bilinear map $V^k \to V^{\otimes k}$, so it extends to a linear map $V^{\otimes k} \to V^{\otimes k}$ defined by
$$\sigma\colon (v_1 \otimes \cdots \otimes v_k) \mapsto (v_{\sigma 1} \otimes \cdots \otimes v_{\sigma k})$$
for any pure tensor. We won't bother to check that this is actually a group action, though it is not a lengthy check. The fixed points of this S_k-action is the symmetric power $\operatorname{Sym}^k(V)$.

Here is a more general notion of invariants.

Lemma 3.24. Fix a regular Lie group G or Lie algebra $\mathfrak{g}$, and let $X^*(G)$ and $X^*(\mathfrak{g})$ denote the set of regular homomorphisms $G \to \mathbb{F}^\times$ and $\mathfrak{g} \to \mathfrak{gl}(\mathbb{F})$, respectively. For the statements, select $R \in \{G, \mathfrak{g}\}$.

(a) Fix some $\chi \in X^*(R)$. For any representation V, the subspace
$$V^\chi := \{v \in V : \rho_V(r)v = \chi(r)v \text{ for all } r \in R\}$$
is a subrepresentation of V.

(b) For distinct characters $\chi_1, \ldots, \chi_k \in X^*(R)$ and any representation V, the subspaces $V^{\chi_1}, \ldots, V^{\chi_k}$ are linearly disjoint.

Proof. Here we go.

(a) Note that V^χ is the kernel of the family of linear maps $V \to V$ defined by $\{v \mapsto \rho_V(r)v - \chi(r)v\}_{r \in R}$, so V^χ is the intersection of linear subspaces and hence a linear subspace. To see that V^χ is preserved by the G-action, we note that any $v \in V^\chi$ and $r \in R$ will have $\rho_V(r)v \in V^\chi$: for any $s \in R$, we see
$$\rho_V(s)\rho_V(r) = \chi(r)\rho_V(s)v = \chi(s)\chi(r)v = \chi(s)\rho_V(r)v.$$

(b) Suppose for the sake of contradiction that there exists a nontrivial relation $v_1 + \cdots + v_k = 0$ where $v_i \in V^{\chi_i}$ for $i \in \{1, \ldots, k\}$. By possibly making k smaller, we may assume that all the $v_\bullet$s are nonzero, and in fact, we may assume that there does not exist such a relation with fewer than k characters $\chi_1, \ldots, \chi_k \in X^*(R)$. Now, if $k = 1$, then we are simply asserting that $v_1 = 0$, so there is nothing to say. Otherwise, we may assume that $k > 1$. Then there is $r \in R$ such that $\chi_k(r) \neq \chi_1(r)$, and we see that multiplying our relation by $\rho_V(r)$ produces the equation

$$\chi_1(r)v_1 + \cdots + \chi_k(r)v_k = 0.$$

But now we can subtract this relation from $\chi_k(r)v_k + \cdots + \chi_k(r)v_k = 0$, which produces a strictly smaller relation with at least one term $(\chi_1(r) - \chi_k(r))v$, which is a contradiction to the minimality of our relation. ■

Remark 3.25. If G is a finite group acting on a vector space V, and χ is a character of G, then we can define an operator $\pi_\chi \colon V \to V$ by

$$\pi_\chi(v) := \frac{1}{|G|} \sum_{g \in G} \chi^{-1}(g)\rho(g)v.$$

We have the following checks on π_χ.

- Note π_χ is a linear map (as the sum of linear maps).
- By rearranging the sum, we see that $\rho_V(h)\pi_\chi(v) = \chi(h)\pi_\chi(v)$ for any $h \in G$, so $\operatorname{im}\pi_\chi \subseteq V^\chi$.
- On the other hand, if $v \in V^\chi$ already, then $\pi_\chi(v)$ is just a sum with $|G|$ copies of v, so π_χ fixes V^χ pointwise.

In conclusion, we see that $\operatorname{im}\pi_\chi = V^\chi$ by Example 3.17. This is an alternate way to see that V^χ is a subrepresentation.

Example 3.26. Suppose V is a representation of a regular Lie group G or Lie algebra $\mathfrak{g}$. Given some nonnegative integer k, we recall that S_k acts on V. Thus, for a character χ of S_k, we note that the map $\pi_\chi \colon V^{\otimes k} \to V^{\otimes k}$ is a projection. In fact, π_χ respects the ambient action on V.

- In the case of a Lie group G, we see that
$$\rho_{V^{\otimes k}}(g)\pi_\chi(v_1 \otimes \cdots \otimes v_k) = \frac{1}{k!}\sum_{\sigma \in S_k} \chi(\sigma)(\rho_V(g)v_{\sigma 1} \otimes \cdots \otimes \rho_V(g)v_{\sigma k}) = \pi_\chi \rho_{V^{\otimes k}}(g)(v_1 \otimes \cdots \otimes v_k),$$
so the equality $\rho_{V^{\otimes k}}(g) \circ \pi_\chi = \pi_\chi \circ \rho_{V^{\otimes k}}(g)$ follows by linearity.
- In the case of Lie algebra $\mathfrak{g}$, we see that
$$\begin{aligned}\rho_{V^{\otimes k}}(X)\pi_\chi(v_1 \otimes \cdots \otimes v_k) &= \frac{1}{k!}\sum_{\sigma \in S_k} \chi(\sigma)\rho_{V^{\otimes k}}(X)(v_{\sigma 1} \otimes \cdots \otimes v_{\sigma k}) \\ &= \frac{1}{k!}\sum_{\sigma \in S_k} \chi(\sigma)(Xv_{\sigma 1} \otimes \otimes \cdots \otimes v_{\sigma k} + \cdots + v_{\sigma 1} \otimes \cdots \otimes Xv_{\sigma k}) \\ &= \pi_\chi \pi_{V^{\otimes k}}(g)(v_1 \otimes \cdots \otimes v_k).\end{aligned}$$

Thus, $\left(V^{\otimes k}\right)^\chi \subseteq V^{\otimes k}$ continues to be a subrepresentation in all cases by Example 3.17. When $\chi = 1$, this is the symmetric power representation $\operatorname{Sym}^k(V)$. When $\chi = \operatorname{sgn}$, this is the alternating representation $\operatorname{Alt}^k(V)$.

3.1.3 Lie's Theorems for Representation Theory

We now discuss how to pass the representation theory for G to the representation theory of $\mathfrak{g}$. We want the following lemma.

Lemma 3.27. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. For a representation $V \in \operatorname{Rep}_k(G)$, give V the natural $\mathfrak{g}$-action via $d\rho$. Further, fix some character $\chi\colon G \to \operatorname{GL}_1(\mathbb{F})$ inducing a character $d\chi\colon \mathfrak{g} \to \mathfrak{gl}(\mathbb{F})$.

(a) We always have $V^\chi \subseteq V^{d\chi}$.

(b) If G is connected, then $V^\chi = V^{d\chi}$.

Proof. Quickly, we reduce to the case where $\chi = 1$ and thus $d\chi = 0$ (because χ is constant). By Example 3.10, we may consider the representation $\chi^{-1}\rho$. On one hand, we see that $v \in V^\chi$ if and only if $\left(\chi^{-1}\rho\right)(g)v = v$ for all $g \in G$; on the other hand, we see similarly that $v \in V^{d\chi}$ if and only if $d\left(\chi^{-1}\rho\right)(X)v = v$ for all $X \in \mathfrak{g}$. Thus, for our arguments, we will take $\chi = 1$ so that we may consider V^G and $V^{\mathfrak{g}}$.

(a) For $v \in V^G$, we must show that $v \in V^{\mathfrak{g}}$. Well, fix any $X \in \mathfrak{g}$, and we would like to show that $d\rho_e(X)(v) = v$. For this, define the path $\gamma\colon \mathbb{F} \to \mathfrak{g}$ given by $\gamma(t) := \exp(tX)$ so that $\gamma(0) = e$ and $\gamma'(0) = X$. Then
$$d\rho_e(X) = d\rho_e(\gamma'(0)) = (\rho \circ \gamma)'(0) \in \mathfrak{gl}(V)$$
by the chain rule, where this last derivative makes technical sense because we are outputting to a Euclidean space. To compute $(\rho \circ \gamma)'(0)(v)$, we note that applying an endomorphism in $\mathfrak{gl}(V)$ to a vector $v \in V$ is a linear map, and linear maps pass through the definition of the derivative, so we find that
$$\begin{aligned}(\rho\circ\gamma)'(0)(v) &= \frac{d}{dt}(\rho\circ\gamma)(t)\bigg|_{t=0}(v)\\ &= \frac{d}{dt}(\rho\circ\gamma)(t)(v)\bigg|_{t=0}\\ &= \frac{d}{dt}\rho(\exp(tX))(v)\bigg|_{t=0}\\ &\overset{*}{=} \frac{d}{dt}v\bigg|_{t=0}\\ &= 0,\end{aligned}$$
where $\overset{*}{=}$ holds because $v \in V^G$.

(b) We already showed one inclusion in (a), so now we just have to show that any $v \in V^{\mathfrak{g}}$ is fully fixed by G. Well, let $H \subseteq G$ be the subgroup of G stabilizing v, which we know to be a closed Lie subgroup. In fact, by our more precise isomorphism theorem, we know that its Lie algebra $\mathfrak{h}$ can be described by
$$\mathfrak{h} = \{X \in \mathfrak{g} : (\rho_* X)_v = 0\}.$$
However, we can compute
$$(\rho_*X)_v f = \frac{d}{dt} f(\exp(-tX)v)\bigg|_{t=0} = df_p((\rho\circ\exp)'(0)v) = df_p(d\rho_e(X)v)$$
for any germ f, but this derivative is of course 0 because $d\rho_e(X)v = 0$ for all $X \in \mathfrak{g}$ by assumption. Thus, $\mathfrak{h} = \mathfrak{g}$, so the exponential map $\exp\colon \mathfrak{h} \to G$ will be a local diffeomorphism. In particular, H contains in an open neighborhood of the identity, so H must equal G because G is connected. Thus, $v \in V^G$. ■

And here is our main result.

Proposition 3.28. Fix a regular Lie group G with Lie algebra $\mathfrak{g}$. Recall the functor $F\colon \operatorname{Rep}(G) \to \operatorname{Rep}(\mathfrak{g})$ sending a representation $\rho\colon G \to \operatorname{GL}(V)$ to $d\rho_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

(a) If G is connected, then F is fully faithful.

(b) If G is connected and simply connected, then F is essentially surjective and hence an equivalence.

Proof. Here we go.

(a) Suppose that G is connected, and we want to show that F is fully faithful. In Lemma 3.2, we showed that F is faithful, so we now must show that F is full. Well, for G-representations V and W, we must show that any linear map $\varphi \in \operatorname{Hom}_{\mathfrak{g}}(V, W)$ is in fact G-invariant. Well, we simply note that

$$\operatorname{Hom}_{\mathfrak{g}}(V, W) = \operatorname{Hom}(V, W)^{\mathfrak{g}},$$

which equals $\operatorname{Hom}(V, W)^G = \operatorname{Hom}_G(V, W)$ by Lemma 3.27, so we are done.

(b) Suppose that G is connected and simply connected. We want to show that F is essentially surjective in order to finish the proof that F is an equivalence of categories. Well, fix a representation of $\mathfrak{g}$ given by some Lie algebra homomorphism $\overline{\rho}\colon \mathfrak{g} \to \mathfrak{gl}(V)$. Then Theorem 2.69 tells us that the differential provides a bijection

$$F\colon \operatorname{Hom}_{\mathrm{LieGrp}}(G, \operatorname{GL}(V)) \cong \operatorname{Hom}_{\mathrm{LieAlg}}(\mathfrak{g}, \mathfrak{gl}(V))$$

because G is simply connected. In particular, there is a Lie algebra homomorphism $\rho\colon G \to \operatorname{GL}(V)$ such that $\overline{\rho} = F\rho$, as required. ■

Remark 3.29. Given any connected Lie group G with universal cover $\widetilde{G}$, one can attempt to recover the representation theory of G from $\widetilde{G}$ via the short exact sequence in Remark 1.132.

Remark 3.30. For any Lie algebra $\mathfrak{g}$ over $\mathbb{R}$, one sees that $\operatorname{Rep}_{\mathbb{C}}(\mathfrak{g}) = \operatorname{Rep}_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}})$. (Another perspective is that we can reduce the complex representation theory of a complex Lie algebra to a real form.) To see this, note that any morphism $\mathfrak{g} \to \mathfrak{gl}(V)$ where V is a complex vector space canonically upgrades to a map $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{gl}(V)$ by taking the tensor product with the canonical inclusion $\mathbb{C} \to \mathfrak{gl}(V)$ given by $c \mapsto c\,\mathrm{id}_V$. In the reverse direction, any representation $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{gl}(V)$ can simply forget about the $\mathbb{C}$ factor to define a representation $\mathfrak{g} \to \mathfrak{gl}(V)$. This remark does not have the space or motivation to check that we have actually defined an equivalence.

3.1.4 Decomposing Representations

With direct sums, we have notions of irreducibility.

Definition 3.31 (indecomposable)**.** Fix a representation V. Then V is *indecomposable* if and only if any direct sum decomposition $V = V_1 \oplus V_2$ must have $V_1 = 0$ or $V_2 = 0$.

Definition 3.32 (irreducible)**.** Fix a representation V. Then V is *irreducible* if and only if V is nonzero, and any subrepresentation $U \subseteq V$ has $U = 0$ or $U = V$.

Example 3.33. The standard representation V of $\operatorname{GL}(V)$ is irreducible. Indeed, any nonzero subrepresentation $U \subseteq V$ has a nonzero vector $v \in U$. But then the orbit of v under $\operatorname{GL}(V)$ is $V \setminus \{0\}$, so U must contain $V \setminus \{0\}$, so $U = V$.

It will also turn out that $\operatorname{Sym}^k(V)$ and $\operatorname{Alt}^k(V)$ are irreducible representations of $\operatorname{GL}(V)$, but this is not so obvious. We will be able to show this with more ease later in the course.

These notions are related but not the same.

Remark 3.34. Any irreducible representation V is indecomposable. Indeed, writing $V = V_1 \oplus V_2$ has $V_1 \subseteq V$, so $V_1 = 0$ or $V_1 = V$.

Example 3.35. Consider the representation $\rho\colon \mathbb{C} \to \mathrm{GL}_2(\mathbb{C})$ given by

$$\rho(x) := \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}.$$

Then $\operatorname{span}\{e_1\}$ is a nontrivial proper subrepresentation of ρ because $\rho(e_1) = e_1$; thus, ρ fails to be irreducible.

However, ρ is indecomposable! Indeed, our vector space is two-dimensional, so a nontrivial decomposition of ρ into $\rho_1 \oplus \rho_2$ must have underlying vector spaces V_1 and V_2 with dimensions $\dim V_1 = \dim V_2 = 1$. But the action of $\mathbb{C}$ on $\mathbb{C}$ must be linear, so V_1 and V_2 must be eigenspaces. As such, we can see from the definition of ρ that all eigenvalues are 1, so ρ_1 and ρ_2 would have to be the trivial representation, meaning that ρ would have to be the sum of trivial representations and hence trivial, which is false because $\rho(1) \neq \mathrm{id}$.

Do note that there is something that we can always do for our decomposition, but it is not always as satisfying as a direct sum.

Remark 3.36 (Jordan–Holder)**.** For any representation V, one can always find a filtration

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_{k-1} \subseteq V_k = V$$

where each quotient V_i/V_{i-1} is irreducible. Indeed, we can proceed by induction on $\dim V$. As a base case, $\dim V = 1$ has nothing to do because V is irreducible for dimension reasons.

If V is already irreducible, then our filtration is $0 \subseteq V$. Otherwise, V is not irreducible, so we can find a nontrivial proper subrepresentation $V' \subseteq V$; choosing a minimal such representation (by dimension) must have V' irreducible. Then we can apply the inductive hypothesis to V/V' (which has smaller dimension than V) to build the required filtration.

In particular, filtrations means that we would have to build representations by short exact sequences, which may be difficult to handle especially when iterated.

We would like to decompose representations into irreducible parts because dealing with filtrations is difficult.

Definition 3.37 (completely reducible)**.** A representation V is *completely reducible* if and only if it is the direct sum of irreducible representations.

Remark 3.38. Technically, we have not required that the decomposition into irreducibles is unique. This is the content of Corollary 3.48.

3.2 October 2

Today we will continue talking about representations.

3.2.1 Schur's Lemma

The following result is our first interesting result about representations.

Proposition 3.39 (Schur's lemma). Fix representations V and W (over a field F) over a Lie group G or Lie algebra $\mathfrak{g}$, denoted R.

(a) If V is irreducible, then any nonzero morphism $\varphi\colon V \to W$ is injective.

(b) If W is irreducible, then any nonzero morphism $\varphi\colon V \to W$ is surjective.

(c) If V and W are both irreducible, then any nonzero morphism $\varphi\colon V \to W$ is an isomorphism.

(d) If V and W are both irreducible, then the endomorphism algebra $\operatorname{End}_R(V)$ is a finite-dimensional division algebra over F. In particular, if F is algebraically closed, then the map $F \cong \operatorname{End}_R(V)$ defined by $\lambda \mapsto \lambda \operatorname{id}_V$ is a ring isomorphism.

Proof. Here we go.

(a) Note that $\ker\varphi \subseteq V$ is a subrepresentation by Examples 3.15 and 3.16. Thus, either $\ker\varphi = 0$ in which case φ is injective, or $\ker\varphi = V$ in which case $\varphi = 0$.

(b) Note that $\operatorname{im}\varphi \subseteq V$ is a subrepresentation by Examples 3.17 and 3.18. Thus, either $\operatorname{im}\varphi = 0$ in which case $\varphi = 0$, or $\operatorname{im}\varphi = W$ in which case φ is surjective.

(c) This follows by combining the previous two parts with Remark 3.1.

(d) Note that $\operatorname{End}_R(V)$ is certainly an algebra (possibly non-commutative). Part (c) explains that all nonzero elements have inverses, so this algebra becomes a division algebra. It remains to check the claim when F is algebraically closed. In fact, we show that any morphism $\varphi\colon V \to V$ must be a scalar, which will complete the proof because it shows that the natural map

$$k \to \operatorname{End}_R(V)$$

given by $c \mapsto c\operatorname{id}_V$ is an isomorphism.[1] Note that φ will have an eigenvector v with eigenvalue λ. Then $\varphi - \lambda\operatorname{id}_V$ is a morphism with a nontrivial kernel, so it must be the zero map because it is not an isomorphism! Thus, we conclude that $\varphi = \lambda\operatorname{id}_V$ is a scalar. ■

This result (and in particular (d)) is important enough to warrant its own subsection. To explain why, here are some interesting corollaries.

Corollary 3.40. Fix an algebraically closed field F.

(a) For any injective irreducible representation $\rho\colon G \to \mathrm{GL}(V)$ of a regular Lie group G, the center of G is
$$Z(G) = \{g \in G : \rho(g) = \lambda\mathrm{id}_V \text{ for some } \lambda \in \mathbb{C}\}.$$

(b) For any injective representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of a Lie algebra $\mathfrak{g}$ over a field F. Then the center of $\mathfrak{g}$ is
$$\mathfrak{z}(\mathfrak{g}) = \{X \in \mathfrak{g} : \rho(X) = \lambda\mathrm{id}_V \text{ for some } \lambda \in k\}.$$

Proof. The point is that living in the center implies commuting with the ambient action, which Proposition 3.39 explains implies the element must be a scalar. The injectivity of the representations implies that this characterizes the center.

[1] Certainly this is a ring map, and it is injective because V is nonzero, so we are really interested in showing that this map is surjective.

(a) In one direction, if $g \in G$ has $\rho(g) = \lambda \operatorname{id}_V$ for some $\lambda \in \mathbb{C}$, then any $h \in G$ has

$$\begin{aligned}\rho\left(hgh^{-1}\right) &= \rho(h)\rho(g)\rho(h)^{-1} \\ &= \rho(h) \circ \lambda \operatorname{id}_V \circ \rho(h)^{-1} \\ &= \lambda\rho(h)\rho(h)^{-1} \\ &= \lambda \operatorname{id}_V \\ &= \rho(g),\end{aligned}$$

so injectivity of ρ implies that $hgh^{-1} = g$; thus, $g \in Z(G)$.

Conversely, suppose $g \in Z(G)$. Then $\rho(g)\colon V \to V$ is an operator on an irreducible representation of G. In fact, $\rho(g)$ commutes with the action of G: for any $h \in G$, we see that

$$\rho(g) \circ \rho(h) = \rho(gh) = \rho(hg) = \rho(h) \circ \rho(g)$$

because $g \in Z(G)$. Thus, Proposition 3.39 implies that $\rho(g) = \lambda \operatorname{id}_V$ for some $\lambda \in \mathbb{C}$.

(b) In one direction, if $X \in \mathfrak{g}$ has $\rho(X) = \lambda \operatorname{id}_V$ for some $\lambda \in \mathbb{C}$, then any $Y \in \mathfrak{g}$ has

$$\begin{aligned}\rho([X,Y]) &= \rho(X) \circ \rho(Y) - \rho(Y) \circ \rho(X) \\ &= \lambda\rho(Y) - \lambda\rho(Y) \\ &= 0,\end{aligned}$$

so the injectivity of ρ implies that $[X,Y] = 0$; thus, $X \in \mathfrak{z}(\mathfrak{g})$.

Conversely, suppose $X \in \mathfrak{z}(\mathfrak{g})$. Then $\rho(X)\colon V \to V$ is an operator on an irreducible representation of $\mathfrak{g}$ which commutes with the $\mathfrak{g}$ action: any $Y \in \mathfrak{g}$ has

$$\rho(X) \circ \rho(Y) - \rho(Y) \circ \rho(X) = \rho([X,Y]) = \rho(0) = 0.$$

Thus, Proposition 3.39 implies that $\rho(X) = \lambda \operatorname{id}_V$ for some $\lambda \in \mathbb{C}$. ■

Example 3.41. By Corollary 3.40, we see that $Z(\mathrm{GL}_n(\mathbb{F}))$ consists of scalar matrices. One can do similar computations for all the classical groups.

Example 3.42. Note that $Z(\mathfrak{sl}_n(\mathbb{F})) = 0$ for $n \geq 2$. (If $n = 1$, then $\mathfrak{sl}_1(\mathbb{F}) = 0$ already.) Indeed, the main point is that the standard representation $\mathfrak{sl}_n(\mathbb{F}) \subseteq \mathrm{gl}_n(\mathbb{F})$ is irreducible. Well, for any nonzero subrepresentation $V \subseteq \mathbb{F}^n$, say $v \in V \setminus \{0\}$, and we may assume that $v = e_1$ upon changing basis. Now, for any $w \in \mathbb{F}^n$, we see that there is a traceless matrix $X \in \mathfrak{sl}_n(\mathbb{F})$ such that $Xv = w$, thus proving that $w \in V$, so $W = \mathbb{R}^n$. Applying this irreducible representation to Corollary 3.40, we conclude that

$$\mathfrak{sl}_n(\mathbb{F}) = \{\lambda 1_n \in \mathfrak{sl}_n(\mathbb{F}) : \lambda \in \mathbb{F}\} = 0$$

because $\operatorname{tr} \lambda 1_n = 0$ requires $\lambda = 0$.

Corollary 3.43. Fix an abelian Lie group G or Lie algebra $\mathfrak{g}$, denoted R. Then all irreducible complex representations are one-dimensional.

Proof. Let V be an irreducible complex representation of R with structure morphism ρ. Then for any $g \in R$, we see that $\rho(g)\colon V \to V$ is an operator commuting with the action of G: for any $h \in R$, we see that

$$\rho(g) \circ \rho(h) = \rho(h) \circ \rho(g)$$

because R is abelian. (Each case with R requires a slightly different argument, but the conclusion is the same: both equal $\rho(gh) = \rho(hg)$ when $R = G$, and the difference equals $\rho([g,h]) = 0$ when $R = \mathfrak{g}$.)

Thus, Proposition 3.39 implies that $\rho(g)$ is a scalar operator $\lambda_g \operatorname{id}_V$ for each $g \in G$. In particular, for any nonzero vector $v \in V$, we see that $\rho(g)$ acts a scalar on $\operatorname{span}\{v\}$ and hence preserves this subspace. Thus, $\operatorname{span}\{v\}$ is a nonzero subrepresentation of V, forcing $V = \operatorname{span}\{v\}$ by irreducibility. ■

Let's compute the representations of some abelian groups/algebras.

Example 3.44. The complex representations of the abelian Lie algebra $\mathbb{F} = \mathfrak{gl}(\mathbb{F})$ are just arbitrary $\mathbb{C}$-vector spaces V with a chosen endomorphism by $\rho\colon \mathfrak{gl}(\mathbb{F}) \to \mathfrak{gl}(V)$. In particular, the irreducible Lie algebra representations $\rho\colon \mathfrak{gl}(\mathbb{F}) \to \mathfrak{gl}(V)$ have $V = \mathbb{C}$ by Corollary 3.43, and then we see we are just asking for a linear map $\mathbb{F} \to \mathbb{C}$.

Now, applying the equivalence of Proposition 3.28, we construct the representation $\rho_\lambda\colon \mathbb{F} \to \mathrm{GL}(\mathbb{C})$ by $\rho_\lambda(t)(v) := \exp(\lambda t v)$ (for any $\lambda \in \mathbb{C}$), and we note that $d(\rho_\lambda)_e\colon \mathbb{F} \to \mathfrak{gl}(\mathbb{C})$ is the multiplication-by-λ irreducible representation of the previous paragraph. In particular, the equivalence of categories establishes these as our irreducible representations of $\mathbb{F}$.

Example 3.45. Consider the real Lie group $G := \mathbb{R}^\times$. Note that $\mathbb{R}^\times \cong \{\pm 1\} \times \mathbb{R}^+$ by the multiplication map, and $\mathbb{R}^+ \cong \mathbb{R}$ by taking the exponential. Now, this Lie group is abelian, so all irreducible representations are one-dimensional, so we can classify irreducible representations $\rho\colon G \to \mathrm{GL}(\mathbb{C})$ as

$$\operatorname{Hom}_{\mathrm{LieGrp}}(\{\pm1\} \times \mathbb{R}, \mathbb{C}^\times) = \operatorname{Hom}_{\mathrm{LieGrp}}(\{\pm1\}, \mathbb{C}^\times) \times \operatorname{Hom}_{\mathrm{LieGrp}}(\mathbb{R}, \mathbb{C}^\times)$$

by tracking the universal property of the product (for both manifolds and groups). Now, $\operatorname{Hom}(\{\pm1\}, \mathbb{C})$ is just looking for elements of $\mathbb{C}^\times$ of order dividing 2, which we know are only $\{\pm1\}$. Continuing, we note $\operatorname{Hom}_{\mathrm{LieGrp}}(\mathbb{R}, \mathbb{C}^\times)$ was classified as the maps $t \mapsto \exp(\lambda t)$ for some $\lambda \in \mathbb{C}$ in Example 3.44 (because such representations must be irreducible by virtue of being one-dimensional). As such, we see that $\operatorname{Hom}_{\mathrm{LieGrp}}(\mathbb{R}^\times, \mathbb{C}^\times)$ consists of the maps $t \mapsto \operatorname{sgn}(t)^\varepsilon\, |t|^\lambda$ for some $\varepsilon \in \{0,1\}$ and $\lambda \in \mathbb{C}$.

Example 3.46. Consider the real Lie group S^1 equipped with the projection $\pi\colon \mathbb{R} \to S^1$ given by $\pi(t) := e^{2\pi i t}$. Then π is a smooth surjection with kernel $\ker \pi = \mathbb{Z}$. Thus, a representation $\rho\colon S^1 \to \mathrm{GL}(\mathbb{C})$ (as usual, all irreducible representations are 1-dimensional by Corollary 3.43) induces a representation $\widetilde{\rho}\colon \mathbb{R} \to \mathrm{GL}(\mathbb{C})$ as $\widetilde{\rho} := \rho \circ \pi$. Now, Example 3.44 tells us that $\widetilde{\rho}(t) = \exp(t\lambda) \in \mathbb{C}^\times$ for some $\lambda \in \mathbb{C}$. However, $\widetilde{\rho}$ must have $\ker \pi = \mathbb{Z}$ in its kernel, so $\exp(\lambda) = 1$, so $\lambda = 2\pi i n$ for some $n \in \mathbb{Z}$. Going back through π, we thus see that $\rho(z) = z^n$ for some $n \in \mathbb{Z}$, and we can check that these are all in fact polynomial (and hence smooth) representations $S^1 \to \mathbb{C}^\times = \mathrm{GL}(\mathbb{C})$.

One can upgrade Proposition 3.39 for arbitrary representations.

Corollary 3.47. Fix a Lie group G or Lie algebra $\mathfrak{g}$, denoted R. Fix complex completely reducible representations

$$V = \bigoplus_{i=1}^n V_i^{\oplus m_i} \qquad \text{and } W \cong \bigoplus_{i=1}^n V_i^{\oplus n_i},$$

where the set $\{V_i\}_{i=1}^n$ consists of pairwise non-isomorphic complex irreducible representations. Then

$$\operatorname{Hom}_R(V, W) = \bigoplus_{i=1}^n \mathbb{C}^{n_i \times m_i}.$$

Proof. Finite sums move outside Hom by the universal properties involved, so

$$\begin{aligned}\operatorname{Hom}_R(V,W) &= \operatorname{Hom}_R\left(\bigoplus_{i=1}^n V_i^{\oplus m_i}, \bigoplus_{i=1}^n V_i^{\oplus n_i}\right) \\ &= \bigoplus_{i,j=1}^n \operatorname{Hom}_R\left(V_i^{\oplus n_i}, V_j^{\oplus m_i}\right) \\ &= \bigoplus_{i,j=1}^n \operatorname{Hom}_R(V_i,V_j)^{n_i\times m_i}.\end{aligned}$$

Now, Proposition 3.39 tells us that $\operatorname{Hom}_R(V_i,V_j)=\mathbb{C}1_{i=j}$, so the result follows. ■

Corollary 3.48. Fix a Lie group G or Lie algebra $\mathfrak{g}$. If V is a completely reducible complex representation, then V has a unique decomposition into irreducibles up to isomorphism and permutation of the factors.

Proof. Because V is finite-dimensional, any two decompositions of V into irreducibles will end up using only finitely many irreducible components, which we can list out as $\{V_1,\ldots,V_n\}$. Then we are given two decompositions

$$\bigoplus_{i=1}^n V_i^{\oplus m_i} \cong V \cong \bigoplus_{i=1}^n V_i^{\oplus n_i}$$

for nonnegative integers $m_\bullet$s and $n_\bullet$s. We want to check that $m_i = n_i$ for each i. Well, for each i, Corollary 3.47 implies that

$$\dim \operatorname{Hom}_G(V_i,V) = \dim \operatorname{Hom}_G\left(V_i, \bigoplus_{i=1}^n V_i^{\oplus m_i}\right) = m_i$$

and similarly $\dim \operatorname{Hom}_G(V_i,V)=n_i$, so $m_i=n_i$ follows. ■

3.2.2 The Unitarization Trick

We would like tools to show that all representations are completely reducible. One place to start is with unitary representations.

Definition 3.49. Fix a Lie group G or Lie algebra $\mathfrak{g}$, denoted R. Then a representation V of R is *unitary* if and only if it has a positive-definite Hermitian inner product $\langle\cdot,\cdot\rangle$ commuting with the R-action. More precisely, we have the following.

- If R is a Lie group, then we want $\langle gv, gw\rangle = \langle v,w\rangle$ for all $g\in G$ and $v,w\in V$.
- If R is a Lie algebra, then we want $\langle Xv,w\rangle + \langle v, Xw\rangle = 0$ for all $X\in\mathfrak{g}$ and $v,w\in V$.

For a general bilinear form $\langle -,-\rangle$ on a representation V, we say that $\langle -,-\rangle$ is *invariant* if and only if the appropriate condition above is satisfied.

Here's a quick coherence check for the definition.

Lemma 3.50. Fix a Lie group G with Lie algebra $\mathfrak{g}$, and let $\rho\colon G \to \mathrm{GL}(V)$ be a complex representation inducing a representation $d\rho_e\colon \mathfrak{g} \to \mathfrak{gl}(V)$. Suppose that V has an $\mathbb{R}$-bilinear product $\langle -, - \rangle\colon V \times V \to \mathbb{C}$ (possibly Hermitian).

(a) For any $X \in \mathfrak{g}$, we have

$$\frac{d}{dt}\langle \exp(tX)v, \exp(tX)w\rangle = \langle Xv, w\rangle + \langle v, Xw\rangle.$$

(b) If $\langle -, - \rangle$ is invariant for G, then it is invariant for $\mathfrak{g}$.

Proof. Here we go.

(a) This is essentially the product rule. In V, we compute

$$\begin{aligned}\frac{d}{dt}\langle \exp(tX)v, \exp(tX)w\rangle &= \lim_{t\to 0}\frac{\langle \exp(tX)v, \exp(tX)w\rangle - \langle v, w\rangle}{t} \\ &= \lim_{t\to 0}\frac{\langle \exp(tX)v, \exp(tX)w\rangle - \langle \exp(tX)v, w\rangle}{t} + \lim_{t\to 0}\frac{\langle v, w\rangle - \langle \exp(tX)v, w\rangle}{t} \\ &= \lim_{t\to 0}\left\langle \exp(tX)v, \frac{\exp(tX)w - w}{t}\right\rangle + \lim_{t\to 0}\left\langle \frac{\exp(tX)v - v}{t}, w\right\rangle.\end{aligned}$$

Now, linearity of the bilinear product implies its continuity, so we can bring the limit inside the bilinear products. Doing so and using Remark 3.3 shows that these limits evaluate to $\langle Xv, w\rangle + \langle v, Xw\rangle$.

(b) If ρ is unitary, then $\langle \exp(tX)v, \exp(tX)w\rangle = \langle v, w\rangle$ always, so the derivative in (a) always vanishes, so $\langle Xv, w\rangle + \langle v, Xw\rangle = 0$ always. ■

We take a moment to acknowledge that, as usual, inner products have applications to duality.

Lemma 3.51. Fix a complex vector space V, and recall that we can define a complex vector space $\overline{V}$ as having the conjugate action. If V has Hermitian inner product $\langle -, - \rangle$, then the map $\overline{V} \to \overline{V^\vee}$ given by $v \mapsto \langle -, v\rangle$ is an isomorphism of vector spaces. If V is a unitary representation over a real Lie group or algebra, then this map is also an isomorphism of representations.

Proof. Well, for any $v \in V$, we see that $\langle -, v\rangle$ is a linear operator $V \to \mathbb{C}$ because $\langle -, - \rangle$ is Hermitian. In fact, this gives an $\mathbb{R}$-linear map $i_\bullet\colon \overline{V} \to V^\vee$ defined by $i_v := \langle -, v\rangle_\bullet$, and it has trivial kernel because v nonzero implies that $\langle v, v\rangle > 0$. Thus, our linear map $V \to V^\vee$ is a vector space isomorphism in light of the fact that $\dim V = \dim V^\vee$. Lastly, we note that this upgrades to an isomorphism of $\mathbb{C}$-vector spaces because

$$\langle -, a\overline{v}\rangle = \overline{a}\langle -, v\rangle.$$

Now, if V is a unitary representation, we need to check that this isomorphism is invariant.

- If V is a representation of a group G, then we note that

$$i_{gv}(w) = \langle gv, w\rangle = \left\langle v, g^{-1}w\right\rangle = (gi_v)(w).$$

- If V is a representation of a Lie algebra $\mathfrak{g}$, then we note that

$$i_{Xv}(w) = \langle Xv, w\rangle = -\langle v, Xw\rangle = (Xi_v)(w).$$

The above checks complete the proof. ■

Lemma 3.52. Fix a complex vector space V. If V has a non-degenerate inner product $\langle -,-\rangle$, then the map $V \to V^\vee$ given by $v \mapsto \langle -,v\rangle$ is an isomorphism of vector spaces. If $\langle -,-\rangle$ is invariant over a Lie group or algebra, then this map is also an isomorphism of representations.

Proof. The first paragraph of Lemma 3.51 (with no over-lines) proves the first sentence. The rest of that proof verbatim shows the second sentence. ■

As another coherence check, we note that the choice of invariant product is more or less unique.

Proposition 3.53. Fix a complex irreducible representation V of a Lie group G. Then there is at most one invariant Hermitian form on V, up to a positive scalar.

Proof. Fixing some invariant Hermitian forms $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ of a representation V, then we claim that we can find a function $\varphi\colon V \to V$ such that

$$\langle v,w\rangle_1 = \langle \varphi(v),w\rangle_2$$

for all $v,w \in V$. Indeed, φ is simply the composite of the representation isomorphisms $\overline{V} \cong V^\vee \cong \overline{V}$ provided by Lemma 3.51.

In particular, taking conjugates, we see that φ is an automorphism of an irreducible representation, so Proposition 3.39 implies that $\varphi = \lambda\,\mathrm{id}_V$ for some $\lambda \in \mathbb{C}$. In particular, we conclude that

$$\langle -,-\rangle_1 = \lambda\langle -,-\rangle_2.$$

It remains to check that λ is a positive real number. Well, choose a nonzero vector $v \in V$, and then we see that $\lambda = \langle v,v\rangle_1/\langle v_v\rangle_2 > 0$. ■

Anyway, here is the reason for defining the notion of unitary.

Proposition 3.54. Let V be a unitary representation of a Lie group G or Lie algebra $\mathfrak{g}$, denoted R. If $W \subseteq V$ is a subrepresentation, then so

$$W^\perp := \{v \in V : \langle v,w\rangle = 0 \text{ for all } w \in W\}.$$

In fact, $V = W \oplus W^\perp$ as representations.

Proof. We run our checks in sequence.

- We claim that $W^\perp \subseteq V$ is a subrepresentation. Well, for each $v \in V$, we note that $w \mapsto \langle w,v\rangle$ is a linear map $V \to \mathbb{C}$ because $\langle -,-\rangle$ is Hermitian, so $W^\perp = \bigcap_{w\in W}\ker\langle -,w\rangle$ is a linear subspace. To see that this is a subrepresentation, we pick up some $v \in W^\perp$, and we want to show that $gv \in W^\perp$ for any $g \in R$. To continue, we do casework on R.
  - If $R = G$, then note that
  $$\langle gv,w\rangle = \langle v,g^{-1}w\rangle = 0$$
  for all $w \in W$ because $g^{-1}w \in W$ as well.
  - If $R = \mathfrak{g}$, then set $X = g$ and note that
  $$\langle Xv,w\rangle = -\langle v,Xw\rangle = 0$$
  for all $w \in W$ because $Xw \in W$ as well.
- We claim that the summation map $W \oplus W^\perp \to V$ is an isomorphism. Because W and $W^\perp$ are both subspaces of V, we certainly have a linear summation map $W \oplus W^\perp \to V$, so it is merely a matter of checking that we have an isomorphism.

- Trivial kernel: suppose that $(w,v) \in W \oplus W^\perp$ has $w+v=0$. Then $w=-v$ lives in $W \cap W^\perp$. In particular, $\langle w,w\rangle = 0$, which implies $w=0$ (and hence $v=0$) because $\langle -,-\rangle$ is Hermitian and non-degenerate.
- Surjective: by a dimension count, it is now enough to check that $\dim V \le \dim W + \dim W^\perp$.[2] Well, the presence of an inner product allows us to begin with an orthonormal basis $\{e_1,\ldots,e_k\}$ of W and then extend it to an orthonormal basis $\{e_{k+1},\ldots,e_n\}$ of V. However, the condition of being orthonormal implies that $\{e_{k+1},\ldots,e_n\} \subseteq W^\perp$, so this orthonormal subset provides a lower bound
$$\dim W^\perp \ge n-k = \dim V - \dim W,$$
as required. ■

Corollary 3.55 (unitarization trick)**.** Let V be a unitary representation of a Lie group G or Lie algebra $\mathfrak{g}$. Then V is completely reducible.

Proof. We induct on $\dim V$ by using Proposition 3.54. If $\dim V = 0$, then $V=0$, so V is the empty sum of irreducible representations. Otherwise, for our inductive step, take $\dim V > 0$. If V is irreducible already, then there is nothing to do. Otherwise, let $W \subseteq V$ be a proper nontrivial subrepresentation of V, and then Proposition 3.54 implies that $V \cong W \oplus W^\perp$. Now, $0 < \dim W, \dim W^\perp < \dim V$, so W and $W^\perp$ are unitary representations with strictly smaller dimension than V, so W and $W^\perp$ are completely irreducible, so V is also completely irreducible (by taking the sum of the decompositions for W and $W^\perp$). ■

Example 3.56. Let $\mathfrak{g}$ be an abelian complex Lie algebra. Then the adjoint representation $\operatorname{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is unitary no matter what Hermitian inner product $\langle -,-\rangle$ we give $\mathfrak{g}$: indeed, we see that
$$\langle \operatorname{ad}_X Y, Z\rangle + \langle Y, \operatorname{ad}_X Z\rangle = \langle [X,Y],Z\rangle + \langle Y,[X,Z]\rangle = 0+0=0$$
by Proposition 2.23. Thus, the adjoint representation is completely reducible by Corollary 3.55.

3.2.3 Compact Lie Groups

The main application of Corollary 3.55 is to compact groups. To explain this, we need some notion of an integration theory. Fix a regular Lie group G of dimension n with Lie algebra $\mathfrak{g}$. Remark 1.140 provides a parallelization of $TG \cong G \times \mathfrak{g}$ by right-invariant vector fields. Choosing a right-invariant global frame $\{\xi_1,\ldots,\xi_n\}$ of TG, we define
$$\omega \coloneqq \xi_1 \wedge \cdots \wedge \xi_n$$
to be a right-invariant top-degree differential form in $\Omega^n G = \wedge^n TG$. Then differential topology explains how to integrate regular compactly supported functions $f \colon G \to \mathbb{F}$ against ω. In particular, tracking through all the definitions, one finds that
$$\int_G (R_g f)\,\omega = \int_G f\,\omega$$
for any $g \in G$ and $f \colon G \to \mathbb{F}$. Indeed, integration is linear, so we may assume that f is supported in a single chart (U,φ) of G. Letting the coordinates be $\varphi = (x_1,\ldots,x_n)$, we see that $\omega = r(x)\,dx_1 \wedge \cdots \wedge dx_n$ for some regular function $r \colon U \to \mathbb{F}$. Then
$$\int_G f\,\omega = \int_U f(x)r(x)\,dx_1 \wedge \cdots \wedge dx_n.$$
However, the G-invariance of ω implies that we can translate everything by g to get the same value of the integral, which is the desired conclusion.

[2] Explicitly, the image of $W + W^\perp \subseteq V$ has dimension $\dim W + \dim W'$ because the summation map already has trivial kernel by the previous point.

The point of having an integration theory is that we are able to take “averages.” The following is our main application.

Proposition 3.57. Fix a complex representation V of a compact Lie group G. Then there is an invariant Hermitian inner product $\langle -,-\rangle$ on V. Thus, Corollary 3.55 implies that representations of a compact Lie group are completely reducible.

Proof. Begin with any Hermitian inner product $\langle -,-\rangle_0$ on V. Then we define $\langle -,-\rangle\colon V\times V\to\mathbb{C}$ by

$$\langle v,w\rangle := \int_G \langle gv,gw\rangle_0\,\omega,$$

where ω is a right-invariant top differential form on G, scaled so that $\int_G\omega=1$.[3] Note that G being compact implies that the integral certainly converges; notably, the function $g\mapsto\langle gv,gw\rangle$ is smooth because $\langle -,-\rangle$ is bilinear, and the representation is regular.

We now claim that $\langle -,-\rangle$ is the required invariant Hermitian inner product.

- Conjugate-symmetric: for any $v,w\in V$, we note that
$$\langle w,v\rangle = \int_G \langle gw,gv\rangle_0\,\omega = \overline{\int_G \langle gv,gw\rangle_0\,\omega} = \overline{\langle v,w\rangle}.$$
- Linear: for any $v,v',w\in V$ and $a,a'\in\mathbb{C}$, we note that
$$\begin{aligned}\langle av+a'v',w\rangle &= \int_G \langle g(av+a'v'),gw\rangle_0\,\omega\\ &= a\int_G\langle gv,gw\rangle_0\,\omega + a'\int_G\langle gv',gw\rangle_0\,\omega\\ &= a\langle v,w\rangle + a'\langle v',w\rangle.\end{aligned}$$
- Non-degenerate: for any $v\in V$, we note that the function $G\to\mathbb{C}$ given by $g\mapsto\langle gv,gv\rangle_0$ is a function which is always positive because $\langle -,-\rangle_0$ is Hermitian. Because G is compact, this function must have a minimum value $m_v>0$, so we conclude that
$$\langle v,v\rangle = \int_G\langle gv,gv\rangle\,\omega \ge m_v > 0.$$
- Invariant: for any $v\in V$ and $h\in G$, we note that
$$\langle hv,hv\rangle = \int_G \langle ghv,ghv\rangle\,\omega \stackrel{*}{=} \int_G\langle gv,gv\rangle\,\omega = \langle v,v\rangle,$$
where $\stackrel{*}{=}$ holds because ω is right-invariant.

The above checks complete the proof. ■

Here is an interesting example.

Example 3.58. The group SU_n is a compact real Lie group, so all its representations are completely reducible by Proposition 3.57. In fact, it is simply connected, so $\mathrm{Rep}_\mathbb{C}(\mathrm{SU}_n)=\mathrm{Rep}_\mathbb{C}(\mathfrak{su}_n)$ by Proposition 3.28. However, $\mathfrak{su}_n$ is also a real form of $\mathfrak{sl}_n(\mathbb{C})$ by Example 2.92, so we can use Remark 3.30 to note that

$$\mathrm{Rep}_\mathbb{C}(\mathfrak{su}_n)=\mathrm{Rep}_\mathbb{C}(\mathfrak{su}_n\otimes_\mathbb{R}\mathbb{C})=\mathrm{Rep}_\mathbb{C}(\mathfrak{sl}_n(\mathbb{C}))=\mathrm{Rep}_\mathbb{C}(\mathfrak{sl}_n(\mathbb{R})\otimes_\mathbb{R}\mathbb{C})=\mathrm{Rep}_\mathbb{C}(\mathrm{SL}_n(\mathbb{R})).$$

Thus, the complex representations of $\mathrm{SL}_n(\mathbb{R})$ are also completely reducible!

[3] Because G is compact, we can cover it in finitely many charts to conclude that $\int_G\omega$ is finite, and then we can scale ω by this integral to conclude that we can choose ω so that $\int_G\omega=1$.

3.3 October 5

Today we classify the representations of $\mathfrak{sl}_2(\mathbb{C})$.

3.3.1 Some Representations of $\mathfrak{sl}_2(\mathbb{C})$

We begin by choosing a basis of $\mathfrak{sl}_2(\mathbb{C})$. By Example 1.153, this Lie algebra consists of the traceless 2×2 matrices and is three-dimensional, so we define

$$e := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad f := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad \text{and} \qquad h := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Note that $\{e, f, h\} \subseteq \mathfrak{sl}_2(\mathbb{C})$ are certainly linearly independent because they occupy disjoint coordinates of these matrices, so we have found a basis. For future use, it will be useful to record the commutator relations

$$\begin{aligned} [e,f] &= ef - fe = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = h, \\ [h,e] &= he - eh = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} = 2e, \\ [h,f] &= hf - fh = (eh - he)^\intercal = -[h,e]^\intercal = -2f. \end{aligned}$$

These commutator relations can be frequently be extending inductively. Here are a couple examples.

Example 3.59. As an example of something we can compute, suppose $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ is some representation. Then we claim that

$$\rho(h) \circ \rho(f)^m = \rho(f)^m \circ (\rho(h) - 2m\mathrm{id}_V)$$

for any $m \geq 0$. Indeed, for $m = 0$, there is nothing to say. For an inductive step, we note

$$\rho(h)\rho(f)^{m+1} = \rho(h)\rho(f)^m\rho(f) = \rho(f)^m \circ (\rho(h) - 2m\mathrm{id}_V)\rho(f) = \rho(f)^{m+1}(\rho(h) - 2(m+1)\mathrm{id}_V),$$

where the key point is that $\rho(h)\rho(f) = \rho(f)(\rho(h) - 2\mathrm{id}_V)$ by the commutator relations.

Example 3.60. Replacing f with e everywhere in Example 3.59 proves that

$$\rho(h) \circ \rho(f)^m = \rho(f)^m \circ (\rho(h) + 2m\mathrm{id}_V),$$

where the point is that $\rho(h)\rho(e) = \rho(e)(\rho(h) + 2\mathrm{id}_V)$ by the commutator relations.

We are going to find many irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$. We will do this fairly geometrically in two ways, starting with $\mathrm{GL}_2(\mathbb{C})$. On one hand, we can start with the standard representation $\rho_{\mathrm{std}} \colon \mathrm{SL}_2(\mathbb{C}) \subseteq \mathrm{GL}_2(\mathbb{C})$ and then define $\rho_n := \mathrm{Sym}^n \rho_{\mathrm{std}}$ for all $n \geq 0$. On the other hand, we can provide a more geometric construction: Note that $\mathrm{SL}_2(\mathbb{C})$ acts on polynomials $\mathbb{C}[x,y]$ by $(\rho(g)p)(x,y) := p\left((x,y)g\right)$. Quickly, we note that this is at least a linear action.

- Identity: $(\rho(1)p)(x,y) = p(x,y)$.
- Associative: $(\rho(gh)p)(x,y) = p\left((x,y)gh\right) = (\rho(g)\rho(h)p)(x,y)$.
- Linear: note $\rho(g)(ap+bq)(x,y) = ap\left((x,y)g\right) + bq\left((x,y)g\right) = (a\rho(g)p + b\rho(g)q)(x,y)$.

Note that $\mathbb{C}[x,y]$ is a ring graded by (total) degree, so we let $V_n \subseteq \mathbb{C}[x,y]$ be the subspace of $\mathbb{C}[x,y]$ spanned by the degree-n monomials. (Namely, V_n consists of 0 and the homogeneous polynomials of degree n.) Then we note that

$$\rho\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)(x^p q^p) = (ax+cy)^p(bx+dy)^q$$

continues to homogeneous of degree $p+q$, so the $\mathrm{SL}_2(\mathbb{C})$-action on $\mathbb{C}[x,y]$ stabilizes the subspaces V_n. Thus, we have produced representations $\rho_n'\colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}(V_n)$. Giving V_n the basis of monomials and expanding about the above formula $(ax+cy)^p(bx+dy)^q$, we see that the matrix coefficients of $\rho_n'(g)$ are polynomial in the matrix coefficients of g, so $\rho_n'(g)$ is indeed a regular representation.

Lemma 3.61. Fix notation as above, and let $\rho_{\mathrm{std}}\colon \mathrm{SL}_2(\mathbb{C}) \subseteq \mathrm{GL}_2(\mathbb{C})$. Then $\rho_n' \cong \operatorname{Sym}^n \rho_{\mathrm{std}}$ for all $n \geq 0$.

Proof. Let $\{v_x, v_y\}$ be the standard basis of $\mathrm{SL}_2(\mathbb{C})$. Then we note that ρ_n' has basis given by the monomials $\{x^py^q\}_{p+q=n}$, and $\operatorname{Sym}^n \rho_{\mathrm{std}}$ has basis given by $\left\{v_x^{\otimes p}v_y^{\otimes q}\right\}_{p+q=n}$ (by some linear algebra).[4]

With this in mind, we define a vector space isomorphism $\varphi\colon \operatorname{Sym}^n \mathbb{C}^2 \to V_n$ by $\varphi\colon v_x^pv_y^q \mapsto x^py^q$. To check that this is $\mathrm{SL}_2(\mathbb{C})$-invariant, it is enough to check on a basis because G-invariance is a linear condition, so we note that $g := \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ has

$$\begin{aligned} \varphi\left(g \cdot v_x^pv_y^q\right) &= \varphi\left((av_x+cv_y)^p(bv_x+dv_y)^q\right) \\ &= (ax+by)^p(bx+dy)^q \\ &= g \cdot x^py^q \\ &= g \cdot \varphi\left(v_x^pv_y^q\right), \end{aligned}$$

as required. ■

Some $V_\bullet$s already have geometric incarnations.

Example 3.62. We see that V_0 is the trivial representation, and V_1 is the standard representation.

Lemma 3.63. The representations $\rho_2\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_2)$ and $\mathrm{ad}_\bullet\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}_2(\mathbb{C}))$ are isomorphic.

Proof. We will construct an explicit isomorphism $\mathfrak{sl}_2(\mathbb{C}) \to V_2$ of representations. For this, we compare the actions of $\rho_2\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_2)$ and $\mathrm{ad}\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}_2(\mathbb{C}))$.

- We compute $\rho_2\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_2)$. Note that $\mathfrak{sl}_2(\mathbb{C})$ acts component-wise on $V_2 = \operatorname{Sym}^2 V \subseteq V \otimes V$. Thus, we see that

$$X(vw) = X \cdot \frac{1}{2}(v \otimes w + w \otimes v) = \frac{1}{2}(Xv \otimes w + v \otimes W + Xw \otimes v + w \otimes Xv) = (Xv)w + v(Xw).$$

  In particular, using the ordered basis $\left\{v_x^2, v_y^2, v_xv_y\right\}$, we compute

$$\begin{aligned} e\left(v_x^2\right) &= 2v_x \cdot e(v_x) = 0, \\ e\left(v_y^2\right) &= 2v_y \cdot e(v_y) = 2v_xv_y, \\ e(v_xv_y) &= v_x \cdot e(v_y) + e(v_x) \cdot v_y = v_x^2, \end{aligned}$$

  and

$$\begin{aligned} f\left(v_x^2\right) &= 2v_x \cdot f(v_x) = 2v_xv_y, \\ f\left(v_y^2\right) &= 2v_y \cdot f(v_y) = 0, \\ f(v_xv_y) &= v_x \cdot f(v_y) + f(v_x) \cdot v_y = v_y^2, \end{aligned}$$

[4] Here, $v_x^{\otimes p}v_y^{\otimes q}$ is the average over all permutations of the vector $v_x^{\otimes p} \otimes v_y^{\otimes q}$. Certainly the permutations of these vectors provide a basis of $\left(\mathbb{C}^2\right)^{\otimes n}$, and taking averages over all permutations of the basis shows that we have in fact given a basis.

and

$$\begin{aligned} h\left(v_x^2\right) &= 2v_x \cdot h(v_x) = 2v_x^2, \\ h\left(v_y^2\right) &= 2v_y \cdot h(v_y) = -2v_y^2, \\ h(v_xv_y) &= v_x \cdot h(v_y) + h(v_x) \cdot v_y = 0 \end{aligned}$$

Thus,

$$\rho_2(e) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \qquad \rho_2(f) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \qquad \text{and} \qquad \rho_2(h) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- We compute $\operatorname{ad}\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}_2(\mathbb{C}))$, where we recall that $\operatorname{ad}_X(Y) = [X, Y]$. Using the commutator relations already computed in the previous homework, we use the ordered basis $\{e, f, h\}$ to see that

$$\operatorname{ad}_e = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \operatorname{ad}_f = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{bmatrix}, \qquad \text{and} \qquad \operatorname{ad}_h = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

With this in mind, we define $\varphi\colon V_2 \to \mathfrak{sl}_2(\mathbb{C})$ by $\varphi\left(v_x^2\right) := -2e$ and $\varphi\left(v_y^2\right) = 2f$ and $\varphi\left(v_xv_y\right) = h$. Then we want to check that $\varphi(\rho_2(X)v) = \operatorname{ad}_X \varphi(v)$ for all $X \in \mathfrak{sl}_2(\mathbb{C})$ and $v \in V_2$ This condition is linear in X and v, so we can check it on bases, where we see that we are asking to check that $\varphi \circ \rho_2(X) \circ \varphi^{-1} = \operatorname{ad}_X$ for any X. However, translating everything into matrices, we see that this comes down to the computations

$$\begin{aligned} \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \\ \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{bmatrix}, \\ \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & & \\ & 2 & \\ & & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

as required. ■

Thus, for the rest of today, we will just work with ρ_n's due to its more concrete description, but we will identify the notations ρ_n and ρ_n'. We now induce a representation $(d\rho_n)_1\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$. We quickly run a few computations; here and throughout the rest of the computations, if we ever find negative exponents, then the relevant expression actually vanishes.

- We claim that $(d\rho_n)_1(e) = x\partial_y$ as operators on V_n. It is enough to check the equality on the basis of monomials of V_n, so we choose (p, q) such that $p + q = n$. Then we note $\exp(te) = \sum_{k=0}^\infty \frac{1}{n!}t^ne^n = \left[\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right]$ because $e^2 = 0$, so we may compute

$$\begin{aligned} (d\rho_n)_1(e)\left(x^py^q\right) &= \frac{d}{dt}\rho_n(\exp(te))\left(x^py^q\right)\bigg|_{t=0} \\ &= \frac{d}{dt}\begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}\left(x^py^q\right)\bigg|_{t=0} \\ &= \frac{d}{dt}x^p(tx+y)^q\bigg|_{t=0} \\ &= qx^{p+1}(tx+y)^{q-1}\bigg|_{t=0} \\ &= qx^{p+1}y^{q-1} \\ &= x\partial_x\left(x^py^q\right). \end{aligned}$$

- We claim that $(d\rho_n)_1(f) = y\partial_x$. Indeed, this follows by switching the roles of x and y everywhere in the previous computation, which effectively exchanges the matrices e and f by interchanging the ordered basis.

- We claim that $h = x\partial_x - y\partial_y$ as operators on V_n. Again, it is enough to check the equality on the basis of monomials of V_n, so we choose (p, q) such that $p + q = n$. Then we note that

$$\exp(th) = \sum_{k=0}^{\infty} \frac{1}{n!} t^n h^n = \sum_{k=0}^{\infty} \frac{1}{n!} \begin{bmatrix} t^n & \\ & (-t)^n \end{bmatrix} = \begin{bmatrix} e^t & \\ & e^{-t} \end{bmatrix},$$

so we may compute

$$\begin{aligned} (d\rho_n)_1(h)\left(x^p y^q\right) &= \frac{d}{dt} \rho_n(\exp(th))\left(x^p y^q\right)\bigg|_{t=0} \\ &= \frac{d}{dt} \begin{bmatrix} e^t & \\ & e^{-t} \end{bmatrix} \left(x^p y^q\right)\bigg|_{t=0} \\ &= \frac{d}{dt} \left(e^t x\right)^p \left(e^{-t} y\right)^q \bigg|_{t=0} \\ &= \frac{d}{dt} e^{(p-q)t} x^p y^q \bigg|_{t=0} \\ &= (p-q) x^p y^q \\ &= \left(x\partial_x - y\partial_y\right)\left(x^p y^q\right). \end{aligned}$$

The rest of the argument forgets about $\mathrm{SL}_2(\mathbb{C})$, so we will replace the notation $(d\rho_n)_1$ with just $d\rho_n \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$. For brevity, we also let v_{pq} be the basis monomial $x^p y^q \in \mathbb{C}[x, y]$, for any $p, q \geq 0$. In particular, the above computations have found that

$$\begin{cases} \rho_n(e) v_{pq} = q v_{p+1,q-1}, \\ \rho_n(f) v_{pq} = p v_{p-1,q+1}, \\ \rho_n(h) v_{pq} = (p-q) v_{pq}. \end{cases} \tag{3.1}$$

In particular, $\rho_n(h)$ acts diagonally on V_n with eigenbasis given by the monomials. We now begin running some checks.

Lemma 3.64. Fix notation as above. Then the representation $\rho_n \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$ is irreducible.

Proof. Fix some nonzero subrepresentation $W \subseteq V_n$, and we want to show that $W = V_n$. Well, W must have some nonzero vector $v := \sum_{p+q=n} a_{pq} v_{pq}$. Supposing that p' is the largest index for p such that $a_{pq} \neq 0$, we see that applying e enough times to v allows us to assume that v is a scalar multiple of v_{n0}. (In particular, $\rho_n(e)^{p'} v$ is a scalar multiple of v_{n0}.) Thus, $v_{n0} \in W$. But then we can apply f inductively to see that

$$\rho_n(f)^k v_{n0} = n \rho_n(f)^{k-1} v_{n-1,1} = \cdots = n(n-1)\cdots(n-k+1) v_{n-k,k},$$

so we see that $v_{n-k,k} \in W$ for each $k \geq 0$. Thus, W contains the monomial basis of V_n, so $W = V_n$. ■

3.3.2 Irreducible Representations of $\mathfrak{sl}_2(\mathbb{C})$

We now classify the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$. An interesting property of the above computations, and one used in the above argument, is that $\rho_n(e)$ and $\rho_n(f)$ are nilpotent operators on V_n. We will eventually show that $\{\rho_n\}_{n \geq 0}$ lists all irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$, so we expect this property to be true in general. Our next step is to prove it.

Lemma 3.65. Let $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ be a complex representation. Then $\rho(e)$ and $\rho(f)$ are nilpotent operators on V.

Proof. Note that (3.1) tells us that we may hope to extract monomials as eigenvectors of h. Thus, we employ the following trick: we let $\sigma(h)$ denote the collection of eigenvalues of h (which is finite because $\dim V < \infty$), and then we let $V[\lambda]$ be the generalized eigenspace for $\rho(h)$ of the eigenvalue $\lambda \in \sigma(h)$. Thus, we have a decomposition

$$V = \bigoplus_{\lambda\in\sigma(h)} V[\lambda].$$

Now, the commutator relations imply that $\rho(h)\rho(e) = \rho(e)\rho(h) + 2\rho(e)$ and $\rho(h)\rho(f) = \rho(f)\rho(h) - 2\rho(f)$, so

$$(\rho(h) - (\lambda+2))^d \circ \rho(e) = \rho(e) \circ (\rho(h) - \lambda)^n \qquad \text{and} \qquad (\rho(h) - (\lambda-2))^d \circ \rho(f) = \rho(f) \circ (\rho(h) - \lambda)^n$$

for all $d \geq 0$ and $\lambda \in \mathbb{C}$. Thus, taking d large enough, we see that $\rho(e)\colon V[\lambda] \to V[\lambda+2]$ and $\rho(f)\colon V[\lambda] \to V[\lambda-2]$. Because $V[\lambda] \neq 0$ for only finitely many λs, we see that $\rho(e)^n$ and $\rho(f)^n$ must be zero for n large enough. For example, any $n > |\sigma(h)|$ will do because then any $\lambda \in \sigma(h)$ and hence $V[\lambda] = 0$ will have $\lambda + 2k \notin \sigma(h)$ for some $k \in \{0, \ldots, n\}$, implying that $\rho(e)^n\colon V[\lambda] \to V[\lambda+2n]$ vanishes for all λ. (A similar argument with $+$ replaced by $-$ shows that $\rho(f)^n$ vanishes.) ■

For technical reasons, we note that the action of h on the V_ns is particularly simple: it diagonalizes. Some algebra with the commutator relations is able to show this in a special case.

Lemma 3.66. Let $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ be a complex representation, and set $U := \ker \rho(e)$. Then $\rho(h)$ preserves U, and $\rho(h)\colon U \to U$ diagonalizes with nonnegative integer eigenvalues.

Proof. We will omit ρ from our actions for brevity. We proceed in steps. Throughout, we fix some $v \in U$.

1. We quickly check that $h(U) \subseteq U$. Indeed, for $v \in U$, we see that $ehv = (he - 2e)v = (h-2)ev = 0$, so $hv \in U$.

2. For any $m \geq 1$, we claim that
$$ef^m v \stackrel{?}{=} f^{m-1} m(h - m + 1)v. \tag{3.2}$$
Indeed, for $m = 1$, we recall that $ef = fe + h$, so $ev = 0$ proves the conclusion. For the inductive step, we take $m \geq 1$ and note
$$\begin{aligned} ef^{m+1}v &= (fe + h)f^m v \\ &= (fef^m + hf^m)\, v \\ &\stackrel{*}{=} (f^m m(h - m + 1) + f^m(h - 2m))\, v \\ &= f^m \left((m+1)h - m^2 - m\right) v \\ &= f^m (m+1)(h - m)v, \end{aligned}$$
where $\stackrel{*}{=}$ holds by Example 3.59.

3. Using the previous step, we conclude that
$$e^m f^m v = e^{m-1} f^{m-1} m(h - (m-1))v.$$
Because h preserves U, we see that $m(h - m + 1)v \in U$ as well. Thus, we may apply the previous step inductively to see that $m \geq 0$ has
$$e^m f^m v \stackrel{?}{=} m! h(h-1)\cdots(h - (m-1))v.$$
Indeed, $m = 0$ has nothing to prove, and for the inductive step, we simply use the previous step.

4. We complete the proof. By Lemma 3.65, we know that there exists $m \geq 0$ such that $\rho(h)^m \colon V \to V$ is the zero operator. For this m, we see

$$0 = e^m f^m v = m! h(h-1) \cdots (h-(m-1))v$$

for all $v \in V$. Thus, the minimal polynomial of h divides $T(T-1)\cdots(T-(m-1))$ and in particular has no repeated roots, so linear algebra implies that h acts diagonally with eigenvalues in $\{0, 1, \ldots, m-1\}$. ■

We now use the eigenvectors we have access to build some subrepresentations by hand.

Lemma 3.67. Fix a nonzero complex representation $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$. Suppose that there is a nonzero eigenvector $v \in \ker \rho(e)$ for $\rho(h)$ with eigenvalue $n \in \mathbb{Z}_{\geq 0}$. Then there exists an embedding $\varphi \colon V_n \to V$ of representations such that $\varphi(x^n) = v$.

Proof. We proceed in steps.

1. We begin by making a motivational remark. With φ as a guide and staring at (3.1), we expect v to be a monomial of the form $x^{n+q}y^q$ because it is an eigenvector for $\rho(h)$ with eigenvalue n, and $\rho(e)v = 0$ suggests that we should have $\varphi(v) = x^n$.

2. We now find other monomials. Namely, note $\operatorname{span}\{v\}$ is not yet stable under the action of $\mathfrak{sl}_2(\mathbb{C})$ because v need not be an eigenvector for $\rho(f)$. Thus, we define

$$v_q \coloneqq \rho(f)^q v$$

for $q \geq 0$ and $v_{-1} \coloneqq 0$, which (3.1) suggests should behave like our monomials with $\varphi(v_q) = n(n-1)\cdots(n-q+1)x^{p-q}y^q$. Indeed, for $q \geq 0$, we have the relations

$$\begin{cases} \rho(e)v_q = q(n-q+1)v_{q-1}, \\ \rho(f)v_q = v_{q+1}, \\ \rho(h)v_q = (n-2q)v_q. \end{cases}$$

Here, the relation for $\rho(e)$ follows from (3.2), and the relation for $\rho(h)$ follows from Example 3.59. In particular, we see that $\rho(e)v_{n+1} = 0$, so $v_{n+1} \in U$, but then $\rho(h)$ acts on v_{n+1} with negative eigenvalue -2, so $v_{n+1} = 0$ is forced by Lemma 3.66; then the $\rho(f)$ relation gives $v_q = 0$ for $q > n$.

3. We construct the map φ. For notational ease, we begin by fixing our collection of monomials by defining

$$w_q \coloneqq \frac{1}{n(n-1)\cdots(n-q+1)} v_q$$

for $q \in \{0, \ldots, n\}$ and $v_q = 0$ for $q \in \mathbb{Z} \setminus \{0, \ldots, n\}$, where now we expect $\varphi(w_q) = x^{n-q}y^q$. Indeed, for $q \in \{0, \ldots, n\}$, we have the relations, we have the relations

$$\begin{cases} \rho(e)w_q = qw_{q-1}, \\ \rho(f)w_q = (n-q)w_{q+1}, \\ \rho(h)w_q = (n-2q)w_q. \end{cases}$$

We now may compare the above relations with (3.1) to see that $\varphi \colon V_n \to V$ defined by $\varphi(x^py^q) \coloneqq w_q$ preserves the $\mathfrak{sl}_2(\mathbb{C})$-action. Indeed, we want to check that $\varphi(\rho_n(X)v) = \rho(X)(\varphi(v))$ for any $X \in \mathfrak{sl}_2(\mathbb{C})$ and $v \in V_n$, for which it suffices to check on the bases $\{e, f, h\} \subseteq \mathfrak{sl}_2(\mathbb{C})$ and $\{x^py^q\}_{p+q=n} \subseteq V_n$; these checks are immediate by comparing the above relations with (3.1).

4. It remains to check that φ is an embedding. We provide two ways of doing this.

   - Note that $\{w_0, \ldots, w_n\} \subseteq V_n$ is a linearly independent set because these vectors have distinct eigenvalues for $\rho(h)$. Thus, φ sends a basis to a linearly independent subset of V_n and hence must be injective.

- Note φ is nonzero because $\varphi(x^n) = w_0 = v$ is nonzero. Because V_n is irreducible by Lemma 3.64, we conclude that φ is injective by Proposition 3.39. ■

At long last, here is our classification result.

Theorem 3.68. Let V_1 be the standard representation $\rho_1 : \mathfrak{sl}_2(\mathbb{C}) \subseteq \mathfrak{gl}_2(\mathbb{C})$, and define $V_n := \operatorname{Sym}^n V_1$ for each $n \geq 0$ so that we have representations $\rho_n : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$. Then

$$\{V_n : n \geq 0\}$$

is exactly the set of irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$, and these are all distinct.

Proof. By Lemma 3.64, we see that the $V_\bullet$s are irreducible, and they are all distinct because their dimensions are all distinct: $\dim V_n = n + 1$.

It remains to check that these are the only irreducible representations. Well, pick up some irreducible representation $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$. Now, we use Lemma 3.67 to get some $n \in \mathbb{Z}_{\geq 0}$ and a nonzero map $\varphi \colon V_n \to V$ of representations. (Note that the existence of the required v is satisfied by Lemma 3.66.) Because V_n and V are both irreducible, Proposition 3.39 implies that φ is an isomorphism, so $V \cong V_n$ is one of the $V_\bullet$s. ■

3.3.3 Complete Reducibility for $\mathfrak{sl}_2(\mathbb{C})$

In this subsection, we provide a purely algebraic proof for the complete reducibility of representations of $\mathfrak{sl}_2(\mathbb{C})$. Namely, we are avoiding the integration theory used in Example 3.58. Technically, this subsection can thus be skipped, but it is instructive because the methods used in this subsection will reappear when we want to show that the complex representations of general semisimple algebras are completely reducible.

Note that the key to the proof of Example 3.58 was the ability to take averages in order to produce some invariant maps. (In particular, we needed to provide an invariant Hermitian form.) Our substitute for being able to take averages is to use the "Casimir" operator

$$C := ef - fe + \frac{1}{2}h^2,$$

which (suitably interpreted) is always an invariant map. Let's check this.

Lemma 3.69. Let $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ be some representation. Then

$$\rho(C) := \rho(e) \circ \rho(f) - \rho(f) \circ \rho(e) + \frac{1}{2}\rho(h) \circ \rho(h)$$

is an $\mathfrak{sl}_2(\mathbb{C})$-invariant morphism $V \to V$.

Proof. Being $\mathfrak{sl}_2(\mathbb{C})$-invariant is linear in $\mathfrak{sl}_2(\mathbb{C})$ and thus can be checked on the standard basis of $\mathfrak{sl}_2$, which is a purely formal computation with the commutator. We will drop the ρs everywhere for brevity. We compute

$$\begin{aligned} Ce &= efe + fee + \frac{1}{2}hhe \\ &= e(ef - h) + (ef - h)e + \frac{1}{2}h(eh + 2e) \\ &= eef - eh + efe + \frac{1}{2}heh \\ &= eef - eh + efe + \frac{1}{2}(eh + 2e)h \\ &= eef + efe + \frac{1}{2}eh^2 \\ &= eC. \end{aligned}$$

Similarly,

$$\begin{aligned}
Cf &= eff + fef + \frac{1}{2}hhf \\
&= (fe+h)f + f(fe+h) + \frac{1}{2}h(fh-2f) \\
&= fef + f(fe+h) + \frac{1}{2}hfh \\
&= fef + f(fe+h) + \frac{1}{2}(fh-2f)h \\
&= fef + ffe + \frac{1}{2}fhh \\
&= fC.
\end{aligned}$$

Lastly,

$$\begin{aligned}
Ch &= efh + feh + \frac{1}{2}hhh \\
&= e(hf+2f) + f(he-2e) + \frac{1}{2}hhh \\
&= ehf + fhe + 2(ef-fe) + \frac{1}{2}hhh \\
&= (he-2e)f + (hf+2f)e + 2(ef-fe) + \frac{1}{2}hhh \\
&= hef + hfe + \frac{1}{2}hhh \\
&= hC.
\end{aligned}$$

■

Remark 3.70. It may appear that our definition of C came out of nowhere, but it turns out that all operators $V \to V$ which can be written as a polynomial in $\{e, f, h\}$ is actually a polynomial in C. We will prove this later once we have talked about the universal enveloping algebra, which is the correct context to talk about polynomials in $\{e, f, h\}$.

Importantly, C provides a basis-free way to distinguish between the irreducible representations $V_\bullet$.

Lemma 3.71. The operator $\rho_n(C)\colon V_n \to V_n$ equals the scalar operator $\frac{n(n+2)}{2}\operatorname{id}_{V_n}$.

Proof. Because V_n is irreducible by Lemma 3.64, and $\rho_n(C)\colon V_n \to V_n$ is $\mathfrak{sl}_2(\mathbb{C})$-invariant by Lemma 3.69, Proposition 3.39 tells us that $\rho(C)$ is a scalar operator, so it only remains to compute what this scalar is. Well, we may just compute $\rho(C)$ on the vector x^n: omitting the ρ_ns everywhere for brevity, we see

$$\begin{aligned}
C(x^n) &= \left(ef + fe + \frac{1}{2}h^2\right)(x^n) \\
&= ef(x^n) + fe(x^n) + \frac{1}{2}h^2(x^n) \\
&= nx^n + 0 + \frac{1}{2}n^2x^n \\
&= \frac{n(n+2)}{2}x^n,
\end{aligned}$$

as required. ■

Remark 3.72. One does not need to use Lemma 3.69 and Proposition 3.39 here; instead, one can simply compute $\rho_n(C)$ on the entire basis $\{x^py^q\}_{p+q=n}$ to see what the scalar should be.

We are now ready for our theorem.

Theorem 3.73. The complex representations of $\mathfrak{sl}_2(\mathbb{C})$ are completely reducible.

Proof. Let $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ be any representation, which we want to be completely reducible. We will induct on $\dim V$, where the base case of $\dim V \in \{0, 1\}$ has no content because $\dim V = 0$ makes V the empty sum of irreducibles, and $\dim V = 1$ makes V already irreducible. For the inductive step, we proceed in steps.

1. We reduce to the case where V is indecomposable: indeed, if V is the direct sum of two nonzero representations $V_1 \oplus V_2$, then $\dim V_1, \dim V_2 < \dim V$, so the induction promises that V_1 and V_2 are the direct sum of irreducible representations, so $V = V_1 \oplus V_2$ is the direct sum of irreducible representations.

   For the rest of the argument, we thus may assume that V is indecomposable.

2. We reduce to the case where $\rho(C)$ has a single generalized eigenvalue on V. Let $\sigma(C)$ be the collection of eigenvalues of V, which is finite because V is finite-dimensional. Then we have a decomposition

$$V = \bigoplus_{\mu \in \sigma(C)} V[\mu]$$

   into generalized eigenspaces, where $V[\mu]$ is the generalized eigenspace with eigenvalue μ.

   We claim that each $V[\mu]$ is a subrepresentation, implying that we must have $V = V[\mu]$ because V is indecomposable, completing this step. Well, $V[\mu]$ is the kernel of $(\rho(C) - \mu\operatorname{id}_V)^d$ for some perhaps large d. Because $\rho(C)$ is $\mathfrak{sl}_2(\mathbb{C})$-invariant by Lemma 3.69, we conclude that $(\rho(C) - \mu\operatorname{id}_V)^d$ is as well, so $V[\mu] \subseteq V$ becomes a subrepresentation by Example 3.16.

3. We place V into a controlled short exact sequence. Let $W \subseteq V$ be an irreducible subrepresentation; to construct one, we can just take a minimal nonzero subrepresentation. We want to show that $V = W$. Note Theorem 3.68 tells us that $W \cong V_n$ for some n, which means that $\rho(C)$ will act on W and thus V as the scalar $\frac{n(n+2)}{2}$ by Lemma 3.71. Now, we have a short exact sequence

$$0 \to V_n \to V \to V/V_n \to 0.$$

   Because $\dim V/V_n < \dim V$, it must also be completely reducible. However, $\rho(C)$ acts as the scalar $\frac{n(n+2)}{2}$ on all irreducible subrepresentations of V, so using Lemma 3.71 backward tells us that V_n is the only permitted irreducible subrepresentation. Thus, $V/V_n \cong V_n^{\oplus(m-1)}$ for some $m \geq 1$, and we are given the short exact sequence

$$0 \to V_n \to V \to V_n^{\oplus(m-1)} \to 0. \tag{3.3}$$

4. We construct morphisms $V_n \to V$, which will eventually produce an isomorphism $V_n^{\oplus m} \to V$. We will use Lemma 3.67, which requires some eigenvectors of $\rho(h)$. Well, let $\sigma(h)$ be the eigenvalues of $\rho(h)$, so we get a decomposition

$$V = \bigoplus_{\lambda \in \sigma(h)} V[\lambda]$$

   into generalized eigenspaces, where $V[\lambda]$ is the generalized eigenspace for $\rho(h)$ with eigenvalue λ. Let λ_0 be an eigenvalue with maximal real part. As in the proof of Lemma 3.65, we know that $\rho(e)\colon V[\lambda] \to V[\lambda + 2]$, so the maximality of λ_0 implies that $\rho(e)\colon V[\lambda_0] \to V[\lambda_0 + 2]$ must be the zero map.

   Thus, $V[\lambda_0] \subseteq \ker \rho(e)$, so Lemma 3.66 tells us that $\rho(h)$ actually acts diagonally on $V[\lambda_0]$, and $\lambda_0 = n'$ for some $n' \in \mathbb{Z}_{\geq 0}$. Each vector in $V[n']$ provides an embedding $V_{n'} \to V$, but we know that all irreducible subrepresentations of V are V_n, so $n' = n$. Quickly, we see that (3.3) tells us that $\dim V[n] = m \dim V_n[n] = m$, where $\dim V_n[n] = 1$ because $V_n[n] = \operatorname{span}\{x^n\}$.

We now note that Lemma 3.67 takes each $u \in V[n]$ and gives an embedding $\varphi\colon V_n \to V$ such that $\varphi(x^n) = u$.

5. We construct an isomorphism $V_n^{\oplus m} \to V$, which completes the proof because it shows that V is completely reducible.[5] Let $\{u_1, \ldots, u_m\}$ be a basis of $V[n]$, which produces embeddings $\varphi_1, \ldots, \varphi_n\colon V_n \to V$ such that $\varphi_i(x^n) = u_i$ for each i. Now, $\bigoplus_i \varphi_i\colon V_n^{\oplus m} \to V$ will be the required morphism.

   We would like to show that $\bigoplus_i \varphi_i$ is an isomorphism. Because $\dim V = \dim V_n^{\oplus m}$ by (3.3), it is enough to check that $\bigoplus_i \varphi_i$ is injective. Because each φ_i commutes with the action of h, it is enough to check that $\bigoplus_i \varphi_i[\lambda]$ is injective for each $\lambda \in V_n^{\oplus m}[\lambda] = V_n[\lambda]^{\oplus m}$.

   We already have the injectivity for $\lambda = n$ because $\{\varphi_i(x^n)\}_{1 \le i \le m}$ is a basis of $V[n]$. We will reduce all of our injectivity checks to this one. Well, recall from (3.1) that $\rho_n(h)$ diagonalizes with eigenvectors $\{-n, -n+2, \ldots, n-2, n\}$, so we are really trying to show that $\bigoplus_i \varphi_i[n-2j]$ is injective for $j \in \{0, 2, \ldots, 2n\}$. We will induct on j, where we already discussed the case of $j = 0$. For the inductive step, take $j \in \{0, \ldots, 2n-2\}$ and note that

$$\begin{tikzcd} V_n[n-2j-2]^{\oplus m} \arrow[r, "\rho_n(e)"] \arrow[d, "\bigoplus_i \varphi_i"'] & V_n[n-2j]^{\oplus m} \arrow[d, "\bigoplus_i \varphi_i"] \\ V[n-2j-2] \arrow[r, "\rho(e)"] & V[n-2j] \end{tikzcd}$$

   commutes, where the horizontal maps are well-defined by the argument of the previous step. By the induction, we may assume that the right map has trivial kernel, and (3.1) tells us that the top map is an isomorphism. Thus, the composite map from the top-left to the bottom-right has trivial kernel, so the left map must have trivial kernel, as required. ■

3.4 October 7

Today we finish classifying the representations of $\mathfrak{sl}_2$.

3.4.1 Applications for Representations of $\mathfrak{sl}_2(\mathbb{C})$

Let's discuss some applications of Theorems 3.68 and 3.73. To begin, we can upgrade the diagonal action of Lemma 3.66.

Corollary 3.74. Fix any complex representation $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$. Then $\rho(h)\colon V \to V$ acts diagonally with eigenvalues in $\mathbb{Z}$.

Proof. By Theorem 3.73, it is enough to check this for irreducible representations V. By Theorem 3.68, we see that $V \cong \operatorname{Sym}^n V_{\text{std}}$, where V_{std} is the standard representation. Then Lemma 3.61 explains that these can be realized as polynomial representations, from which the required diagonalization of $\rho(h)$ follows from its computation on the monomial basis given in (3.1). ■

Corollary 3.75 (Jacobson–Morozov)**.** For any nilpotent operator $N\colon V \to V$ on a finite-dimensional complex vector space V, there exists a (unique up to isomorphism) structure of $\mathfrak{sl}_2(\mathbb{C})$-representation on V such that $e \mapsto N$. More precisely, we have the following.

(a) There exists a representation $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ such that $\rho(e) = N$.

(b) If $\rho^1, \rho^2\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ have $\rho^1(e) = \rho^2(e)$, then $\rho^1 \cong \rho^2$.

[5] Because V is assumed to be indecomposable, we actually know that $m = 1$, but we do not need this to conclude the proof.

Proof. We will proceed with the claims separately.

(a) By giving V a basis, we may identify it with $\mathbb{C}^d$; using the Jordan normal form, we are able to choose this basis so that V is the direct sum of Jordan blocks of the form

$$J_n = \underbrace{\begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}}_{n+1} \in \mathbb{C}^{(n+1)\times(n+1)}.$$

Decomposing $V = V_1 \oplus \cdots \oplus V_m$ so that N decomposes into these Jordan blocks as $N = J_{n_1} \oplus \cdots \oplus J_{n_m}$, we see that we may assume that $N = J_n$ for some n: if we can find $\rho_i \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_i)$ such that $\rho_i(e) = J_{n_i}$ for each i, then $\rho := \rho_1 \oplus \cdots \oplus \rho_m$ will be a representation $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ satisfying $\rho(e) = N$.

We are thus reduced to the case where $N = J_n$ for some $n \geq 0$; note then that $\dim V = n+1$, so we expect to be able to take $\rho = \rho_n$. Let $\{v_0, \ldots, v_n\}$ be the given basis of V, which we will adjust to fit $\rho(e) = J_n$. With this in mind, define $\varphi \colon V \to V_n$ by $\varphi(v_q) := q!^{-1} x^{n-q} y^q$; this sends a basis to a basis, so φ is an isomorphism of vector spaces. Further, we claim that $\varphi \circ N = \rho_n(e) \circ \varphi$: it is enough to check this on the basis $\{v_0, \ldots, v_n\}$, so we use (3.1) to compute $\varphi(N v_0) = 0 = \rho_n(e) x^n$ and

$$\begin{aligned} \varphi(N v_q) &= \varphi(v_{q-1}) \\ &= (q-1)!^{-1} x^{n-q+1} y^{q-1} \\ &= \rho_n(e) \left(q!^{-1} x^{n-q} y^q \right) \\ &= \rho_n(e) \varphi(v_q) \end{aligned}$$

for $q \geq 1$. Thus, we may define $\rho \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ by $\rho(X) := \varphi^{-1} \circ \rho_n(X) \circ \varphi$ for all $X \in \mathfrak{sl}_2(\mathbb{C})$. Adjusting by conjugating φ makes it so that ρ succeeds by a representation, and we checked that $\varphi \circ N = \rho_n(e) \circ \varphi$, so $\rho(e) = N$, which is what we wanted.

(b) We proceed directly. We will read the structure of ρ^1 and ρ^2 directly off of N. Fix some $i \in \{1, 2\}$. By using Theorems 3.68 and 3.73, we may decompose

$$\rho^i \cong \bigoplus_{n \geq 0} \rho_n^{\oplus a_n^i}$$

for some nonnegative integers $a_n^i \geq 0$. We will show that $a_m^1 = a_m^2$ for each $m \geq 0$, which will complete the proof by comparing the two decompositions.

For this, we use the dimension of $\ker \rho^i(e)^m$ for various $m \geq 0$. In particular, (3.1) gives

$$\dim \ker \rho_n(e)^m = \dim \operatorname{span}\left\{ x^n, y x^{n-1}, \ldots, x^{n-m+1} y^{m-1} \right\} = \min\{m, n+1\},$$

so

$$\dim \ker N^m = \dim \ker \rho^i(e)^m = \sum_{n \geq 0} \dim \ker \left((\rho_n(e)^{\oplus a_n^i} \right)^m = \sum_{n \geq 0} a_n^i \max\{m, n+1\}.$$

We now use this to read off the values of a_n^i: for any $m \geq 1$, we see

$$\sum_{n=0}^{m-1} a_n^i = \sum_{n \geq 0} a_n^i \max\{m+1, n+1\} - \sum_{n \geq 0} a_n^i \max\{m, n+1\} = \dim \ker N^{m+1} - \dim \ker N^m,$$

so

$$a_m^i = \sum_{n=0}^{m} a_n^i - \sum_{n=0}^{m-1} a_n^i = \dim \ker N^{m+2} - \dim \ker N^m.$$

Thus, a_m^i is independent of i, so $a_m^1 = a_m^2$ for all $m \geq 0$. ■

Example 3.76. We may hope that $\rho_1 = \rho_2$ on the nose, but this is not true in general. For example, one can use an inner automorphism of ρ fixing $\rho(e)$ to produce an isomorphic representation $\rho' : \mathfrak{sl}_2(\mathbb{C}) \to V$ with $\rho(e) = \rho'(e)$. Concretely, take $V = \mathbb{C}^2$ and $N := \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$. Then we could define $\rho = \rho_{\text{std}}$ and

$$\rho'(X) := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \rho(X) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1}$$

so that $\rho'(e) = \rho(e)$ while $\rho'(h) \neq \rho(h)$.

3.4.2 Character Theory of $\mathfrak{sl}_2(\mathbb{C})$

By analogy with the representation theory of finite groups, we may want a notion of characters for representations of $\mathfrak{sl}_2(\mathbb{C})$. Our classification allows to do this cleanly.

Definition 3.77. Fix a finite-dimensional complex representation $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to V$. For any $n \in \mathbb{Z}$, let $V[n]$ be the eigenspace of $\rho(h)$ with eigenvalue n. Then the *character* of V is the rational polynomial

$$\chi_\rho(T) := \sum_{n \in \mathbb{Z}} \dim V[n] T^n.$$

This is a polynomial in $\mathbb{Z}\left[T, T^{-1}\right]$ because only finitely many of the $V[n]$ may be nonzero because $\dim V$ is finite. We will write χ_V for χ_ρ when no confusion is possible.

Remark 3.78. This definition does not lose any information by merely considering the given $V[n]$s because Corollary 3.74 tells us that $\rho(h)$ diagonalizes with integral eigenvalues.

Remark 3.79. To relate this definition with the characters of finite groups, we claim

$$\chi_V\left(e^t\right) \stackrel{?}{=} \operatorname{tr} \exp(t\rho(h)).$$

Indeed, using Corollary 3.74, we may write $\rho(h) = \operatorname{diag}(n_1, \ldots, n_d)$ where $\{n_1, \ldots, n_d\}$ are integers. (Technically, we do not need to know that the action is diagonal for the subsequent argument.) Then

$$\operatorname{tr} \exp(t\rho(h)) = \operatorname{tr} \exp(\operatorname{diag}(tn_1, \ldots, tn_d)) = \sum_{i=1}^d e^{tn_1} = \sum_{i=1}^d \left(e^t\right)^{n_i}.$$

Grouping the $n_\bullet$s by multiplicity, we conclude that this equals $\chi_V\left(e^t\right)$: note $\dim V[n] = \#\{i : n_i = n\}$.

Example 3.80. Using (3.1), we see that $n \geq 0$ has

$$\chi_{V_n}(T) = T^{-n} + T^{-n+2} + \cdots + T^{n-2} + T^n = \frac{T^{n+1} - T^{-n-1}}{T - T^{-1}}.$$

For example, we see that $\chi_{V_n}(T)$ is the first of the χ_ns with nonzero coefficient on T^n, so the collection $\{\chi_n : n \geq 0\}$ is $\mathbb{C}$-linearly independent. Explicitly, any nontrivial expression $\sum_{n \geq 0} a_n \chi_{V_n}$ (with $a_n = 0$ for all but finitely many n) will have some largest N for which $a_N \neq 0$, but then the monomial $a_N T^N$ lives in $\sum_{n \geq 0} a_n \chi_{V_n}$, so $\sum_{n \geq 0} a_n \chi_{V_n} \neq 0$.

Here are some easy checks on our characters.

Lemma 3.81. Fix complex representations $\rho_V : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ and $\rho_W : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(W)$.

(a) We have $\chi_{V\oplus W} = \chi_V + \chi_W$.

(b) We have $\chi_{V\otimes W} = \chi_V \cdot \chi_W$.

(c) We have $\chi_{V^\vee}(T) = \chi_V\left(T^{-1}\right)$.

Proof. These checks are purely formal.

(a) Because $\rho_{V\oplus W}(h) = \rho_V(h) \oplus \rho_W(h)$, we can split up our eigenspaces for $n \in \mathbb{Z}$ by

$$(V \oplus W)[n] = V[n] \oplus W[n],$$

so

$$\begin{aligned}\chi_{V\oplus W}(T) &= \sum_{n\in\mathbb{Z}} \dim(V \oplus W)[n]T^n \\ &= \sum_{n\in\mathbb{Z}} \dim(V[n] \oplus W[n])T^n \\ &= \sum_{n\in\mathbb{Z}} \dim V[n]T^n + \sum_{n\in\mathbb{Z}} \dim W[n]T^n \\ &= \chi_V(T) + \chi_W(T).\end{aligned}$$

(b) Let $\{v_1, \ldots, v_k\}$ and $\{w_1, \ldots, w_\ell\}$ be eigenbases for the operators $\rho_V(h) \colon V \to V$ and $\rho_W(h) \colon W \to W$ with eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ and $\{\mu_1, \ldots, \mu_\ell\}$, respectively. Then $\{v_i \otimes w_j\}$ is a basis for $V \otimes W$, and in fact it is an eigenbasis for $\rho_{V\otimes W}(h)$: note

$$\begin{aligned}\rho_{V\otimes W}(h)(v_i \otimes w_j) &= \rho_V(h)v_i \otimes w_j + v_i \otimes \rho_W(h)w_j \\ &= \lambda_i v_i \otimes w_j + w_i \otimes \mu_j w_j \\ &= (\lambda_i + \mu_j)(v_i \otimes w_j).\end{aligned}$$

Thus, for any $z \in \mathbb{C}$, we see that

$$\dim(V \otimes W)[n] = \#\{(i,j) : \lambda_i + \mu_j = n\},$$

so

$$\begin{aligned}\chi_{V\otimes W}(T) &= \sum_{n\in\mathbb{Z}} \dim(V \otimes W)[n]T^n \\ &= \sum_{n\in\mathbb{Z}} \#\{(i,j) : \lambda_i + \mu_j = n\}T^n \\ &= \sum_{n\in\mathbb{Z}} \sum_{a+b=z} (\dim V[a] \dim W[b])T^n \\ &= \sum_{a,b\in\mathbb{Z}} (\dim V[a] \dim W[b])T^{a+b} \\ &= \left(\sum_{a\in\mathbb{Z}} \dim V[a]T^a\right)\left(\sum_{b\in\mathbb{Z}} \dim V[b]T^b\right) \\ &= \chi_V(T)\chi_W(T),\end{aligned}$$

(c) Let $\{v_1, \ldots, v_k\}$ be an eigenbasis for the operator $\rho_V(h) \colon V \to V$ with eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$. Then we claim that the dual basis $\{v_1^\vee, \ldots, v_k^\vee\}$ is an eigenbasis for $\rho_{V^\vee}(h)$: for any $v_i^\vee$ and v_j, we see

$$\begin{aligned}\left(\rho_{V^\vee}(h)v_i^\vee\right)(v_j) &= -v_i^\vee\left(\rho_V(h)v_j\right) \\ &= -\lambda_j v_i^\vee(v_j) \\ &= -\lambda_j 1_{i=j},\end{aligned}$$

so $\rho_{V^\vee}(h)v_i^\vee = -\lambda_i v_i^\vee$. Thus, gathering multiplicities, we see that $\dim V[n] = \dim V^\vee[-n]$ for any $n \in \mathbb{Z}$, so

$$\begin{aligned}\chi_{V^\vee}(T) &= \sum_{n\in\mathbb{Z}} \dim V^\vee[n]T^n \\ &= \sum_{n\in\mathbb{Z}} \dim V[-n]T^n \\ &= \sum_{n\in\mathbb{Z}} \dim V[n]\left(T^{-1}\right)^n \\ &= \chi_V\left(T^{-1}\right),\end{aligned}$$

as required. ■

Importantly, we can use characters to determine representations.

Proposition 3.82. Fix complex representations $\rho_V \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$ and $\rho_W \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(W)$. If $\chi_V = \chi_W$, then $V \cong W$.

Proof. By Theorems 3.68 and 3.73, we have decompositions

$$V \cong \bigoplus_{n\ge 0} V_n^{\oplus a_n} \qquad \text{and} \qquad W \cong \bigoplus_{n\ge 0} V_n^{\oplus b_n}.$$

We will show that $a_n = b_n$ for all n, which will complete the proof upon comparing the decompositions. Well, Lemma 3.81 tells us that

$$\begin{aligned}0 &= \chi_V(T) - \chi_W(T) \\ &= \sum_{n\ge0} a_n\chi_{V_n}(T) - \sum_{n\ge0} b_n\chi_{V_n}(T) \\ &= \sum_{n\ge0}(a_n - b_n)\chi_{V_n}(T).\end{aligned}$$

This relation is enough to imply $a_n = b_n$ for all n by the linear independence given in Example 3.80. ■

Example 3.83. We claim that $V \cong V^\vee$ for any complex representation $\rho_V \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$. By Proposition 3.82, we may check this on characters. Using the complete reducibility of Theorem 3.73 with Lemma 3.81, it is enough to check this for irreducible V (notably, $(V \oplus W)^\vee \cong V^\vee \oplus W^\vee$). Thus, Theorem 3.68 lets us assume that $V = V_n$ for some $n \ge 0$, so we are left to show that

$$\chi_{V_n}(T) \stackrel{?}{=} \chi_{V_n}\left(T^{-1}\right)$$

by Lemma 3.81. This is true by the explicit computation of Example 3.80.

Example 3.84. We claim that $V_2 \otimes V_3 \cong V_1 \oplus V_3 \oplus V_5$. By Proposition 3.82, it is enough to show an equality of characters. For this, we use Example 3.80 with Lemma 3.81 to see

$$\begin{aligned}\chi_{V_2\otimes V_3}(T) &= \left(T^{-2}+1+T^2\right)\left(T^{-3}+T^{-1}+T+T^3\right) \\ &= T^{-5} + 2T^{-3} + 3T^{-1} + 3T + 2T^3 + T^5 \\ &= \chi_{V_5}(T) + \chi_{V_3}(T) + \chi_{V_1}(T) \\ &= \chi_{V_1\oplus V_3\oplus V_5}(T).\end{aligned}$$

Here is the general case of Example 3.84.

Proposition 3.85 (Clebsch–Gordan rule)**.** Let V_1 be the standard representation $\rho_1 \colon \mathfrak{sl}_2(\mathbb{C}) \subseteq \mathfrak{gl}_2(\mathbb{C})$, and define $V_n := \operatorname{Sym}^n V_1$ for each $n \ge 0$ so that we have representations $\rho_n \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$. Then

$$V_m \otimes V_n \cong \bigoplus_{i=0}^{\min\{m,n\}} V_{|m-n|+2i}.$$

Proof. By symmetry, we may assume that $m \le n$. We generalize the argument of Example 3.84. By Proposition 3.82, it is enough to compare the characters of both sides, for which we use Lemma 3.81 with Example 3.80. Now, we compute

$$\begin{aligned}
\chi_{V_m \otimes V_n}(T) &= \chi_{V_m}(T)\chi_{V_n}(T) \\
&= \frac{T^{m+1} - T^{-m-1}}{T - T^{-1}} \cdot \frac{T^{n+1} - T^{-n-1}}{T - T^{-1}}, \\
\chi_{\bigoplus_{i=0}^m V_{n-m+2i}}(T) &= \sum_{i=0}^m \chi_{V_{n-m+2i}}(T) \\
&= \sum_{i=0}^m \frac{T^{n-m+2i+1} - T^{-(n-m-2i+1)}}{T - T^{-1}}.
\end{aligned}$$

Thus, it remains to show the combinatorial identity

$$\frac{T^{m+1} - T^{-m-1}}{T - T^{-1}} \cdot \frac{T^{n+1} - T^{-n-1}}{T - T^{-1}} \stackrel{?}{=} \sum_{i=0}^m \frac{T^{n-m+2i+1} - T^{-(n-m+2i+1)}}{T - T^{-1}}.$$

Multiplying both sides by $T - T^{-1}$, we would like to show that

$$\frac{\left(T^{m+1} - T^{-m-1}\right)\left(T^{n+1} - T^{-n-1}\right)}{T - T^{-1}} \stackrel{?}{=} \sum_{i=0}^m \left(T^{n-m+2i+1} - T^{-(n-m+2i+1)}\right).$$

Well,

$$\sum_{i=0}^m T^{\pm(n-m+2i+1)} = \frac{T^{\pm(n+m+2+1)} - T^{\pm(n-m+1)}}{T^{\pm 2} - 1},$$

so

$$\begin{aligned}
\sum_{i=0}^m \left(T^{n-m+2i+1} - T^{-(n-m+2i+1)}\right) &= \frac{T^{(n+m+2+1)} - T^{(n-m+1)}}{T^2 - 1} - \frac{T^{-(n+m+2+1)} - T^{-(n-m+1)}}{T^{-2} - 1} \\
&= \frac{T^{n+m+2} - T^{n-m}}{T - T^{-1}} - \frac{T^{-n-m-2} - T^{-n+m}}{T^{-1} - T} \\
&= \frac{T^{n+m+2} - T^{n-m} - T^{-n+m} + T^{-n-m-2}}{T - T^{-1}} \\
&= \frac{\left(T^{m+1} - T^{-m-1}\right)\left(T^{n+1} - T^{-n-1}\right)}{T - T^{-1}},
\end{aligned}$$

as desired. ∎

3.4.3 Dual Representations of $\mathfrak{sl}_2(\mathbb{C})$

Example 3.83 showed that $V_n \cong V_n^\vee$ for each $n \ge 0$. We would like to be able to provide an explicit such map. Well, recalling Lemma 3.52, we see that we would like to give an invariant bilinear inner product. Here is the result.

Proposition 3.86. Let V_1 be the standard representation $\rho_1\colon \mathfrak{sl}_2(\mathbb{C}) \subseteq \mathfrak{gl}_2(\mathbb{C})$, and define $V_n := \operatorname{Sym}^n V_1$ for each $n \geq 0$ so that we have representations $\rho_n\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_n)$. Then V_n admits an $\mathfrak{sl}_2(\mathbb{C})$-invariant inner product $\langle -, - \rangle$ which is symmetric when n is even and skew-symmetric when n is odd.

Proof. We begin by defining the bilinear form. Set $V := V_{\text{std}}$ for brevity. Now, for any $n \geq 0$, we define $\langle -, - \rangle\colon V_n \times V_n \to \mathbb{C}$ on by

$$\big\langle (v_1, \ldots, v_n), (w_1, \ldots, w_n) \big\rangle := \prod_{i=1}^n \det \begin{bmatrix} | & | \\ v_i & w_i \\ | & | \end{bmatrix}.$$

Because determinants (and products of determinants) are multilinear, we see that this produces a multilinear map $V^{\otimes n} \otimes V^{\otimes n} \to \mathbb{C}$. Thus, we have produced a bilinear form. We now run checks in sequence.

1. Note that this bilinear form on $V^{\otimes n}$ is $\mathrm{SL}_2(\mathbb{C})$-invariant, which implies that it will be $\mathfrak{sl}_2(\mathbb{C})$-invariant upon passing to the differential representation by Lemma 3.50. Well, for any $g \in \mathrm{SL}_2(\mathbb{C})$, it is enough to check the invariance on pure tensors by the ambient bilinearity, so we compute

$$\begin{aligned}
&\big\langle g(v_1 \otimes \cdots \otimes v_n), g(w_1 \otimes \cdots \otimes w_n) \big\rangle \\
&= \big\langle (gv_1 \otimes \cdots \otimes gv_n), (gw_1 \otimes \cdots \otimes gw_n) \big\rangle \\
&= \prod_{i=1}^n \det \begin{bmatrix} | & | \\ gv_i & gw_i \\ | & | \end{bmatrix} \\
&= \prod_{i=1}^n \det g \begin{bmatrix} | & | \\ v_i & w_i \\ | & | \end{bmatrix} \\
&= \underbrace{(\det g)}_{1}{}^n \prod_{i=1}^n \begin{bmatrix} | & | \\ v_i & w_i \\ | & | \end{bmatrix} \\
&= \big\langle g(v_1 \otimes \cdots \otimes v_n), g(w_1 \otimes \cdots \otimes w_n) \big\rangle \\
&= \big\langle (v_1 \otimes \cdots \otimes v_n), (w_1 \otimes \cdots \otimes w_n) \big\rangle,
\end{aligned}$$

as required.

2. We now restrict $\langle -, - \rangle$ to the subrepresentation $\operatorname{Sym}^n V \subseteq V^{\otimes n}$. As such, we see that

$$\begin{aligned}
\langle v_1 \cdots v_n, w_1 \cdots w_n \rangle &= \frac{1}{(n!)^2} \sum_{\sigma, \tau \in S_n} \prod_{i=1}^n \det \begin{bmatrix} | & | \\ v_{\sigma i} & w_{\tau i} \\ | & | \end{bmatrix} \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^n \det \begin{bmatrix} | & | \\ v_{\sigma i} & w_i \\ | & | \end{bmatrix}
\end{aligned}$$

by rearranging. Before doing the non-degeneracy check, we verify that $\langle -, - \rangle$ is symmetric when n is even and skew-symmetric when n is odd. This condition is linear in all the coordinates, so it is enough to check the (skew-)symmetry on the spanning set of tensors of the form $v_1 \cdots v_n$. Well,

$$\begin{aligned}\langle v_1 \cdots v_n, w_1 \cdots w_n \rangle &= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^n \det \begin{bmatrix} | & | \\ v_{\sigma i} & w_i \\ | & | \end{bmatrix} \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^n - \det \begin{bmatrix} | & | \\ w_{\sigma i} & v_i \\ | & | \end{bmatrix} \\ &= (-1)^n \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^n \det \begin{bmatrix} | & | \\ w_i & v_{\sigma i} \\ | & | \end{bmatrix} \\ &= (-1)^n \langle v_1 \cdots v_n, w_1 \cdots w_n \rangle,\end{aligned}$$

as required.

3. We now check that $\langle -, - \rangle$ is non-degenerate. We will do this by computing it on a basis. Let V have basis $\{v_x, v_y\}$, and then we note that $\operatorname{Sym}^n V$ has basis $\{v_x^p v_y^q\}_{p+q=n}$. For bookkeeping reasons, for each pair (p,q), define the function $v_{pq} \colon \{1, \ldots, n\} \to V$ by

$$v_{pq}(i) := \begin{cases} v_x & \text{if } i \le p, \\ v_y & \text{if } i > p. \end{cases}$$

In particular, $v_x^p v_y^q = v_{pq}(1) \cdots v_{pq}(n)$. We now choose two basis vectors $v_x^p v_y^q$ and $v_x^{p'} v_y^{q'}$ and compute

$$\begin{aligned}\left\langle v_x^p v_y^q, v_x^{p'} v_y^{q'} \right\rangle &= \langle v_{pq}(1) \cdots v_{pq}(n), v_{p'q'}(1) \cdots v_{p'q'}(n) \rangle \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^n \det \begin{bmatrix} | & | \\ v_{pq}(\sigma i) & v_{p'q'}(i) \\ | & | \end{bmatrix} \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^{p'} \det \begin{bmatrix} | & | \\ v_{pq}(\sigma i) & v_x \\ | & | \end{bmatrix} \prod_{i=p'+1}^{n} \det \begin{bmatrix} | & | \\ v_{pq}(\sigma i) & v_y \\ | & | \end{bmatrix}\end{aligned}$$

Now, the only way for a product of these determinants to not vanish is to have $v_{pq}(\sigma i) = v_y$ for $i \le p'$ and $v_{pq}(\sigma i) = v_x$ for $i > p'$. In particular, by counting the number of v_xs and v_ys, we see that all terms of the sum vanish unless $(p,q) = (q',p')$. In the case where $(p,q) = (q',p')$, we have

$$\left\langle v_x^p v_y^q, v_x^q v_y^p \right\rangle = \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^{q} \det \begin{bmatrix} | & | \\ v_{pq}(\sigma i) & v_x \\ | & | \end{bmatrix} \prod_{i=q+1}^{n} \det \begin{bmatrix} | & | \\ v_{pq}(\sigma i) & v_y \\ | & | \end{bmatrix},$$

so we see that each nonzero term will evaluate to $(-1)^q$, and the number of nonzero terms is the number of $\sigma \in S_n$ such that σ carries $\{1, \ldots, q\}$ to $\{p+1, \ldots, n\}$. There are $p!q!$ total such permutations, so we conclude that

$$\left\langle v_x^p v_y^q, v_x^q v_y^p \right\rangle = (-1)^q \frac{p!q!}{n!} \ne 0.$$

Thus, we see that the matrix given by the bilinear form on the basis $\left\{v_x^p v_y^q\right\}_{p+q=n}$ is anti-diagonal with all nonzero anti-diagonal entries, so it is invertible. Thus, the relevant bilinear form is non-degenerate. ■

3.4.4 The Universal Enveloping Algebra

We have shown that the category $\operatorname{Rep}_{\mathbb{C}} \mathfrak{g}$ is abelian for any Lie algebra $\mathfrak{g}$, so we may expect to be able to realize this category as a category of modules over some (possibly non-commutative) ring. This is the role of the universal enveloping algebra.

To start, we begin with a universal algebra which does not remember the Lie bracket.

Definition 3.87 (universal tensor algebra)**.** Fix a vector space $\mathfrak{g}$ over a field F. Then the *universal tensor algebra* is the vector space

$$T\mathfrak{g} := \bigoplus_{k=0}^{\infty} \mathfrak{g}^{\otimes k}.$$

We turn $T\mathfrak{g}$ into an F-algebra by defining multiplication $T\mathfrak{g} \otimes_F T\mathfrak{g} \to T\mathfrak{g}$ on the components $\mathfrak{g}^{\otimes k} \otimes \mathfrak{g}^{\otimes \ell} \to \mathfrak{g}^{\otimes (k+\ell)}$ by the natural "concatenation" isomorphism. We will let $\iota_{\mathfrak{g}} \colon \mathfrak{g} \to T\mathfrak{g}$ denote the canonical map.

Remark 3.88. We won't bother to check that $T\mathfrak{g}$ is in fact an associative F-algebra, but we do note that the multiplication is F-linear automatically by the construction.

Example 3.89. Fix a basis $\{X_1, \ldots, X_n\}$ of some $\mathfrak{g}$. Then $T\mathfrak{g}$ is (by its definition) the free (non-commutative!) polynomial algebra $F\langle X_1, \ldots, X_n\rangle$.

Quickly, we note that this construction is functorial.

Lemma 3.90. Fix a field F. The universal tensor algebra defines a functor $\mathrm{Vec}(F) \to \mathrm{Alg}(F)$.

Proof. To begin, note that a linear map $f \colon \mathfrak{g} \to \mathfrak{h}$ of F-vector spaces induces an F-algebra map $Tf \colon T\mathfrak{g} \to T\mathfrak{h}$. Namely, note there is certainly an F-linear map $f \colon \mathfrak{g}^{\otimes k} \to \mathfrak{h}^{\otimes k}$ on the components of $T\mathfrak{g}$ by functoriality of the tensor product (more precisely, note that the map $\mathfrak{g}^k \to \mathfrak{h}^{\otimes k}$ given by $(v_1, \ldots, v_k) \mapsto f(v_1) \otimes \cdots \otimes f(v_k)$ if F-multilinear), so we get a linear map $Tf \colon T\mathfrak{g} \to T\mathfrak{h}$. Then we check that Tf is an algebra map: it is enough to check the vanishing on a spanning subset of $T\mathfrak{g}$, for which we see that pure tensors span by the definition of $T\mathfrak{g}$, so it is enough to compute

$$\begin{aligned} Tf((v_1 \otimes \cdots \otimes v_k) \cdot (w_1 \otimes \cdots \otimes w_\ell)) &= f(v_1) \otimes \cdots \otimes f(v_k) \otimes f(w_1) \otimes \cdots \otimes f(w_\ell) \\ &= Tf(v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_\ell). \end{aligned}$$

Quickly, we note that Tf is in fact the unique map making the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "f"] \arrow[d, "\iota_{\mathfrak{g}}"'] & \mathfrak{h} \arrow[d, "\iota_{\mathfrak{h}}"] \\ T\mathfrak{g} \arrow[r, "Tf"] & T\mathfrak{h} \end{tikzcd}$$

commute. Above we checked that Tf is a well-defined algebra morphism, and its definition gives $Tf(\iota_{\mathfrak{g}} v) = \iota - \mathfrak{h} f(v)$, so the diagram commutes. For the uniqueness, note that any algebra morphism Tf is determined by a spanning subset of $T\mathfrak{g}$, for which we can take the pure tensors; however, being an algebra morphism then requires that

$$Tf(v_1 \otimes \cdots \otimes v_k) = f(v_1) \otimes \cdots \otimes f(v_k)$$

on pure tensors, fully determining Tf.

It remains to run some functoriality checks.

- Identity: the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{id}_{\mathfrak{g}}"] \arrow[d, "\iota_{\mathfrak{g}}"'] & \mathfrak{g} \arrow[d, "\iota_{\mathfrak{g}}"] \\ T\mathfrak{g} \arrow[r, "\mathrm{id}_{T\mathfrak{g}}"] & T\mathfrak{g} \end{tikzcd}$$

commutes, so the uniqueness of $T\,\mathrm{id}_{\mathfrak{g}}$ requires $T\mathrm{id}_{\mathfrak{g}} = \mathrm{id}_{T\mathfrak{g}}$.

- Associativity: for morphisms $f\colon \mathfrak{g} \to \mathfrak{h}$ and $g\colon \mathfrak{h} \to \mathfrak{k}$, we see that the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "f"] \arrow[d, "\iota_{\mathfrak{g}}"'] & \mathfrak{h} \arrow[r, "g"] \arrow[d, "\iota_{\mathfrak{h}}"'] & \mathfrak{k} \arrow[d, "\iota_{\mathfrak{k}}"] \\ T\mathfrak{g} \arrow[r, "Tf"'] & T\mathfrak{h} \arrow[r, "Tg"'] & T\mathfrak{k} \end{tikzcd}$$

  commutes, so the uniqueness of $T(g \circ f)$ forces $T(g \circ f) = Tg \circ Tf$. ■

Here is the universal property.

Lemma 3.91. Fix a vector space $\mathfrak{g}$ over a field F, and let $\iota\colon \mathfrak{g} \to T\mathfrak{g}$ be the natural map.

(a) For any F-algebra A, restriction provides a natural bijection

$$\operatorname{Hom}_{\mathrm{Alg}(F)}(T\mathfrak{g}, A) \to \operatorname{Hom}_F(\mathfrak{g}, A).$$

(b) The category $\operatorname{Mod}_F(T\mathfrak{g})$ is equivalent to the category $\operatorname{Mod}_F(\mathfrak{g})$, where a $\mathfrak{g}$-module is an F-vector space V with a morphism $\mathfrak{g} \to \mathfrak{gl}(V)$. The functor $\operatorname{Mod}_F(T\mathfrak{g}) \to \operatorname{Mod}_F(\mathfrak{g})$ sends algebra morphisms $\varphi\colon T\mathfrak{g} \to \mathfrak{gl}(V)$ to linear morphisms $(\varphi \circ \iota)\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

Proof. We run our checks separately.

(a) Of course, one can restrict an F-algebra morphism $T\mathfrak{g} \to A$ to a morphism $\mathfrak{g} \to A$ via the inclusion $\iota\colon \mathfrak{g} \subseteq T\mathfrak{g}$. Here are our checks on this construction.

- Linear: given $a_1, a_2 \in F$ and $\varphi_1, \varphi_2\colon T\mathfrak{g} \to A$, we see that

$$(a_1\varphi_1 + a_2\varphi_2) \circ \iota = a_1(\varphi \circ \iota) + a_2(\varphi \circ \iota)$$

  because all maps in sight are linear.
- Injective: it is enough to show that we have trivial kernel, so suppose that $\varphi\colon T\mathfrak{g} \to A$ has $\varphi \circ \iota = 0$, and we want to check that $\varphi = 0$. Well, $T\mathfrak{g}$ is spanned by its components $\mathfrak{g}^{\otimes k}$, so it is enough to check that $\varphi|_{\mathfrak{g}^{\otimes k}}$. Further, $\mathfrak{g}^{\otimes k}$ is spanned by pure tensors, so it is enough to check that the linear map φ vanishes on pure tensors in $\mathfrak{g}^{\otimes k}$. Well, any pure tensor looks like $v_1 \otimes \cdots \otimes v_k$ for some vectors $v_1, \ldots, v_k \in \mathfrak{g}$, for which we note

$$\varphi(v_1 \otimes \cdots \otimes v_k) = \varphi(v_1) \cdots \varphi(v_k) = 0 \cdots 0 = 0.$$

- Surjective: given any linear map $\psi\colon \mathfrak{g} \to A$, we must extend it to an algebra map $\varphi\colon T\mathfrak{g} \to A$. We begin by defining the linear map φ, and then we will show that it is actually a map of F-algebras. Well, it is enough to define φ on each of the components $\mathfrak{g}^{\otimes k}$ and then take the direct sum; thus, we need to define an F-multilinear map $\varphi\colon \mathfrak{g}^k \to A$, for which we take

$$\varphi(v_1, \ldots, v_k) \coloneqq \psi(v_1) \cdots \psi(v_k).$$

  Because A is an F-algebra, this map is in fact F-multilinear, so we descend to a linear map $\varphi\colon \mathfrak{g}^{\otimes k} \to A$, which we can then sum together to produce an F-linear map $\varphi\colon T\mathfrak{g} \to A$.
  It remains to check that φ is actually multiplicative. Well, the condition that $\varphi(x)\varphi(y) - \varphi(xy) = 0$ for all $x, y \in T\mathfrak{g}$ is equivalent to the vanishing of the corresponding bilinear functional $T\mathfrak{g} \times T\mathfrak{g} \to A$. Thus, we are checking if some linear functional $T\mathfrak{g} \otimes_F T\mathfrak{g} \to A$ vanishes, which we can check on a spanning subset of $T\mathfrak{g} \otimes_F T\mathfrak{g}$. Well, we see that $T\mathfrak{g}$ is spanned by pure tensors of the form $(v_1 \otimes \cdots \otimes v_k)$, so it suffices to compute that

$$\varphi(v_1 \otimes \cdots \otimes v_k)\varphi(w_1 \otimes \cdots \otimes w_\ell) = \psi(v_1) \cdots \psi(v_k)\psi(w_1) \cdots \psi(w_\ell) = \varphi((v_1 \otimes \cdots \otimes v_k) \cdot (w_1 \otimes \cdots \otimes w_\ell)),$$

  as required.

- Natural in A: we claim that the given morphism is natural in A. Indeed, we note that any F-algebra map $f\colon A \to B$ makes the diagram

$$\begin{tikzcd} \operatorname{Hom}_{\operatorname{Alg}(F)}(T\mathfrak{g}, A) \arrow[r, "\iota"] \arrow[d, "f"'] & \operatorname{Hom}_F(\mathfrak{g}, A) \arrow[d, "f"] \\ \operatorname{Hom}_{\operatorname{Alg}(F)}(T\mathfrak{g}, B) \arrow[r, "\iota"] & \operatorname{Hom}_F(\mathfrak{g}, B) \end{tikzcd} \qquad \begin{tikzcd} \varphi \arrow[r, mapsto] \arrow[d, mapsto] & (\varphi \circ \iota) \arrow[d, mapsto] \\ (f \circ \varphi) \arrow[r, mapsto] & (f \circ \varphi \circ \iota) \end{tikzcd}$$

  commute.
- Natural in $\mathfrak{g}$: we claim that the given morphism is natural in $\mathfrak{g}$. Indeed, for any linear map $f\colon \mathfrak{g} \to \mathfrak{h}$, we note that the diagram

$$\begin{tikzcd} \operatorname{Hom}_{\operatorname{Alg}(F)}(T\mathfrak{g}, A) \arrow[r, "\iota_{\mathfrak{g}}"] & \operatorname{Hom}_F(\mathfrak{g}, A) \\ \operatorname{Hom}_{\operatorname{Alg}(F)}(T\mathfrak{h}, A) \arrow[r, "\iota_{\mathfrak{h}}"] \arrow[u, "Tf"] & \operatorname{Hom}_F(\mathfrak{h}, A) \arrow[u, "f"'] \end{tikzcd} \qquad \begin{tikzcd} \varphi \circ Tf \arrow[r, mapsto] & \varphi \circ (Tf \circ \iota_{\mathfrak{g}}) \\ \varphi \arrow[r, mapsto] \arrow[u, mapsto] & \varphi \circ \iota_{\mathfrak{h}} \arrow[u, mapsto] \end{tikzcd}$$

  commutes because $\iota_{\mathfrak{h}} \circ f = \iota_{\mathfrak{g}} \circ Tf$.

(b) Here are the checks on this functor. Throughout, V and W denote F-vector spaces with structure morphisms $\varphi_V : T\mathfrak{g} \to \mathfrak{gl}(V)$ and $\varphi_W : T\mathfrak{g} \to \mathfrak{gl}(W)$ if they are in $\operatorname{Mod}_F(T\mathfrak{g})$ and with structure morphisms $\psi_V : \mathfrak{g} \to \mathfrak{gl}(V)$ and $\psi_W : \mathfrak{g} \to \mathfrak{gl}(W)$ if they are in $\operatorname{Mod}_F(\mathfrak{g})$.

- Functorial: if $f\colon V \to W$ is a map in $\operatorname{Mod}_F(T\mathfrak{g})$, we claim that this is also a map in $\operatorname{Mod}_F(\mathfrak{g})$. Indeed, we are being given that

$$f \circ \varphi_V(X) = \varphi_W(X) \circ f$$

  for any $X \in T\mathfrak{g}$. Restricting our attention to $X \in \mathfrak{g}$, we see that

$$f \circ (\varphi_V \circ \iota)(X) = (\varphi_W \circ \iota)(X) \circ f$$

  for any $X \in \mathfrak{g}$, so we are done.
  Because the induced maps

$$\operatorname{Hom}_{T\mathfrak{g}}(V, W) \to \operatorname{Hom}_{\mathfrak{g}}(V, W)$$

  are just inclusion maps, we see that our mapping is automatically functorial and faithful.
- Full: given $V, W \in \operatorname{Mod}_{T\mathfrak{g}}(A)$, we need to show that any morphism $f\colon V \to W$ preserved by $\mathfrak{g}$ is fully preserved by $T\mathfrak{g}$. Explicitly, we are given that

$$f \circ \varphi_V(X) = \varphi_W(X) \circ f$$

  for any $X \in \mathfrak{g}$, which we would like to extend to all $T\mathfrak{g}$. However, the condition that $f \circ \varphi_V(v) = \varphi_W(v) \circ f$ for all v is linear in $v \in T\mathfrak{g}$, so we may check it on a spanning set. Pure tensors span $T\mathfrak{g}$, so we may assume that $v = v_1 \otimes \cdots \otimes v_k$ for some $v_1, \ldots, v_k \in \mathfrak{g}$, for which we note that

$$\begin{aligned} f \circ \varphi_V(v_1 \otimes \cdots \otimes v_k) &= f \circ \varphi_V(v_1) \circ \cdots \circ \varphi_V(v_k) \\ &\stackrel{*}{=} \varphi_W(v_1) \circ \cdots \circ \varphi_W(v_k) \circ f \\ &= \varphi(v_1 \otimes \cdots \otimes v_k) \circ f, \end{aligned}$$

  where $\stackrel{*}{=}$ holds by applying the hypothesis inductively.
- Essentially surjective: given any linear map $\psi\colon \mathfrak{g} \to \mathfrak{gl}(V)$, we need to extend it to an algebra map $\varphi\colon T\mathfrak{g} \to \mathfrak{gl}(V)$. This is the content of (a). ■

We now take a quotient to remember the Lie bracket.

Definition 3.92 (universal enveloping algebra)**.** Fix a Lie algebra $\mathfrak{g}$ over a field F. We define the *universal enveloping algebra* $U\mathfrak{g}$ as the quotient of $T\mathfrak{g}$ by the two-sided ideal $L\mathfrak{g}$ generated by the elements

$$\{X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g}\}.$$

We continue to denote the natural linear map $\mathfrak{g} \to U\mathfrak{g}$ by $\iota_{\mathfrak{g}}$, though we frequently omit writing this morphism for brevity (and will as such treat elements of $\mathfrak{g}$ as already living in $U\mathfrak{g}$).

Remark 3.93. Technically speaking, the construction of $U\mathfrak{g}$ did not require that $[-,-]$ be a Lie bracket. In particular, one can imagine constructing a version of $U\mathfrak{g}$ for any bilinear form $[-,-]$ on a vector space $\mathfrak{g}$ by taking the quotient of $T\mathfrak{g}$ by the two-sided ideal generated by

$$\{X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g}\}.$$

As before, we quickly note that this construction is functorial.

Lemma 3.94. Fix a field F. The universal enveloping algebra defines a functor $\operatorname{LieAlg}(F) \to \operatorname{Alg}(F)$.

Proof. We begin by defining U on morphisms. We claim that there is a unique morphism Uf making the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "f"] \arrow[d, "\iota_{\mathfrak{g}}"'] & \mathfrak{h} \arrow[d, "\iota_{\mathfrak{h}}"] \\ U\mathfrak{g} \arrow[r, "Uf"] & U\mathfrak{h} \end{tikzcd}$$

commute. For the uniqueness, we note that the required morphism, being a morphism of algebras, will be defined on a generating subset of $U\mathfrak{g}$. The subset $\mathfrak{g} \subseteq T\mathfrak{g}$ generates, and $T\mathfrak{g} \twoheadrightarrow U\mathfrak{g}$ surjects, so it is enough to define Uf on $\mathfrak{g}$. But the diagram dictates $Uf(\iota_{\mathfrak{g}}X) = \iota_{\mathfrak{h}}f(X)$, so Uf is uniquely determined.

For existence, we note that we already have a morphism $Tf\colon T\mathfrak{g} \to T\mathfrak{h}$ of algebras, which we would like to descend to a quotient morphism $Uf\colon U\mathfrak{g} \to U\mathfrak{h}$. For this, it is enough to check that $Tf(L\mathfrak{g}) \subseteq L\mathfrak{h}$. Because Tf is an algebra morphism, it is enough to check the inclusion on generators of $L\mathfrak{g}$, for which we note that the elements $X \otimes Y - Y \otimes X - [X, Y]$ generate, so we compute

$$\begin{aligned} Tf(X \otimes Y - Y \otimes X - [X, Y]) &= f(X) \otimes f(Y) - f(Y) \otimes f(X) - f([X, Y]) \\ &= f(X) \otimes f(Y) - f(Y) \otimes f(X) - [f(X), f(Y)] \end{aligned}$$

is a generating element of $L\mathfrak{h}$, so we are done. We now note that the functoriality checks for $U\mathfrak{g}$ are identical to the functoriality checks for $T\mathfrak{g}$ done at the end of the proof of Lemma 3.90. ■

Here is the universal property.

Lemma 3.95. Fix a Lie algebra $\mathfrak{g}$ over a field F, and let $\iota\colon \mathfrak{g} \to U\mathfrak{g}$ be the natural map.

(a) For any F-algebra A, restriction provides a natural bijection

$$\operatorname{Hom}_{\operatorname{Alg}(F)}(U\mathfrak{g}, A) \to \operatorname{Hom}_{\operatorname{LieAlg}(\mathrm{F})}(\mathfrak{g}, A).$$

(b) The category $\operatorname{Mod}_F(U\mathfrak{g})$ is equivalent to the category $\operatorname{Mod}_F(\mathfrak{g})$.

Proof. We run our checks separately.

(a) We begin by checking that the given map is well-defined. Namely, if $\varphi\colon U\mathfrak{g} \to A$ is a morphism of algebras, then $(\varphi \circ \iota_{\mathfrak{g}})\colon \mathfrak{g} \to A$ is a morphism of Lie algebras. Well, for any $X, Y \in \mathfrak{g}$, we compute

$$\begin{aligned}(\varphi \circ \iota_{\mathfrak{g}})([X,Y]) &= \varphi(\iota_{\mathfrak{g}}[X,Y]) \\ &\overset{*}{=} \varphi(X \otimes Y - Y \otimes X) \\ &= \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X),\end{aligned}$$

where the key step is $\overset{*}{=}$ where we used the construction of $U\mathfrak{g}$.

The linearity, injectivity, and naturality checks are now all exactly the same as in Lemma 3.91 (merely replace T with U throughout), so it only remains to check surjectivity. Namely, given a Lie algebra morphism $\psi\colon \mathfrak{g} \to A$, we must extend it to an algebra map $\overline{\varphi}\colon U\mathfrak{g} \to A$. Well, Lemma 3.91 provides some algebra map $\varphi\colon T\mathfrak{g} \to A$, which we would like to show descends to the quotient to give a morphism $\overline{\varphi}\colon U\mathfrak{g} \to A$. To descend to the quotient, we want to check that $L\mathfrak{g} \subseteq \ker\varphi$, for which we note that it is really enough to show that a generating subset of $L\mathfrak{g}$ is contained in $\ker\varphi$. For this, we compute

$$\varphi(X \otimes Y - Y \otimes X - [X,Y]) = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) - \varphi([X,Y])$$

vanishes because $\varphi\colon \mathfrak{g} \to A$ is a morphism of Lie algebras.

(b) The exact same proof as in Lemma 3.91 goes through after replacing T with U throughout. ■

In order to actually compute $U\mathfrak{g}$, one can fix a basis $\{X_1, \ldots, X_n\}$ of $\mathfrak{g}$ and then note that $T\mathfrak{g}$ will be the free polynomial ring in these variables, so we can take a quotient to recover $U\mathfrak{g}$.

Example 3.96. If $\mathfrak{g}$ is any vector space, we can upgrade $\mathfrak{g}$ to a Lie algebra with an abelian Lie bracket (namely, $[X,Y] = 0$ for all $X, Y \in \mathfrak{g}$). Giving $\mathfrak{g}$ the usual basis, we get the polynomial ring

$$S\mathfrak{g} := U\mathfrak{g} = \frac{T\mathfrak{g}}{(X_i \otimes X_j - X_j \otimes X_i : 1 \le i, j \le n)} = F[X_1, \ldots, X_n].$$

Even if $\mathfrak{g}$ is a Lie algebra, we can forget about its Lie bracket and replace it with the abelian Lie bracket to produce a commutative F-algebra $S\mathfrak{g}$.

Example 3.97. We see that

$$U(\mathfrak{sl}_2(\mathbb{C})) = \frac{\mathbb{C}\langle e, f, h\rangle}{(ef - fe - h, he - eh - 2e, hf - fh + 2f)}.$$

3.4.5 The Adjoint Action

We close this initial discussion of $U\mathfrak{g}$ by noting that $U\mathfrak{g}$ has an adjoint action.

Lemma 3.98. Fix a Lie algebra $\mathfrak{g}$ over a field F. For $A\mathfrak{g} \in \{T\mathfrak{g}, U\mathfrak{g}, S\mathfrak{g}\}$, there are unique Lie algebra morphisms $\operatorname{ad}_\bullet\colon \mathfrak{g} \to \mathfrak{gl}(A\mathfrak{g})$ (with infinite-dimensional target!) satisfying the following.

- $\operatorname{ad}_X(\iota_{\mathfrak{g}} Y) = \iota_{\mathfrak{g}}([X,Y])$ for all $X, Y \in \mathfrak{g}$.
- Leibniz rule: $\operatorname{ad}_X(ab) = (\operatorname{ad}_X a)b + a(\operatorname{ad}_X b)$.

For $U\mathfrak{g}$, we have $\operatorname{ad}_X(a) = (\iota_{\mathfrak{g}} X)a - a(\iota_{\mathfrak{g}} X)$.

Proof. Throughout, we write $A\mathfrak{g}$ to denote either $U\mathfrak{g}$ or $S\mathfrak{g}$ when the proof works for both. We begin by checking the uniqueness of $\operatorname{ad}_\bullet$. This means that we must define $\operatorname{ad}_X\colon A\mathfrak{g} \to A\mathfrak{g}$ for each $X \in \mathfrak{g}$. For this, it is enough to show that the definition of ad_X is determined by a spanning subset of $U\mathfrak{g}$. Well, the surjection $T\mathfrak{g} \twoheadrightarrow A\mathfrak{g}$ tells us that pure tensors $Y_1 \otimes \cdots \otimes Y_k$ span $A\mathfrak{g}$ because they span $A\mathfrak{g}$. We claim that the definition of ad_X is uniquely defined on these pure tensors from the given rules for each k, which we prove by induction on k. For $k = 0$, there is nothing to do because $\mathrm{ad}_X(0) = 0$. For the inductive step, we note that the Leibniz rule requires

$$\mathrm{ad}_X(Y_1 \otimes \cdots \otimes Y_k \otimes Y_{k+1}) = \mathrm{ad}_X(Y_1 \otimes \cdots \otimes Y_k)\iota_{\mathfrak{g}} Y_{k+1} + (Y_1 \otimes \cdots \otimes Y_k)\iota_{\mathfrak{g}} \mathrm{ad}_X(Y_{k+1}),$$

and the right-hand side is uniquely determined by the inductive hypothesis.

It remains to show that there exists a Lie algebra morphism $\mathrm{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(A\mathfrak{g})$ satisfying the given properties. Well, we begin by defining the map on the level of $T\mathfrak{g}$. To merely define the map, it is enough to define it on the components $\mathfrak{g}^{\otimes k}$, so we need an F-multillinear map $\mathfrak{g}^k \to \mathfrak{g}^{\otimes k}$, for which we take motivation from the Leibniz rule to write

$$\mathrm{ad}_X(Y_1, \ldots, Y_k) := \sum_{i=1}^k Y_1 \otimes \cdots \otimes [X, Y_i] \otimes \cdots \otimes X_k$$

for any $Y_1, \ldots, Y_k \in \mathfrak{g}$. This map is of course F-multilinear because the tensor product and Lie bracket are both multilinear, so we have induced a map $\mathfrak{g}^{\otimes k} \to \mathfrak{g}^{\otimes k}$, which then sums to a map $\mathrm{ad}_X \colon T\mathfrak{g} \to T\mathfrak{g}$. Here are some checks on this action.

- We note that $\mathrm{ad}_X \colon T\mathfrak{g} \to T\mathfrak{g}$ is F-linear by its construction.
- For any $Y \in \mathfrak{g}$, we note that $\mathrm{ad}_X Y = [X, Y]$ by construction.
- Leibniz rule: we claim that $\mathrm{ad}_X(ab) = (\mathrm{ad}_X a)b + a(\mathrm{ad}_X b)$ for any $X \in \mathfrak{g}$ and $a, b \in T\mathfrak{g}$. This corresponds to an equality of F-bilinear maps $T\mathfrak{g} \times T\mathfrak{g} \to T\mathfrak{g}$, so the equality can be checked on a spanning subset of $T\mathfrak{g} \otimes_F T\mathfrak{g}$. For this, we note that $T\mathfrak{g}$ has a spanning subset of pure tensors, so it is enough to compute

$$\begin{aligned}
\mathrm{ad}_X((Y_1 \otimes \cdots \otimes Y_k)(Z_1 \otimes \cdots \otimes Z_\ell)) &= \mathrm{ad}_X((Y_1 \otimes \cdots \otimes Y_k) \otimes (Z_1 \otimes \cdots \otimes Z_\ell)) \\
&= \sum_{i=1}^k (Y_1 \otimes \cdots \otimes [X, Y_i] \otimes \cdots \otimes Y_k) \otimes (Z_1 \otimes \cdots \otimes Z_\ell) \\
&\quad + \sum_{j=1}^\ell (Y_1 \otimes \cdots \otimes Y_k) \otimes (Z_1 \otimes \cdots \otimes [X, Z_j] \otimes \cdots \otimes Z_\ell) \\
&= \mathrm{ad}_X(Y_1 \otimes \cdots \otimes Y_k)(Z_1 \otimes \cdots \otimes Z_\ell) \\
&\quad + (Y_1 \otimes \cdots \otimes Y_k)\,\mathrm{ad}_X(Z_1 \otimes \cdots \otimes Z_\ell),
\end{aligned}$$

  as required.
- Lie algebra morphism: we claim that $\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \circ \mathrm{ad}_Y - \mathrm{ad}_Y \circ \mathrm{ad}_X$. This corresponds to equality of morphisms $T\mathfrak{g} \to T\mathfrak{g}$, so we can check it on a spanning subset of $T\mathfrak{g}$, for which we use the pure

tensors. As such, we compute

$$\begin{aligned}
&(\mathrm{ad}_X \circ \mathrm{ad}_Y - \mathrm{ad}_Y \circ \mathrm{ad}_X)(Z_1 \otimes \cdots \otimes Z_k) \\
=\ & \mathrm{ad}_X \mathrm{ad}_Y (Z_1 \otimes \cdots \otimes Z_k) - \mathrm{ad}_Y \mathrm{ad}_X (Z_1 \otimes \cdots \otimes Z_k) \\
=\ & \sum_{i=1}^k \mathrm{ad}_X (Z_1 \otimes \cdots \otimes [Y, Z_i] \otimes \cdots \otimes Z_k) \\
& \quad - \sum_{j=1}^k \mathrm{ad}_Y (Z_1 \otimes \cdots \otimes [X, Z_i] \otimes \cdots \otimes Z_k) \\
=\ & \sum_{i=1}^k (Z_1 \otimes \cdots \otimes [X, [Y, Z_i]] \otimes \cdots \otimes Z_k) \\
& \quad + \sum_{\substack{1 \le i,j \le k \\ i \ne j}}^k (Z_1 \otimes \cdots \otimes [Y, Z_i] \otimes \cdots \otimes [X, Z_j] \otimes \cdots \otimes Z_k) \\
& \quad - \sum_{j=1}^k (Z_1 \otimes \cdots \otimes [Y, [X, Z_i]] \otimes \cdots \otimes Z_k) \\
& \quad - \sum_{\substack{1 \le i,j \le k \\ i \ne j}}^k (Z_1 \otimes \cdots \otimes [Y, Z_i] \otimes \cdots \otimes [X, Z_j] \otimes \cdots \otimes Z_k) \\
=\ & \sum_{i=1}^k (Z_1 \otimes \cdots \otimes ([X, [Y, Z_i]] - [Y, [X, Z_i]]) \otimes \cdots \otimes Z_k) \\
\overset{*}{=}\ & \sum_{i=1}^k (Z_1 \otimes \cdots \otimes [[X, Y], Z_i] \otimes \cdots \otimes Z_k) \\
=\ & \mathrm{ad}_{[X,Y]} (Z_1 \otimes \cdots \otimes Z_k),
\end{aligned}$$

where we have used the Jacobi identity at $\overset{*}{=}$.

We now descend this definition of ad_X to a map $A\mathfrak{g} \to A\mathfrak{g}$ for $A\mathfrak{g} \in \{U\mathfrak{g}, S\mathfrak{g}\}$. Let I be the kernel of the natural projection $T\mathfrak{g} \to A\mathfrak{g}$, and we want to check that $\mathrm{ad}_X(I) \subseteq I$. Note that it suffices to check this on generators of I: if $J \subseteq I$ is the subspace such that $\mathrm{ad}_X(a) \subseteq I$ for each $a \in J$, we claim that J is an ideal. Indeed, we note that J is certainly a linear subspace (it is the pre-image of a subspace under a linear map), and for any $a \in T\mathfrak{g}$ and $b \in J$, we see that $ab \in I$ and

$$\mathrm{ad}_X(ab) = \mathrm{ad}_X(a)b + a \underbrace{\mathrm{ad}_X(b)}_{\in I} \in I,$$

so J is closed under multiplication by $T\mathfrak{g}$. Thus, to check that $J = I$, it is enough to check that J contains the generators of I.

- In the case where $A\mathfrak{g} = S\mathfrak{g}$, we see that the elements $Y \otimes Z - Z \otimes Y$ generate I, so we compute that

$$\begin{aligned}
\mathrm{ad}_X(Y \otimes Z - Z \otimes Y) &= ([X, Y] \otimes Z + Y \otimes [X, Z]) - ([X, Z] \otimes Y - Z \otimes [X, Y]) \\
&= ([X, Y] \otimes Z - Z \otimes [X, Y]) + (Y \otimes [X, Z] - [X, Z] \otimes Y)
\end{aligned}$$

  lives in I.

- In the case where $A\mathfrak{g} = U\mathfrak{g}$, we see that the elements $Y \otimes Z - Z \otimes Y - [Y, Z]$ generate I, so we compute that

$$\begin{aligned}
\mathrm{ad}_X(Y \otimes Z - Z \otimes Y - [Y, Z]) &= Y \otimes [X, Z] + [X, Y] \otimes Z - [X, Z] \otimes Y - Z \otimes [X, Y] - [X, [Y, Z]] \\
&= (Y \otimes [X, Z] - [X, Z] \otimes Y - [Y, [X, Z]]) \\
&\quad + ([X, Y] \otimes Z - Z \otimes [X, Y] - [[X, Y], Z]),
\end{aligned}$$

Now, the checks that $\operatorname{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(A\mathfrak{g})$ is a Lie algebra morphism satisfying the required properties follow because $\operatorname{ad}_X \colon A\mathfrak{g} \to A\mathfrak{g}$ is a quotient of the map $\operatorname{ad}_X \colon T\mathfrak{g} \to T\mathfrak{g}$. In particular, all the checks we needed to do amounted to checking some equalities of functions from some tensor power of $T\mathfrak{g}$ to $T\mathfrak{g}$, and these maps all quotient appropriately down to $A\mathfrak{g}$.

Lastly, we must check that $\operatorname{ad}_X(a) = Xa - aX$ in the case where $A\mathfrak{g} = U\mathfrak{g}$. As usual, we note that we are checking the equality of some linear maps $Y\mathfrak{g} \to U\mathfrak{g}$, so it is enough to check this on a spanning subset of $U\mathfrak{g}$, for which we use the pure tensors: we compute

$$\begin{aligned}
\operatorname{ad}_X(Y_1 \cdots Y_k) &= \sum_{i=1}^{k} (Y_1 \cdots [X, Y_i] \cdots Y_k) \\
&= \sum_{i=1}^{k} (Y_1 \cdots Y_{i-1} X Y_i Y i + 1 \cdots Y_k - Y_1 \cdots Y_{i-1} Y_i X Y i + 1 \cdots Y_k) \\
&= X(Y_1 \cdots Y_k) - (Y_1 \cdots Y_k) X,
\end{aligned}$$

where the last equality uses the observation that the given sum telescopes. ■

3.5 October 9

Today we continue discussing the universal enveloping algebra.

3.5.1 Gradings and Filtrations

We would like to "temper" the infinite-dimensional representation $\operatorname{ad}_\bullet \, \mathfrak{g} \to \mathfrak{gl}(U\mathfrak{g})$ given by Lemma 3.98. For $T\mathfrak{g}$, one has a grading.

Definition 3.99 (graded algebra). Fix a field F. A *grading* on an F-algebra A is a decomposition $A = \bigoplus_{i=0}^{\infty} A_i$ where each $A_\bullet \subseteq A$ is a subspace, and

$$A_i \cdot A_j \subseteq A_{i+j}$$

for all $i, j \geq 0$. An F-algebra equipped with a grading is a *graded algebra*. An element of some A_i is *homogeneous*.

Example 3.100. Any F-algebra A has a trivial grading given by $A_0 = A$ and $A_i = 0$ for all $i > 0$. Indeed, $A_i \cdot A_j \subseteq A_{i+j}$ has $A_i A_j = 0$ unless $i = j = 0$, in which case we see that the inclusion reads $A_0 A_0 = A = A_0$.

Example 3.101. Fix a vector space $\mathfrak{g}$. Then the universal tensor algebra $T\mathfrak{g}$ is graded by $T\mathfrak{g} = \bigoplus_{i \geq 0} \mathfrak{g}^{\otimes i}$. Indeed, $\mathfrak{g}^{\otimes i} \subseteq T\mathfrak{g}$ has

$$\mathfrak{g}^{\otimes i} \cdot \mathfrak{g}^{\otimes j} \to \mathfrak{g}^{\otimes (i+j)}$$

for all $i, j \geq 0$ by definition of our multiplication.

Gradings sometimes descend quotients.

Lemma 3.102. Fix a graded F-algebra $A = \bigoplus_{i=0}^\infty A_i$. If $I \subseteq A$ is a two-sided ideal generated by homogeneous elements, then we have a decomposition

$$\frac{A}{I} = \bigoplus_{i=0}^\infty \frac{A_i}{I \cap A_i}$$

making A/I into a graded F-algebra.

Proof. Note that we have maps $A_i \subseteq A \twoheadrightarrow A/I$ with kernel $I \cap A_i$, so we induce injections $A_i/(I \cap A_i) \hookrightarrow A/I$. Summing over all i, we get an F-linear map

$$\bigoplus_{i=0}^\infty \frac{A_i}{I \cap A_i} \to \frac{A}{I}.$$

We claim that this map is an isomorphism of F-vector spaces providing the required grading. Here are our checks.

- Surjective: note that any $a + I \in A/I$ can decompose $a = \sum_{i=0}^\infty a_i$ where $a_i \in A_i$ for each i, so $(a_i + (I \cap A_i))_i$ maps to $\sum_i (a_i + I) = a + I$.
- Injective: suppose that $(a_i + (I \cap A_i))_i$ maps to 0 in A/I. Namely, we see that

  $$\sum_{i=0}^\infty a_i \in I,$$

  which we must prove implies $a_i \in I$ for each i. It suffices to show that

  $$I \stackrel{?}{=} \left\{ \sum_{i=0}^\infty a_i \in A : a_i \in (I \cap A_i) \text{ for all } i \right\}.$$

  Let the right-hand side be J. Certainly, $J \subseteq I$: note $a_i \in I$ for each i implies that $\sum_{i=0}^\infty a_i \in I$ because I is an ideal. To show $I \subseteq J$, we use the hypothesis that I is generated by homogeneous elements.[6] In other words, I is generated by the subsets $\{I \cap A_i\}_{i=0}^\infty$, which are all contained in J, so we will know that $I \subseteq J$ as soon as we check that J is actually a two-sided ideal.

  Well, the construction of J implies that J is certainly an F-subspace, so it remains to do the ideal checks. By symmetry, it will be enough to just check that J is a left ideal. Thus, we want to check that $AJ \subseteq J$. By linearity of this condition (and because J is already an F-subspace), it is enough to merely check this on a spanning subset of A, for which we take the homogeneous elements. Namely, it is enough to check that $A_i J \subseteq J$ for each i, so pick up some $b_j \in A_j$ and $\sum_i a_i \in J$, and we see

  $$b_j \cdot \sum_{i=0}^\infty a_i = \sum_{i=0}^\infty b_j a_i.$$

  Now, $b_j a_i \in A_{i+j}$ by the grading, and $b_j a_i \in I$ because I is an ideal, so we see $b_j \cdot \sum_{i=0}^\infty a_i \in J$ follows.
- Grading: we claim that $\frac{A_i}{I \cap A_i} \cdot \frac{A_j}{I \cap A_j} \subseteq \frac{A_{i+j}}{I \cap A_{i+j}}$. Well, for any $a_i + (I \cap A_i)$ and $a_j + (I \cap A_j)$, we see that the product has $(a_i a_j + I) \in A/I$, which comes from $a_i a_J + (I \cap A_{i+j}) = A_{i+j}/(I \cap A_{i+j})$. ■

Example 3.103. Let $\mathfrak{g}$ be an F-vector space. Then $S\mathfrak{g}$ is the quotient of $T\mathfrak{g}$ by the ideal generated by the homogeneous elements $X \otimes Y - Y \otimes X$, so Lemma 3.102 tells us that the grading on $T\mathfrak{g}$ descends to a grading on $S\mathfrak{g}$.

[6] This is the only place where we need this hypothesis, which explains why this check must be somewhat difficult.

However, the grading $T\mathfrak{g}$ need not descend to the quotient $U\mathfrak{g}$. For example, given $X, Y \in \mathfrak{g}$ which we expect to degree 1, we have

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

but the left-hand side is expected to live in degree 2 while the right-hand side is expected to live in degree 1. Of course, if $\mathfrak{g}$ is abelian, then everything here vanishes, so this is okay, but in general, there is no reason to believe that this would be okay.

Thus, we want a version of grading which descends along quotients, which leads to the following notion.

Definition 3.104 (filtered algebra)**.** Fix a field F. A *filtration* on an F-algebra A is a sequence of ascending F-subspaces

$$0 = \mathcal{F}_{-1}A \subseteq \mathcal{F}_0 A \subseteq \mathcal{F}_1 A \subseteq \cdots \subseteq \mathcal{F}_{n-1}A \subseteq \mathcal{F}_n A \subseteq \mathcal{F}_{n+1}A \subseteq \cdots$$

such that $A = \bigcup_{i \geq 0} \mathcal{F}_i A$, and $\mathcal{F}_i A \cdot \mathcal{F}_j A \to \mathcal{F}_{i+j}A$ for all $i, j \geq 0$.

Here's a sanity check.

Lemma 3.105. Fix a graded F-algebra $A = \bigoplus_{i=0}^{\infty} A_i$. Then

$$\mathcal{F}_j A := \bigoplus_{i=0}^{j} A_i$$

provides a filtration of A.

Proof. Here are our checks.

- Certainly each $\mathcal{F}_j A$ is an F-subspace.
- The construction implies $\mathcal{F}_j A \subseteq \mathcal{F}_j A \oplus A_{j+1} = \mathcal{F}_{j+1} A$ for each j.
- Note $\bigcup_{j \geq 9} \mathcal{F}_j A = \bigoplus_{i \geq 0} A_i = A$. Explicitly, any $a \in A$ can be decomposed into $a = \bigcup_{i \geq 0} a_i$ with $a_i \in A_i$ for each i; but then there is the largest n for which $a_n \neq 0$, and we see $a \in \bigoplus_{i=0}^{n} a_i = \mathcal{F}_n A$.
- We check that $\mathcal{F}_i A \cdot \mathcal{F}_j A \subseteq \mathcal{F}_{i+j} A$. Everything here is linear, so is enough to check this on spanning subsets of $\mathcal{F}_i A$ and $\mathcal{F}_j A$, for which we use the homogeneous elements. Well, choose some $a_k \in A_k \subseteq \mathcal{F}_i A$ and $a_\ell \in A_\ell \subseteq \mathcal{F}_j A$ where $k \leq i$ and $\ell \leq j$. Then the product $a_k a_\ell$ lives in $A_{k+\ell}$, which is contained in $\mathcal{F}_{i+j} A$ because $k + \ell \leq i + j$. ■

The up-shot of filtrations is that they do descend to quotients.

Lemma 3.106. Fix a filtered F-algebra A with filtration $\{\mathcal{F}_i A\}_{i=0}^{\infty}$. If $I \subseteq I$ is any two-sided ideal, then A/I is a filtered F-algebra with filtration

$$\left\{ \frac{\mathcal{F}_i A}{I \cap \mathcal{F}_i A} \right\}_{i=0}^{\infty}.$$

Proof. Here are our checks.

- Well-defined: note that the inclusions $\mathcal{F}_i A \subseteq A$ turn into inclusions $\frac{\mathcal{F}_i A}{I \cap \mathcal{F}_i A} \subseteq \frac{A}{I}$ because $I \cap \mathcal{F}_i A$ is the kernel of the composite $\mathcal{F}_i A \subseteq A \twoheadrightarrow A/I$.

- Covers: we claim that
$$\bigcup_{i=0}^{\infty} \frac{\mathcal{F}_i A}{I \cap \mathcal{F}_i A} = \frac{A}{I}.$$
Well, any $a + I$ has $a \in \mathcal{F}_i A$ for some i, so $a + I$ is in the image of $a + (I \cap \mathcal{F}_i A) \in \frac{\mathcal{F}_i A}{I \cap \mathcal{F}_i A}$.

- Filtration: for any i and j, we must check that $\frac{\mathcal{F}_i A}{I \cap \mathcal{F}_i A} \cdot \frac{\mathcal{F}_j A}{I \cap \mathcal{F}_j A} \to \frac{\mathcal{F}_{i+j} A}{I \cap \mathcal{F}_{i+j} A}$. Well, for any $a_i + (I \cap \mathcal{F}_i A)$ and $a_j + (I \cap \mathcal{F}_j A)$ where $a_i \in \mathcal{F}_i A$ and $a_j \in \mathcal{F}_j A$, we see that $a_i a_j \in \mathcal{F}_{i+j} A$, so the product
$$(a_i + (I \cap \mathcal{F}_i A)) \cdot (a_j + (I \cap \mathcal{F}_j A))$$
is $(a_i a_j + I) \in A/I$, which comes from $(a_i a_j + (I \cap \mathcal{F}_{i+j} A))$, as needed. ■

Example 3.107. The grading on $T\mathfrak{g}$ of Example 3.101 produces a filtration via Lemma 3.105, which then descends to the quotient $U\mathfrak{g}$ via Lemma 3.106. We will denote the induced filtration by $\{\mathcal{F}_i U\mathfrak{g}\}_{i=0}^{\infty}$. Notably, looking at the induced quotient filtration of Lemma 3.106, we see that $\mathcal{F}_i U\mathfrak{g}$ is simply the image of $\mathcal{F}_i T\mathfrak{g}$ in $U\mathfrak{g}$. (This formally follows from the well-definedness check in Lemma 3.106.) Importantly, each graded component of $T\mathfrak{g}$ is finite-dimensional, so $\mathcal{F}_i T\mathfrak{g}$ is finite-dimensional, so its image $\mathcal{F}_i U\mathfrak{g}$ is also finite-dimensional.

Here is a concrete consequence of the intuition that the filtration "tempers" $U\mathfrak{g}$.

Lemma 3.108. Fix a Lie algebra $\mathfrak{g}$ over a field F. For $A\mathfrak{g} \in \{T\mathfrak{g}, U\mathfrak{g}, S\mathfrak{g}\}$, the adjoint action descends to a Lie algebra morphism $\operatorname{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(\mathcal{F}_i A\mathfrak{g})$ for each i.

Proof. The main point is to check that $\operatorname{ad}_X \colon A\mathfrak{g} \to A\mathfrak{g}$ restricts to $\operatorname{ad}_X \colon \mathcal{F}_i A\mathfrak{g} \to \mathcal{F}_i A\mathfrak{g}$ for each $i \geq 0$ and $X \in \mathfrak{g}$. Well, it suffices to check this on a spanning subset of $\mathcal{F}_i A\mathfrak{g}$, for which we use pure tensors of length less than or equal to i. Namely, for any $Y_1 \cdots Y_j \in A\mathfrak{g}$ for $Y_1, \ldots, Y_j \in \mathfrak{g}$, we see that
$$\operatorname{ad}_X(Y_1 \cdots Y_j) = \sum_{k=1}^{j} Y_1 \cdots Y_{k-1}[X, Y_k]Y_{k+1} \cdots Y_j$$
by definition of $\operatorname{ad}_\bullet$ (or alternatively by repeated application of the product rule). Now, we see that the summand $Y_1 \cdots Y_{k-1}[X, Y_k]Y_{k+1} \cdots Y_j$ lives in the image of $\mathfrak{g}^{\otimes j} \subseteq \mathcal{F}_i T\mathfrak{g} \twoheadrightarrow \mathcal{F}_i A\mathfrak{g}$, so we conclude that ad_X does in fact restrict.

We now describe the other checks. Note ad_X was linear on $A\mathfrak{g}$, so it continues to be linear on $\mathcal{F}_i A\mathfrak{g}$. Additionally, the equalities
$$\operatorname{ad}_{c_1 X_1 + c_2 X_2} = c_1 \operatorname{ad}_{X_1} + c_2 \operatorname{ad}_{X_2}$$
and
$$\operatorname{ad}_{[X,Y]} = \operatorname{ad}_X \circ \operatorname{ad}_Y - \operatorname{ad}_Y \circ \operatorname{ad}_X$$
held as equalities of linear maps $A\mathfrak{g} \to A\mathfrak{g}$, so they will continue to be true on their restrictions to $\mathcal{F}_i A\mathfrak{g}$. This completes the proof that $\operatorname{ad}_\bullet \colon \mathfrak{g} \to \mathfrak{gl}(\mathcal{F}_i A\mathfrak{g})$ is a representation of Lie algebras. ■

The statement of Lemma 3.105 tells us that we may hope to recover a grading from a filtration by using the quotients $\mathcal{F}_{j+1} A / \mathcal{F}_j A$. This leads to the following definition.

Definition 3.109 (associated graded algebra)**.** Fix a filtered F-algebra A with filtration $\{\mathcal{F}_i A\}_{i=0}^{\infty}$. Then the *associated graded algebra* is
$$\operatorname{gr} A := \bigoplus_{i=0}^{\infty} \frac{\mathcal{F}_i A}{\mathcal{F}_{i-1} A}.$$
Here, $\mathcal{F}_{-1} A$ is understood to be 0.

Here are the checks on this definition.

Lemma 3.110. Fix a filtered F-algebra A with filtration $\{\mathcal{F}_iA\}_{i=0}^\infty$. Then $\operatorname{gr} A$ is a graded F-algebra with grading given by its definition.

Proof. Here are our checks.

- Note that $\operatorname{gr} A$ is certainly a vector space over F as the sum of vector spaces over F.
- We must define a multiplication for $\operatorname{gr} A$. We are trying to define an F-bilinear map $\operatorname{gr} A \times \operatorname{gr} A \to \operatorname{gr} A$, so we are trying to define a linear map $\operatorname{gr} A \otimes_F \operatorname{gr} A \to \operatorname{gr} A$. Swapping the tensor product and sum, it is enough to define a map
$$\frac{\mathcal{F}_iA}{\mathcal{F}_{i-1}A} \otimes_F \frac{\mathcal{F}_jA}{\mathcal{F}_{j-1}A} \to \frac{\mathcal{F}_{i+j}A}{\mathcal{F}_{i+j-1}A} \subseteq \operatorname{gr} A$$
for $i, j \geq 0$. For this, we note the multiplication on A defines an F-bilinear map $\mathcal{F}_iA \times \mathcal{F}_jA \to \mathcal{F}_{i+j}A$, but $\mathcal{F}_{i-1}A\mathcal{F}_jA \subseteq \mathcal{F}_{i+j-1}A$ and $\mathcal{F}_iA\mathcal{F}_{j-1}A \subseteq \mathcal{F}_{i+j-1}A$ means that our map descends to an F-bilinear map
$$\frac{\mathcal{F}_iA}{\mathcal{F}_{i-1}A} \times \frac{\mathcal{F}_jA}{\mathcal{F}_{j-1}A} \to \frac{\mathcal{F}_{i+j}A}{\mathcal{F}_{i+j-1}A},$$
so this descends to a map on the tensor product, as needed. Note that this definition of multiplication automatically respects the "grading" in the definition of $\operatorname{gr} A$.
- Thus, it only remains that we actually have a well-defined multiplication. Well, the multiplication is F-bilinear by its construction, so it satisfies the usual distributivity and F-linear requirements. It remains to check associativity. Well, the associativity condition $(ab)c = a(bc)$ corresponds to the vanishing of some multilinear functional on $(\operatorname{gr} A)^3$, which corresponds to the vanishing of some functional on $(\operatorname{gr} A)^{\otimes 3}$, which can be checked on the spanning subset of homogeneous elements. As such, we pick up three homogeneous elements $a_i + \mathcal{F}_{i-1}A$ and $a_j + \mathcal{F}_{j-1}A$ and $a_k + \mathcal{F}_{k-1}A$ and use the above definition to compute
$$\begin{aligned}
&((a_i + \mathcal{F}_{i-1}A)(a_j + \mathcal{F}_{j-1}A))(a_k + \mathcal{F}_{k-1}A) \\
&= (a_ia_j + \mathcal{F}_{i+j-1}A)(a_k + \mathcal{F}_{k-1}A) \\
&= (a_ia_ja_k + \mathcal{F}_{i+j+k-1}A) \\
&= (a_i + \mathcal{F}_{i-1}A)(a_ja_k + \mathcal{F}_{j+k-1}A) \\
&= (a_i + \mathcal{F}_{i-1}A)((a_j + \mathcal{F}_{j-1}A)(a_k + \mathcal{F}_{k-1}A)),
\end{aligned}$$
as required. ■

Before ending our discussion, we note the following property of $\operatorname{gr} A$.

Lemma 3.111. Fix a filtered F-algebra A with filtration $\{\mathcal{F}_iA\}_{i=0}^\infty$. If $\operatorname{gr} A$ has no zero divisors, then A has no zero divisors.

Proof. We proceed by contraposition. Suppose A has zero divisors so that we have nonzero elements $a, b \in A$ such that $ab = 0$. Let i be the smallest nonnegative element $a \in \mathcal{F}_iA$ but $a \notin \mathcal{F}_{i-1}A$. (If $i = 0$, then certainly $a \notin \mathcal{F}_{i-1}A$ because $\mathcal{F}_{i-1}A = 0$, and $a \neq 0$.) Similarly, we find $j \geq 0$ such that $b \in \mathcal{F}_jA$ but $b \notin \mathcal{F}_{j-1}A$. Then $a + \mathcal{F}_{i-1}A$ and $b + \mathcal{F}_{j-1}A$ are nonzero elements in $\operatorname{gr} A$ which multiply to $0 + \mathcal{F}_{i+j-1}A$, so $\operatorname{gr} A$ has zero divisors. ■

3.5.2 The Poincaré–Birkoff–Witt Theorem

To understand the theorem of this subsection, we make the following observation.

Lemma 3.112. Fix a Lie algebra $\mathfrak{g}$ over a field F with basis $\{X_1, \ldots, X_n\}$. Then the ordered monomials

$$\left\{X_1^{d_1} \cdots X_n^{d_n} : a_1, \ldots, a_n \geq 0\right\}$$

span $U\mathfrak{g}$.

Proof. The point is to use the equality $XY - YX = [X, Y]$ in $U\mathfrak{g}$ to slowly reorder unordered monomials, one transposition at a time. Let's be more explicit. The monomials $X_{i_1} \otimes \cdots \otimes X_{i_m}$ spans $\mathfrak{g}^{\otimes m}$ (here, $i_1, \ldots, i_m \in \{1, \ldots, n\}$), so they will span $T\mathfrak{g}$ as m varies. Thus, it is enough to show that any such monomial $X_{i_1} \cdots X_{i_m} \in U\mathfrak{g}$ lives in the span of the above ordered monomials. We will do this via a nested induction. To begin, we induct on m, for which the base cases $m = 0$ and $m = 1$ have no content because the monomial is already ordered. Thus, for the inductive step, we may assume the result for any monomial with less than m terms.

Next up, we note that $i_\bullet \colon \{1, \ldots, m\} \to \{1, \ldots, n\}$ is some function, so we choose some permutation $\sigma \in S_m$ so that

$$i_{\sigma(1)} \leq i_{\sigma(2)} \leq \cdots \leq i_{\sigma(m)}.$$

Now, $\sigma \in S_m$ can be written as a product of transpositions of the form $(j, j+1)$, say $\sigma = (j_1, j_1+1) \cdots (j_\ell, j_\ell+1)$. We next induct on ℓ, where the base case of $\ell = 0$ has no content because it means that σ is the identity so that the $i_\bullet$s are already ordered so that our monomial is already ordered. For the inductive step, we pick up our monomial $X_{i_1} \cdots X_{i_m}$ with $\sigma = (j_1, j_1 + 1) \cdots (j_\ell, j_\ell + 1)$ where $\ell > 0$, and we note that

$$\begin{aligned}
X_{i_1} \cdots X_{i_m} &= X_{i_1} \cdots X_{i_{j_\ell}} X_{i_{j_\ell+1}} \cdots X_{i_m} \\
&\stackrel{*}{=} X_{i_1} \cdots \left(X_{i_{j_\ell+1}} X_{i_{j_\ell}} + [X_{i_{j_\ell}}, X_{i_{j_\ell+1}}]\right) \cdots X_{i_m} \\
&= X_{i_1} \cdots X_{i_{j_\ell-1}} X_{i_{j_\ell+1}} X_{i_{j_\ell}} X_{i_{j_\ell+2}} \cdots X_{i_m} + X_{i_1} \cdots X_{i_{j_\ell-1}} [X_{i_{j_\ell}}, X_{i_{j_\ell+1}}] X_{i_{j_\ell+2}} \cdots X_{i_m}.
\end{aligned}$$

Here, the key step is the application of $XY - YX = [X, Y]$ in $U\mathfrak{g}$ in the equality $\stackrel{*}{=}$. Anyway, the second term has fewer than m terms, so the first inductive hypothesis implies that it is the span of our ordered monomials. The first term still has m terms, but the permutation $\sigma' \in S_m$ required to reorder the $i_\bullet$s can simply be taken to be $(j_1, j_1 + 1) \cdots (j_{\ell-1}, j_{\ell-1} + 1)$, which has length smaller than ℓ, so the second inductive hypothesis implies that it too is in the span of our ordered monomials. ■

Remark 3.113. The above nested induction can be turned into an algorithm: simply use the relation $XY - YX = [X, Y]$ to gradually reorder the consecutive terms of any monomial until it becomes ordered. The complications in this algorithm is that the "error term" $[X, Y]$ spawn lower-order monomials which must be dealt with recursively.

Remark 3.114. Let's note a consequence of the proof of Lemma 3.112: in $U\mathfrak{g}$, the difference

$$X_{i_1} \cdots X_{i_m} - X_{\sigma(i_1)} \cdots X_{\sigma(i_m)}$$

is a linear combination of ordered monomials of degree than smaller m. Indeed, the exact same inductive argument as in the proof of Lemma 3.112 shows this, where the point is that the produced error term

$$X_{i_1} \cdots X_{i_{j_\ell-1}} [X_{i_{j_\ell}}, X_{i_{j_\ell+1}}] X_{i_{j_\ell+2}} \cdots X_{i_m}$$

has fewer than m monomials after expanding $[X_{i_{j_\ell}}, X_{i_{j_\ell+1}}]$.

With the lemma in mind, here is our statement of the theorem.

Theorem 3.115 (Poincaré–Birkoff–Witt)**.** Fix a Lie algebra $\mathfrak{g}$ over a field F with basis $\{X_1, \ldots, X_n\}$. Then the ordered monomials

$$\left\{X_1^{d_1} \cdots X_n^{d_n} : a_1, \ldots, a_n \geq 0\right\}$$

are linearly independent in $U\mathfrak{g}$.

Proof. We will define a linear map $\varphi\colon T\mathfrak{g} \to S\mathfrak{g}$ satisfying the following properties.

(a) We have $\varphi\left(X_1^{d_1} \cdots X_n^{d_n}\right) = X_1^{d_1} \cdots X_n^{d_n}$ for any ordered monomial $X_1^{d_1} \cdots X_n^{d_n}$.

(b) We have $L\mathfrak{g} \subseteq \ker \varphi$.

Let's see why this will complete the proof. Note (b) tells us that φ descends to a linear map $U\mathfrak{g} \to S\mathfrak{g}$. But then because the ordered monomials are linearly independent in $S\mathfrak{g}$, we see that the ordered monomials of $U\mathfrak{g}$ (which go the ordered monomials in $S\mathfrak{g}$ by (a)) must also be linearly independent.

We will define φ inductively by taking a union of linear maps $\varphi_k : \mathcal{F}_k T\mathfrak{g} \to S\mathfrak{g}$ satisfying the following.

(a') We have $\varphi_k\left(X_1^{d_1} \cdots X_n^{d_n}\right) = X_1^{d_1} \cdots X_n^{d_n}$ for any ordered monomial $X_1^{d_1} \cdots X_n^{d_n}$ of total degree at most k.

(b') We have

$$\sum_{a+b+2 \leq k} \left\{A(XY - YX - [X,Y])B : X, Y \in \mathfrak{g}, A \in \mathcal{F}_a T\mathfrak{g}, B \in \mathcal{F}_b T\mathfrak{g}\right\} \subseteq \ker \varphi_k$$

for each k. (We are not going to directly argue that the above equals $L\mathfrak{g} \cap \mathcal{F}_k T\mathfrak{g}$.)

(c') $\varphi_{k+1}|_{\mathcal{F}_k T\mathfrak{g}} = \varphi_k$.

Let's quickly explain how these conditions imply the result. Here are these checks.

- Quickly, note that (c') tells us that our $\varphi_\bullet$s at least assemble into a function $\varphi\colon T\mathfrak{g} \to S\mathfrak{g}$.
- Linear: we use the linearity of the $\varphi_\bullet$s. Indeed, for any $c_1, c_2 \in F$ and $a_1, a_2 \in T\mathfrak{g}$, find some k large enough so that $a_1, a_2 \in \mathcal{F}_k T\mathfrak{g}$, and then the linearity check

$$\varphi(c_1 a_1 + c_2 a_2) \stackrel{?}{=} c_1\varphi(a_1) + c_2\varphi(a_2)$$

immediately reduces to the corresponding equality for φ_k.
- Note that φ satisfies (a) because the $\varphi_\bullet$s satisfy (a').
- For (b), we need to show that $\ker \varphi$ contains $L\mathfrak{g}$. Well, we know that $\ker\varphi$ is a subspace of $T\mathfrak{g}$, so it suffices to check that an F-spanning subset of $L\mathfrak{g}$ is contained in $\ker\varphi$. As such, it is enough to check that

$$A(XY - YX - [X,Y]) \in \ker \varphi$$

for any $A \in \mathcal{F}_k T\mathfrak{g}$ and $X, Y \in \mathfrak{g}$. (Technically we must also check this for elements of the form $(XY - YX - [X,Y])A$, which follows by a symmetric argument.) However, the above element is in the kernel of φ_{k+2} by (b'), so it will be in the kernel of φ.

We will construct the $\varphi_\bullet$s inductively. For $k = -1$, there is nothing to do because $\mathcal{F}_{-1} T\mathfrak{g} = 0$, and for $k = 0$, there is still nothing to do because $\mathcal{F}_0 T\mathfrak{g} = F$, which must land identically to $F \subseteq S\mathfrak{g}$.

We now proceed with the induction to define the $\varphi_\bullet$s. Suppose that we have defined φ_k already, and we would like to define φ_{k+1}. Because monomials $X := X_{i_1} \cdots X_{i_{k+1}}$ of total degree $k+1$ form a basis of $\mathcal{F}_{k+1} T\mathfrak{g} = \mathfrak{g}^{\otimes(k+1)}$, we merely have to define φ_{k+1} on such unordered monomials. Well, start with an ordering of the indices $i_1 \leq \cdots \leq i_{k+1}$, and we want to define $\varphi(\sigma(X_{i_1} \cdots X_{i_{k+1}}))$ for each $\sigma \in S_{k+1}$, where S_{k+1} acts on $\mathcal{F}_{k+1} T\mathfrak{g}$ by permuting the basis.

For motivation, note that it is clearest what to do on transpositions σ: for $Y_1 \cdots Y_{k+1} \in \mathfrak{g}$ and a transposition $(j, j+1) \in S_{k+1}$, we see that we need

$$\begin{aligned}\varphi(Y_1 \cdots Y_{k+1}) &= Y_1 \cdots Y_{j-1}(Y_{j+1}Y_j + [Y_j, Y_{j+1}])Y_{j+2} \cdots Y_{k+1} \\ &= Y_1 \cdots Y_{j-1}Y_{j+1}Y_jY_{j+2} \cdots Y_{k+1} \\ &\qquad + Y_1 \cdots Y_{j-1}[Y_j, Y_{j+1}]Y_{j+2} \cdots Y_{k+1}\end{aligned}$$

in order for $L\mathfrak{g} \subseteq \ker \varphi$.

We now iterate the computation of the previous paragraph to define φ on general $\sigma(X_{i_1} \cdots X_{i_{k+1}})$. Because the transpositions $s_j := (j, j+1) \in S_{k+1}$ generate S_{k+1}, we may decompose

$$\sigma = s_{j_\ell} \cdots s_{j_1}.$$

We now define

$$\varphi_{k+1}\left(\sigma(X_{i_1} \cdots X_{i_{k+1}})\right) := X_{\sigma(i_1)} \cdots X_{\sigma(i_{k+1})} - \sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big),$$

where the notation $[,]_j$ applied to a monomial is given by

$$[,]_j(Y_1 \cdots Y_{k+1}) := Y_1 \cdots Y_{j-1}[Y_j, Y_{j+1}]Y_{j+2} \cdots Y_{k+1},$$

which we note lives in $\mathcal{F}_kT\mathfrak{g}$. There are many checks we must do to see that this definition works.

- We check that our definition does not depend on the choice of decomposition $\sigma = (j_1, j_1+1) \cdots (j_\ell, j_\ell+1)$. This is somewhat technical. To this end, we note that viewing S_n as a Coxeter group with the reflection generators $s_j := (j, j+1)$, we see that S_{k+1} is generated by $\{s_1, \ldots, s_k\}$ with the generators

$$\begin{cases} s_j^2 & \text{if } j \in \{1, \ldots, k\}, \\ s_js_{j'} = s_{j'}s_j & \text{if } |j - j'| \geq 2, \\ s_js_{j+1}s_j = s_{j+1}s_js_{j+1} & \text{if } j \in \{1, \ldots, k-1\}. \end{cases}$$

  (The first relation holds because it squares a transposition. The second relation holds because s_j and $s_{j'}$ are disjoint permutations. Computing $s_js_{j+1}s_j = (j, j+2) = s_{j+1}s_js_{j+1}$ gives the third relation. We will not check that these relations generate all relations.) Thus, any two decompositions may apply the above three rules to be shown equal. We will use this for our well-defined check as follows.

  - First relation: after the dust settles, this follows by the skew-symmetry of the Lie bracket. Suppose our decompositions differ as $\sigma = \sigma_1\sigma_2 = \sigma_1 s_j^2 \sigma_2$, where $\sigma_1 = s_{j_\ell} \cdots s_{j_{r+1}}$ and $\sigma_2 = s_{j_r} \cdots s_{j_1}$; label the second decomposition as $s_{j'_{\ell+2}} \cdots s_{j'_1}$. We now expand

$$\begin{aligned}&\sum_{m=0}^{\ell+1} \varphi_k\big([,]_{j'_{m+1}}(s_{j'_m} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) - \sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ = &\sum_{m=0}^{r-1} \varphi_k\big([,]_{j'_{m+1}}(s_{j'_m} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) - \sum_{m=0}^{r-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &+ \varphi_k\big([,]_{j'_{r+1}}(s_{j'_r} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) - \varphi_k\big([,]_{j_{r+1}}(s_{j_r} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &+ \varphi_k\big([,]_{j'_{r+1+1}}(s_{j'_{r+1}} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) + \varphi_k\big([,]_{j'_{r+2+1}}(s_{j'_{r+2}} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &+ \sum_{m=r+3}^{\ell+1} \varphi_k\big([,]_{j'_{m+1}}(s_{j'_m} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) - \sum_{m=r+1}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big).\end{aligned}$$

    Now, we see that the sums in each row are identically equal: in the top row, the summands are both contained in σ_2, and in the bottom row, the summands are both contained in $\sigma_1\sigma_2$. Thus, it remains to show that the middle rows cancel out. Note $s_{j'_{r+1}} \cdots s_{j'_1} = s_j \sigma_2$, and $s_{j'_{r+2}} \cdots s_{j'_1} = s_j \sigma_2 = s_j s_j \sigma_2$, so we see are looking at

$$\begin{aligned}
&\varphi_k\big([,]_{j'_{r+1}}(s_{j'_r} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) - \varphi_k\big([,]_{j_{r+1}}(s_{j_r} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\
&\quad + \varphi_k\big([,]_{j'_{r+1+1}}(s_{j'_{r+1}} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) + \varphi_k\big([,]_{j'_{r+2+1}}(s_{j'_{r+2}} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\
&= \varphi_k\big([,]_j \sigma_2 X_{i_1} \cdots X_{i_{k+1}}\big) - \varphi_k\big([,]_{j_{r+1}} \sigma_2 X_{i_1} \cdots X_{i_{k+1}}\big) \\
&\quad + \varphi_k\big([,]_j s_j \sigma_2 X_{i_1} \cdots X_{i_{k+1}}\big) + \varphi_k\big([,]_{j_{r+1}} \sigma_2 X_{i_1} \cdots X_{i_{k+1}}\big).
\end{aligned}$$

The two right terms now vanish, so we want the two left terms to cancel. Well, we see that

$$[,]_j \sigma_2 X_{i_1} \cdots X_{i_{k+1}} = X_{\sigma_2(i_1)} \cdots X_{\sigma_2(i_{j-1})}[X_{\sigma_2(i_j)}, X_{\sigma_2(i_{j+1})}]X_{\sigma_2(i_{j+2}} \cdots X_{\sigma_2(i_{k+1})},$$

but

$$[,]_j s_j \sigma_2 X_{i_1} \cdots X_{i_{k+1}} = X_{\sigma_2(i_1)} \cdots X_{\sigma_2(i_{j-1})}[X_{\sigma_2(i_{j+1})}, X_{\sigma_2(i_j)}]X_{\sigma_2(i_{j+2}} \cdots X_{\sigma_2(i_{k+1})},$$

so the two terms cancel by skew-symmetry of the Lie bracket!

– Second relation: suppose our decompositions differ as $\sigma = \sigma_1 s_{j_{r+1}} s_{j_{r+2}} \sigma_2 = \sigma_1 s_{j_{r+2}} s_{r+1} \sigma_2$ where $\sigma_1 = s_{j_\ell} \cdots s_{j_{r+3}}$ and $\sigma_2 = s_{j_r} \cdots s_{j_1}$; label the second decomposition as $s_{j'_\ell} \cdots s_{j'_1}$. We now note that the sums

$$\sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j'_{m+1}}(s_{j'_m} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) \qquad \text{and} \qquad \sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big)$$

differ only at terms $m \in \{r, r+1\}$: otherwise, the permutations applied to $X_{i_1} \cdots X_{i_{k+1}}$ are the same, we are taking the bracket at the same location. To deal with the terms $m \in \{r, r+1\}$ without too much pain, assume $(j, j') := (j_{r+1}, j_{r+2})$ has $j < j'$ (without loss of generality) and write $\sigma_2(X_{i_1} \cdots X_{i_{k+1}}) = AY_jY_{j+1}BY_{j'}Y_{j'+1}C$, where A, B, C are monomials, and the $Y_\bullet$s are at the corresponding location. Examining the terms for $m \in \{r, r+1\}$, we would like

$$\begin{aligned}
&\varphi_k(AY_jY_{j+1}B[Y_{j'}, Y_{j'+1}]C) + \varphi_k(A[Y_j, Y_{j+1}]BY_{j'+1}Y_{j'}C) \\
&\stackrel{?}{=} \varphi_k(A[Y_j, Y_{j+1}]BY_{j'}Y_{j'+1}C) + \varphi_k(AY_{j+1}Y_jB[Y_{j'}, Y_{j'+1}]C).
\end{aligned}$$

Well, by the inductive hypothesis on φ_k, we see

$$\begin{aligned}
&\varphi_k(AY_jY_{j+1}B[Y_{j'}, Y_{j'+1}]C) - \varphi_k(AY_{j+1}Y_jB[Y_{j'}, Y_{j'+1}]C) \\
&= \varphi_k(A(Y_jY_{j+1} - Y_{j+1}Y_j)B[Y_{j'}, Y_{j'+1}]C) \\
&= \varphi_k(A[Y_j, Y_{j+1}]B[Y_{j'}, Y_{j'+1}]C) \\
&= \varphi_k(A[Y_j, Y_{j+1}]B(Y_{j'}Y_{j'+1} - Y_{j'+1} - Y_{j'})C) \\
&= \varphi_k(A[Y_j, Y_{j+1}]BY_{j'}Y_{j'+1}C) - \varphi_k(A[Y_j, Y_{j+1}]BY_{j'+1}Y_{j'}C),
\end{aligned}$$

so we are done.

– Third relation: suppose our decompositions differ as $\sigma = \sigma_1 s_j s_{j+1} s_j \sigma_2 = \sigma_1 s_{j+1} s_j s_{j+1} \sigma_2$, where $\sigma_2 = s_{j_r} \cdots s_{j_1}$ and $\sigma_1 = \sigma_{j_\ell} \cdots s_{j_{r+4}}$. Labeling the two decompositions as $s_{j_\ell} \cdots s_{j_1}$ and $s_{j'_\ell} \cdots s_{j'_1}$, respectively, we as above notice that the two sums

$$\sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j'_{m+1}}(s_{j'_m} \cdots s_{j'_1})X_{i_1} \cdots X_{i_{k+1}}\big) \qquad \text{and} \qquad \sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big)$$

only differ at terms $m \in \{r, r+1, r+2\}$, for otherwise we are applying to the same location of Lie bracket to the same permuted monomial. As such, we write $\sigma_2(X_{i_1} \cdots X_{i_{k+1}}) = AY_jY_{j+1}Y_{j+2}B$

for some monomials A and B, where the $Y_\bullet$s are at the corresponding location. Comparing terms $m \in \{r, r+1, r+2\}$, we would like for

$$\begin{aligned}&\varphi_k(A[Y_j, Y_{j+1}]Y_{j+2}B) + \varphi_k(AY_{j+1}[Y_j, Y_{j+2}]B) + \varphi_k(A[Y_{j+1}, Y_{j+2}]Y_jB)\\ &\stackrel{?}{=} \varphi_k(AY_j[Y_{j+1}, Y_{j+2}]B) + \varphi_k(A[Y_j, Y_{j+2}]Y_{j+1}B) + \varphi_k(AY_{j+2}[Y_j, Y_{j+1}]B).\end{aligned}$$

Well, modulo $\ker \varphi_k$, we see that

$$\begin{aligned}&A[Y_j, Y_{j+1}]Y_{j+2}B + AY_{j+1}[Y_j, Y_{j+2}]B + A[Y_{j+1}, Y_{j+2}]Y_jB\\ &= A\big(Y_jY_{j+1}Y_{j+2} - Y_{j+1}Y_jY_{j+2} + Y_{j+1}Y_jY_{j+2} - Y_{j+1}Y_{j+2}Y_j + Y_{j+1}Y_{j+2}Y_j - Y_{j+2}Y_{j+1}Y_j\big)B\\ &= A\big(Y_jY_{j+1}Y_{j+2} - Y_{j+2}Y_{j+1}Y_j\big)B,\end{aligned}$$

and

$$\begin{aligned}&AY_j[Y_{j+1}, Y_{j+2}]B + A[Y_j, Y_{j+2}]Y_{j+1}B + AY_{j+2}[Y_j, Y_{j+1}]B\\ &= A\big(Y_jY_{j+1}Y_{j+2} - Y_jY_{j+2}Y_{j+1} + Y_jY_{j+2}Y_{j+1} - Y_{j+2}Y_jY_{j+1} + Y_{j+2}Y_jY_{j+1} - Y_{j+2}Y_{j+1}Y_j\big)B\\ &= A\big(Y_jY_{j+1}Y_{j+2} - Y_{j+2}Y_{j+1}Y_j\big)B,\end{aligned}$$

so we will get the same result out after applying φ_k.

Because the three relations generate all equalities in S_{k+1}, we see that our φ does not depend on the decomposition of σ.

- We check that the definition of $\varphi_{k+1}\left(\sigma(X_{i_1} \cdots X_{i_{k+1}})\right)$ does not depend on the choice of σ. Indeed, if $\sigma(X_{i_1} \cdots X_{i_{k+1}}) = \sigma'(X_{i_1} \cdots X_{i_{k+1}})$ for two $\sigma, \sigma' \in S_{k+1}$, then $\sigma^{-1}\sigma'$ is a permutation which can only swap $X_{i_{j_1}}$ and $X_{i_{j_2}}$ when $i_{j_1} = i_{j_2}$. The collection of legal permutations is thus some product of symmetric groups (there is a permutation group for each X_i because we are allowed to swap two copies of X_i in different places). However, all the copies of X_i in $X_{i_1} \cdots X_{i_{k+1}}$ are located right next to each other because this is an ordered monomial, so the legal permutations live in a product of symmetric groups of intervals of integers. Each of these individual symmetric groups of intervals of integers are generated by permutations s_j, so we can write

$$\sigma' = \sigma s_{j_\ell} \cdots s_{j_1},$$

where $s_{j_1}, \ldots, s_{j_\ell}$ have $X_{i_{j_\bullet}} = X_{i_{j_\bullet+1}}$ for each $j_\bullet$.

Thus, to check that our definition does not depend on the choice of σ, it is enough (by induction) to check that we can replace σ with some σs_{j_1} when $X_{i_{j_1}} = X_{i_{j_1+1}}$. As such, we decompose $\sigma = s_{j_\ell} \cdots s_{j_2}$, and we want to show

$$\sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \stackrel{?}{=} \sum_{m=1}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_2})X_{i_1} \cdots X_{i_{k+1}}\big).$$

These sums are term-wise equal for $m \geq 1$ because it is the same Lie bracket applied to the same location of the same monomial (notably, s_{j_1} fixes $X_{i_1} \cdots X_{i_{k+1}}$ by its construction). Thus, we want to show that

$$\varphi_k\big([,]_{j_1}X_{i_1} \cdots X_{i_{k+1}}\big) \stackrel{?}{=} 0.$$

Well,

$$[,]_{j_1}X_{i_1} \cdots X_{i_{k+1}} = X_{i_1} \cdots X_{i_{j-1}}[X_{i_j}, X_{i_{j+1}}]X_{i_{j+1}} \cdots X_{i_{k+1}},$$

which vanishes because $X_{i_j} = X_{i_{j+1}}$.

- We now know that our definition of φ_k is well-defined. It remains to check that our linear map satisfies (c′) for φ_{k+1}. Thus, fix some pair (a, b) with $a + b + 2 \leq k$; we may assume equality for otherwise this is true by the inductive hypothesis. The condition is linear in A and B, so we may check it on a spanning subset of $\mathcal{F}_aT\mathfrak{g}$ and $\mathcal{F}_bT\mathfrak{g}$, for which we take the monomials; write $A = X_{i'_1} \cdots X_{i'_a}$ and

$B = X_{i'_{a+3}} \cdots X_{i'_{k+1}}$. Similarly, the condition is linear in X and Y, so we may assume that they are basis vectors $X_{i'_{a+1}}$ and $X_{i'_{a+2}}$. Rearranging, we would like to show that

$$\varphi_k(s_a(X_{i'_1} \cdots X_{i'_{k+1}})) \stackrel{?}{=} \varphi_k(X_{i'_1} \cdots X_{i'_{k+1}}) - \varphi_k([,]_a X_{i'_1} \cdots X_{i'_{k+1}}).$$

Well, let $i_1 \leq \cdots \leq i_{k+1}$ be an ordering, and let $\sigma \in S_{k+1}$ be the permutation so that $\sigma(i_j) = i'_{\sigma(j)}$. Thus, to compute φ_k, we expand $\sigma = s_{j_\ell} \cdots s_{j_1}$ and set $j_{\ell+1} \coloneqq a$ so that

$$\begin{aligned} \varphi_k(s_a\sigma(X_{i_1} \cdots X_{i_{k+1}})) &= X_{i_1} \cdots X_{i_{k+1}} - \sum_{m=0}^{\ell} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &= X_{i_1} \cdots X_{i_{k+1}} - \sum_{m=0}^{\ell-1} \varphi_k\big([,]_{j_{m+1}}(s_{j_m} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &\quad - \varphi_k\big([,]_{j_{\ell+1}}(s_{j_\ell} \cdots s_{j_1})X_{i_1} \cdots X_{i_{k+1}}\big) \\ &= \varphi_k(\sigma(X_{i_1} \cdots X_{i_{k+1}})) - \varphi_k\big([,]_a\sigma(X_{i_1} \cdots X_{i_{k+1}})\big), \end{aligned}$$

as required. ■

3.5.3 Consequences of the Poincaré–Birkoff–Witt Theorem

Theorem 3.115 has important consequences for the structure of $U\mathfrak{g}$. Let's see some.

Corollary 3.116. Fix a Lie algebra $\mathfrak{g}$ over a field F. Then the inclusion $\mathfrak{g} \to U\mathfrak{g}$ is actually injective.

Proof. Let $\{X_1, \ldots, X_n\}$ be a basis of $\mathfrak{g}$. Then Theorem 3.115 implies that the monomials $\{X_1, \ldots, X_n\} \subseteq U\mathfrak{g}$ are linearly independent. In particular, we see that any nonzero $X \in \mathfrak{g}$ can be expanded as $\sum_{i=1}^n a_iX_i$, where $a_i \neq 0$ for some i, meaning that

$$\sum_{i=1}^n a_iX_i \in U\mathfrak{g}$$

is also nonzero by the aforementioned linear independence. Thus, the map $\mathfrak{g} \to U\mathfrak{g}$ has trivial kernel. ■

Corollary 3.117. Fix a Lie algebra $\mathfrak{g}$ over a field F with Lie subalgebras $\mathfrak{g}_1, \ldots, \mathfrak{g}_n \subseteq \mathfrak{g}$ such that we have a vector space decomposition $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$. Then the "multiplication" map $\mu \colon U\mathfrak{g}_1 \otimes_F \cdots \otimes_F U\mathfrak{g}_n \to U\mathfrak{g}$, given by

$$\mu(a_1 \otimes \cdots \otimes a_n) \coloneqq a_1 \cdots a_n$$

on pure tensors, is an isomorphism of vector spaces.

Proof. The point is that μ should send a basis of ordered monomials to a basis of ordered monomials. Quickly, observe that functoriality produces maps $U\mathfrak{g}_\bullet \to U\mathfrak{g}$ from the inclusions $\mathfrak{g}_\bullet \subseteq \mathfrak{g}$. As such, we may note that the map $U\mathfrak{g}_1 \times \cdots \times U\mathfrak{g}_n \to U\mathfrak{g}$

$$(a_1, \ldots, a_n) \mapsto a_1 \cdots a_n$$

is F-multilinear because $U\mathfrak{g}$ is an F-algebra, so we have indeed induced a unique morphism $U\mathfrak{g}_1 \otimes \cdots \otimes U\mathfrak{g}_n \to U\mathfrak{g}$ of vector spaces over F.

It remains to check that this map is an isomorphism. Well, it is enough to check that it sends a basis. For this, we give each $\mathfrak{g}_i$ a basis $\{X_{i1}, \ldots, X_{in_i}\}$ so that the concatenation of these bases provides a basis of $\mathfrak{g}$. Then the combination of Lemma 3.112 and Theorem 3.115 tells us that the ordered monomials

$$X_{i1}^{d_{i1}} \cdots X_{in_i}^{d_{in_i}}$$

form a basis of $\mathfrak{g}_i$ for each i, so by taking the tensor product, we see that the ordered monomials

$$X_{11}^{d_{11}}\cdots X_{1n_1}^{d_{1n_1}}\otimes\cdots\otimes X_{n1}^{d_{n1}}\cdots X_{nn_n}^{d_{nn_n}}$$

form a basis of $U\mathfrak{g}_1\otimes_F\cdots\otimes_F U\mathfrak{g}_n$. On the other hand, the elements

$$\left(X_{11}^{d_{11}}\cdots X_{1n_1}^{d_{1n_1}}\right)\cdots\left(X_{n1}^{d_{n1}}\cdots X_{nn_n}^{d_{nn_n}}\right)$$

are the ordered monomials of $U\mathfrak{g}$, so they also form a basis. Thus, μ sends a basis to a basis, so we are done. ■

Remark 3.118. The above proposition does not require the $\mathfrak{g}_\bullet$s to commute within $\mathfrak{g}$. Namely, it is not at all required that the vector space decomposition is actually a Lie algebra decomposition.

We next move on to results which sharpen Theorem 3.115 in various ways. To begin, we will exhibit an isomorphism between $S\mathfrak{g}$ and $\operatorname{gr} U\mathfrak{g}$. Let's begin with a couple lemmas.

Lemma 3.119. Fix a Lie algebra $\mathfrak{g}$ over a field F. Then $\operatorname{gr} U\mathfrak{g}$ is a commutative F-algebra.

Proof. The point is to use Remark 3.114 after making enough reductions. We would like to check that $ab=ba$ for all $a,b\in\operatorname{gr}U\mathfrak{g}$. This condition corresponds to an equality of some multilinear maps $\operatorname{gr}U\mathfrak{g}\times\operatorname{gr}U\mathfrak{g}\to\operatorname{gr}U\mathfrak{g}$, so it can be checked on spanning subsets of $\operatorname{gr}U\mathfrak{g}$. Namely, we may assume that $a\in\mathcal{F}_kU\mathfrak{g}/\mathcal{F}_{k-1}U\mathfrak{g}$ and $b\in\mathcal{F}_\ell U\mathfrak{g}/\mathcal{F}_{\ell-1}U\mathfrak{g}$. Now, by definition of $\mathcal{F}_kU\mathfrak{g}$ in Example 3.107 means that elements of $\mathcal{F}_kU\mathfrak{g}$ can be written as homogeneous polynomials of degree n in a basis $\{X_1,\ldots,X_n\}$ of $\mathfrak{g}$. Thus, using the multilinearity, we may assume that $a:=X_{i_1}\cdots X_{i_k}$ and $b:=X_{j_1}\cdots X_{j_\ell}$ are monomials. We would like to show that $ab-ba=0$, which corresponds to showing that

$$ab-ba\in\mathcal{F}_{k+\ell-1}U\mathfrak{g},$$

which means that $ab-ba$ must be a linear combination of monomials of degree less than $k+\ell$. Well, let $\sigma,\tau\in S_{k+\ell}$ be permutations which order the monomials ab and ba, which will produce the same ordered monomial $X_1^{d_1}\cdots X_n^{d_n}$ from both ab and ba (because we are reordering the same multiset $\{i_1,\ldots,i_k\}\sqcup\{j_1,\ldots,j_\ell\}$). But now Remark 3.114 tells us that $ab-X_1^{d_1}\cdots X_n^{d_n}$ and $ba-X_1^{d_1}\cdots X_n^{d_n}$ both live in $\mathcal{F}_{k+\ell-1}U\mathfrak{g}$, so $ab-ba\in\mathcal{F}_{k+\ell-1}U\mathfrak{g}$, as required. ■

Proposition 3.120. Fix a Lie algebra $\mathfrak{g}$ over a field F with basis $\{X_1,\ldots,X_n\}$. Then the map $S\mathfrak{g}\to\operatorname{gr}U\mathfrak{g}$ given by

$$X_1^{d_1}\cdots X_n^{d_n}\mapsto X_1^{d_1}\cdots X_n^{d_n}\in\frac{\mathcal{F}_{d_1+\cdots+d_n}U\mathfrak{g}}{\mathcal{F}_{d_1+\cdots+d_n-1}U\mathfrak{g}}$$

is an isomorphism of graded F-algebras.

Proof. Quickly, we note that there is a unique F-linear map $S\mathfrak{g}\to\operatorname{gr}U\mathfrak{g}$ because ordered monomials form a basis of $S\mathfrak{g}$: recall from Example 3.96 that $S\mathfrak{g}=F[X_1,\ldots,X_n]$ is a commutative polynomial ring.

To show that such an F-algebra morphism exists, we use the universal property of such polynomial rings is that an F-algebra map out of $S\mathfrak{g}$ is determined exactly by choosing where the elements $\{X_1,\ldots,X_n\}$ go, provided the target is a commutative F-algebra. Thus, we define $\varphi\colon S\mathfrak{g}\to\operatorname{gr}U\mathfrak{g}$ by defining $\varphi(X_i):=X_i$ for each i. (Note $\operatorname{gr}U\mathfrak{g}$ is commutative by Lemma 3.119.) Here are the required checks on this map.

- On ordered monomials $X_1^{d_1}\cdots X_n^{d_n}$, we see that

$$\varphi\left(X_1^{d_1}\cdots X_n^{d_n}\right)=\varphi(X_1)^{d_1}\cdots\varphi(X_n)^{d_n}=X_1^{d_1}\cdots X_n^{d_n},$$

  which we note lives in the degree-d piece $\mathcal{F}_{d_1+\cdots+d_n}U\mathfrak{g}/\mathcal{F}_{d_1+\cdots+d_n-1}U\mathfrak{g}$ because each $X_i\in\operatorname{gr}U\mathfrak{g}$ has degree 1.

- Graded: we claim that if $p(X_1,\ldots,X_n)\in S\mathfrak{g}$ is in the degree-d graded piece, then $\varphi(p)\in\operatorname{gr}U\mathfrak{g}$ is in the degree-d graded piece as well. The graded pieces are F-linear subspaces, so we may check this on a basis of the degree-d graded piece of $S\mathfrak{g}$, for which we use the ordered monomials $X_1^{d_1}\cdots X_n^{d_n}$ of total degree d. But then the previous check already verified that $\varphi\left(X_1^{d_1}\cdots X_n^{d_n}\right)$ lands in the correct graded piece.

- Surjective: we use Lemma 3.112. It is enough to check that the spanning subset of homogeneous elements of $\operatorname{gr}U\mathfrak{g}$ are in the image of φ. Namely, we want a component $\mathcal{F}_kU\mathfrak{g}/\mathcal{F}_{k-1}U\mathfrak{g}$ to be in $\operatorname{im}\varphi$. However, $\mathcal{F}_kU\mathfrak{g}$ consists of the sum of homogeneous polynomials in $\{X_1,\ldots,X_n\}$ of degree k, so it is enough to check that homogeneous polynomials in $\{X_1,\ldots,X_n\}$ of degree k in $\mathcal{F}_kU\mathfrak{g}/\mathcal{F}_{k-1}U\mathfrak{g}$ are in $\operatorname{im}\varphi$. (Explicitly, $\mathcal{F}_{k-1}U\}$ kills homogeneous polynomials of degree less than k.) However, such homogeneous polynomials are spanned by the monomials

$$X_1^{d_1}\cdots X_n^{d_n}$$

of total degree k, which we note equals $\varphi\left(X_1^{d_1}\cdots X_n^{d_n}\right)\in\mathcal{F}_kU\mathfrak{g}/\mathcal{F}_{k-1}$ and thus is in the image of φ.

- Injective: we use Theorem 3.115 to show $\ker\varphi=0$. Suppose that some polynomial $p(X_1,\ldots,X_n)\in S\mathfrak{g}$ vanishes in $\operatorname{gr}U\mathfrak{g}$. Splitting up p up into graded pieces by the degree, the fact that φ preserves grading allows us to assume that p is homogeneous of degree d. Thus, say that

$$\sum_{\substack{(d_1,\ldots,d_n)\in\mathbb{N}^n\\ d_1+\cdots+d_n=d}} a_{(d_1,\ldots,d_n)}X_1^{d_1}\cdots X_n^{d_n}\in\ker\varphi,$$

and we want to check that this element actually vanishes. Well, unwrapping the definition of φ, we see that

$$\sum_{\substack{(d_1,\ldots,d_n)\in\mathbb{N}^n\\ d_1+\cdots+d_n=d}} a_{(d_1,\ldots,d_n)}X_1^{d_1}\cdots X_n^{d_n}\in\mathcal{F}_{d_1+\cdots+d_n-1}U\mathfrak{g}.$$

Namely, we have a homogeneous polynomial of degree d which can be written as an F-linear combination of homogeneous polynomials of strictly smaller degree, so we may write

$$\sum_{\substack{(d_1,\ldots,d_n)\in\mathbb{N}^n\\ d_1+\cdots+d_n=d}} a_{(d_1,\ldots,d_n)}X_1^{d_1}\cdots X_n^{d_n}=\sum_{\substack{(d_1,\ldots,d_n)\in\mathbb{N}^n\\ d_1+\cdots+d_n<d}} a_{(d_1,\ldots,d_n)}X_1^{d_1}\cdots X_n^{d_n}$$

for some new coefficients $a_{(d_1,\ldots,d_n)}$. However, the monomials $X_1^{d_1}\cdots X_n^{d_n}$ over all total degrees are linearly independent in $U\mathfrak{g}$ by Theorem 3.115, so we see that all coefficients $a_{(d_1,\ldots,d_n)}$ must vanish, so in particular $p=0$. ■

Corollary 3.121. Fix a Lie algebra $\mathfrak{g}$ over a field F. Then the algebra $U\mathfrak{g}$ has no zero divisors.

Proof. By Lemma 3.111, it is enough to check that $\operatorname{gr}U\mathfrak{g}$ has no zero divisors. But Proposition 3.120 tells us that $\operatorname{gr}U\mathfrak{g}$ is isomorphic to $S\mathfrak{g}$ as rings, and the commutative polynomial ring $S\mathfrak{g}$ certainly does not have zero divisors. ■

Proposition 3.120 may be unsatisfying because it requires us to pass through $\operatorname{gr}U\mathfrak{g}$ even though it is really $U\mathfrak{g}$ which interests us. Of course, we cannot expect $U\mathfrak{g}$ and $S\mathfrak{g}$ to be isomorphic as F-algebras because $U\mathfrak{g}$ is not commutative in general (indeed, $XY-YX=[X,Y]$ only vanishes when $\mathfrak{g}$ is abelian). However, we can get some structure preserved.

Proposition 3.122. Fix a Lie algebra $\mathfrak{g}$ over a field F. Define $\operatorname{sym}\colon S\mathfrak{g} \to U\mathfrak{g}$ by

$$\operatorname{sym}(Y_1 \cdots Y_k) := \frac{1}{k!} \sum_{\sigma \in S_k} Y_{\sigma(1)} \cdots Y_{\sigma(k)}$$

for any $Y_1, \ldots, Y_k \in \mathfrak{g}$. Then sym is a well-defined isomorphism of $\mathfrak{g}$-modules, where we are using the adjoint action by $\mathfrak{g}$.

Proof. This is on the homework. ■

Corollary 3.123. Fix a Lie algebra $\mathfrak{g}$ over a field F. The map $\operatorname{sym}\colon S\mathfrak{g} \to U\mathfrak{g}$ of Proposition 3.122 restricts to an isomorphism

$$(S\mathfrak{g})^{\mathfrak{g}} \to Z(U\mathfrak{g}).$$

Proof. Because sym is an isomorphism of $\mathfrak{g}$-modules, we see that $\mathfrak{g}$ will restrict to an isomorphism of $\mathfrak{g}$-invariants. On one hand, the $\mathfrak{g}$-invariants of $S\mathfrak{g}$ make the space $(S\mathfrak{g})^{\mathfrak{g}}$. On the other hand, we claim $(U\mathfrak{g})^{\mathfrak{g}} = Z(U\mathfrak{g})$, which will complete the proof. In one direction, note $a \in Z(U\mathfrak{g})$ implies that $Xa = aX$ for all $X \in \mathfrak{g}$, so $\operatorname{ad}_X(a) = 0$ by Lemma 3.98. In the other direction, if $a \in (U\mathfrak{g})^{\mathfrak{g}}$, we merely know $\operatorname{ad}_X(a) = Xa - aX$ vanishes for all $X \in \mathfrak{g}$. Thus, we define

$$C(a) := \{b \in U\mathfrak{g} : ab = ba\}.$$

Note that $C(a) \subseteq U\mathfrak{g}$ is the kernel of the linear map $b \mapsto ab - ba$, so it is a linear subspace. Additionally, $C(a)$ is closed under multiplication: if $b, b' \in C(a)$, then $abb' = bab' = bb'a$, so $bb' \in C(a)$. Now, we would like to check that $C(a) = U\mathfrak{g}$, for which it is enough to check that $C(a)$ contains monomials (because these span $U\mathfrak{g}$), for which it is enough to check that $\mathfrak{g} \subseteq C(a)$ (because monomials are products of elements of $\mathfrak{g}$). ■

Theme 4

Semisimple Structure Theory

4.1 October 11

I missed class due to a minor leg injury. I thank Justin for access to his notes.

4.1.1 Ideals and Commutants

We will spend some time focusing on Lie algebras in their own right, instead of studying their representations. As such, we pick up where we left off in section 2.2.3 on this topic. Here are a few ways to build Lie ideals.

Lemma 4.1. Fix a Lie algebra $\mathfrak{g}$, and let $\{I_\alpha\}_{\alpha\in\lambda}$ be a collection of Lie ideals. Then the sum

$$\sum_{\alpha\in\lambda} I_\alpha := \left\{\sum_{\alpha\in\lambda} X_\alpha : X_\alpha \in I_\alpha \text{ and } X_\alpha = 0 \text{ for all but finitely many } \alpha\in\lambda\right\}$$

is also a Lie ideal.

Proof. Linear algebra tells us that $\sum_{\alpha\in\lambda} I_\alpha$ is at least a subspace. Now, for any $X \in \mathfrak{g}$ and $\sum_{\alpha\in\lambda} X_\alpha \in \sum_{\alpha\in\lambda} I_\alpha$, we see that

$$\left[X, \sum_{\alpha\in\lambda} X_\alpha\right] = \sum_{\alpha\in\lambda}[X, X_\alpha],$$

where we are allowed to use the bilinearity of the bracket here because $\sum_{\alpha\in\lambda} X_\alpha$ is actually a finite sum. Now, $[X, X_\alpha] \in I_\alpha$ for each α because I_α is a Lie ideal, so we conclude that $\left[X, \sum_{\alpha\in\lambda} X_\alpha\right] \in \sum_{\alpha\in\lambda} I_\alpha$. ■

Lemma 4.2. Fix Lie ideals I and J of a Lie algebra $\mathfrak{g}$. Then

$$[I, J] := \operatorname{span}\{[X, Y] : X \in I, Y \in J\}$$

is a Lie ideal of $\mathfrak{g}$. In fact, $[I, J] \subseteq I \cap J$.

Proof. We have taken a span of some vectors, so $[I, J]$ is certainly a subspace. To check that it is a Lie ideal, we must check that $[W, Z] \in [I, J]$ for any $W \in \mathfrak{g}$ and $Z \in [I, J]$. Well, $[W, -]\colon \mathfrak{g} \to \mathfrak{g}$ is a linear map, so the pre-image of $[I, J]$ is a linear subspace; to check that this pre-image contains $[I, J]$, it is thus enough to check that the pre-image contains a spanning subset, for which we use the elements of the form $[X, Y]$ where $X \in I$ and $Y \in J$. As such, we want to see $[W, [X, Y]] \in [I, J]$, for which we use the Jacobi identity to write

$$[W, [X, Y]] = -[X, [Y, W]] - [Y, [W, X]] = -[X, [Y, W]] + [[W, X], Y].$$

Now, $[Y, W] \in J$ because J is a Lie ideal, and $[W, X] \in I$ because I is a Lie ideal, so we see that $-[X, [Y, W]] + [[W, X], Y] \in [I, J]$.

Lastly, to check that $[I, J] \subseteq I \cap J$, we note that the latter is a Lie ideal by Remark 2.40 and hence a subspace, so it is enough to check the inclusion on a spanning subset of $[I, J]$, for which we take the elements of the form $[X, Y] = -[Y, X]$ for $X \in I$ and $Y \in J$. Well, $[X, Y] \in I$ and $-[Y, X] \in J$ because I and J are Lie ideals, so we are done. ■

We will get a lot of utility out of the above lemma. For example, we can make the following definition.

Definition 4.3 (commutator)**.** The *commutator* of a Lie algebra $\mathfrak{g}$ is the Lie ideal $[\mathfrak{g}, \mathfrak{g}]$.

Lemma 4.4. Fix a Lie algebra $\mathfrak{g}$. For any Lie ideal $I \subseteq \mathfrak{g}$, the quotient $\mathfrak{g}/I$ is abelian if and only if I contains $[\mathfrak{g}, \mathfrak{g}]$. For example, $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is abelian.

(a) The quotient $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is an abelian Lie algebra.

(b)

Proof. Note that the last sentence follows from the previous one (take $I = [\mathfrak{g}, \mathfrak{g}]$), so it only remains to prove the second sentence. We show the two implications separately.

- Suppose $I \subseteq \mathfrak{g}$, and we will show $\mathfrak{g}/I$ is abelian. Well, for any $X + I, Y + I \in \mathfrak{g}/I$, we compute

  $$[X + I, Y + I] = [X, Y] + I,$$

  which we note equals $0 + I$ because $[X, Y] \in [\mathfrak{g}, \mathfrak{g}] \subseteq I$.

- Suppose $\mathfrak{g}/I$ is abelian, and we will show $[\mathfrak{g}, \mathfrak{g}] \subseteq I$. Well, $[\mathfrak{g}, \mathfrak{g}]$ and I are both subspaces, so it is enough to check that a spanning subset of $[\mathfrak{g}, \mathfrak{g}]$ is contained in I, for which we use the elements of the form $[X, Y]$. Then we see that

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

  where the last equality holds because $\mathfrak{g}/I$ is abelian. Thus, $[X, Y] \in I$, as required. ■

Exercise 4.5. Fix a field F. Then $[\mathfrak{gl}_n(F), \mathfrak{gl}_n(F)] = \mathfrak{sl}_n(F)$.

Proof. Here are our inclusions.

- To show $[\mathfrak{gl}_n(F), \mathfrak{gl}_n(F)] = \mathfrak{sl}_n(F)$, it is enough to check that a spanning subset of $[\mathfrak{gl}_n(F), \mathfrak{gl}_n(F)]$ lives in $\mathfrak{sl}_n(F)$, for which we note that elements of the form $[X, Y]$ have $\operatorname{tr}[X, Y] = \operatorname{tr} XY - \operatorname{tr} YX = 0$ and thus $[X, Y] \in \mathfrak{sl}_n(F)$.

- To show $\mathfrak{sl}_n(F) \subseteq [\mathfrak{gl}_n(F), \mathfrak{gl}_n(F)]$, we should show that $[\mathfrak{gl}_n(F), \mathfrak{gl}_n(F)]$ contains a spanning subset of $\mathfrak{sl}_n(F)$. Let E_{ij} be the matrix with a 1 in the (i, j) component and a 0 everywhere else. Then

  $$[E_{ij}, E_{ji}] = E_{ij}E_{ji} - E_{ji}E_{ij} = E_{ii} - E_{jj},$$

  and

  $$[E_{ii} - E_{jj}, E_{ij}] = E_{ii}E_{ij} - E_{jj}E_{ij} - E_{ij}E_{ii} + E_{ij}E_{jj} = E_{ij} - 0 - 0 + E_{ij} = 2E_{ij}$$

  for any $i \neq j$. These elements span $\mathfrak{sl}_n(F)$ (we have any off-diagonal entry, and a traceless diagonal matrix can be written as a sum of $(E_{ii} - E_{nn})$s), so we conclude. ■

As with group theory, commutators allow us to define solvability.

Definition 4.6 (derived series)**.** Fix a Lie algebra $\mathfrak{g}$. Then the *derived series* is a sequence $\{D^i\mathfrak{g}\}_{i\geq 0}$ of Lie ideals defined inductively by $D^0\mathfrak{g} := \mathfrak{g}$ and

$$D^{i+1}\mathfrak{g} := \left[D^i\mathfrak{g}, D^i\mathfrak{g}\right].$$

Remark 4.7. Inductively applying Lemma 4.2 shows that each $D^i\mathfrak{g}$ is in fact a Lie ideal. Furthermore, we see Lemma 4.2 implies each $i \geq 0$ has

$$D^{i+1}\mathfrak{g} = \left[D^i\mathfrak{g}, D^i\mathfrak{g}\right] \subseteq D^i\mathfrak{g} \cap D^i\mathfrak{g} = D^i\mathfrak{g}.$$

Lemma 4.8. Fix a Lie algebra $\mathfrak{g}$. Then the following conditions are equivalent.

(a) $D^m\mathfrak{g} = 0$ for some m.

(b) $D^m\mathfrak{g} = 0$ for all sufficiently large m.

(c) There is a descending chain
$$\mathfrak{g} = \mathfrak{g}^0 \supseteq \mathfrak{g}^1 \supseteq \cdots \supseteq \mathfrak{g}^m = 0$$
of Lie ideals such that $\mathfrak{g}/\mathfrak{g}_{i+1}$ is abelian.

Proof. We show our implications separately.

- Note that (a) implies (b) because $D^{i+1}\mathfrak{g} \subseteq D^i\mathfrak{g}$ for each $i \geq 0$ (by Remark 4.7), so $D^m\mathfrak{g} = 0$ implies (inductively) that $D^{m+i}\mathfrak{g} = 0$ for all $i \geq 0$.
- We show that (b) implies (c) by taking $\mathfrak{g}^i := D^i\mathfrak{g}$ for all $i \in \{0, \ldots, m\}$ for some sufficiently large m. This is a descending chain of Lie ideals by Remark 4.7, and it has $\mathfrak{g}^m = 0$ by hypothesis. Lastly, we note that
$$\mathfrak{g}^i/\mathfrak{g}^{i+1} = \mathfrak{g}^i/\left[\mathfrak{g}^i, \mathfrak{g}^i\right]$$
is abelian by Lemma 4.2.
- We show that (c) implies (a). We claim that $D^i\mathfrak{g} \subseteq \mathfrak{g}^i$ for each $i \in \{0, \ldots, m\}$, which will imply that $D^m\mathfrak{g} = 0$ and thus complete the proof. Well, we show the claim by induction, for which the base case of $i = 0$ has nothing to show. For the inductive step, we note that $\mathfrak{g}^i/\mathfrak{g}^{i+1}$ being abelian implies that
$$\mathfrak{g}^{i+1} \subseteq \left[\mathfrak{g}^i, \mathfrak{g}^i\right].$$
The spanning subset of commutators defining $\left[\mathfrak{g}^i, \mathfrak{g}^i\right]$ is a subset of those defining $D^{i+1}\mathfrak{g} = \left[D^i\mathfrak{g}, D^i\mathfrak{g}\right]$ by the inductive hypothesis, so we conclude that $\mathfrak{g}^{i+1} \subseteq D^{i+1}\mathfrak{g}$. ■

Definition 4.9 (solvable)**.** A Lie algebra $\mathfrak{g}$ is *solvable* if and only if

Remark 4.10. One can see that $\mathfrak{g}$ is solvable if and only if $\mathfrak{g} \otimes_F \overline{F}$ is solvable.

One can check that sums of ideals are ideals. Also, one sees that $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is the maximal abelian quotient: if $I \subseteq \mathfrak{g}$ is an ideal with $\mathfrak{g}/I$ abelian, then we must have $[\mathfrak{g}, \mathfrak{g}] \subseteq I$.

Example 4.11. One can check that $[\mathfrak{gl}_n, \mathfrak{gl}_n] \subseteq \mathfrak{sl}_n$ because the trace of $XY - YX$ is zero for any $X, Y \in \mathfrak{gl}_n$. In fact, this is an equality, which one can check by hand.

These commutants provide a derived series: we define $\{D^i\mathfrak{g}\}_{i\ge 0}$ inductively by $D^0\mathfrak{g} := \mathfrak{g}$ and

$$D^{i+1}\mathfrak{g} := [D^i\mathfrak{g}, D^i\mathfrak{g}]$$

for all $i \geq 0$. This derived series plays the role of derived series in group theory. For example, one can use this to define solvability.

Proposition 4.12. Fix a Lie algebra $\mathfrak{g}$. Then the following are equivalent.

(a) $D^n\mathfrak{g} = 0$ for n sufficiently large.

(b) There exists a sequence of subalgebras

$$\mathfrak{g} = \mathfrak{a}^0 \supseteq \mathfrak{a}^1 \supseteq \cdots \supseteq \mathfrak{a}^k = 0$$

such that $\mathfrak{a}^{i+1}$ is an ideal in $\mathfrak{a}^i$ with abelian quotient.

(c) For every n sufficiently large and sequence of elements $\{x_1, \ldots, x_{2^n}\} \subseteq \mathfrak{g}$, the n-fold commutator

$$[\cdots[[x_1, x_2], [x_3, x_4]], \cdots]$$

vanishes.

Proof. The equivalence of (a) and (c) has no content. Note that (a) implies (b) because one may take $\mathfrak{a}^i = D^i\mathfrak{g}$. One achieves (b) implies (a) by showing that $\mathfrak{a}^i \supseteq D^i\mathfrak{g}$ inductively. ∎

4.2 October 14

Today we finish up some structure theory of Lie algebras.

4.2.1 Engel's Theorem

Last time we proved the following theorem.

Theorem 4.13 (Lie)**.** If $\mathfrak{g}$ is a complex solvable Lie algebra, then any irreducible representation V of $\mathfrak{g}$ is 1-dimensional. For any representation of V of $\mathfrak{g}$, there is a basis in which $\mathfrak{g}$ acts by upper-triangular matrices.

Today we begin by proving Engel's theorem. Here are some lemmas.

Lemma 4.14. Fix a Lie algebra $\mathfrak{g}$ over a field F, and let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation into an F-vector space V.

(a) If all elements of $\mathfrak{g}$ act by nilpotent operators, then there is nonzero $v \in V$ such that

$$Xv = 0$$

for all $X \in \mathfrak{g}$.

(b) There exists a basis of V in which all matrices in $\mathfrak{g}$ are strictly upper-triangular.

Proof. Note that (b) follows from (a) inductively, so we focus on proving (a). For this, we induct on $\dim\mathfrak{g}$, where the case of $\dim\mathfrak{g} = 0$ means that $\mathfrak{g}$ vanishes, so there is nothing to do.

To induct downwards, we would like to find a Lie ideal of $\mathfrak{g}$. Well, we claim that any maximal proper subalgebra $\mathfrak{h}$ is a Lie ideal of codimension 1. Of course, $\mathfrak{h}$ has codimension 1 because otherwise we could add a vector to it, violating maximality. We will Now, consider the adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$, which we note will descend to an adjoint action of $\mathfrak{h}$ on $\mathfrak{g}/\mathfrak{h}$. Notably, all $X \in \mathfrak{h}$ are nilpotent (this is true for $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ even), so one can check that the operator $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is nilpotent by some repeated applications, so $\operatorname{ad}_X : (\mathfrak{g}/\mathfrak{h}) \to (\mathfrak{g}/\mathfrak{h})$ is a nilpotent operator, so the inductive hypothesis tells us that we can find nonzero $Y + \mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$ such that $\operatorname{ad}_X(Y) \in \mathfrak{h}$ for all X. Thus, we see that we can write $\mathfrak{g} = \mathfrak{h} + kY$ is a subalgebra, so we conclude that $\mathfrak{h}$ is an ideal of codimension 1.

For the inductive step, we let W by the collection of $\mathfrak{h}$-invariants of W. The inductive hypothesis applied to $\mathfrak{h}$ tells us that $V^{\mathfrak{h}}$ is nonzero. Note that W is in fact a subrepresentation for $\mathfrak{g}$ because any $X \in \mathfrak{h}$ and $w \in W$ has

$$XYw + YXw + [X,Y]w = 0,$$

so Yw is fixed by $\mathfrak{h}$, and we are done. To complete the proof, we choose some nonzero $w \in W$ and apply Y enough times until $Y^{k-1}w \neq 0$ but $Y^k w = 0$. Then $Y^{k-1}w$ works. ■

Remark 4.15. This result works over any field F.

Definition 4.16. An element $X \in \mathfrak{g}$ is nilpotent if and only if $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is a nilpotent operator.

Corollary 4.17. A finite-dimensional complex Lie algebra $\mathfrak{g}$ is nilpotent if and only if each $X \in \mathfrak{g}$ is nilpotent.

Proof. Of course, if $\mathfrak{g}$ is nilpotent, then each element is nilpotent by considering some adjoint composite with the lower central series. Conversely, if each element of $\mathfrak{g}$ is nilpotent, then the lemma above allows us to strictly upper-triangularize our operators ad_X, which means that they are nilpotent. ■

4.2.2 Semisimple Lie Algebras

We now give a central definition of this subject: semisimple.

Lemma 4.18. Fix a Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}$ contains a maximal solvable Lie ideal containing all solvable Lie ideals.

Proof. The main point it to show that the sum of two solvable ideals is again solvable, for which we note that we have short exact sequences

$$0 \to I \to I + J \to \frac{I+J}{I} \to 0$$

and

$$0 \to (I \cap J) \to J \to \frac{I+J}{I} \to 0,$$

so we are able to conclude that $I + J$ is solvable. Everything in sight is finite-dimensional, so we can just sum over all solvable ideals to get the required ideal. ■

The lemma allows us to define semisimple.

Definition 4.19 (radical, simple, semisimple)**.** Fix a Lie algebra $\mathfrak{g}$.

(a) The *radical* $\operatorname{rad}\mathfrak{g}$ of $\mathfrak{g}$ is the maximal solvable ideal of $\mathfrak{g}$.

(b) $\mathfrak{g}$ is *semisimple* if and only if $\operatorname{rad}\mathfrak{g} = 0$.

(c) $\mathfrak{g}$ is *simple* if and only if it is not abelian and the only ideals of $\mathfrak{g}$ are 0 and $\mathfrak{g}$.

Some remarks are in order to make sure that these definitions make sense.

Remark 4.20. One can see that $\mathfrak{g}$ is simple if and only if $\mathrm{ad}_\bullet : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is an irreducible representation.

Remark 4.21. If $\mathfrak{g}$ is simple and $\dim \mathfrak{g} > 1$, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. Indeed, $[\mathfrak{g}, \mathfrak{g}]$ is some Lie ideal, and it must be nonzero: if $[\mathfrak{g}, \mathfrak{g}] = 0$, then $\mathfrak{g}$ is abelian, which is not permitted.

Remark 4.22. If $\mathfrak{g}$ is simple, then $\mathfrak{g}$ is semisimple. Indeed, if $\mathfrak{g}$ is simple, then the only Lie ideals available are 0 and $\mathfrak{g}$. However, $\mathfrak{g}$ is not solvable because $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

Remark 4.23. We claim that

$$\mathrm{rad}(\mathfrak{g}_1 \oplus \mathfrak{g}_2) = \mathrm{rad}\,\mathfrak{g}_1 \oplus \mathrm{rad}\,\mathfrak{g}_2.$$

This implies that the direct sum of semisimple Lie algebras is semisimple.

Example 4.24. The Lie algebras $\mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{so}_3(\mathbb{C})$ are simple. This can be checked directly. We will show later that $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{so}_n(\mathbb{C})$ and $\mathfrak{sp}_{2n}(\mathbb{C})$ are all semisimple.

In general, one can always reduce our Lie algebras to semisimple ones.

Lemma 4.25. The Lie algebra $\mathfrak{g}_{\mathrm{ss}} := \mathfrak{g}/\mathrm{rad}\,\mathfrak{g}$ is semisimple. In fact, if $\mathfrak{h} \subseteq \mathfrak{g}$ is solvable and has $\mathfrak{g}/\mathfrak{h}$ semisimple, then $\mathfrak{h} = \mathrm{rad}\,\mathfrak{g}$.

Proof. Solvable ideals lift from the quotient. Namely, $I \subseteq \mathfrak{g}/\mathrm{rad}\,\mathfrak{g}$ being solvable implies that its pre-image in $\mathfrak{g}$ is solvable, so it will be contained in the radical. For the second part, we note that $\mathfrak{h} \subseteq \mathrm{rad}\,\mathfrak{g}$, so $\mathfrak{g}/\mathfrak{h} \twoheadrightarrow \mathfrak{g}/\mathrm{rad}\,\mathfrak{g}$. ■

Remark 4.26. It is a theorem that $\mathrm{char}\,F = 0$ makes the short exact sequence

$$0 \to \mathrm{rad}\,\mathfrak{g} \to \mathfrak{g} \to \mathfrak{g}_{\mathrm{ss}} \to 0$$

splits. We will not prove this today.

Example 4.27. One can show that the group of rigid motions preserving orientation is a semidirect product of rotation group $\mathrm{SO}_3(\mathbb{R})$ with the translation group $\mathbb{R}^3$.

4.3 October 16

Today we continue with our structure theory.

4.3.1 Some Semisimple Lie Algebras

We continue with our study of $\mathrm{rad}(\mathfrak{g})$.

Proposition 4.28. Fix an algebraically closed field F of characteristic 0. If V is an irreducible representation of a Lie algebra $\mathfrak{g}$ (over F), then $\mathrm{rad}(\mathfrak{g})$ acts on V by scalars, and $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ vanishes.

Proof. By Theorem 4.13, we get a common eigenvector $v \in V$ for $\operatorname{rad}(\mathfrak{g})$, so there is a linear functional $\lambda\colon \operatorname{rad}(\mathfrak{g}) \to F$ so that $Xv = \lambda(X)v$ for all $X \in \operatorname{rad}(\mathfrak{g})$. Now, for any $X \in \mathfrak{g}$, we define

$$\mathfrak{g}_X \coloneqq \operatorname{rad}(\mathfrak{g}) + FX.$$

Note that $\mathfrak{g}_X$ is a subalgebra because $\operatorname{rad}(\mathfrak{g})$ is a Lie ideal. Now, for any $a \in \operatorname{rad}(\mathfrak{g})$, we see that

$$aX^n v = \lambda(a)X^n v + \sum_{i=1}^{n} c_i X^{n-i} v$$

for some constants $c_1, \ldots, c_n \in F$ where $c_i \coloneqq \lambda([X,a]\ldots)$. As such, we may let W be the span of the $x^\bullet v$s, so W is stable under $\mathfrak{g}_X$, and a only has the eigenvalue λ. Thus, for all $[X,a] \in [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})]$, we see $\lambda([X,a]) = 0$ because $[X,a]$ on W has vanishing trace. Thus, we actually see that $aXv = \lambda(a)Xv$. As such, the λ-eigenspace V_λ of V is actually stable under all $X \in \mathfrak{g}$, so $V_\lambda = V$ by the irreducibility. ■

We are now ready to define reductive.

Definition 4.29 (reductive)**.** A Lie algebra $\mathfrak{g}$ is reductive if and only if $\operatorname{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$.

Remark 4.30. Note $\mathfrak{g}$ is reductive if and only if $[\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] = 0$. Indeed, this will imply that $\operatorname{rad}(\mathfrak{g}) \subseteq \mathfrak{z}(\mathfrak{g})$, but of course $\mathfrak{z}(\mathfrak{g})$ is solvable, so the other inclusion holds as well.

Remark 4.31. Intuitively, one can say that being reductive means being a direct sum of semisimple and center.

To check that certain Lie algebras are semisimple or reductive, it will be helpful to have access to invariant inner products.

Definition 4.32 (invariant)**.** Fix a Lie algebra $\mathfrak{g}$ over a field F. A bilinear form $\langle -, - \rangle$ on $\mathfrak{g}$ is *invariant* if and only if

$$B([X,Y],Z) = B(X,[Y,Z])$$

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

Example 4.33. For any representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of a Lie algebra $\mathfrak{g}$, the form

$$B_V(X,Y) \coloneqq \operatorname{tr}(\rho(X)\rho(Y)).$$

(Technically, we ought to write B_ρ, but we will write B_V or even B when confusion cannot arise.)

As usual, inner products allow us to take complements.

Proposition 4.34. Fix a symmetric invariant bilinear form B on a Lie algebra $\mathfrak{g}$. For any ideal $I \subseteq \mathfrak{g}$, the orthogonal complement

$$I^\perp \coloneqq \{X \in \mathfrak{g} : \langle X, Y \rangle = 0 \text{ for all } Y \in I\}$$

is also a Lie ideal.

Proof. Check it. ■

Proposition 4.35. Fix a Lie algebra $\mathfrak{g}$. Suppose there is a representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ for which B_ρ is non-degenerate. Then $\mathfrak{g}$ is reductive.

Proof. Let $\{V_1, \ldots, V_n\}$ be the irreducible factors of V, counted with multiplicity. Then ρ can be upper-triangularized appropriately to see that

$$B_V = \sum_{i=1}^{k} B_{V_i}.$$

Now, for each $X \in [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})]$, we see that $\rho_{V_i}(X) = 0$, so $B_{V_i}(X, Y) = 0$ for any $Y \in \mathfrak{g}$, so $X = 0$ because B_V is non-degenerate. ■

Example 4.36. For any classical Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}(V)$, the standard representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ has B_ρ non-degenerate. Thus, $\mathfrak{g}$ is reductive. If further $\mathfrak{z}(\mathfrak{g}) = 0$, then we see that $\mathfrak{g}$ is semisimple. Perhaps one should be worried about positive characteristic.

4.3.2 The Jordan Decomposition

We will require a notion of Jordan decomposition in the sequel.

Proposition 4.37. Fix a perfect field F, and let V be a finite-dimensional F-vector space. Then any $A \in \mathfrak{gl}(V)$ can be written uniquely in the form

$$A = A_s + A_n$$

for $A_s, A_n \in \mathfrak{gl}(V_{\overline{F}})$ satisfying the following.

- A_s is diagonalizable over $\overline{F}$.
- A_n is nilpotent.
- $A_s A_n = A_n A_s$.

It turns out that $A_s, A_n \in \mathfrak{gl}(V)$ and that A_s can be expressed as a polynomial in A.

Proof. We begin by finding A_s over an algebraic closure. Here, we let $\{\lambda_1, \ldots, \lambda_e\} \subseteq \overline{F}$ be the roots of the characteristic polynomial $\chi_A(T)$ of A, where λ_i occurs with multiplicity m_i. Then we note $\overline{F}[T]$ is a principal ideal domain with maximal given by $\{(T - \lambda) : \lambda \in \overline{F}\}$, so we may decompose

$$V_{\overline{F}} \cong \frac{\overline{F}[T]}{(\chi_A(T))} \simeq \bigoplus_{i=1}^{e} \frac{\overline{F}[T]}{(T - \lambda_i)^{m_i}},$$

where $V_{\overline{F}}$ has been given the structure of an $\overline{F}[T]$-module via $T \mapsto A$. Now, the Chinese remainder theorem grants us a polynomial P such that

$$P(T) \equiv \lambda_i \pmod{(T - \lambda_i)^{m_i}}$$

for each i. In particular, we see that

$$P(A) - \lambda_i \mathrm{id}_V \equiv (A - \lambda_i \mathrm{id}_V)^{m_i} Q_i(A)$$

for some polynomial Q_i. In particular, evaluating in $\mathfrak{gl}(V)$, we see that $P(A)$ acts as $\lambda \,\mathrm{id}_V$ on $V[\lambda]$ for each λ_i, so $A_s := P(A)$ is semisimple, and one can check that $A_n := A - A_s$ has all eigenvalues equal to 0 and hence is nilpotent.

We now argue that our decomposition is unique. Suppose we have another such decompositions $A = A'_s + A'_n$. Then A commuting with A'_s means that $A_s = P(A)$ commutes with A'_s, so A_s and A'_s can be simultaneously diagonalized. As such, A_n and A'_n can commute by taking the differences, so we see that

$$A_s - A'_s = A'_n - A_n$$

is a matrix which is both diagonalizable and nilpotent and hence must be zero. The uniqueness allows us to see that $A_s, A_n \in \mathfrak{gl}(V)$ because $A = A_s + A_n$ implies that $A = \sigma(A_s) + \sigma(A_n)$ for all $\sigma \in \operatorname{Gal}(\overline{F}/F)$, so $A_s = \sigma(A_s)$ and $A_n = \sigma(A_n)$ for each σ. ■

Here are some remarks and computations.

Remark 4.38. For $A \in \mathfrak{gl}_n(F)$, we claim that $(\operatorname{ad} A)_s = \operatorname{ad} A_s$ and $(\operatorname{ad} A)_n = \operatorname{Ad}_n$. Indeed, this follows by the uniqueness everywhere in sight.

Remark 4.39. If $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathfrak{gl}_n(F)$ is diagonal, then $\operatorname{ad} A$ is diagonalizable with eigenvectors given by the matrices E_{ij} each with eigenvalue $\lambda_i - \lambda_j$. This follows by a direct commutator computation.

4.4 October 18

Today we prove the Cartan criteria.

4.4.1 Cartan Criteria

We now define a special invariant form.

Definition 4.40 (Killing form)**.** Fix a Lie algebra $\mathfrak{g}$. Then the *Killing form* is the invariant form

$$B_{\mathfrak{g}}(X, Y) := \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y).$$

We will write $K_{\mathfrak{g}}$ for this form or simply K if no confusion can arise.

We will prove two theorems about this.

Theorem 4.41 (Cartan criterion of solvability)**.** Fix a Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. Then $\mathfrak{g}$ is solvable if and only if $[\mathfrak{g}, \mathfrak{g}] \subseteq \ker K$.

Theorem 4.42 (Cartan criterion of semisimplicity)**.** Fix a Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. Then $\mathfrak{g}$ is semisimple if and only if K is non-degenerate.

We begin with a lemma.

Lemma 4.43. Fix an algebraically closed field F of characteristic 0, and choose some Lie subalgebra $\mathfrak{g} \subseteq \mathfrak{gl}_n(F)$. Suppose that all $X \in [\mathfrak{g}, \mathfrak{g}]$ and $Y \in \mathfrak{g}$ have

$$\operatorname{tr} XY = 0.$$

Then $\mathfrak{g}$ is solvable.

Proof. To show $\mathfrak{g}$ is solvable, it is enough to show that $[\mathfrak{g},\mathfrak{g}]$ is nilpotent, so we may hope to show that all the eigenvalues of any element of $[\mathfrak{g},\mathfrak{g}]$ has vanishing eigenvalues. (We are using Engel's theorem.)

Well, pick up some $X \in [\mathfrak{g},\mathfrak{g}]$, and we let $\{\lambda_1,\ldots,\lambda_n\}$ be the eigenvalues of X counted with multiplicity. We will show that

$$\Lambda \coloneqq \operatorname{span}_{\mathbb{Q}}\{\lambda_1,\ldots,\lambda_n\}$$

is the 0 vector space, for which we will show that all of its linear functionals $\varphi\colon \Lambda \to \mathbb{Q}$ vanish. Well, we extend φ to an operator on V by acting on each generalized eigenspace $V[\lambda_i]$ by $\varphi(\lambda_i)$. Then we can compute $\operatorname{ad}\varphi$ is diagonalizable with eigenvalues $\varphi(\lambda_i-\lambda_j)$; similarly, $\operatorname{ad}X_s$ is diagonalizable with eigenvalues $\lambda_i-\lambda_j$. As such, we can choose a polynomial $Q(T) \in F[T]$ such that $Q(\lambda_i-\lambda_j) = \varphi(\lambda_i-\lambda_j)$ for each $\lambda_i-\lambda_j$. In particular, we see that

$$\operatorname{ad} b = Q(\operatorname{ad} X_s).$$

Similarly, we know that there is a polynomial P such that $\operatorname{ad}X_s = P(\operatorname{ad}X)$. Thus, we see that

$$\operatorname{ad} b = (Q(P(\operatorname{ad} X))).$$

As an example computation, note that $P(0) = Q(0) = 0$ because 0 is an eigenvalue. Now, $X \in [\mathfrak{g},\mathfrak{g}]$ can be written as $\sum_{i=1}^m [Y_i,Z_i]$, so on one hand,

$$\operatorname{tr}(\varphi \circ X) = \sum_{i=1}^m \varphi(\lambda_i)n_i,$$

where $n_i = \dim V[\lambda_i]$. On the other hand,

$$\operatorname{tr}\left(\varphi \circ \sum_{i=1}^m [Y_i,Z_i]\right) = \operatorname{tr}\left(\sum_{i=1}^m \operatorname{ad}_\varphi(Y_i)Z_i\right),$$

which we see must vanish because $R(0) = 0$. Thus, we get the condition

$$\sum_{i=1}^m n_i\varphi(\lambda_i)\lambda_i = 0,$$

so applying φ again allows us to conclude that $\varphi = 0$. ■

We are now ready to prove Theorem 4.41.

Proof of Theorem 4.41. There are two implications.

- If $\mathfrak{g}$ is solvable, then we must show that $[\mathfrak{g},\mathfrak{g}]$ lives in the kernel of K. Well, for all $X \in \mathfrak{g}$, we note that solvability implies that ad_X is strictly upper-triangular, so $[\mathfrak{g},\mathfrak{g}]$ continues to be strictly upper-triangular (by Lie's theorem), so $K(x,y) = 0$ whenever $X \in [\mathfrak{g},\mathfrak{g}]$.
- Suppose that $[\mathfrak{g},\mathfrak{g}]$ lives in the kernel of K. Then $\operatorname{im}\operatorname{ad}$ inside $\mathfrak{gl}(\mathfrak{g})$ is solvable by the above lemma, so $\mathfrak{g}/\mathfrak{z}(\mathfrak{g}) \cong \operatorname{im}\operatorname{ad}$ is solvable, but $\mathfrak{z}(\mathfrak{g}) \subseteq \mathfrak{g}$ is a solvable ideal, so we conclude. ■

And here is the proof of Theorem 4.42.

Proof of Theorem 4.42. There are two implications.

- If $\mathfrak{g}$ is semisimple, then the kernel of K is an ideal I of $\mathfrak{g}$. However, $K|_I = K_{\mathfrak{g}}|_I$ will vanish,[1] so I is solvable as discussed, so $I = 0$, so $K_{\mathfrak{g}}$ is non-degenerate.
- If K is non-degenerate, then $\mathfrak{g}$ is reductive by a result from last class. However, $\operatorname{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$ would be contained in the kernel of K, so we see that $\mathfrak{z}(\mathfrak{g}) = 0$, so $\mathfrak{g}$ is in fact semisimple. ■

[1] This property holds for general ideals.

Corollary 4.44. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. Then $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g} \otimes_F \overline{F}$ is semisimple.

Proof. The Killing form is stable under field extension, so this is immediate from Theorem 4.42. ■

Remark 4.45. It is false that being simple is preserved by restriction: any simple Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ has $\mathfrak{g}|_{\mathbb{R}}$ split into two Lie algebras (given by the "realification" Lie ideal, defined by the fixed points of the conjugation action).

Corollary 4.46. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. For any ideal $I \subseteq \mathfrak{g}$, there exists an ideal $J \subseteq \mathfrak{g}$ such that $\mathfrak{g} = I \oplus J$.

Proof. Let $J := I^\perp$ be the orthogonal complement of I with respect to the Killing form K on $\mathfrak{g}$. One can check that J is an ideal, and $I \cap J = 0$ because K is non-degenerate (namely, K vanishes on $I \cap J$, so $I \cap J$ is solvable, so $I \cap J = 0$ because $\mathfrak{g}$ is semisimple). ■

Corollary 4.47. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. Then $\mathfrak{g}$ is a direct sum of simple Lie algebras.

Proof. Induct with the previous corollary. Namely, if $\mathfrak{g}$ fails to be simple, we can decompose it into two smaller pieces. ■

Here is a more powerful version of the above result.

Proposition 4.48. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0, and write $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ as a sum of simple Lie algebras. Then any ideal $I \subseteq \mathfrak{g}$ is of the form

$$\bigoplus_{i \in S} \mathfrak{g}_i,$$

where $S \subseteq \{1, \ldots, k\}$ is some subset.

Proof. Induct on k. If $k \in \{0, 1\}$, there is nothing to do. For the induction, write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{k+1}$. Then consider the projection $\pi_k \colon \mathfrak{g} \to \mathfrak{g}_{k+1}$. There are two cases.

- If $\pi_k(I) = 0$, then $I \subseteq \mathfrak{h}$, so we are done by the inductive hypothesis.
- If $\pi_k(I) = \mathfrak{g}_k$, then we note that $[\mathfrak{g}_{k+1}, I] = \mathfrak{g}_{k+1}$ because $\mathfrak{g}_{k+1}$ is simple, so $I = I' \oplus \mathfrak{g}_{k+1}$ for some other ideal I', for which we again use the inductive hypothesis. ■

Corollary 4.49. Any ideal in a semisimple Lie algebra is semisimple. Any quotient of a semisimple Lie algebra is semisimple.

4.5 October 21

Today we begin talking about representations of semisimple Lie algebras.

4.5.1 More on Derivations

Here is a nice aside.

Proposition 4.50. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. Then the map $\operatorname{ad}_\bullet\colon \mathfrak{g} \to \operatorname{Der}(\mathfrak{g})$ is a bijection.

Proof. We checked earlier that $\operatorname{ad}_\bullet$ at least outputs to derivations. It is injective because the kernel is the abelian Lie ideal $\mathfrak{z}(\mathfrak{g})$, which is trivial because $\mathfrak{g}$ is semisimple. It remains to check surjectivity.

For example, we check that $\mathfrak{g} \subseteq \operatorname{Der}(\mathfrak{g})$ is a Lie ideal. Indeed, for any $X \in \mathfrak{g}$ and $a \in \operatorname{Der}(\mathfrak{g})$, we would like to check that $[a, \operatorname{ad}_X]$ is still of the form $\operatorname{ad}_\bullet$. Well, for any $Y \in \mathfrak{g}$, we compute

$$[a, \operatorname{ad}_X](Y) = a([X,Y]) - [X, a(Y)] = [a(X), Y] = \operatorname{ad}_{a(X)} Y,$$

where we have used the Jacobi identity.

The moral of the story is that the Killing form K of $\operatorname{Der}(\mathfrak{g})$ will restrict to $\mathfrak{g}$ as its own Killing form $K_\mathfrak{g}$. Because $\mathfrak{g}$ is semisimple, we know that $K|_\mathfrak{g}$ is thus non-degenerate. Now, let $I := \mathfrak{g}^\perp$, which we would like to vanish. Well, $I \subseteq \operatorname{Der}(\mathfrak{g})$ is an ideal satisfying $[I, \mathfrak{g}] = 0$, which means that any $a \in I$ and $X \in \mathfrak{g}$ makes $[a, \operatorname{ad}_X] = \operatorname{ad}_{a(X)}$ (see above) vanish, meaning $a(X) = 0$ by injectivity, so $a = 0$. ■

Corollary 4.51. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field $F \in \{\mathbb{R}, \mathbb{C}\}$. Then the Lie algebra of the group $G := \operatorname{Aut}(\mathfrak{g})$ is $\mathfrak{g}$.

Proof. We already know that $\operatorname{Aut}(\mathfrak{g})$ has Lie algebra $\operatorname{Der}(\mathfrak{g})$. ■

This provides a clean way to produce a Lie group for a semisimple Lie algebra.

4.5.2 Motivating H^1

We are going to show that representations of semisimple Lie algebras are completely reducible. For this, as in Theorem 3.73, the key point is to show that irreducible subrepresentations have a complement. Thus, for our Lie algebra $\mathfrak{g}$, we would like to know when a short exact sequence

$$0 \to U \to V \to W \to 0$$

of representations will split into a direct sum $V = U \oplus W$. Well, we can surely decompose $V = U \oplus W$ as vector spaces, but this need not make the $\mathfrak{g}$-action-commute: in general, there will be some function $a\colon \mathfrak{g} \to \operatorname{Hom}_F(W, U)$ such that

$$\rho_V(X)(u, w) = (\rho_U(X)u + a(X)w, \rho_W(w))$$

by properties of the short exact sequence, and one can check that a is a linear map satisfying

$$([X,Y](u,w) - XY + YX)(u,w) = (a([X,Y]), 0),$$

so one finds that

$$a([X,Y]) = [X, a(Y)] + [a(X), Y],$$

where we interpret $[X, a(Y)]$ as $\rho_U(X) \circ a(Y) - a(Y) \circ \rho_V(X)$ and similar for $[a(X), Y]$. This notation is actually okay because one can see that $[X, a]$ is the natural action of $\mathfrak{g}$ on $\operatorname{Hom}_F(W, U)$.

This motivates the following definition.

Definition 4.52 ($Z^1(\mathfrak{g}, E)$)**.** Fix a representation E of a Lie algebra $\mathfrak{g}$. Then the 1-*cocycle* group $Z^1(\mathfrak{g}, E)$ consists of the group of morphisms $a\colon \mathfrak{g} \to E$ satisfying

$$a([X,Y]) = X \cdot a(Y) - Y \cdot a(X).$$

Example 4.53. One sees that $Z^1(\mathfrak{g}, F)$ consists of functionals $\mathfrak{g} \to F$ which vanish on commutators, so this space is simply $(\mathfrak{g}/[\mathfrak{g},\mathfrak{g}])^\vee$. For example, if $\mathfrak{g}$ is semisimple, then $Z^1(\mathfrak{g}, F)$ vanishes.

Example 4.54. One sees that $Z^1(\mathfrak{g}, \mathfrak{g})$ consists exactly of the derivations $\operatorname{Der}(\mathfrak{g})$.

Thus far we have showed that any extension produces a cocycle. In fact, one can check that a 1-cocycle $a \in Z^1(\mathfrak{g}, \operatorname{Hom}_F(W, U))$ produces a representation $\rho_a \colon \mathfrak{g} \to \mathfrak{gl}(U \oplus W)$ by

$$\rho_a(X)(u, w) := (\rho_U(X)u + a(X)w, \rho_W(X)w)$$

which sits in the short exact sequence

$$0 \to U \to U \oplus W \to W \to 0.$$

In the sequel, we may call this ρ_a by simply V_a. Of course, if $a = 0$, then the short exact sequence splits, but this condition is too strong for our purposes.

To set ourselves up, recall that a morphism of extensions U and W is a diagram of the following form.

$$\begin{tikzcd} 0 \arrow[r] & U \arrow[r] \arrow[d, equal] & V_a \arrow[r] \arrow[d] & W \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & U \arrow[r] & V_b \arrow[r] & W \arrow[r] & 0 \end{tikzcd}$$

We are interested in classifying the extensions up to isomorphism. Thus far, we have found a way to list out all extensions as $V_\bullet$, but we still need to check when they are isomorphic. Well, if $f \colon V_a \to V_b$ is an isomorphism of short exact sequences, then following around the diagram means that

$$f(u, w) = (u + \varphi(w), w)$$

for some $\varphi \colon W \to U$. Note that any such morphism is automatically an isomorphism of vector spaces: its inverse is $(u, w) \mapsto (u - \varphi(w), w)$. Anyway, to check that φ is a morphism of representations, we must check that $f \circ \rho_a(X) = \rho_b(X) \circ f$ for all $X \in \mathfrak{g}$, which upon expansion yields

$$(Xu + X\varphi w + b(X)w, Xw) \stackrel{?}{=} (Xu + a(X)w + \varphi Xw, Xw),$$

so we see that we end up asking for

$$a(X) - b(X) \stackrel{?}{=} X\varphi - \varphi X = [X, \varphi].$$

Thus, we see that $V_a \cong V_b$ if and only if $a - b$ lives in the space of homomorphisms of the form $X \mapsto (X\varphi - \varphi X)$ for some $\varphi \in \operatorname{Hom}_F(W, U)$. This motivates the following definition.

Definition 4.55 ($B^1(\mathfrak{g}, E)$)**.** Fix a representation E of a Lie algebra $\mathfrak{g}$. Then the 1-*coboundary* group $B^1(\mathfrak{g}, E)$ consists of the group of morphisms $a \colon \mathfrak{g} \to E$ for which there exists a vector $v \in E$ such that

$$a(X) = X \cdot v.$$

Example 4.56. We can compute that $B^1(\mathfrak{g}, \mathfrak{g})$ consists of the maps $a \colon \mathfrak{g} \to \mathfrak{g}$ of the form $a = \operatorname{ad}_Y$ for some $Y \in \mathfrak{g}$. Thus, $B^1(\mathfrak{g}, \mathfrak{g})$ consists of the inner derivations.

Remark 4.57. One can check that $B^1(\mathfrak{g}, E) \subseteq Z^1(\mathfrak{g}, E)$.

Thus, we have the following definition.

Definition 4.58 ($H^1(\mathfrak{g}, E)$)**.** Fix a representation E of a Lie algebra $\mathfrak{g}$. Then the *first cohomology group* $H^1(\mathfrak{g}, E)$ is the quotient

$$H^1(\mathfrak{g}, E) := \frac{Z^1(\mathfrak{g}, E)}{B^1(\mathfrak{g}, E)}.$$

For convenience, we also define $\operatorname{Ext}^1_{\mathfrak{g}}(W, U) := H^1(\mathfrak{g}, \operatorname{Hom}(W, U))$.

Example 4.59. Above we showed that $\operatorname{Ext}^1_{\mathfrak{g}}(W, U)$ classifies extensions

$$0 \to U \to V \to W \to 0$$

of representations.

Example 4.60. Thus, we see that Proposition 4.50 has showed that

$$H^1(\mathfrak{g}, \mathfrak{g}) = 0.$$

Remark 4.61. There is a general procedure to define $H^n(\mathfrak{g}, E)$ for any $n \geq 0$. It can be constructed as the cohomology of the "Chevalley complex" defined by

$$0 \to E \to \operatorname{Hom}_k(\mathfrak{g}, E) \to \operatorname{Hom}_k\left(\operatorname{Alt}^2 \mathfrak{g} E\right) \to \cdots,$$

where

$$df(X_1 \wedge \cdots \wedge X_k) := \sum_{i=1}^k (-1)^{i-1} X_i f(X_1 \wedge \cdots \widehat{X_i} \wedge \cdots X_k) - \sum_{i,j=1}^k (-1)^{i-j+1} f([X_i, X_j] \wedge \cdots \wedge \wedge \widehat{X_i} \wedge \widehat{X_j}).$$

Alternatively, we can simply see this as the Ext groups for the ring $U\mathfrak{g}$. For example, one can compute that $H^0(\mathfrak{g}, E)$ consists of the $\mathfrak{g}$-fixed points of E.

4.6 October 21

Today we show that representations of semisimple Lie algebras are completely reducible.

4.6.1 Complete Reducibility

Complete reducibility basically amounts to showing that any short exact sequence

$$0 \to U \to V \to W \to 0$$

of $\mathfrak{g}$-representation splits, which last time we checked is equivalent to asking for $\operatorname{Ext}^1(W, U)$ to vanish. Thus, we are after a result for vanishing of cohomology, so we pick up some theory around cohomology.

Lemma 4.62. Fix a Lie algebra $\mathfrak{g}$. Then a short exact sequence

$$0 \to U \to V \to W \to 0$$

of representations of $\mathfrak{g}$ gives rise to a longer exact sequence

$$0 \to U^{\mathfrak{g}} \to V^{\mathfrak{g}} \to W^{\mathfrak{g}} \xrightarrow{\delta} H^1(\mathfrak{g}, U) \to H^1(\mathfrak{g}, V) \to H^1(\mathfrak{g}, W).$$

Proof. Label our maps by

$$0 \to U \overset{\alpha}{\to} V \overset{\beta}{\to} W \to 0.$$

In the sequel, we may view the embedding $\alpha \colon U \to V$ as an identification.

The exactness of

$$0 \to U^{\mathfrak{g}} \to V^{\mathfrak{g}} \to W^{\mathfrak{g}}$$

follows because this is simply a restriction of the short exact sequence. Explicitly, exactness at $V^{\mathfrak{g}}$ has a little content: for some $v \in V^{\mathfrak{g}}$ which is in the kernel of the map $V^{\mathfrak{g}} \to W^{\mathfrak{g}}$, we know that there is some $u \in U$ such that $\alpha \colon u \mapsto v$; however, this is a morphism of representations, so $X\alpha \colon u \mapsto Xv$ vanishes for all $X \in \mathfrak{g}$, so the injectivity of α requires $Xu = 0$ for all X, so $u \in U^{\mathfrak{g}}$.

Our next step is define the map δ. Well, we take some $w \in W^{\mathfrak{g}}$. We need to get all the way to U, so we begin by pulling this element back to some $v \in V$. Then we define the 1-cocycle $c_v \colon \mathfrak{g} \to U$ by

$$c_v(X) := Xv,$$

which we note will vanish when mapped to W (because $w \in W^{\mathfrak{g}}$) and hence can be identified with an element of U. Additionally, we note that this is actually a 1-cocycle because

$$c_v([X,Y]) = [X,Y]v = XYv - YXv = Xc_v(Y) - Yc_v(X).$$

Quickly, note that this map $w \mapsto c_v$ is well-defined up to cohomology class: namely, if we choose a different v' lifting w, then the difference $c_v - c_{v'}$ is a coboundary in $B^1(\mathfrak{g}, U)$. Namely, there is some $u \in U$ such that $v' = u + v$, and we can compute that

$$c_{v'}(X) = c_v(X) + Xu,$$

and the mapping $u \mapsto Xu$ is a 1-cobounary.

We now check the remaining exactness points.

- Exact at $W^{\mathfrak{g}}$: note that any $v \in V^{\mathfrak{g}}$ has $\beta(v)$ lifting to $v \in V^{\mathfrak{g}}$, which has $c_v = 0$, so $\delta(\beta(v)) = 0$. On the other hand, for any $c \in H^1(\mathfrak{g}, U)$ which vanishes in $H^1(\mathfrak{g}, V)$, we are being told that c is a 1-coboundary in $H^1(\mathfrak{g}, V)$, so there exists $v \in V$ such that $c(X) = Xv$ for all v, so $c = c_v = \delta(\beta(v))$ is in the image from $W^{\mathfrak{g}}$.
- Exact at $H^1(\mathfrak{g}, V)$ follows by restriction of the short exact sequence again. In one direction, any $c \in H^1(\mathfrak{g}, U)$ vanishes in $H^1(\mathfrak{g}, W)$ because $\alpha(\beta(c))(X)$ vanishes always. In the other direction, if $c \in H^1(\mathfrak{g}, V)$ vanishes under β, then $c(X) \in U$ for all $X \in \mathfrak{g}$, so c actually defines an element of $H^1(\mathfrak{g}, U)$. ■

We are after some vanishing result for H^1, as follows.

Theorem 4.63. Fix a semisimple Lie algebra $\mathfrak{g}$ over a field F of characteristic 0. For any representation V of $\mathfrak{g}$, we have $H^1(\mathfrak{g}, V)$.

Proof. We proceed in steps.

1. We begin by reducing to the case where V is an irreducible representation. This is by induction on $\dim V$. Indeed, suppose we have the result for irreducible representations. For any representation V, find an irreducible subrepresentation $U \subseteq V$. Then Lemma 4.62 produces the exact sequence

$$H^1(\mathfrak{g}, U) \to H^1(\mathfrak{g}, V) \to H^1(\mathfrak{g}, V/U).$$

The left term vanishes because U is irreducible, and the right term vanishes because V/U has smaller dimension than V, so we may apply an inductive hypothesis. Thus, the middle term vanishes.

2. For the rest of the proof, we will assume that V is an irreducible representation; in fact, $H^1(\mathfrak{g}, F) = 0$ by Example 4.53, so we may assume that $V \neq F$. Imitating the idea of Theorem 3.73, we are interested in $Z(U\mathfrak{g})$. Well, suppose that there is $c \in Z(U\mathfrak{g})$ such that $\rho_V(c)$ is a nonzero scalar $\lambda\,\mathrm{id}_V$ but $\rho_F(c) = 0$. Then we claim that $H^1(\mathfrak{g}, V) = \mathrm{Ext}^1_{\mathfrak{g}}(F, V)$ vanishes. Well, suppose that we have some extension

$$0 \to V \overset{\alpha}{\to} W \overset{\beta}{\to} F \to 0$$

which we would like to split. To make this split, we will eventually define a function $s\colon F \to W$ splitting β, but for this, we should look for $s(1)$. Well, by scaling and using c suitably, we can find $w \in W$ such that $\beta(w) = 1$. Now, we note that we can adjust W by some w by some $\lambda^{-1}cw$ in order to further get $cw = 0$.

We now use w to define the required splitting. Define $s\colon F \to W$ by $s(1) := w$. Notably, $Fw \subseteq W$ is a subrepresentation: for any $X \in \mathfrak{g}$, we see that $\beta(Xw) = X\beta(w) = 0$, so $Xw \in V$, but then $\lambda Xw = cXw = Xcw = 0$, so $Xw = 0$. Thus, s is the required splitting.

3. We are now on the hunt for the desired $c \in Z(U\mathfrak{g})$. Let $B_V(x, y) := \operatorname{tr} \rho_V(X)\rho_V(Y)$. Note that B_V needs to be nonzero on the image of $\mathfrak{g} \to \mathfrak{gl}(V)$: Lemma 4.43 would then imply that the image of $\mathfrak{g}$ would be solvable, which is a problem because $\mathfrak{g}$ has no nonzero solvable quotients (as these would lift to solvable ideals of $\mathfrak{g}$), meaning that V is the trivial representation.

Now, $\mathfrak{g}$ will split into $K \oplus \mathfrak{g}'$, where K is the kernel of B_V. Then we can select a basis $\{x_1, \ldots, x_n\}$ of $\mathfrak{g}'$, and then B_V provides a dual basis $\{x_1^\vee, \ldots, x_n^\vee\}$ of $\mathfrak{g}'$. We now define

$$C := \sum_{i=1}^n x_i x_i^\vee,$$

which is some element of $U\mathfrak{g}$. One can check that this does not depend on the choice of basis and lives in $Z(U\mathfrak{g})$, and in fact $C|_F = 0$. Further, we see that $\operatorname{tr} \rho_V(C) = \dim \mathfrak{g}$ because $B_V(x_i, x_i^\vee) = 1$, so $\rho_V(C) \neq 0$; further, one can check that C is a scalar when passing to the algebraic closure, so C must just be a scalar. Thus, we are done by the previous step. ■

From the theorem, we get the desired result.

Corollary 4.64. Fix a semisimple Lie algebra $\mathfrak{g}$ over an algebraically closed field F of characteristic 0. Then any representation V of $\mathfrak{g}$ is completely reducible.

Proof. Induct on $\dim V$. For $\dim V \in \{0, 1\}$, there is nothing to do. For the inductive step, find an irreducible subrepresentation $U \subseteq V$, and then we see that the short exact sequence

$$0 \to U \to V \to (V/U) \to 0$$

must split by Theorem 4.63, so complete reducibility for V follows from the inductive hypothesis applied to U and V/U. ■

We can use this result to prove the Levi decomposition for reductive groups.

Corollary 4.65 (Levi decomposition)**.** A reductive Lie algebra $\mathfrak{g}$ over an algebraically closed field F of characteristic 0 is the direct sum of an abelian and semisimple Lie algebra.

Proof. We want to show that the exact sequence

$$0 \to \mathfrak{z}(\mathfrak{g}) \to \mathfrak{g} \to \mathfrak{g}_{\mathrm{ss}} \to 0$$

splits. Well, the action of $\mathfrak{g}$ on $\mathfrak{z}(\mathfrak{g})$ actually descends to make $\mathfrak{z}(\mathfrak{g})$ into a representation of $\mathfrak{g}_{\mathrm{ss}}$. Thus, this is a short exact sequence of representations of $\mathfrak{g}_{\mathrm{ss}}$, so this splits as a sequence of representations of $\mathfrak{g}_{\mathrm{ss}}$. But then we only need to add the center back in to see that this implies the splitting of Lie algebras. ■

4.7 October 21

Today we begin discussing root decompositions.

4.7.1 Jordan Decomposition for Semisimple Lie Algebras

Fix a Lie algebra $\mathfrak{g}$ over a field F. We begin with some motivational discussion. We would like to make sense of decomposing $\mathfrak{g}$ into eigenspaces; here is one such avatar of this.

Lemma 4.66. Fix a Lie algebra $\mathfrak{g}$ over an algebraically closed field F. For any $X \in \mathfrak{g}$, let $\mathfrak{g}_\lambda$ be the generalized eigenspace for the operator $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ with eigenvalue λ. Then the decomposition

$$\mathfrak{g} = \bigoplus_{\lambda \in F} \mathfrak{g}_\lambda$$

has $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$ for any $\lambda, \mu \in F$.

Proof. Choose $Y \in \mathfrak{g}_\lambda$ and $Z \in \mathfrak{g}_\mu$ so that we want to show $(\operatorname{ad}_X - (\lambda+\mu))^\bullet [Y, Z]$ vanishes for large enough power. Well, selecting some $N > 0$, we see

$$(\operatorname{ad}_X - (\lambda + \mu))^N = \sum_{a+b+c=N} \binom{N}{a, b, c} \operatorname{ad}_X^a (-\lambda)^b (-\mu)^n,$$

which when applied to $[Y, Z]$ will rearrange into

$$\sum_{k+\ell=N} \binom{N}{k} \left[(\operatorname{ad}_X - \lambda)^k Y, (\operatorname{ad}_X - \mu)^\ell Z \right].$$

This will vanish for N large enough, so we are done. ■

In order to make this decomposition a diagonalization, we pick up the following definitions.

Definition 4.67 (semisimple, nilpotent)**.** Fix a Lie algebra $\mathfrak{g}$. Then $X \in \mathfrak{g}$ is *semisimple* if and only if $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is a semisimple operator. Further, $X \in \mathfrak{g}$ is *nilpotent* if and only if $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is nilpotent.

Remark 4.68. Note $X \in \mathfrak{g}$ is both semisimple and nilpotent if and only if ad_X vanishes, which is equivalent to $X \in \mathfrak{z}(\mathfrak{g})$.

The Jordan decomposition grants the following.

Proposition 4.69. Fix a Lie algebra $\mathfrak{g}$ over a perfect field F. Then each $X \in \mathfrak{g}$ admits a unique decomposition

$$X = X_s + X_n$$

where X_s is semisimple, X_n is nilpotent, and $[X_s, X_n] = 0$. In fact, if $Y \in \mathfrak{g}$ has $[X, Y] = 0$, then $[X_s, Y] = 0$.

Proof. Because $\mathfrak{g}$ is semisimple, we may embed $\mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g})$ via $\operatorname{ad}_\bullet$. Now, in $\mathfrak{gl}(\mathfrak{g})$, we do have a decomposition $X = X_s + X_n$ where $X_s, X_n \in \mathfrak{gl}(\mathfrak{g})$ are semisimple and nilpotent (respectively) and satisfy $X_s X_n = X_n X_s$ (so that $[X_s, X_n] = 0$), and we know that this is the only possible decomposition. Furthermore, we see that X_s, being a polynomial in X in $\mathfrak{gl}(\mathfrak{g})$, implies that $[X, Y] = 0$ forces $[X_s, Y] = 0$.

It remains to check that these elements actually live in $\mathfrak{g}$. Well, for each $\lambda \in \mathfrak{g}$, define $\mathfrak{g}_\lambda$ as in Lemma 4.66, and our construction of X_s yields $[X_s, Y] = \lambda Y$ for each $Y \in \mathfrak{g}_\lambda$. Thus, the conclusion of Lemma 4.66 allows us to check that $\operatorname{ad}_{X_s} : \mathfrak{g} \to \mathfrak{g}$ is a derivation! However, all derivations come from $\mathfrak{g}$ because $\mathfrak{g}$ is semisimple, so $X_s \in \mathfrak{g}$ after all. Thus, $X_n = X - X_s$ lives in $\mathfrak{g}$ as well, so we are done. ■

Remark 4.70. For $\mathfrak{sl}_n(F)$, one can check that the above Jordan decomposition coincides with the usual one.

Corollary 4.71. Suppose $\mathfrak{g}$ is a nonzero semisimple Lie algebra over a perfect field F. Then there exists a nonzero semisimple element X.

Proof. If not, Proposition 4.69 forces all elements of $\mathfrak{g}$ to be nilpotent, so $\mathfrak{g}$ is nilpotent, so $\mathfrak{g} = 0$ because $\mathfrak{g}$ is semisimple. ■

4.7.2 The Root Decomposition

We are now interested in repeating our eigenvalue decomposition for semisimple elements.

Definition 4.72 (toral). Fix a Lie algebra $\mathfrak{g}$. A subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is *toral* if and only if $\mathfrak{h}$ is abelian, and all elements of $\mathfrak{h}$ are semisimple.

Example 4.73. For $\mathfrak{sl}_n(F)$, we see that there is a subalgebra $\mathfrak{h}$ of the diagonal matrices.

We now decompose with respect to toral subalgebras.

Proposition 4.74. Fix a semisimple Lie algebra $\mathfrak{g}$ over an algebraically closed field F, and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a toral subalgebra. Suppose B is a non-degenerate symmetric bilinear form on $\mathfrak{g}$.

(a) For each functional $\alpha\colon \mathfrak{h} \to F$, let $\mathfrak{g}_\alpha$ be the eigenspace of $\mathfrak{g}$ with the corresponding eigenvalue α. Then

$$\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^\vee} \mathfrak{g}_\alpha.$$

(b) For each $\alpha, \beta \in \mathfrak{h}^\vee$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

(c) We have $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) \neq 0$ if and only if $\alpha + \beta = 0$.

(d) The bilinear form B is non-degenerate when restricted to $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}$ for any $\alpha \in \mathfrak{h}^\vee$.

Proof. Note (a) and (b) are immediate from Lemma 4.66, where we use the fact that $\mathfrak{h}$ has semisimple elements in order to diagonalize, and we use the fact that $\mathfrak{h}$ is abelian to simultaneously diagonalize. For (c), we choose $Y \in \mathfrak{g}_\alpha$ and $Z \in \mathfrak{g}_\beta$ and $X \in \mathfrak{h}$, and we see

$$\alpha(X)B(Y,Z) = B(\operatorname{ad}_X Y, Z) = B(Y, \operatorname{ad}_X Z) = \beta(X)B(Y,Z)$$

by the invariance of the bilinear form. Thus, $(\alpha+\beta)(X) = 0$ for all $X \in \mathfrak{h}$, or $B(Y,Z) = 0$ for all $Y \in \mathfrak{g}_\alpha$ and $Z \in \mathfrak{g}_\beta$, which proves (c). Now, (d) follows because B must be non-degenerate, and (c) tells us that B vanishes except on the given subspaces. ■

Remark 4.75. Note that $\mathfrak{g}_0$ is the commutator of $\mathfrak{h}$ by its definition.

Remark 4.76. In fact, we claim that $\mathfrak{g}_0 \subseteq \mathfrak{g}$ is reductive because the invariant bilinear form B restricts to be non-degenerate on $\mathfrak{g}$.

Remark 4.77. Furthermore, we see that $X \in \mathfrak{g}_0$ will imply that $X_s \in \mathfrak{g}_0$ because X commutes with $\mathfrak{h}$ implies that X_s commutes with $\mathfrak{g}$; thus, we are also able to say that $X_n \in \mathfrak{h}$.

In order to profit the most from our toral subalgebra, we would like for it to be large.

Definition 4.78 (Cartan subalgebra)**.** Fix a Lie algebra $\mathfrak{g}$. Then a toral subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is *Cartan* if and only if $\mathfrak{h}$ equals its own centralizer.

Example 4.79. One can check by hand that the diagonal torus of $\mathfrak{sl}_n(F)$ is Cartan.

Let's check that such things exist.

Lemma 4.80. Fix a semisimple Lie algebra $\mathfrak{g}$ over a perfect field F of characteristic 0. Then a toral subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is a Cartan subalgebra if and only if $\mathfrak{h}$ is a maximal toral subalgebra.

Proof. We show our implications separately. Let $C(\mathfrak{h})$ be the centralizer of $\mathfrak{h}$.

- If $\mathfrak{h}$ is Cartan, then we note that $\mathfrak{h}$ is a maximal torus. Indeed, if we have a toral subalgebra $\mathfrak{h}'$ such that $\mathfrak{h} \subseteq \mathfrak{h}'$, then $\mathfrak{h}'$ is contained in the centralizer of $\mathfrak{h}$, so $\mathfrak{h}' \subseteq \mathfrak{h}$ because $\mathfrak{h}$ is Cartan.
- Suppose that $\mathfrak{h}$ is a maximal toral subalgebra, and choose some X commuting with $\mathfrak{h}$, and we want to show that $X \in \mathfrak{h}$. Well, Proposition 4.69 implies X_s also commutes with $\mathfrak{h}$, so $\mathfrak{h} + FX_s$ is an abelian algebra consisting of semisimple elements (these elements commute, so sums of them will be simultaneously diagonalizable, so this algebra still has semisimple elements), so $X_s \in \mathfrak{h}$ by the maximality of $\mathfrak{h}$.

  Now, $C(\mathfrak{h})$ is reductive by an above remark, and we see that all $X \in C(\mathfrak{h})$ will have ad_{X_s} vanish on $C(\mathfrak{h})$, so ad_X must be nilpotent on $C(\mathfrak{h})$. Thus, $C(\mathfrak{h})$ is nilpotent as well by Engel's theorem, so $C(\mathfrak{h})$ must be abelian.

  We would like to show that $C(\mathfrak{h})$ is furthermore consisting of semisimple elements, which will complete the proof by the ambient maximality. Well, for any nilpotent $X \in C(\mathfrak{h})$, we see that any $Y \in C(\mathfrak{h})$ has $[Y, X] = 0$ by the commutativity, so the composite $\operatorname{ad}_X \circ \operatorname{ad}_Y$ is nilpotent as an operator $\mathfrak{g} \to \mathfrak{g}$ (because X is nilpotent, and these operators commute), so

  $$\operatorname{tr}_{\mathfrak{g}}(\operatorname{ad}_X \circ \operatorname{ad}_Y) = 0,$$

  so X is in the kernel of the Killing form of $C(\mathfrak{h})$, so $X = 0$ because the Killing form is non-degenerate. ■

We are now ready to state our root decompositions.

Corollary 4.81 (root decomposition)**.** Fix a semisimple Lie algebra $\mathfrak{g}$ over an algebraically closed field F of characteristic 0 and choose a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. Let Φ be the collection of nonzero $\alpha \in \mathfrak{h}^\vee$ such that $\mathfrak{g}_\alpha \neq 0$. Then

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha.$$

Proof. Immediate from Proposition 4.74. ■

Definition 4.82 (root system)**.** Fix a semisimple Lie algebra $\mathfrak{g}$ over an algebraically closed field F of characteristic 0 and choose a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. Then the collection Φ of nonzero $\alpha \in \mathfrak{h}^\vee$ such that $\mathfrak{g}_\alpha \neq 0$ is called the *root system*.

Example 4.83. For $\mathfrak{sl}_n(F)$, we choose $\mathfrak{h}$ to be the diagonal subalgebra. Let $e_i \in \mathfrak{h}^\vee$ be the projection onto the (i,i) coordinate. Then we compute that $\operatorname{ad}_X(E_{ij}) = (X_i - X_j)E_{ij}$ for any $X \in \mathfrak{h}$, so our root system consists of the functionals $e_i - e_j$ for each $i \neq j$.

Here is a result on decompositions.

Proposition 4.84. Fix a semisimple Lie algebra $\mathfrak{g}$ over an algebraically closed field F of characteristic 0. Write $\mathfrak{g}$ as a sum $\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$ of simple algebras.

(a) For each i, choose a Cartan subalgebra $\mathfrak{h}_i \subseteq \mathfrak{g}_i$, producing a root system Φ_i. Then $\mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$ is a Cartan subalgebra $\mathfrak{g}$, and the corresponding root system Φ is disjoint union $\Phi_1 \sqcup \cdots \sqcup \Phi_n$.

(b) Any Cartan subalgebra of $\mathfrak{g}$ is a direct sum of Cartan subalgebras of the $\mathfrak{g}_\bullet$s.

Proof. Omitted. This is a matter of running around with the direct sums everywhere. ■

BIBLIOGRAPHY

[Hat01] Allen Hatcher. *Algebraic Topology*. Cambridge, 2001.

[Kir08] Alexander Kirillov Jr. *An Introduction to Lie Groups and Lie Algebras*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[Lee13] John M. Lee. *Introduction to smooth manifolds*. Second. Vol. 218. Graduate Texts in Mathematics. Springer, New York, 2013, pp. xvi+708. ISBN: 978-1-4419-9981-8.

[Shu16] Neal Shusterman. *Scythe*. Arc of a Scythe. Simon & Schuster, 2016.

[Vak17] Ravi Vakil. *The Rising Sea: Foundations of Algebraic Geometry*. 2017. URL: http://math.stanford.edu/~vakil/216blog/FOAGnov1817public.pdf.

[Elb22] Nir Elber. *Introduction to Topology and Analysis*. 2022. URL: https://dfoiler.github.io/notes/202A/notes.pdf.

[Elb23] Nir Elber. *Algebraic Topology*. 2023. URL: https://dfoiler.github.io/notes/215A/notes.pdf.

[Elb24] Nir Elber. *Differential Topology*. 2024. URL: https://dfoiler.github.io/notes/214/notes.pdf.

LIST OF DEFINITIONS